Supporting information for:

Halogen Bonding Increases the Potency and Isozyme-selectivity of Protein Arginine Deiminase 1 Inhibitors

Santanu Mondal,[a] Xuefeng Gong,[b] Xiaoqian Zhang,[b] Ari J. Salinger,[a][c] Li Zheng,[a] Sudeshna Sen,[a] Eranthie Weerapana,[c] Xuesen Zhang,[b] and Paul R. Thompson*[a]

[a]Department of Biochemistry and Molecular Pharmacology, UMass Medical School, 364 Plantation Street, Worcester, MA 01605, USA

[b]State Key Laboratory of Reproductive Medicine, Nanjing Medical University, Nanjing 211166, China

[c]Department of Chemistry, Boston College, Chestnut Hill, MA 02467, USA

*Author to whom correspondence should be addressed: Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, LRB 826, 364 Plantation Street,

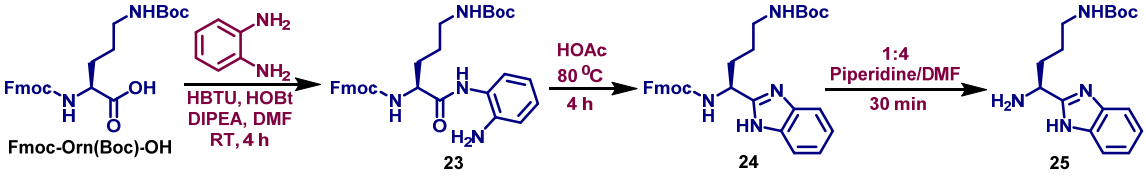
Worcester MA 01605 tel: 508-856-8492; fax: 508-856-6215; e-mail: paul.thompson@umassmed.edu.

Materials and Methods:

Fmoc-Orn(Boc)-OH, HBTU, HOBt and N-iodosuccinimide were bought from Chem-Impex International, Inc. Fluoroacetonitrile was purchased from Reagent World, Inc. 3-iodosalicylic acid and 4-fluoro-2-hydroxy-3,5-diiodobenzoic acid were obtained from Enamine and Aurum Pharmatech, respectively. Methyl 4-bromo-2-hydroxybenzoate was bought from Ark Pharm. Salicylic acid, 3,5-diiodosalicylic acid, 3,5-dibromosalicylic acid, 3,5-dichlorosalicylic acid, 3,5-isopropylsalicylic acid, 5-iodosalicylic acid, 3,5-diiodobenzoic acid, DIPEA, anhydrous methanol, anhydrous DMF, anhydrous dichloromethane, piperidine, triethylamine, trifluoroacetic acid, chloroacetonitrile and HPLC-grade acetonitrile were purchased from Sigma-Aldrich. TCEP, streptavidin agarose beads (catalogue no. 20353) and Halt protease inhibitor cocktail (EDTA-free) were obtained from Thermo Scientific. Precoated silica gel plates were bought from Merck. Deuterated solvents were purchased from Cambridge Isotope Laboratories. Plasmid purification kit was bought from Bio Basic Canada Inc. Rabbit polyclonal anti-PAD1 (catalogue no. ab181762), mouse monoclonal anti-histone H3 (catalogue no. ab10799) and rabbit polyclonal anti-H3 (Citrulline R2, R8 and R17) (catalogue no. ab5103) were obtained from Abcam. ^{1}H and ^{13}C NMR spectra were recorded in d_4-MeOH or d_6-DMSO as solvent using a Bruker 500 MHz NMR spectrometer. Chemical shift values are cited with respect to $SiMe_4$ (TMS) as the internal standard. Column chromatography was performed in glass columns. All the inhibitors and probes (compounds **1**-**22**) were purified by reverse-phase HPLC using a semi-preparative C18 column (Agilent, 21.2 × 250 mm, 10 μm) and a water/acetonitrile gradient supplemented with 0.05% trifluoroacetic acid. Fluorographs were recorded using a Typhoon scanner with excitation/emission maxima of ~546/579, respectively. PADs 1, 2, 3 and 4 were expressed and purified as reported earlier.[1]

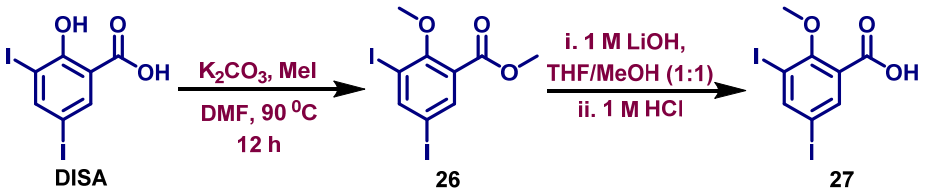
Synthesis.

Synthesis of 25.



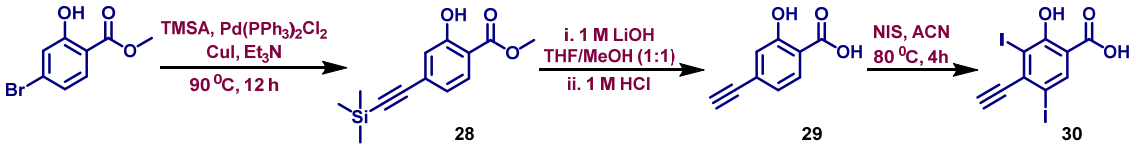
Compound **25** was synthesized as reported earlier with minor modifications.[2] Briefly, diisopropyl ethylamine (DIPEA) (1.2 mL, 6.6 mmol), HBTU (1.3 g, 3.3 mmol) and HOBt (297 mg, 2.2 mmol) were added sequentially to a solution of Fmoc-Orn(Boc)-OH (1 g, 2.2 mmol) and 1,2-phenylenediamine (238 mg, 2.2 mmol) in anhydrous dimethylformamide (DMF) and the mixture was stirred for 4 h at 25 ^{0}C under nitrogen atmosphere. Then the reaction mixture was poured into water to precipitate compound **23** and it was recovered by vacuum filtration, washed with water and dried *in vacuo*. Crude **23** was then dissolved in glacial acetic acid (20 mL) and was refluxed for 4 h. Then the mixture was evaporated *in vacuo* to afford a gummy brown liquid which was poured into brine, neutralized with sodium bicarbonate and extracted with excess dichloromethane. The organic extract was washed thoroughly with water, brine, dried over anhydrous sodium sulphate and concentrated *in vacuo* to afford compound **24**. Compound **24** was treated with 1:4 piperidine/DMF (v/v) for 30 min to remove the Fmoc-group and the mixture was vigorously stirred with excess hexane. The hexane layer was decanted off and this procedure was repeated for several times until most of the DMF was removed. Fmoc-removal afforded compound **25** as gummy brown oil which was used in subsequent steps without further purification.

Synthesis of 27.



Methyl iodide (0.18 mL, 2.9 mmol) was added dropwise to a suspension of 3,5-diiodosalicylic acid (DISA) (0.5 g, 1.3 mmol) and potassium carbonate (0.45 g, 3.3 mmol) in anhydrous DMF. The mixture was heated in a sealed tube at 90 ^{0}C for 12 h followed by pouring into water to precipitate compound **26**, which was recovered by vacuum filtration, washed with water and dried *in vacuo*. Crude compound **26** (0.22 g, 0.53 mmol) was then dissolved in 1:1 THF/MeOH (10 mL) and was treated with 1 M aqueous lithium hydroxide solution (5.3 mL). The mixture was allowed to stir for 2 h at 25 ^{0}C and the organic solvents were removed under reduced pressure. The mixture was then acidified with 2 M aqueous hydrochloric acid and was extracted twice with excess diethyl ether. The combined ether extracts were washed with water, dried over anhydrous sodium sulphate and concentrated *in vacuo* to afford **27** as an off-white solid. Crude **27** was purified by column chromatography using hexane/ethyl acetate as the eluent and was characterized by ^{1}H, ^{13}C NMR spectroscopy and mass spectrometry. Overall yield: 60%. ^{1}H NMR (DMSO-d_6) δ (ppm): 13.4 (s, 1H), 8.3 (s, 1H), 7.95 (s, 1H), 3.76 (s, 3H); ^{13}C NMR (DMSO-d_6) δ (ppm): 165.6, 158.6, 149.7, 139.7, 129.1, 97.0, 89.6, 62.3; ESI-MS (m/z) calculated for $C_8H_6I_2O_3$ $[M + H]^+$: 404.85, found 404.8.

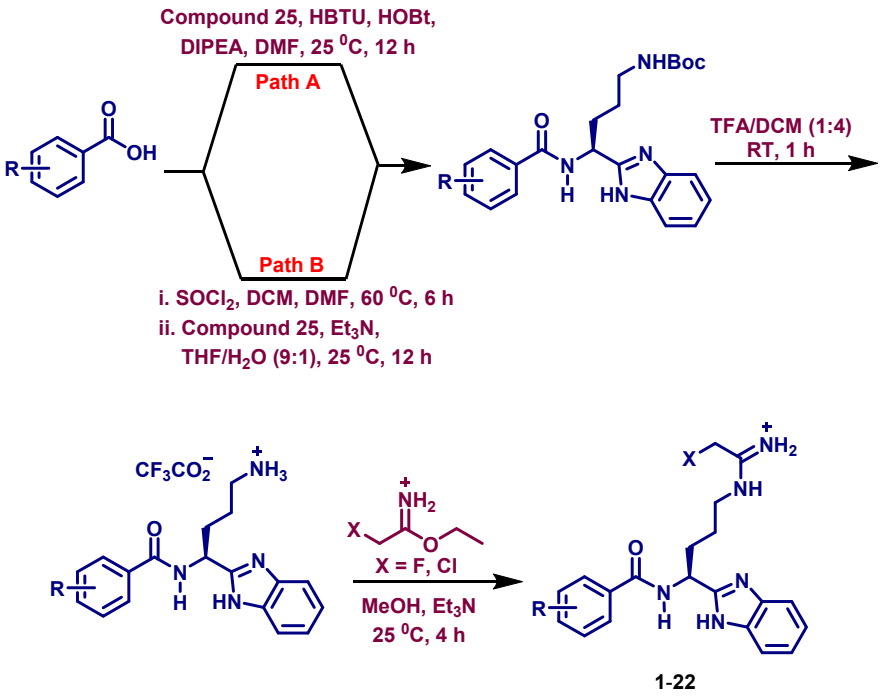
Synthesis of 30.



Deoxygenated triethylamine (20 mL) and trimethylsilyl acetylene (TMSA) (2.4 mL, 17.3 mmol) was added to methyl 4-bromo-2-hydroxy benzoate (1g, 4.3 mmol), $Pd(PPh_3)_2Cl_2$ (304 mg, 0.4 mmol) and copper (I) iodide (83 mg, 0.4 mmol), and the mixture was refluxed for 12 h under nitrogen atmosphere. Excess triethylamine was evaporated under reduced pressure and residue was resuspended in water. Then the dark brown mixture was extracted several times with excess dichloromethane and the combined dichloromethane extracts were filtered under vacuum, washed with water and brine, dried over anhydrous sodium sulphate and concentrated *in vacuo* to yield a dark brown gummy liquid. **28** was purified from this crude mixture by column chromatography using hexane/ethyl acetate as eluent and was characterized by ^{1}H NMR spectroscopy and mass spectrometry. Yield: 70%. ^{1}H NMR ($CDCl_3$) δ (ppm): 10.53 (s, 1H), 7.57 (d, J = 10 Hz, 1H), 6.88 (d, J = 1.5 Hz, 1H), 6.77 (dd, J = 8.2 Hz, 1H), 3.76 (s, 3H), 0.07 (s, 9H); ESI-MS (m/z) calculated for $C_{13}H_{16}O_3Si$ $[M + H]^+$: 249.09, found 249.2. Compound **28** (0.5 g, 2.0 mmol) was dissolved in 1:1 THF/MeOH (10 mL) and was treated with a 1 M aqueous solution of lithium hydroxide (20 mL). The mixture was stirred at room temperature for 2 h and the organic solvents were evaporated under reduced pressure followed by acidification with 1 M hydrochloric acid. The mixture was then extracted twice with excess diethyl ether and the combined organic extract was washed with water, dried over anhydrous sodium sulphate and concentrated *in vacuo*. Compound **29** was purified from the crude mixture by column chromatography using hexane/ethyl acetate as mobile phase and was characterized by ^{1}H NMR spectroscopy and mass spectrometry. Yield: 90%. ^{1}H NMR ($CDCl_3$) δ (ppm): 7.77 (d, J = 8.2 Hz, 1H), 7.04 (d, J = 1.2 Hz, 1H), 6.93 (dd,

J = 8.2 Hz, 1H), 3.16 (s, 1H); ESI-MS (m/z) calculated for $C_9H_6O_3$ $[M + H]^+$: 163.04, found 163.2. N-iodosuccinimide (417 mg, 1.9 mmol) was added to a solution of **29** (100 mg, 0.6 mmol) in acetonitrile (ACN) (15 mL) and the mixture was heated at 80 ^{0}C in a sealed tube for 4 h. Excess acetonitrile was then evaporated under reduced pressure and the residue was resuspended in dichloromethane, washed with water, dried over anhydrous sodium sulphate and concentrated *in vacuo* to yield **30** as off-white solid. **30** was purified from the crude mixture by column chromatography using hexane/ethyl acetate as the eluent and was characterized by ^{1}H NMR spectroscopy and mass spectrometry. Yield: 78%. ^{1}H NMR (DMSO-d_6) δ (ppm): 8.07 (s, 1H), 4.87 (s, 1H); ESI-MS (m/z) calculated for $C_9H_4I_2O_3$ $[M + H]^+$: 414.83, found 414.8.

Synthesis of 1-22.



Compounds **1**-**22** were synthesized following three steps. In the first step, compound **25** was coupled with commercially available carboxylic acids or compound **27** (for compounds **3** and **4**) or compound **30** (for compounds **21** and **22**) by following either of the two pathways (Path

A or Path B) to afford amides with Boc-protected ornithine side chain. In the second step, the Boc-protecting group was removed from the ornithine side chain to produce the free amine. The haloacetamidine warhead was then installed in the third step using an ethyl haloacetimidate hydrochloride. General procedures for these three steps are given below.

General Procedure for Path A.

DIPEA (3 eq), HBTU (2 eq) and HOBt (2 eq) were added sequentially to a solution of carboxylic acid (1 eq) and compound **25** (1 eq) in anhydrous DMF. The mixture was allowed to stir at room temperature for 12 h under a nitrogen atmosphere. Then the reaction mixture was poured into water to precipitate the desired amide, which was recovered by vacuum filtration, washed with water and dried *in vacuo*. The amide was used in the subsequent step without further purification.

General Procedure for Path B.

Thionyl chloride (3 eq) was added dropwise to a suspension of carboxylic acid (1 eq) in anhydrous dichloromethane with vigorous stirring. Then a catalytic amount of DMF (100 μL for 1 g of carboxylic acid) was added and the reaction mixture was refluxed for 6 h. The organic solvent and excess thionyl chloride were evaporated under reduced pressure to afford the acyl chloride. Triethylamine (5 eq) was added to a suspension of compound **25** (1 eq) and acyl chloride (2 eq) in THF, and the resulting precipitate was dissolved by dropwise addition of water. The mixture was allowed to stir at room temperature for 12 h. THF was evaporated under reduced pressure and the residue was extracted twice with dichloromethane. The combined organic extract was washed with water, dried over anhydrous sodium sulphate and concentrated *in vacuo* to afford the amide, which was used in the subsequent step without further purification.

General Procedure for Boc-deprotection and warhead installation.

A Boc-protected compound was dissolved in 1:4 trifluoroacetic acid/dichloromethane (v/v) (10 mL for 1 g of Boc-protected compound) and the mixture was stirred at room temperature for 1 h. Excess trifluoroacetic acid/dichloromethane was evaporated under reduced pressure to afford the free amine as a gummy liquid. Triethylamine (4 eq) and ethyl haloacetimidate hydrochloride (2 eq) were added sequentially to a solution of free amine (1 eq) in anhydrous methanol. The mixture was allowed to stir at room temperature for 4 h. Methanol and excess triethylamine were evaporated under reduced pressure. Compounds **1**-**22** were then purified as trifluoroacetate salts by reverse phase HPLC using a pre-packed C18 column and a water/acetonitrile (supplemented with 0.05% TFA) gradient as the eluent.

Compounds **1**-**22** were thoroughly characterized by ^{1}H and ^{13}C NMR spectroscopy and mass spectrometry. These data are given in Figures S14-S57. The purity of compounds **1**-**22** was determined by ^{1}H NMR spectroscopy and LC-MS analysis. All the tested compounds were ≥95% pure.

Compound 1. This compound was synthesized from commercially available 3,5-diiodosalicylic acid following Path B. Yield: 32% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.18 (d, *J* = 2 Hz, 1H), 8.13 (d, *J* = 2 Hz, 1H), 7.62-7.64 (m, 2H), 7.41-7.43 (m, 2H), 5.55-5.58 (m, 1H), 5.21 (s, 1H), 5.12 (s, 1H), 3.37 (t, *J* = 7.3 Hz, 2H), 2.25-2.32 (m, 1H), 2.16-2.24 (m, 1H), 1.80-1.88 (m, 1H), 1.71-1.79 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.3, 163.1, 163.0, 161.2, 160.9, 160.1, 153.2, 151.0, 135.8, 133.0, 125.3, 115.7, 113.9, 86.9, 79.6, 78.3, 76.9, 41.4, 28.9, 23.7; ESI-MS (m/z) calculated for $C_{20}H_{20}N_5O_2F_1I_2$ $[M + H]^+$: 635.97, found 636.0.

Compound 2. This compound was synthesized from commercially available 3,5-diiodosalicylic acid following Path B. Yield: 26% over three steps. ^{1}H NMR (CD_3OD) δ

(ppm): 8.16 (d, J = 2 Hz, 1H), 8.13 (d, J = 2 Hz, 1H), 7.59-7.62 (m, 2H), 7.37-7.40 (m, 2H), 5.53-5.56 (m, 1H), 4.27 (s, 2H), 3.33-3.36 (m, 2H), 2.25-2.32 (m, 1H), 2.14-2.22 (m, 1H), 1.80-1.86 (m, 1H), 1.71-1.78 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.4, 163.4, 161.6, 161.3, 160.1, 153.1, 151.0, 135.8, 132.7, 125.5, 115.6, 113.9, 86.9, 79.6, 53.4, 41.9, 38.7, 28.9, 23.6; ESI-MS (m/z) calculated for $C_{20}H_{20}N_5O_2Cl_1I_2$ $[M + H]^+$: 651.95, found 652.0.

Compound 3. This compound was synthesized from compound **27** following Path B. Yield: 35% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.19 (d, J = 2.2 Hz, 1H), 7.89 (d, J = 2.2 Hz, 1H), 7.65-7.67 (m, 2H), 7.43-7.45 (m, 2H), 5.47 (t, J = 6.8 Hz, 1H), 5.22 (s, 1H), 5.13 (s, 1H), 3.67 (s, 3H), 3.36-3.4 (m, 2H), 2.17-2.22 (m, 2H), 1.83-1.93 (m, 1H), 1.70-1.78 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 165.7, 163.2, 163.0, 161.5, 161.3, 157.5, 153.2, 149.8, 138.7, 132.7, 130.4, 125.5, 113.9, 93.4, 87.8, 78.3, 76.9, 61.7, 41.3, 29.3, 23.7; ESI-MS (m/z) calculated for $C_{21}H_{22}N_5O_2F_1I_2$ $[M + H]^+$: 649.99, found 649.8.

Compound 4. This compound was synthesized from compound **27** following Path B. Yield: 29% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.20 (d, J = 2.2 Hz, 1H), 7.90 (d, J = 2.2 Hz, 1H), 7.63-7.66 (m, 2H), 7.40-7.44 (m, 2H), 5.44-5.48 (m, 1H), 4.28 (s, 2H), 3.68 (s, 3H), 3.34-3.37 (m, 2H), 2.17-2.22 (m, 2H), 1.83-1.91 (m, 1H), 1.70-1.78 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 165.7, 163.4, 161.5, 161.2, 157.5, 153.2, 149.8, 138.7, 133.0, 130.4, 125.3, 114.0, 93.4, 87.8, 61.7, 41.9, 38.7, 29.3, 23.6; ESI-MS (m/z) calculated for $C_{21}H_{22}N_5O_2Cl_1I_2$ $[M + H]^+$: 665.96, found 665.8.

Compound 5. This compound was synthesized from commercially available 3,5-diiodobenzoic acid following Path A. Yield: 37% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.19-8.21 (m, 3H), 7.65-7.67 (m, 2H), 7.45-7.47 (m, 2H), 5.48-5.51 (m, 1H), 5.21 (s, 1H), 5.12 (s, 1H), 3.35-3.39 (m, 2H), 2.19-2.25 (m, 2H), 1.81-1.89 (m, 1H), 1.70-1.77 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 165.8, 163.1, 163.0, 161.9, 161.6, 161.4, 161.1, 153.5,

148.4, 136.1, 135.8, 131.7, 125.9, 113.8, 94.0, 78.3, 76.9, 41.3, 29.1, 23.7; ESI-MS (m/z) calculated for $C_{20}H_{20}N_5O_1F_1I_2$ $[M + H]^+$: 619.98, found 620.0.

Compound 6. This compound was synthesized from commercially available 3,5-diiodobenzoic acid following Path A. Yield: 32% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.19-8.21 (m, 3H), 7.65-7.67 (m, 2H), 7.45-7.47 (m, 2H), 5.48-5.51 (m, 1H), 4.28 (s, 2H), 3.33-3.36 (m, 2H), 2.17-2.27 (m, 2H), 1.82-1.88 (m, 1H), 1.70-1.77 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 165.8, 163.4, 161.8, 161.6, 161.3, 161.0, 153.5, 148.4, 136.1, 135.8, 131.8, 125.9, 113.8, 94.1, 41.9, 38.7, 29.0, 23.6; ESI-MS (m/z) calculated for $C_{20}H_{20}N_5O_1Cl_1I_2$ $[M + H]^+$: 635.95, found 636.0.

Compound 7. This compound was synthesized from commercially available salicylic acid following Path A. Yield: 33% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 7.81 (dd, J = 7.9 Hz, 1H), 7.64-7.65 (m, 2H), 7.44-7.45 (m, 2H), 7.33-7.36 (m, 1H), 6.84-6.87 (m, 2H), 5.54-5.57 (m, 1H), 5.21 (s, 1H), 5.11 (s, 1H), 3.35-3.41 (m, 2H), 2.23-2.28 (m, 2H), 1.83-1.93 (m, 1H), 1.72-1.80 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.4, 163.1, 163.0, 159.3, 153.9, 134.0, 130.5, 128.6, 127.6, 124.7, 119.1, 117.0, 115.5, 114.0, 78.3, 76.8, 41.4, 29.4, 23.7; ESI-MS (m/z) calculated for $C_{20}H_{22}N_5O_2F_1$ $[M + H]^+$: 384.18, found 384.4.

Compound 8. This compound was synthesized from commercially available salicylic acid following Path A. Yield: 29% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 7.81 (dd, J = 8 Hz, 1H), 7.62-7.64 (m, 2H), 7.42-7.43 (m, 2H), 7.32-7.36 (m, 1H), 6.83-6.87 (m, 2H), 5.53-5.56 (m, 1H), 4.27 (s, 2H), 3.34-3.38 (m, 2H), 2.23-2.29 (m, 2H), 1.82-1.90 (m, 1H), 1.72-1.80 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.4, 163.4, 159.3, 153.9, 134.2, 134.0, 130.5, 128.5, 127.6, 124.5, 119.1, 117.0, 115.5, 114.0, 42.0, 38.7, 37.5, 29.5, 23.6; ESI-MS (m/z) calculated for $C_{20}H_{22}N_5O_2Cl_1$ $[M + H]^+$: 400.15, found 400.4.

Compound 9. This compound was synthesized from commercially available 3,5-diisopropylsalicylic acid following Path B. Yield: 35% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 7.66-7.69 (m, 2H), 7.51 (d, *J* = 2.1 Hz, 1H), 7.47-7.52 (m, 2H), 7.18 (d, *J* = 2.1 Hz, 1H), 5.59-5.62 (m, 1H), 5.21 (s, 1H), 5.12 (s, 1H), 3.37-3.40 (m, 2H), 2.76-2.81 (m, 2H), 2.25-2.31 (m, 2H), 1.87-1.92 (m, 1H), 1.73-1.80 (m, 1H), 1.18-1.24 (m, 1H), 1.15 (d, *J* = 7 Hz, 6H), 1.10-1.12 (m, 6H); ^{13}C NMR (CD_3OD) δ (ppm): 172.0, 163.5, 163.1, 163.0, 161.6, 161.3, 161.1, 160.8, 157.1, 153.7, 138.8, 136.8, 131.5, 129.4, 126.0, 121.6, 112.3, 78.3, 76.8, 41.4, 33.7, 28.7, 26.5, 23.8, 23.2, 23.1, 21.4; ESI-MS (m/z) calculated for $C_{26}H_{34}N_5O_2F_1$ $[M + H]^+$: 468.28, found 468.4.

Compound 10. This compound was synthesized from commercially available 3,5-diisopropylsalicylic acid following Path B. Yield: 30% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 7.65-7.67 (m, 2H), 7.50 (d, *J* = 2.1 Hz, 1H), 7.45-7.47 (m, 2H), 7.19 (d, *J* = 2.1 Hz, 1H), 5.60 (t, J = 7.5 Hz, 1H), 4.27 (s, 2H), 3.34-3.38 (m, 2H), 2.75-2.81 (m, 1H), 2.26-2.31 (m, 2H), 1.84-1.93 (m, 1H), 1.73-1.80 (m, 1H), 1.19-1.23 (m, 1H), 1.15 (d, *J* = 6.9 Hz, 6H), 1.10-1.12 (m, 6H); ^{13}C NMR (CD_3OD) δ (ppm): 171.9, 163.4, 161.8, 161.5, 161.2, 160.9, 157.1, 153.7, 138.8, 136.8, 132.0, 129.3, 125.8, 121.6, 113.8, 112.4, 46.8, 41.9, 38.7, 33.7, 28.8, 26.5, 23.7, 23.2, 23.1, 21.4; ESI-MS (m/z) calculated for $C_{26}H_{34}N_5O_2Cl_1$ $[M + H]^+$: 484.25, found 484.4.

Compound 11. This compound was synthesized from commercially available 3,5-dibromosalicylic acid following Path B. Yield: 34% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.04 (d, *J* = 2.3 Hz, 2H), 7.79 (d, *J* = 2.3 Hz, 1H), 7.64-7.66 (m, 2H), 7.43-7.45 (m, 2H), 5.56-5.59 (m, 1H), 5.21 (s, 1H), 5.12 (s, 1H), 3.37 (t, *J* = 7.3 Hz, 2H), 2.19-2.30 (m, 2H), 1.82-1.89 (m, 1H), 1.73-1.79 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.4, 163.1, 163.0, 162.0, 161.7, 161.4, 161.1, 157.1, 153.1, 139.4, 132.3, 129.3, 125.7, 116.3, 113.8, 112.3,

110.0, 78.3, 76.8, 41.3, 28.9, 23.7; ESI-MS (m/z) calculated for $C_{20}H_{20}N_5O_2F_1Br_2$ $[M + H]^+$: 540.00, found 541.8.

Compound 12. This compound was synthesized from commercially available 3,5-dibromosalicylic acid following Path B. Yield: 30% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.04 (d, *J* = 2.3 Hz, 2H), 7.80 (d, *J* = 2.3 Hz, 1H), 7.64-7.66 (m, 2H), 7.43-7.45 (m, 2H), 5.56-5.59 (m, 1H), 4.27 (s, 2H), 3.33-3.36 (m, 2H), 2.18-2.32 (m, 2H), 1.82-1.89 (m, 1H), 1.72-1.79 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.3, 163.4, 161.8, 161.6, 161.3, 161.0, 157.1, 153.1, 139.4, 132.5, 129.3, 125.6, 116.3, 113.9, 112.4, 110.0, 41.9, 38.7, 28.9, 23.6; ESI-MS (m/z) calculated for $C_{20}H_{20}N_5O_2Cl_1Br_2$ $[M + H]^+$: 555.97, found 557.8.

Compound 13. This compound was synthesized from commercially available 3,5-dichlorosalicylic acid following Path B. Yield: 36% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 7.87 (d, *J* = 2.5 Hz, 2H), 7.65-7.67 (m, 2H), 7.53 (d, *J* = 2.5 Hz, 1H), 7.45-7.46 (m, 2H), 5.56-5.59 (m, 1H), 5.21 (s, 1H), 5.12 (s, 1H), 3.38 (t, *J* = 7.2 Hz, 2H), 2.21-2.31 (m, 2H), 1.84-1.90 (m, 1H), 1.72-1.79 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.2, 163.1, 163.0, 161.5, 161.2, 155.5, 153.1, 133.7, 132.1, 125.9, 125.8, 123.3, 123.2, 116.3, 113.8, 78.3, 76.8, 41.3, 28.9, 23.7; ESI-MS (m/z) calculated for $C_{20}H_{20}N_5O_2F_1Cl_2$ $[M + H]^+$: 452.11, found 452.0.

Compound 14. This compound was synthesized from commercially available 3,5-dichlorosalicylic acid following Path B. Yield: 31% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 7.86 (d, *J* = 2.5 Hz, 2H), 7.64-7.66 (m, 2H), 7.53 (d, *J* = 2.5 Hz, 1H), 7.43-7.45 (m, 2H), 5.56-5.59 (m, 1H), 4.27 (s, 2H), 3.34-3.37 (m, 2H), 2.21-2.31 (m, 2H), 1.84-1.88 (m, 1H), 1.73-1.78 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.2, 163.4, 161.9, 161.6, 161.3, 155.5, 153.1, 133.6, 132.4, 125.9, 125.6, 123.3, 123.2, 116.4, 113.9, 41.9, 38.7, 29.0, 23.6; ESI-MS (m/z) calculated for $C_{20}H_{20}N_5O_2Cl_3$ $[M + H]^+$: 468.08, found 468.0.

Compound 15. This compound was synthesized from commercially available 5-iodosalicylic acid following Path B. Yield: 32% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.14 (d, *J* = 2.2 Hz, 1H), 7.65-7.66 (m, 2H), 7.61 (dd, *J* = 8.7 Hz, 1H), 7.45-7.47 (m, 2H), 6.70 (d, *J* = 8.7 Hz, 1H), 5.54-5.57 (m, 1H), 5.21 (s, 1H), 5.12 (s, 1H), 3.35-3.40 (m, 2H), 2.20-2.29 (m, 2H), 1.82-1.89 (m, 1H), 1.71-1.79 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 168.1, 163.1, 163.0, 161.5, 161.2, 159.0, 153.6, 142.6, 137.3, 132.1, 125.7, 119.4, 117.8, 113.8, 79.7, 78.3, 76.8, 46.5, 41.3, 29.0, 23.7; ESI-MS (m/z) calculated for $C_{20}H_{21}N_5O_2F_1I_1$ $[M + H]^+$: 510.08, found 510.2.

Compound 16. This compound was synthesized from commercially available 5-iodosalicylic acid following Path B. Yield: 26% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.14 (d, *J* = 2.2 Hz, 1H), 7.60-7.63 (m, 3H), 7.41-7.42 (m, 2H), 6.69 (d, *J* = 8.7 Hz, 1H), 5.52-5.55 (m, 1H), 4.27 (s, 2H), 3.33-3.37 (m, 2H), 2.19-2.29 (m, 2H), 1.80-1.88 (m, 1H), 1.71-1.78 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 168.0, 163.4, 161.2, 158.9, 153.6, 142.5, 137.3, 132.9, 125.3, 119.4, 117.9, 113.9, 79.7, 41.9, 38.7, 29.2, 23.6; ESI-MS (m/z) calculated for $C_{20}H_{21}N_5O_2Cl_1I_1$ $[M + H]^+$: 526.05, found 526.2.

Compound 17. This compound was synthesized from commercially available 3-iodosalicylic acid following Path A. Yield: 31% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 7.82-7.87 (m, 2H), 7.62-7.66 (m, 2H), 7.42-7.45 (m, 2H), 6.64 (t, *J* = 7.9 Hz, 1H), 5.57-5.60 (m, 1H), 5.21 (s, 1H), 5.12 (s, 1H), 3.36-3.39 (m, 2H), 2.19-2.32 (m, 2H), 1.82-1.89 (m, 1H), 1.72-1.79 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 170.6, 163.1, 163.0, 161.9, 161.6, 161.3, 160.2, 153.4, 144.1, 132.6, 127.2, 125.5, 120.2, 113.9, 113.9, 85.3, 78.3, 76.8, 41.4, 28.9, 23.7; ESI-MS (m/z) calculated for $C_{20}H_{21}N_5O_2F_1I_1$ $[M + H]^+$: 510.08, found 510.0.

Compound 18. This compound was synthesized from commercially available 3-iodosalicylic acid following Path A. Yield: 25% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 7.82-7.86

(m, 2H), 7.63-7.64 (m, 2H), 7.41-7.43 (m, 2H), 6.64 (t, *J* = 7.9 Hz, 1H), 5.57-5.60 (m, 1H), 4.27 (s, 1H), 3.33-3.37 (m, 2H), 2.20-2.32 (m, 2H), 1.82-1.89 (m, 1H), 1.72-1.79 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 170.6, 163.4, 160.2, 153.3, 144.1, 132.8, 127.2, 125.4, 120.2, 113.9, 113.9, 85.3, 41.9, 38.7, 28.9, 23.6; ESI-MS (m/z) calculated for $C_{20}H_{21}N_5O_2Cl_1I_1$ $[M + H]^+$: 526.05, found 526.0.

Compound 19. This compound was synthesized from commercially available 4-Fluoro-2-hydroxy-3,5-diiodobenzoic acid following Path B. Yield: 34% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.33 (d, *J* = 6.9 Hz, 1H), 7.62-7.64 (m, 2H), 7.41-7.43 (m, 2H), 5.55-5.58 (m, 1H), 5.21 (s, 1H), 5.12 (s, 1H), 3.37 (t, *J* = 7.3 Hz, 2H), 2.18-2.30 (m, 2H), 1.82-1.89 (m, 1H), 1.71-1.77 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.0, 165.0, 163.3, 163.2, 163.1, 163.0, 153.2, 137.7, 137.7, 133.0, 125.0, 113.9, 112.3, 78.3, 76.8, 73.3, 73.0, 66.2, 66.0, 41.4, 29.0, 23.7, 21.7; ESI-MS (m/z) calculated for $C_{20}H_{19}N_5O_2F_2I_2$ $[M + H]^+$: 653.97, found 653.6.

Compound 20. This compound was synthesized from commercially available 4-Fluoro-2-hydroxy-3,5-diiodobenzoic acid following Path B. Yield: 28% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.34 (d, *J* = 6.9 Hz, 1H), 7.64-7.66 (m, 2H), 7.44-7.46 (m, 2H), 5.57-5.60 (m, 1H), 4.27 (s, 1H), 3.31-3.36 (m, 2H), 2.22-2.30 (m, 2H), 1.83-1.89 (m, 1H), 1.72-1.78 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 169.1, 165.0, 163.4, 163.2, 163.2, 163.1, 153.1, 137.8, 137.7, 132.2, 125.7, 113.8, 112.3, 73.2, 73.0, 66.3, 66.0, 41.9, 38.7, 28.9, 23.6; ESI-MS (m/z) calculated for $C_{20}H_{19}N_5O_2F_1Cl_1I_2$ $[M + H]^+$: 669.94, found 669.6.

Compound 21. This compound was synthesized from compound **30** following Path A. Yield: 25% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.35 (s, 1H), 7.59 (s, 2H), 7.34-7.36 (m, 2H), 5.54-5.56 (m, 1H), 5.21 (s, 1H), 5.12 (s, 1H), 4.32 (s, 1H), 3.37 (t, *J* = 7.1 Hz, 2H), 2.27-2.29 (m, 1H), 2.15-2.19 (m, 1H), 1.70-1.83 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 168.8,

163.1, 160.7, 139.3, 136.6, 124.6, 114.9, 114.1, 92.7, 88.3, 87.2, 84.8, 78.3, 76.9, 41.4, 29.1, 28.9, 23.7; ESI-MS (m/z) calculated for $C_{22}H_{20}N_5O_2F_1I_2$ $[M + H]^+$: 659.98, found 659.8.

Compound 22. This compound was synthesized from compound **30** following Path A. Yield: 20% over three steps. ^{1}H NMR (CD_3OD) δ (ppm): 8.36 (s, 1H), 7.59-7.61 (m, 2H), 7.36-7.38 (m, 2H), 5.54-5.57 (m, 1H), 4.32 (s, 1H), 4.27 (s, 2H), 3.33-3.36 (m, 2H), 2.26-2.33 (m, 1H), 2.14-2.22 (m, 1H), 1.80-1.85 (m, 1H), 1.72-1.77 (m, 1H); ^{13}C NMR (CD_3OD) δ (ppm): 168.8, 163.4, 160.7, 153.2, 139.3, 136.6, 133.9, 124.8, 114.8, 114.0, 92.7, 88.3, 87.2, 84.8, 42.0, 38.7, 29.1, 23.6; ESI-MS (m/z) calculated for $C_{22}H_{20}N_5O_2Cl_1I_2$ $[M + H]^+$: 675.95, found 675.8.

Inactivation kinetics. The rates of enzyme inactivation for compounds **1**-**22** were determined using previously described methods.[3] Briefly, PAD (2 μM for PAD1, 2 and 4; 5 μM for PAD3) was added to a prewarmed (10 min at 37 ^{0}C) inactivation mixture (100 mM TRIS pH 7.4, 50 mM NaCl, 10 mM $CaCl_2$, and 2 mM DTT with a final volume of 50 μL) containing various concentrations of inhibitors. At various time points, 6 μL of this inactivation mixture was removed and was added to a prewarmed (10 min at 37 ^{0}C) reaction mixture (100 mM TRIS pH 7.4, 50 mM NaCl, 10 mM $CaCl_2$, 2 mM DTT, and 10 mM BAEE (for PAD1, 2 and 4) or 10 mM BAA (for PAD3) with a final volume of 60 μL). After 15 min, reactions were quenched with liquid nitrogen and the production of citrulline was quantified using the COLDER assay.[4] The time-dependence of PAD inhibition was fit to the following equation 1,

$$\nu = \nu_0 \, e^{-kt} \qquad (1),$$

using Grafit, version 5.0.11, where ν is velocity, ν_0 is initial velocity, k (or k_{obs} is the pseudo-first-order rate constant of inactivation, and t is time. Upon reaching saturation, the concentration dependence of k_{obs} was fit to equation 2,

$$k_{obs} = k_{inact}\,[I] / (K_I + [I]) \quad (2),$$

using Grafit, version 5.0.11, where k_{inact} is the maximal rate of inactivation, K_I is the concentration of inhibitor that affords half-maximal inactivation, and [I] is the concentration of inhibitor. When k_{obs} with [I] varied linearly, k_{inact}/K_I was determined from the slope of the line. All the experiments were performed at least in duplicate.

RFA labelling. RFA labelling of PAD1 in the presence of inhibitors followed a protocol similar to that established for PAD2.[5] Briefly, PAD1 (2 μM final) was added to prewarmed (10 min at 37 ^{0}C) reaction mixture (100 mM TRIS pH 7.4, 50 mM NaCl, 10 mM $CaCl_2$, and 2 mM DTT in a final volume of 30 μL) containing RFA (2 μM final) and DMSO or various concentrations of the inhibitor. After incubating at 37 ^{0}C for 2 h, the reaction mixture was quenched with 5X SDS-PAGE loading dye, incubated at 95 ^{0}C for 15 min and then loaded onto a 10% SDS-PAGE gel. In-gel fluorescence of the protein bands was recorded using a typhoon scanner (excitation/emission maxima of ~546/579, respectively) and was quantified using ImageJ software. IC_{50} values of the inhibitors were obtained by plotting the relative decrease in the fluorescence intensity of protein bands against the concentration of inhibitor. All the experiments were performed at least in duplicate.

Determination of log P. Compounds of interest (1 mM) were suspended in a 1:1 mixture of n-octanol and 20 mM TRIS-HCl (pH 7.4) (1 mL), and the solvents were constantly mixed on an end over end shaker at room temperature for 1h. The emulsion was then centrifuged (8000 rpm, 1 min) to separate the n-octanol and aqueous layers. The concentration of the compound in each layer was determined by LC-MS analysis and the log P values are expressed as the

ratio of compound concentration in n-octanol to aqueous layer. All the experiments were performed at least in triplicate.

Preparation and Maintenance of HEK293TPAD1 Cells. The human PADI1 gene was amplified from the pET28-PAD1 vector by PCR using the following forward (encoding a 5′ KpnI site along with a FLAG tag) and reverse (encoding a 3′ XhoI site) primers.

Forward primer: 5′-CATGGTACCATGGACTACAAGGACGACGACGACAAGATGGCCCCAAAGAGAGTTGT-3′

Reverse primer: 5′-GTCTCGAGTCAGGGCACCATGTTCCACCATT-3′

The PCR product was purified using a PCR purification kit. The purified PCR product was digested with KpnI and XhoI, and the mixture was then run on a 1% agarose gel. The desired DNA band was cut from the gel and was extracted using a gel purification kit (Bio Basic Canada Inc.). The DNA was then ligated into pcDNA 3.1 Hygro (+), which had also been digested with KpnI and XhoI, using T4 DNA ligase. The ligation mixture was transformed into competent *E. Coli* XL1-Blue cells and penicillin-resistant single colonies were grown by culturing at 37 ^{0}C for 12 h. pcDNA3.1-PAD1 plasmid was then purified from the cultured cells using a plasmid purification kit (Bio Basic Canada Inc.). HEK293T cells (ATCC) were grown to 50-60% confluence in Dulbecco's Modified Eagle's Medium (DMEM) (supplemented with 10% heat-inactivated fetal bovine serum, 100 units/mL penicillin and 100 µg/mL streptomycin) under 5% CO_2 atmosphere and were transfected with 20 µg of the pcDNA3.1-PAD1 construct using Lipofectamine 2000 according to the manufacturer's protocol. Cells were further cultured for 48 h post-transfection and then selected with 0.2 mg/mL hygromycin B to establish stable transfection. Overexpression of PAD1 in HEK293TPAD1 cells was confirmed by Western blotting. HEK293TPAD1 cells were maintained in DMEM (supplemented with 10% heat-inactivated fetal bovine serum, 100

units/mL penicillin, 100 µg/mL streptomycin and 0.15 mg/mL hygromycin B) under 5% CO_2 atmosphere.

Histone H3 Citrullination in HEK293TPAD1 Cells. HEK293TPAD1 cells were seeded on 6-well plates at a density of ~1 × 10^6 cells/well and were allowed to grow in DMEM (supplemented with 10% heat-inactivated fetal bovine serum, 100 units/mL penicillin, 100 µg/mL streptomycin and 0.15 mg/mL hygromycin B) for 12 h. The monolayer of cells was washed gently with 2 mL 1X HBS (140 mM NaCl, 0.7 mM $Na_2HPO_4.2H_2O$, 20 mM HEPES pH 7.4) and each well was filled with 2 mL of 1X HBS (pH 7.4). Then the cells were treated with 1 mM $CaCl_2$, 5 µM ionomycin (calcium ionophore) and DMSO or various concentrations of inhibitor for 3 h. The final concentration of DMSO in each well was 1%. Cells in each well were then collected using a cell scraper, harvested by centrifugation at 3,000 rpm for 5 min and were resuspended in 1X HBS (100 µL). Triton X-100 (1% final) was added to each sample and the cells were lysed using a probe sonicator (pulse amplitude of 10 for 10 s, oscillating between 1 s on and 1 s off). Lysates were cleared by centrifugation at 21,000g for 15 min and soluble proteins were quantified by DC-assay (Bio-Rad). 2 µg of total protein was loaded onto a 4-15% gradient SDS-PAGE gel and separated by electrophoresis followed by transferring to a PVDF membrane (Bio-Rad) at 80 V for 50 min. The membrane was blocked by treating with 5% BSA in phosphate buffered saline supplemented with 0.1% Tween-20 (PBST). The membrane was then treated with mouse monoclonal anti-histone H3 (1:1000) and rabbit polyclonal anti-histone H3Cit 2,8,17 (1:1000) antibodies in blocking solution for 12 h at 4 ^{0}C. The membrane was washed with PBST (3 times, 10 min for each wash) and was incubated with anti-mouse as well as anti-rabbit IgG Licor conjugates (1:5000) in PBST and 5% BSA for 1 h at 25 ^{0}C. Then the membrane was washed with PBST (3 times, 10 min for each wash) and imaged by Licor analysis. All the experiments were performed at least in duplicate.

Cytotoxicity Studies. HEK293TPAD1 cells were seeded (2×10^4 cells/well) on 96-well plate and were allowed to grow in DMEM (supplemented with 10% heat-inactivated fetal bovine serum, 100 units/mL penicillin, 100 μg/mL streptomycin and 0.15 mg/mL hygromycin B) for 24 h. Cells were then treated with DMSO or various concentrations of inhibitor at 37 ^{0}C for 24 h. Activated XTT reagent (ATCC) was added to each well and cells were further incubated at 37 ^{0}C for 5 h. Cell viability was measured by recording the absorbance at 475 nm and 660 nm. Equation 3,

$$Y = \text{Bottom} + (\text{Top-Bottom}) / [1 + 10^{((\log EC50-X)*\text{Hillslope})}] \quad (3),$$

was used to fit an eight-point dose-response curve to determine the EC_{50} values for inhibition of cell-proliferation using GraphPad Prism 7.03. Top and Bottom are plateaus of the dose-response curve, X is the log of inhibitor-concentration, Hillslope is the slope factor or Hill slope. All the experiments were performed at least in triplicate.

In Vitro Fertilization. Female ICR mice (4–6 weeks) were injected intraperitoneally with 10 international units (IU) of pregnant mare serum gonadotropin (PMSG) and after 46 h, 10 IU of human chorionicgonadotropin (hCG). After 16 h, cumulus masses were released from oviducts into HEPES-buffered HTF culture medium (HTF) (InVitro Care, Frederick, MD). Cauda epididymal mouse spermatozoa were released and placed in drops of HTF culture medium (InVitro Care). Then the sperm suspensions were added to the cumulus mass drops with a final concentration of approximately 1–3 $\times$ 10^5/mL. Zygotes were collected and transferred to KSOM medium containing 100 μM PAD1 inhibitor or equal volume of DMSO as a control. Then 4-cell and blastocyst stage embryos were collected after 48 and 100 h of culture, respectively.

Immunofluorescent and Confocal Microscopy. Embryos were fixed in 4% paraformaldehyde in PBS (pH 7.4) for 30 min and then permeabilized in 0.5% Triton X-100 for 30 min at room temperature. After blocking in 1% BSA/PBS at room temperature for 1 h, samples were incubated with anti-H3Cit2/8/17 antibody (Catalogue no. ab77164, Abcam) (1:100) at 4 °C for 12 h. After five washes, samples were incubated with FITC-conjugated secondary antibody for 1 h at room temperature. Nuclei were counterstained with DAPI for 5 min. Embryos were mounted on glass slides in a drop of antifade medium (Vectashield, Burlingame, CA, USA) and were examined under confocal laser scanning microscope (LSM 700; Zeiss, Oberkochen, Germany). All the experiments were performed at least in triplicate.

Dose-dependent *in vitro* Labelling of Recombinant PADs with 21 and 22. A PAD (1.4 μM) was added to a prewarmed (10 min at 37 ^{0}C) reaction mixture (25 mM TRIS pH 7.4, 10 mM NaCl, and 5 mM $CaCl_2$) containing various concentrations of **21** or **22** (or equal volume of DMSO for control). The mixture was incubated at 37 ^{0}C for 15 min. Then TAMRA-N_3 (50 μM), TCEP (1 mM), TBTA (0.3 mM) and freshly prepared copper (II) sulphate (1 mM) were sequentially added to the reaction mixture and the tubes were gently tumbled at room temperature for 2 h. Then the reactions were quenched with 5X SDS-PAGE loading dye, heated at 95 ^{0}C for 15 min and loaded onto 4-15% gradient SDS-PAGE gel. Protein bands were visualized by recording in-gel fluorescence using a typhoon scanner (approximate excitation/emission maxima ~546/579, respectively). All the experiments were performed at least in duplicate.

Limit of Detection for *in vitro* Labelling of Recombinant PADs with 21 and 22. Various concentrations of PAD were added to prewarmed (10 min at 37 ^{0}C) reaction mixtures (25 mM TRIS pH 7.4, 10 mM NaCl, 5 mM $CaCl_2$) containing 10 μM of **21** or **22**. The mixture was incubated at 37 ^{0}C for 15 min followed by sequential addition of TAMRA-N_3 (50 μM),

TCEP (1 mM), TBTA (0.3 mM) and freshly prepared copper (II) sulphate (1 mM). The tubes were then gently tumbled at room temperature for 2 h. Reactions were quenched with 5X SDS-PAGE loading dye, heated at 95 ^{0}C for 15 min and loaded onto 4-15% gradient SDS-PAGE gel. Protein bands were visualized by recording in-gel fluorescence using a typhoon scanner (approximate excitation/emission maxima ~546/579, respectively). All the experiments were performed at least in duplicate.

Dose-dependent Labelling of PAD1 in HEK293TPAD1 cells with 21. HEK293TPAD1 cells were plated in 6-well plates at a density of ~1 × 10^6 cells/well and were allowed to grow in DMEM (supplemented with 10% heat-inactivated fetal bovine serum, 100 units/mL penicillin, 100 μg/mL streptomycin and 0.15 mg/mL hygromycin B) for 12 h. Then the cells were washed gently with serum-free DMEM (supplemented with 100 units/mL penicillin and 100 μg/mL streptomycin) and each well was filled with 2 mL of serum-free DMEM. The cells were treated with 1 mM $CaCl_2$, 5 μM ionomycin (calcium ionophore) and DMSO or various concentrations of inhibitor for 3 h. Cells were collected using a cell scraper, harvested by centrifugation at 3,000 rpm for 5 min and were resuspended in 1X PBS (80 μL). Then the cells were lysed using a probe sonicator and lysates were cleared by centrifugation at 21000g for 15 min. Soluble proteins in the lysate were quantified by DC-assay (Bio-Rad). Then the lysate (1 mg/mL, 50 μL total) was treated with TAMRA-N_3 (50 μM), TCEP (1.5 mM), TBTA (0.3 mM) and freshly prepared copper (II) sulphate (4 mM). The tubes were then gently tumbled at room temperature for 2 h. The precipitated proteins were collected by centrifugation at 21000g for 10 min. The protein pellets were resuspended in 5X SDS-PAGE loading dye, heated at 95 ^{0}C for 15 min and separated by SDS-PAGE. Protein bands were visualized by recording in-gel fluorescence using a typhoon scanner (approximate excitation/emission maxima ~546/579, respectively). All the experiments were performed at least in duplicate.

Time-dependent Labelling of PAD1 in HEK293TPAD1 cells with 21. This study followed a procedure similar to that described for the dose-dependent labelling of PAD1 in HEK293TPAD1 cells with **21**. The only difference is that the cells were treated with 1 mM $CaCl_2$, 5 μM ionomycin and 5 μM of **21** in serum-free DMEM for various times (0-6 h). At various time points, the labelling reaction was quenched by freezing the harvested cells in liquid nitrogen. Cells were lysed and lysates were clicked with TAMRA-N_3 in the presence of TCEP, TBTA and $CuSO_4$ according to the methods described above. Protein bands were separated by SDS-PAGE and were visualized by recording in-gel fluorescence. All the experiments were performed at least in duplicate.

Target Engagement Assay - Labelling of PAD1 in HEK293TPAD1 cells with 21 in the presence of 1. This study followed a similar procedure to that described for the dose-dependent labelling of PAD1 in HEK293TPAD1 cells with **21**. Briefly, HEK293TPAD1 cells were treated with 1 mM $CaCl_2$, 5 μM ionomycin, 5 μM of **21**, and increasing concentrations (0-50 μM) of **1** in serum-free DMEM for 3 h. Cells were scraped, harvested by centrifugation at 3,000 rpm for 5 min and resuspended in 1X PBS. Lysis was carried out by sonicating the cells with a probe sonicator and lysates were clicked with TAMRA-N_3 in the presence of TCEP, TBTA and $CuSO_4$ as described above. Protein bands were separated by SDS-PAGE and were visualized by recording in-gel fluorescence. Fluorescence intensities of the protein bands were quantified by ImageJ software and were plotted against the concentration of **1** to afford the EC_{50} value. All the experiments were performed at least in duplicate.

Enrichment of Proteins Labelled by 21 in HEK293TPAD1 Cells on Streptavidin-agarose. HEK293TPAD1 cells were cultured in 175 cm^2 T-175 flasks in DMEM (supplemented with 10% heat-inactivated fetal bovine serum, 100 units/mL penicillin, 100 μg/mL streptomycin and 0.15 mg/mL hygromycin B). Upon reaching ~90% confluence, cells were treated with 1 mM $CaCl_2$, 5 μM ionomycin and 5 μM of **21** (or equal volume of DMSO

for control) in serum-free DMEM for 3 h. Then the cells were scraped, harvested by centrifugation at 3,000 rpm for 5 min and resuspended in 1X PBS. Cell lysis was performed by probe sonication and soluble proteins in the lysates were quantified by DC Assay (Bio-Rad). Endogenous biotinylated proteins were precleared by incubating the lysates with streptavidin-agarose beads (25 µL for 1 mg of total protein in the lysate) at room temperature for 1 h with constant mixing on an end-over-end shaker. The mixture was centrifuged (1200 g, 3 minute) to separate the beads and the supernatant (pre-cleared lysate). Pre-cleared lysates (2 mg/mL, 500 µL final) were clicked with Biotin-N_3 (100 µM), TCEP (1.5 mM), TBTA (0.3 mM) and freshly prepared $CuSO_4$ (4 mM), and the tubes were gently tumbled at room temperature for 2 h. The precipitated proteins were collected by centrifugation at 5,000 rpm for 10 min. The protein pellets were washed with ice-cold methanol, dried at room temperature for 5 min and resuspended in 1.2% SDS in PBS. SDS-solubilized proteins were then diluted with PBS to a final SDS concentration of 0.2% and were incubated with streptavidin-agarose beads (170 µL for 2 mg of total protein) at 4 ^{0}C for 16 h on an end-over-end shaker. The solutions were then incubated at room temperature for 3 h. The streptavidin beads were collected by centrifugation at 1200g for 3 min and were washed with 2 M urea (2 × 5 mL), 0.2% SDS in PBS (2 × 5 mL), PBS (3 × 5 mL) and water (3 × 5 mL). The beads were pelleted by centrifugation at 1200g for 3 min between washes. The washed beads were resuspended in 6 M urea (500 µL) and were treated with dithiothreitol (10 mM) at 65 ^{0}C for 20 min. Iodoacetamide (20 mM final) was then added to the mixture and the beads were further incubated at 37 ^{0}C for 30 min. The beads were collected by centrifugation at 1200g for 3 min and were treated with a premixed solution of 2 M urea in PBS (200 µL), 100 mM $CaCl_2$ (2 µL) and trypsin (4 µL of 20 µg reconstituted in 40 µL of trypsin buffer) at 37 ^{0}C for 12 h. The digested peptides were separated from the beads by centrifugation and the beads were washed twice with water (50 µL). The digest was then desalted using a Pierce C18 spin

column (catalogue no 89870) according to the manufacturer' protocol and resuspended in 100 µL 100 mM triethylammonium bicarbonate pH 8.5. 4 µL of 20% $D^{13}CDO$ (heavy formaldehyde) and 2.5 µL of 37% HCHO (light formaldehyde) were added to the **21**-treated and control samples, respectively. 20 µL of 0.6 M sodium cyanoborohydride was then added to both the **21**-treated and control samples, and the samples were incubated at room temperature for 2 h. The samples were cooled on ice and the reaction quenched with 4 µL of 20% ammonium hydroxide. 8 µL of formic acid was then added to the samples. Heavy (**21**-treated) and light (control) samples were mixed together and were stored at -20 ^{0}C for proteomic analysis. All the experiments were performed at least in duplicate.

Proteomic Analysis and Data Processing. Mass Spectrometry was performed using a Thermo Fisher LTQ Orbitrap Discovery mass spectrometer in conjunction with an Agilent 1200 series HPLC. Labeled peptide samples were pressure-loaded onto a 250 µm fused silica desalting column packed with 4 cm of Aqua C18 reverse phase resin (Phenomenex). Peptides were eluted onto a biphasic 100 µm fused silica column with a 5 µm tip packed with 4 cm Partisphere strong cation exchange resin (SCX, Whatman) followed by 10 cm of C18 resin. The peptides were separated from the SCX onto the C18 resin using 5 distinct salt pulses (95% high-purity water, 5% Optima-grade acetonitrile, 0.1% formic acid and 500 mM ammonium acetate) as outlined in Weerapana *et. al.*[6] C18 elution took place using a gradient from 5% Buffer A (95% water, 5% acetonitrile, 0.1% formic acid), to 100% Buffer B (20% water, 80% acetonitrile, 0.1% formic acid). The flow rate through the column was 0.25 mL/min, with a spray voltage of 2.75 kV. With dynamic exclusion enabled, each full MS scan (400-1800 MW) was followed by 8, data-dependent scans of the n^{th} most intense ion.

The tandem MS data was analyzed by the SEQUEST algorithm[7] using a static Cys modification (+57 Da), and a static modification on Lys and the N-termini of either +28 Da (light) or +34 Da (heavy). Data was searched against a human reverse-concatenated non-redundant FASTA database applying Uniprot identifiers. MS2 spectra were assembled using DTASelect 2.0,[8] with the –trypstat option applied. Quantification of Heavy/Light ratios were calculated using the Cimage quantification package.[9] Ratios across peptides were averaged to generate relative protein abundance ratios.

Table S1. k_{inact}/K_I values and selectivity of compounds 1-22 for inhibition of PAD1-4.

Compound	k_{inact}/K_I ($M^{-1}min^{-1}$) [*Fold PAD selectivity*]			
	PAD1	**PAD2**	**PAD3**	**PAD4**
1	10400 ± 2380* [*74*]	140 ± 20† [*1*]	200 ± 40† [*1*]	270 ± 40† [*2*]
2	58860 ± 3640* [*19*]	8110 ± 1020* [*3*]	3050 ± 1150* [*1*]	11060 ± 2880* [*4*]
3	5870 ± 790* [*8*]	1650 ± 70† [*2*]	2160 ± 80† [*3*]	750 ± 70† [*1*]
4	66020 ± 2870* [*15*]	9730 ± 2210* [*2*]	4400 ± 140‡ [*1*]	24110 ± 90* [*5*]
5	7280 ± 960* [*9*]	770 ± 80† [*1*]	2470 ± 260† [*3*]	1070 ± 180† [*1*]
6	59480 ± 14000* [*18*]	3230 ± 790† [*1*]	3300 ± 70‡ [*1*]	35190 ± 4970* [*11*]
7	120 ± 20† [*1*]	120 ± 20† [*1*]	1620 ± 360* [*14*]	2530 ± 200* [*21*]
8	600 ± 70‡ [*1*]	2800 ± 210‡ [*5*]	10590 ± 3980† [*18*]	41040 ± 630* [*68*]
9	190 ± 20† [*1*]	400 ± 30† [*2*]	580 ± 120† [*3*]	240 ± 20† [*1*]
10	5080 ± 210* [*1*]	4190 ± 900* [*1*]	10060 ± 1850* [*2*]	9690 ± 480* [*2*]
11	3280 ± 210* [*25*]	130 ± 10† [*1*]	420 ± 30† [*3*]	250 ± 20† [*2*]
12	40740 ± 2960* [*3*]	14210 ± 110* [*1*]	21560 ± 2430* [*2*]	18810 ± 970* [*1*]
13	1290 ± 90* [*10*]	130 ± 20† [*1*]	550 ± 40† [*4*]	290 ± 10† [*2*]
14	5300 ± 210‡ [*2*]	3400 ± 70‡ [*1*]	5500 ± 300‡ [*2*]	2600 ± 70‡ [*1*]
15	1500 ± 30* [*13*]	120 ± 30† [*1*]	500 ± 50† [*4*]	640 ± 90† [*5*]
16	3900 ± 100‡ [*2*]	1700 ± 70‡ [*1*]	5000 ± 280‡ [*3*]	3600 ± 70‡ [*2*]
17	300 ± 30† [*3*]	110 ± 10† [*1*]	240 ± 30† [*2*]	180 ± 20† [*2*]
18	4100 ± 300‡ [*4*]	2210 ± 270† [*2*]	1170 ± 40† [*1*]	3160 ± 550† [*3*]
19	13250 ± 2870* [*44*]	300 ± 60† [*1*]	700 ± 70† [*2*]	300 ± 30† [*1*]
20	94360 ± 14870* [*4*]	34290 ± 1970* [*1*]	70000 ± 15000* [*3*]	24310 ± 1060* [*1*]
21	4480 ± 200* [*34*]	130 ± 30† [*1*]	310 ± 80† [*2*]	220 ± 30† [*2*]
22	72160 ± 10970* [*9*]	31230 ± 1440* [*4*]	8200 ± 30* [*1*]	25000 ± 1370* [*3*]

*k_{inact} and K_I were determined from a nonlinear fit of the k_{obs} versus [I] data. ‡k_{inact}/K_I was determined from a linear fit of the k_{obs} versus [I] data. †A single k_{obs} was determined

Table S2. k_{inact} and K_I values of inhibition of PAD1-4 by compounds 1-22.

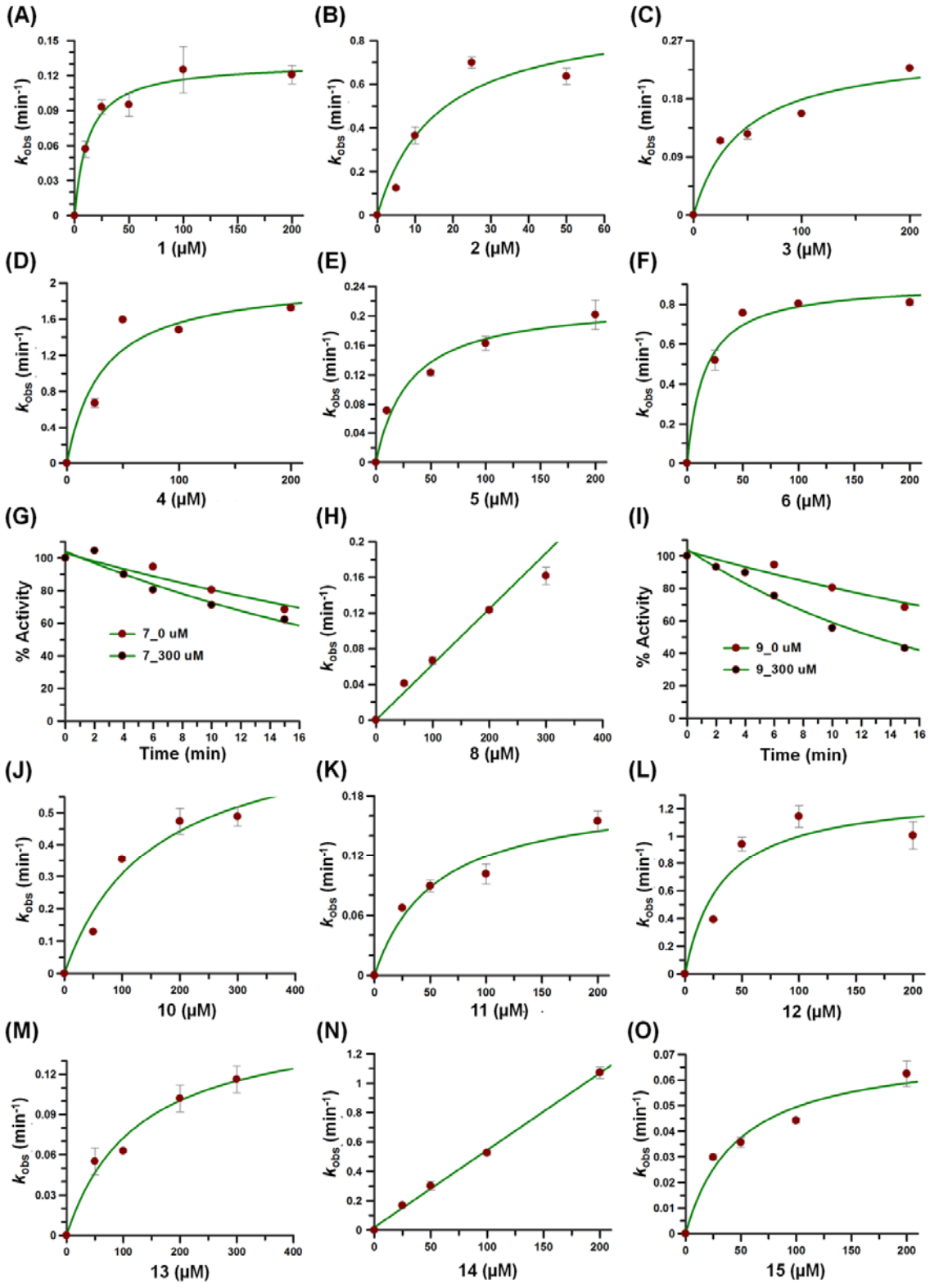
Compounds	PAD1		PAD2		PAD3		PAD4	
	k_{inact} (min^{-1})	K_I (μM)	k_{inact} (min^{-1})	K_I (μM)	k_{inact} (min^{-1})	K_I (μM)	k_{inact} (min^{-1})	K_I (μM)
1	0.13 ± 0.007	12.5 ± 2.9	NA†	NA†	NA†	NA†	NA†	NA†
2	0.93 ± 0.02	15.8 ± 3.3	0.76 ± 0.07	93.7 ± 17.8	0.77 ± 0.009	252.7 ± 5.3	0.68 ± 0.01	61.5 ± 11.0
3	0.26 ± 0.01	44.3 ± 8.5	NA†	NA†	NA†	NA†	NA†	NA†
4	2.04 ± 0.01	30.9 ± 1.6	2.02 ± 0.38	207.6 ± 61.6	NA‡	NA‡	3.05 ± 0.01	126.5 ± 1.06
5	0.22 ± 0.01	30.2 ± 7.6	NA†	NA†	NA†	NA†	NA†	NA†
6	0.91 ± 0.03	15.3 ± 4.0	NA†	NA†	NA‡	NA‡	2.27 ± 0.002	64.5 ± 9.4
7	NA†	NA†	NA†	NA†	0.12 ± 0.01	73.9 ± 10.3	0.25 ± 0.004	98.4 ± 7.7
8	NA‡	NA‡	NA‡	NA‡	NA†	NA†	1.03 ± 0.04	25.1 ± 1.3
9	NA†	NA†	NA†	NA†	NA†	NA†	NA†	NA†
10	0.78 ± 0.007	153.6 ± 4.9	0.55 ± 0.04	131.2 ± 38.9	0.96 ± 0.03	95.4 ± 14.3	1.58 ± 0.08	163.1 ± 16.7
11	0.18 ± 0.005	54.8 ± 5.5	NA†	NA†	NA†	NA†	NA†	NA†
12	1.32 ± 0.2	32.4 ± 7.0	1.30 ± 0.08	91.5 ± 6.7	2.32 ± 0.15	107.6 ± 19.6	1.46 ± 0.28	77.6 ± 18.8
13	0.16 ± 0.02	124.5 ± 7.4	NA†	NA†	NA†	NA†	NA†	NA†
14	NA‡	NA‡	NA‡	NA‡	NA‡	NA‡	NA‡	NA‡
15	0.07 ± 0.006	46.7 ± 3.8	NA†	NA†	NA†	NA†	NA†	NA†
16	NA‡	NA‡	NA‡	NA‡	NA‡	NA‡	NA‡	NA‡
17	NA†	NA†	NA†	NA†	NA†	NA†	NA†	NA†
18	NA‡	NA‡	NA†	NA†	NA†	NA†	NA†	NA†
19	0.4 ± 0.07	30.2 ± 12.7	NA†	NA†	NA†	NA†	NA†	NA†
20	1.84 ± 0.18	19.5 ± 4.2	0.60 ± 0.01	17.5 ± 1.4	1.19 ± 0.06	17.0 ± 4.0	0.44 ± 0.008	18.1 ± 0.2
21	0.31 ± 0.005	69.2 ± 4.6	NA†	NA†	NA†	NA†	NA†	NA†
22	2.54 ± 0.05	35.2 ± 4.7	0.79 ± 0.01	25.3 ± 1.6	0.8 ± 0.03	97.6 ± 3.9	0.77 ± 0.01	30.8 ± 1.1

NA: Not applicable. †A single k_{obs} was determined to calculate k_{inact}/K_I. ‡k_{inact}/K_I was determined from a linear fit of the k_{obs} versus [I] data.

Table S3. ***Calculated and experimental log P values of compounds 1,7, 11, 13 and 19.***

Compound	log P	
	Calculated†	Experimental
1	4.76	3.89 ± 0.07
19	4.92	3.59 ± 0.22
11	3.71	1.95 ± 0.06
13	3.44	1.89 ± 0.01
7	2.30	0.87 ± 0.06

†log P values were calculated using ChemBioDraw Ultra 13.0.


(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
(I)
(J)
(K)
(L)
(M)
(N)
(O)
k_{obs} (min^{-1})
1 (μM)
2 (μM)
3 (μM)
4 (μM)
5 (μM)
6 (μM)
% Activity
7_0 uM
7_300 uM
Time (min)
8 (μM)
9_0 uM
9_300 uM
10 (μM)
11 (μM)
12 (μM)
13 (μM)
14 (μM)
15 (μM)

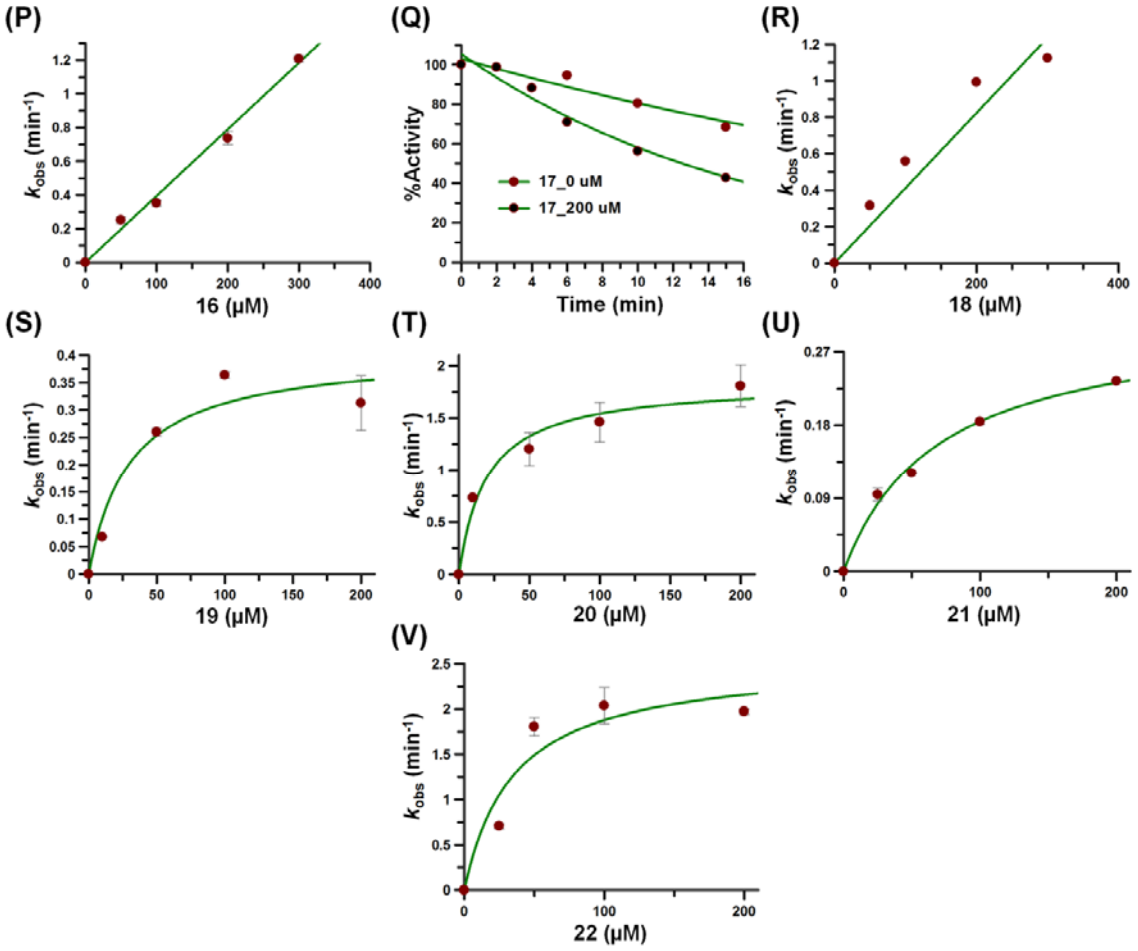
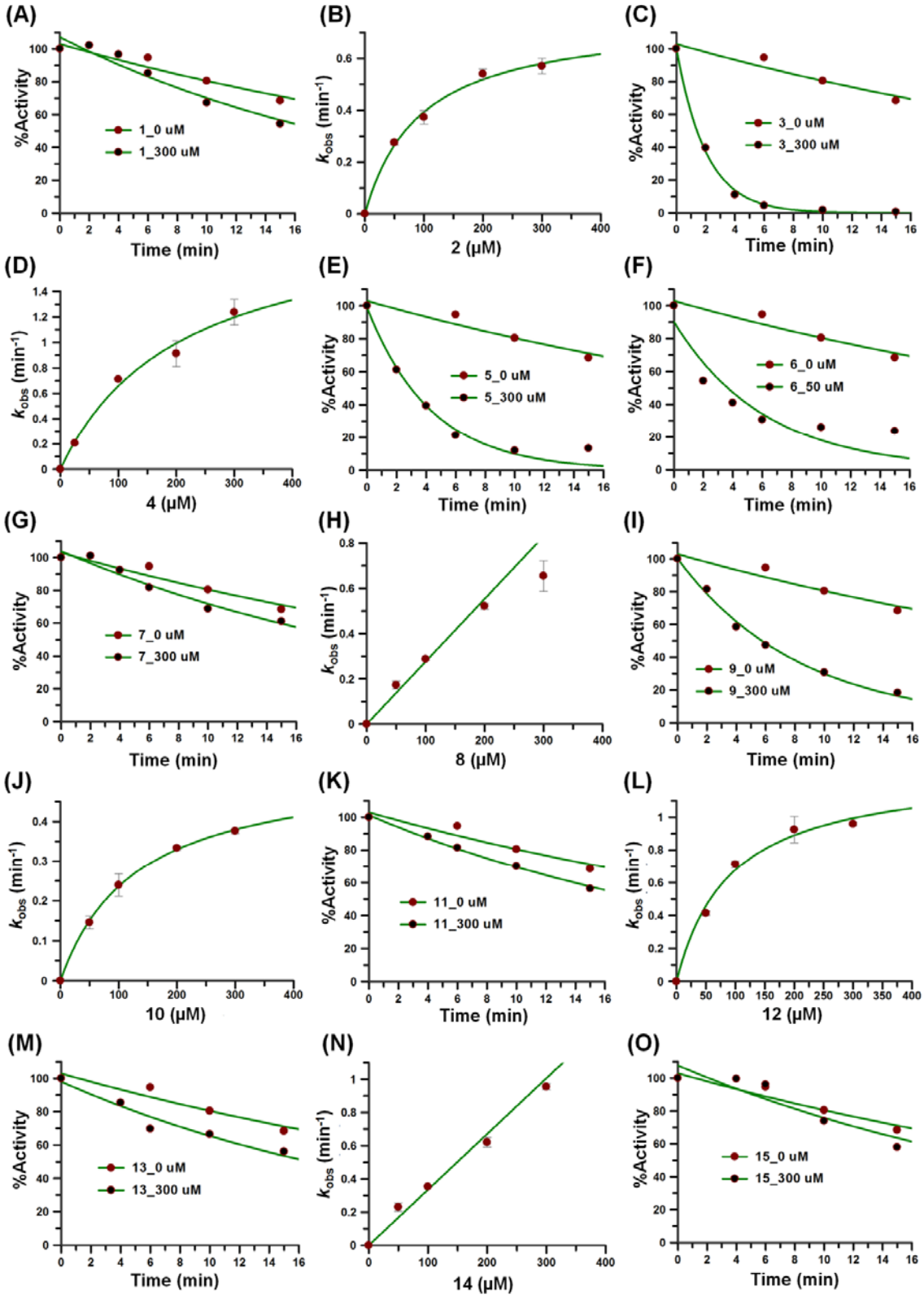


Figure S1. Inhibition of PAD1 by compounds **1**-**22**. **1** (A), **2** (B), **3** (C), **4** (D), **5** (E), **6** (F), **7** (G), **8** (H), **9** (I), **10** (J), **11** (K), **12** (L), **13** (M), **14** (N), **15** (O), **16** (P), **17** (Q), **18** (R), **19** (S), **20** (T), **21** (U) and **22** (V).


(A)
%Activity
1_0 uM
1_300 uM
Time (min)
(B)
k_{obs} (min^{-1})
2 (μM)
(C)
%Activity
3_0 uM
3_300 uM
Time (min)
(D)
k_{obs} (min^{-1})
4 (μM)
(E)
%Activity
5_0 uM
5_300 uM
Time (min)
(F)
%Activity
6_0 uM
6_50 uM
Time (min)
(G)
%Activity
7_0 uM
7_300 uM
Time (min)
(H)
k_{obs} (min^{-1})
8 (μM)
(I)
%Activity
9_0 uM
9_300 uM
Time (min)
(J)
k_{obs} (min^{-1})
10 (μM)
(K)
%Activity
11_0 uM
11_300 uM
Time (min)
(L)
k_{obs} (min^{-1})
12 (μM)
(M)
%Activity
13_0 uM
13_300 uM
Time (min)
(N)
k_{obs} (min^{-1})
14 (μM)
(O)
%Activity
15_0 uM
15_300 uM
Time (min)

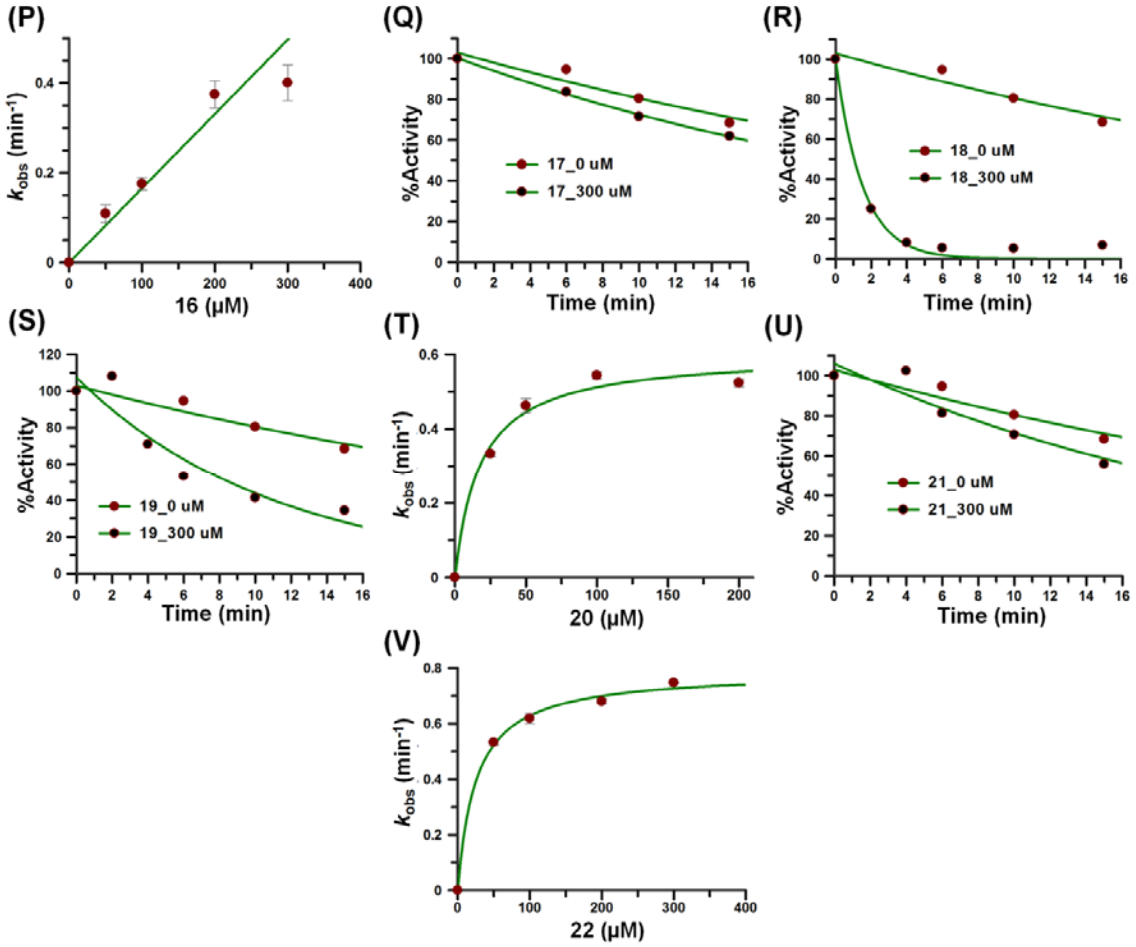
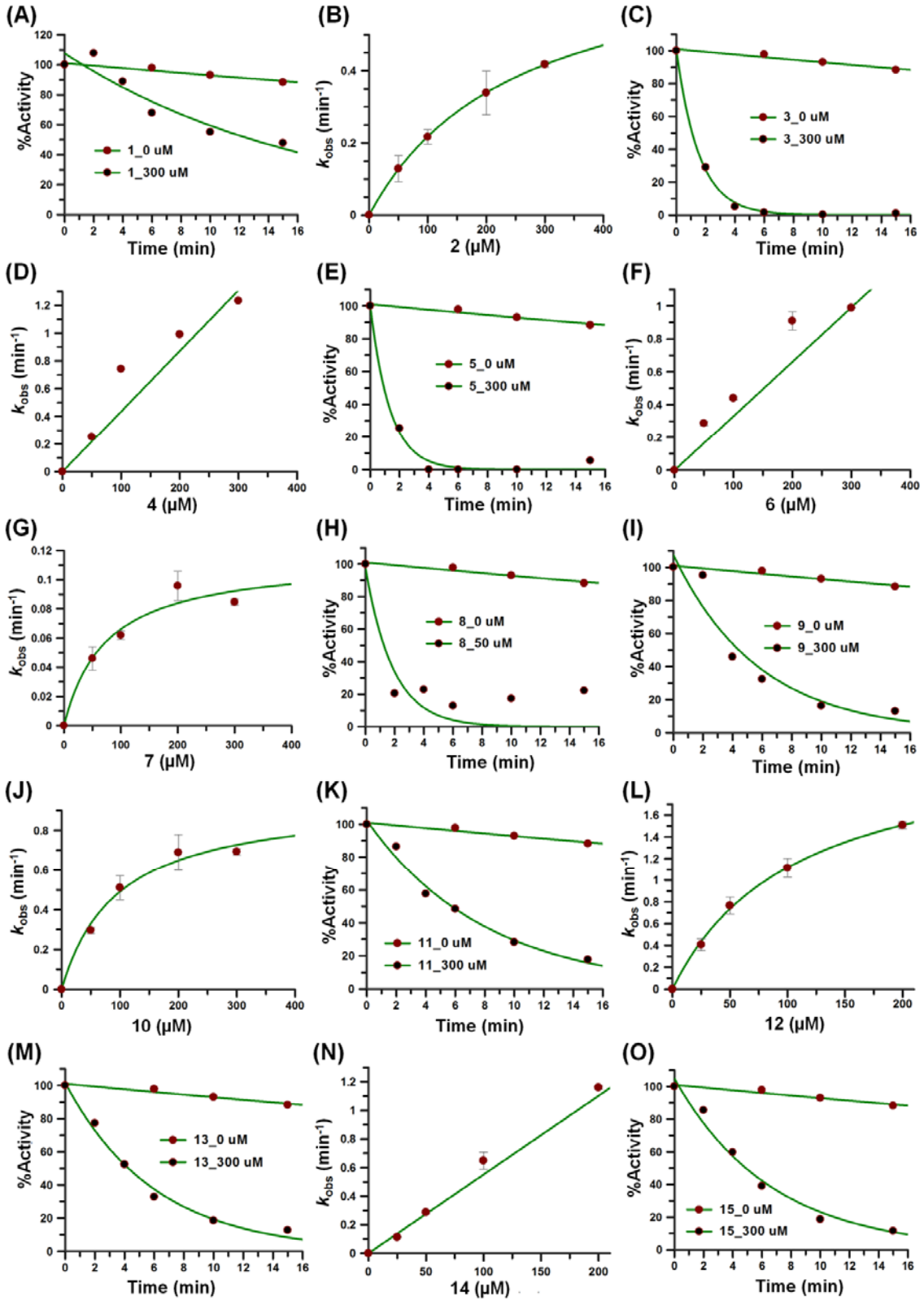


Figure S2. Inhibition of PAD2 by compounds **1**-**22**. **1** (A), **2** (B), **3** (C), **4** (D), **5** (E), **6** (F), **7** (G), **8** (H), **9** (I), **10** (J), **11** (K), **12** (L), **13** (M), **14** (N), **15** (O), **16** (P), **17** (Q), **18** (R), **19** (S), **20** (T), **21** (U) and **22** (V).


(A)
%Activity
Time (min)
1_0 uM
1_300 uM
(B)
k_{obs} (min^{-1})
2 (μM)
(C)
%Activity
Time (min)
3_0 uM
3_300 uM
(D)
k_{obs} (min^{-1})
4 (μM)
(E)
%Activity
Time (min)
5_0 uM
5_300 uM
(F)
k_{obs} (min^{-1})
6 (μM)
(G)
k_{obs} (min^{-1})
7 (μM)
(H)
%Activity
Time (min)
8_0 uM
8_50 uM
(I)
%Activity
Time (min)
9_0 uM
9_300 uM
(J)
k_{obs} (min^{-1})
10 (μM)
(K)
%Activity
Time (min)
11_0 uM
11_300 uM
(L)
k_{obs} (min^{-1})
12 (μM)
(M)
%Activity
Time (min)
13_0 uM
13_300 uM
(N)
k_{obs} (min^{-1})
14 (μM)
(O)
%Activity
Time (min)
15_0 uM
15_300 uM

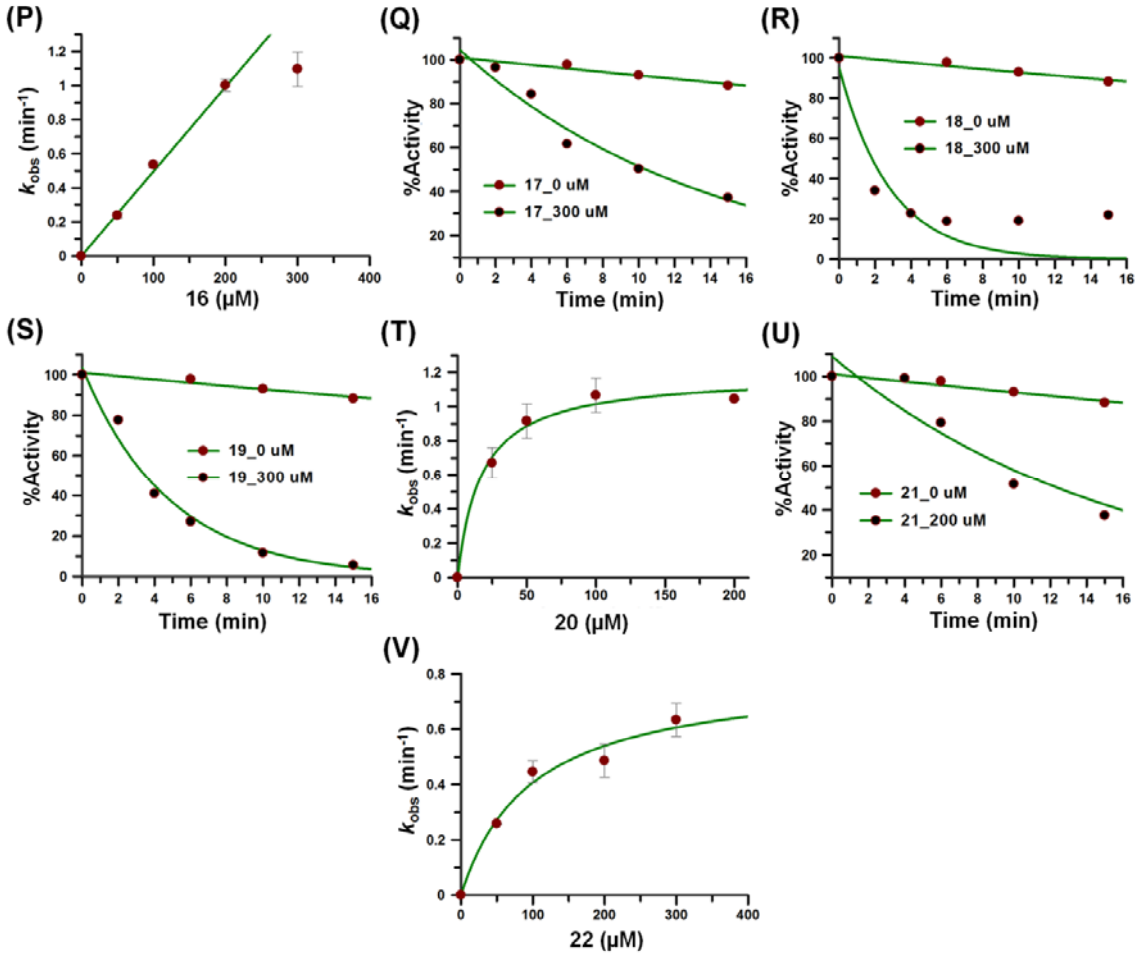
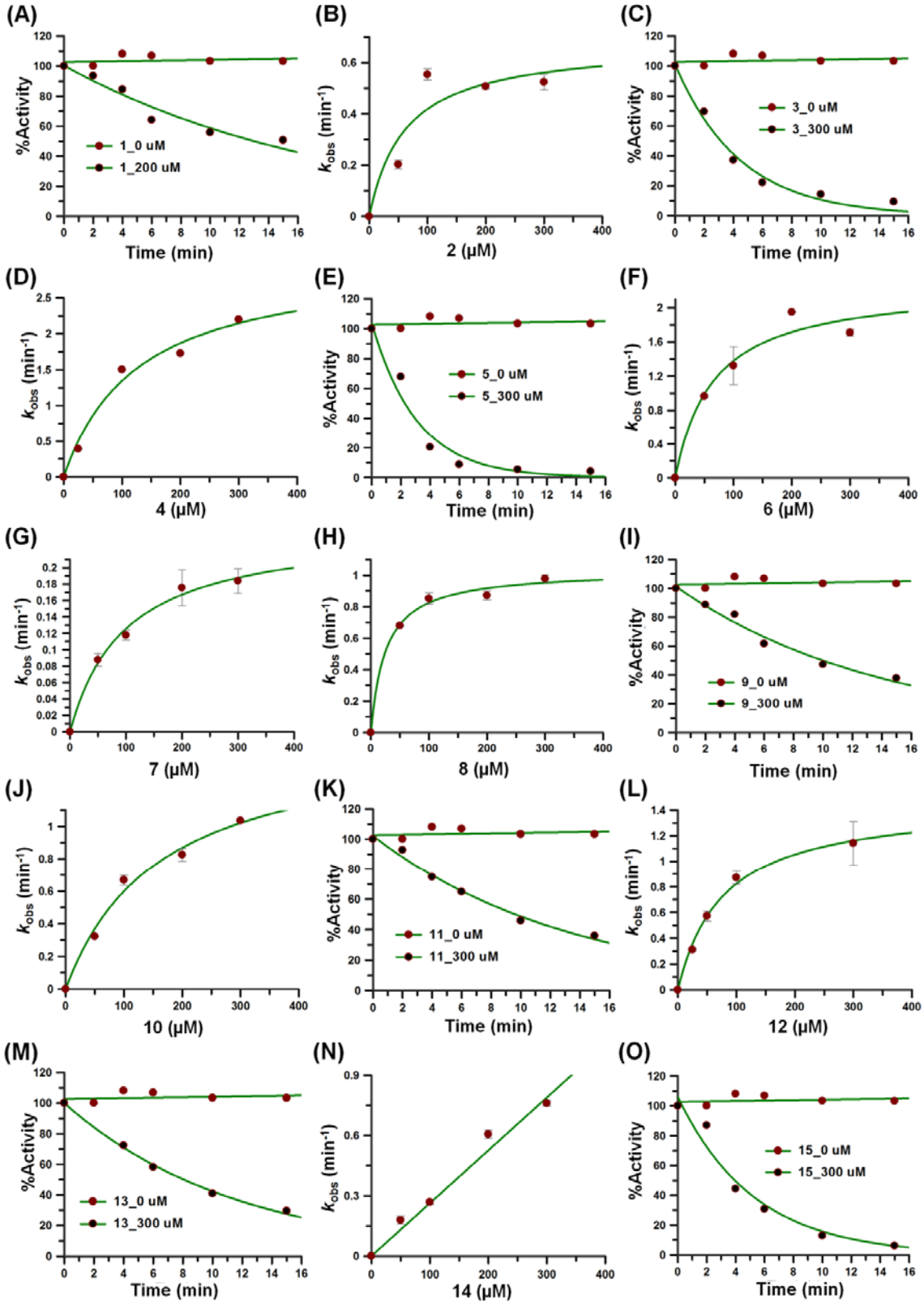


Figure S3. Inhibition of PAD3 by compounds **1**-**22**. **1** (A), **2** (B), **3** (C), **4** (D), **5** (E), **6** (F), **7** (G), **8** (H), **9** (I), **10** (J), **11** (K), **12** (L), **13** (M), **14** (N), **15** (O), **16** (P), **17** (Q), **18** (R), **19** (S), **20** (T), **21** (U) and **22** (V).


(A)
%Activity
Time (min)
1_0 uM
1_200 uM
(B)
k_{obs} (min^{-1})
2 (μM)
(C)
%Activity
Time (min)
3_0 uM
3_300 uM
(D)
k_{obs} (min^{-1})
4 (μM)
(E)
%Activity
Time (min)
5_0 uM
5_300 uM
(F)
k_{obs} (min^{-1})
6 (μM)
(G)
k_{obs} (min^{-1})
7 (μM)
(H)
k_{obs} (min^{-1})
8 (μM)
(I)
%Activity
Time (min)
9_0 uM
9_300 uM
(J)
k_{obs} (min^{-1})
10 (μM)
(K)
%Activity
Time (min)
11_0 uM
11_300 uM
(L)
k_{obs} (min^{-1})
12 (μM)
(M)
%Activity
Time (min)
13_0 uM
13_300 uM
(N)
k_{obs} (min^{-1})
14 (μM)
(O)
%Activity
Time (min)
15_0 uM
15_300 uM

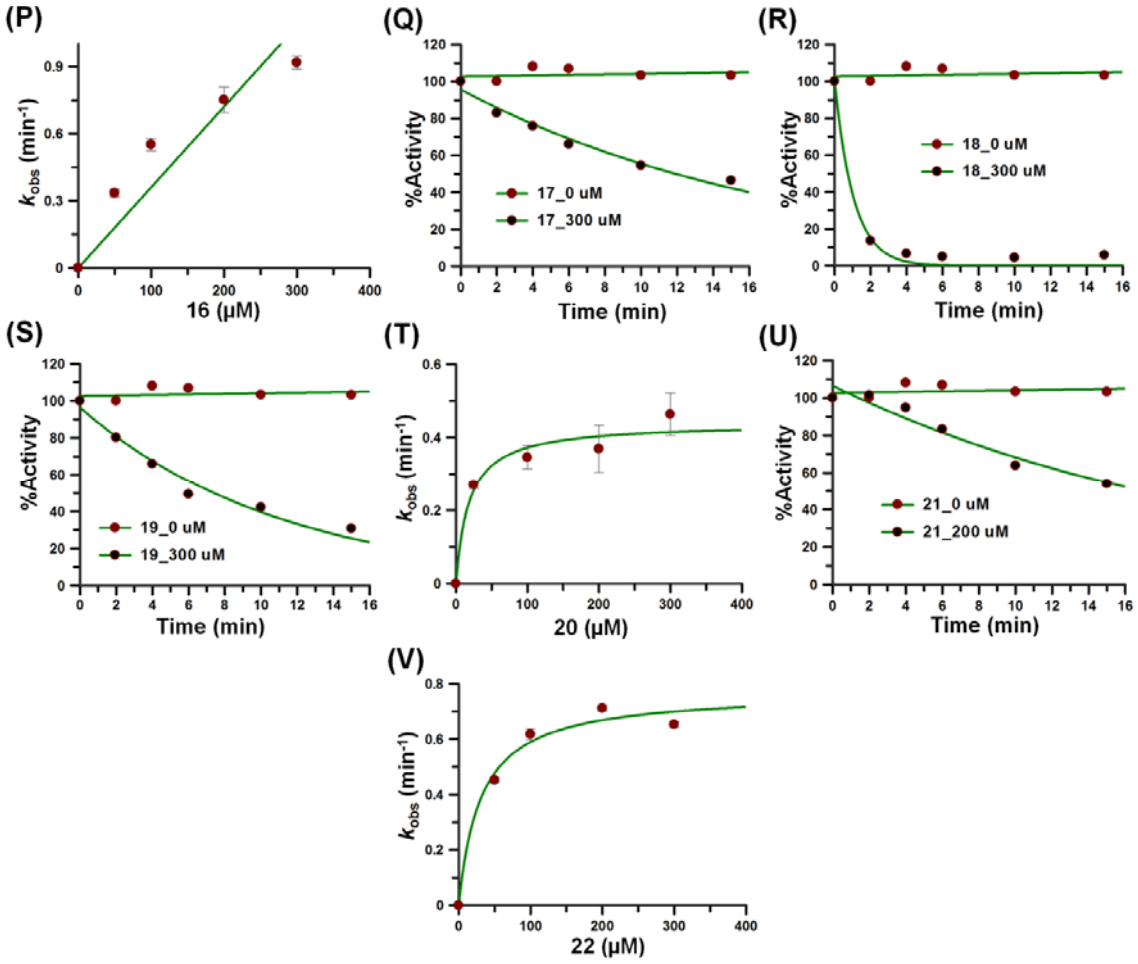
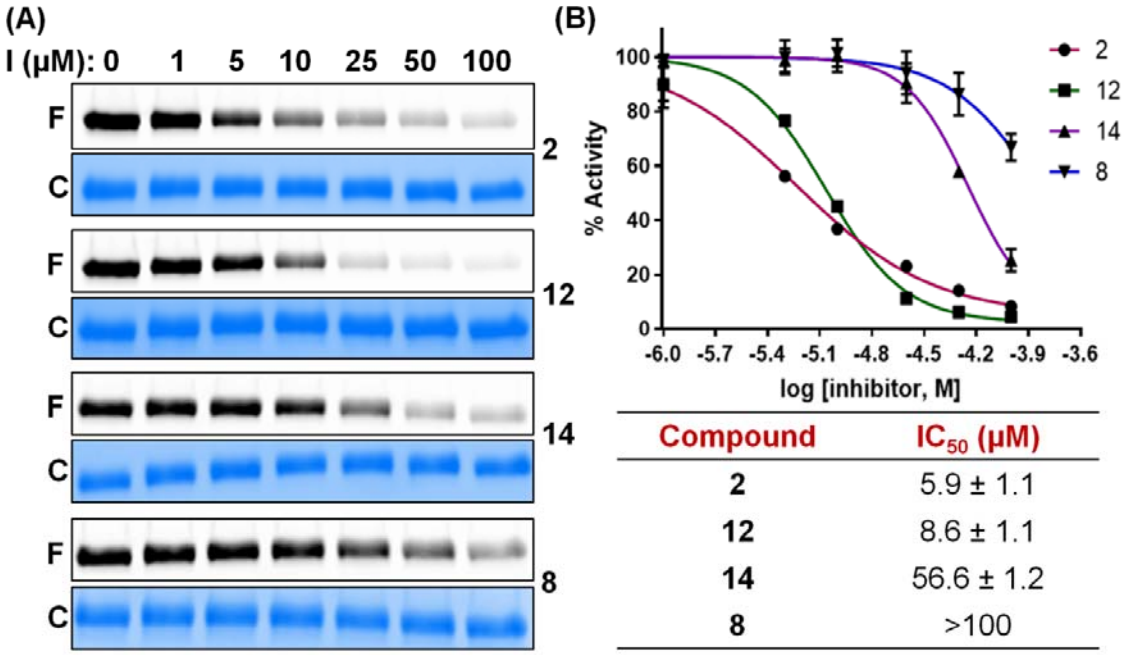


Figure S4. Inhibition of PAD4 by compounds **1**-**22**. **1** (A), **2** (B), **3** (C), **4** (D), **5** (E), **6** (F), **7** (G), **8** (H), **9** (I), **10** (J), **11** (K), **12** (L), **13** (M), **14** (N), **15** (O), **16** (P), **17** (Q), **18** (R), **19** (S), **20** (T), **21** (U) and **22** (V).



Compound	IC_{50} (μM)
2	5.9 ± 1.1
12	8.6 ± 1.1
14	56.6 ± 1.2
8	>100

Figure S5. (A) Labelling of PAD1 by RFA in the presence of compounds **2**, **12**, **14** and **8**. I, F and C refers to the inhibitor, fluorograph and coomassie stain, respectively. (B) Determination of IC_{50} of compounds **2**, **12**, **14** and **8** from the inhibition of RFA-labelling of PAD1.

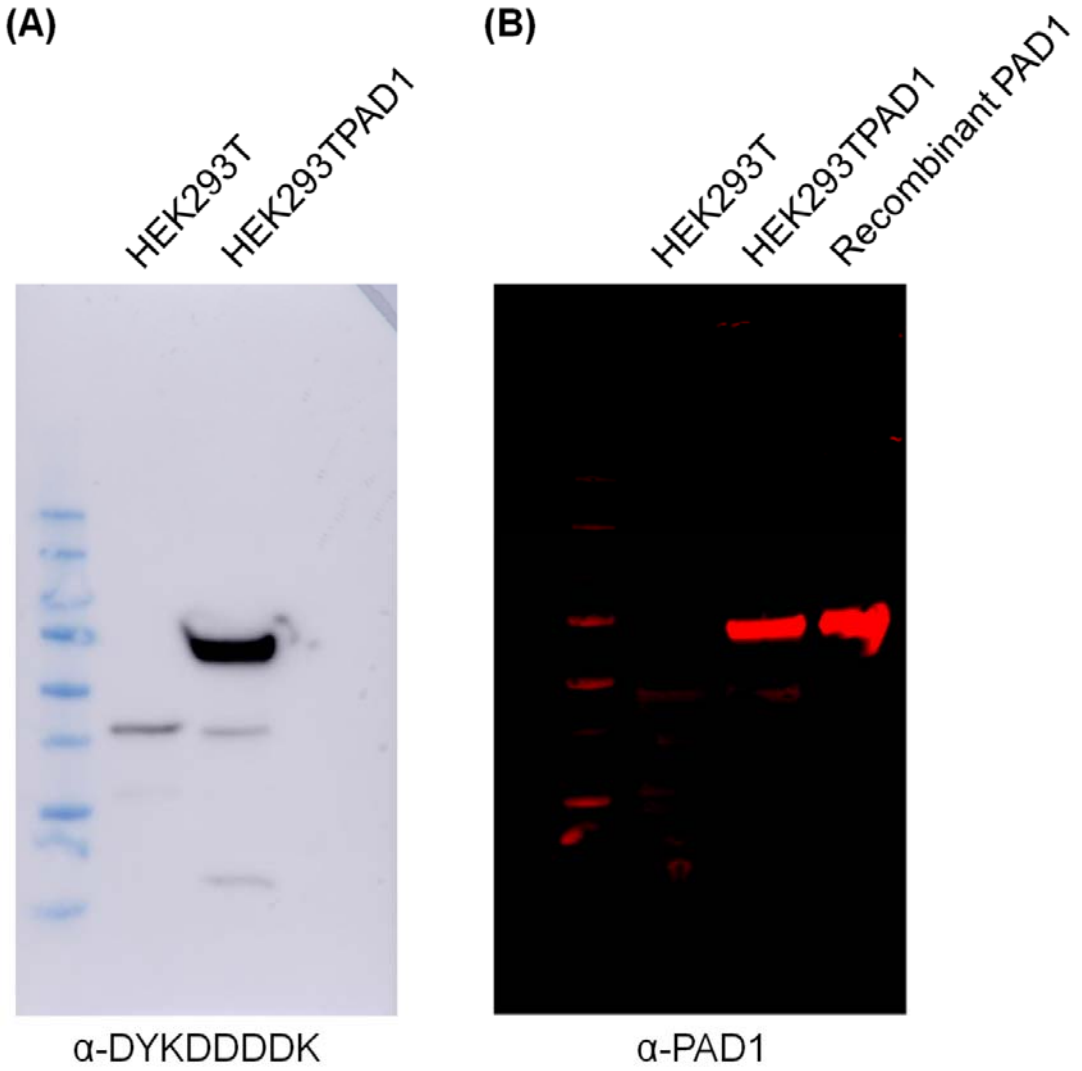


Figure S6. Overexpression of PAD1 in HEK293TPAD1 cells as detected by Western blot analysis of the cell lysates using anti-FLAG (α-DYKDDDDK) (A) and anti-PAD1 (B) antibodies. HEK293T cell lysates and recombinant PAD1 were used as negative and positive controls, respectively.

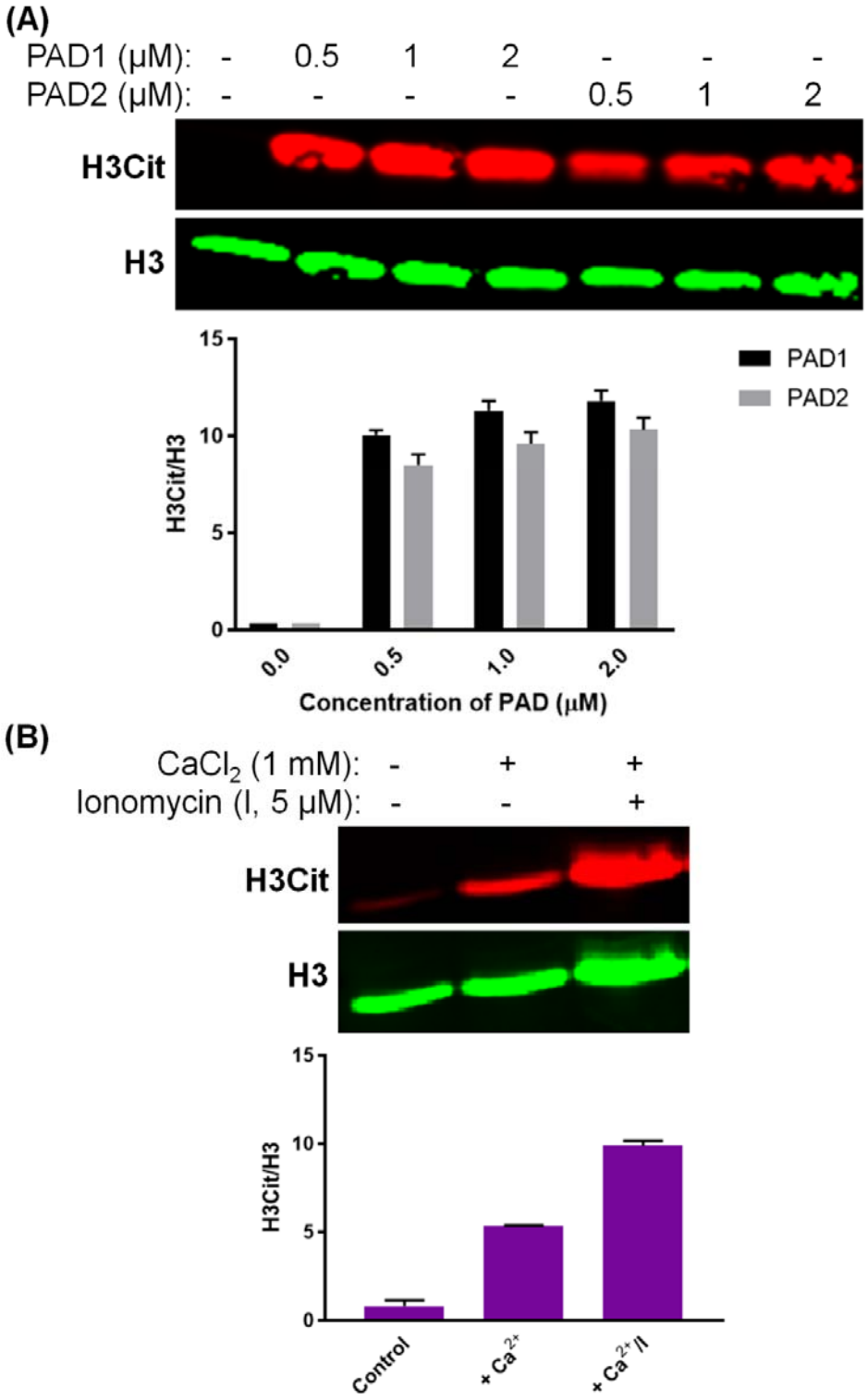


Figure S7. (A) Histone H3 citrullination in HEK293T cell lysates spiked with various concentrations of PAD1 and PAD2. Assay conditions: Lysate (2 mg/mL), PAD (desired concentration), 100 mM TRIS, 50 mM NaCl, 10 mM $CaCl_2$, pH 7.4, 37 ^{0}C, 3 h. PAD2 was used as a positive control. (B) Histone H3 citrullination in HEK293TPAD1 cells in the presence of calcium and plus/minus a calcium ionophore, ionomycin (I). All the experiments were performed at least in duplicate.

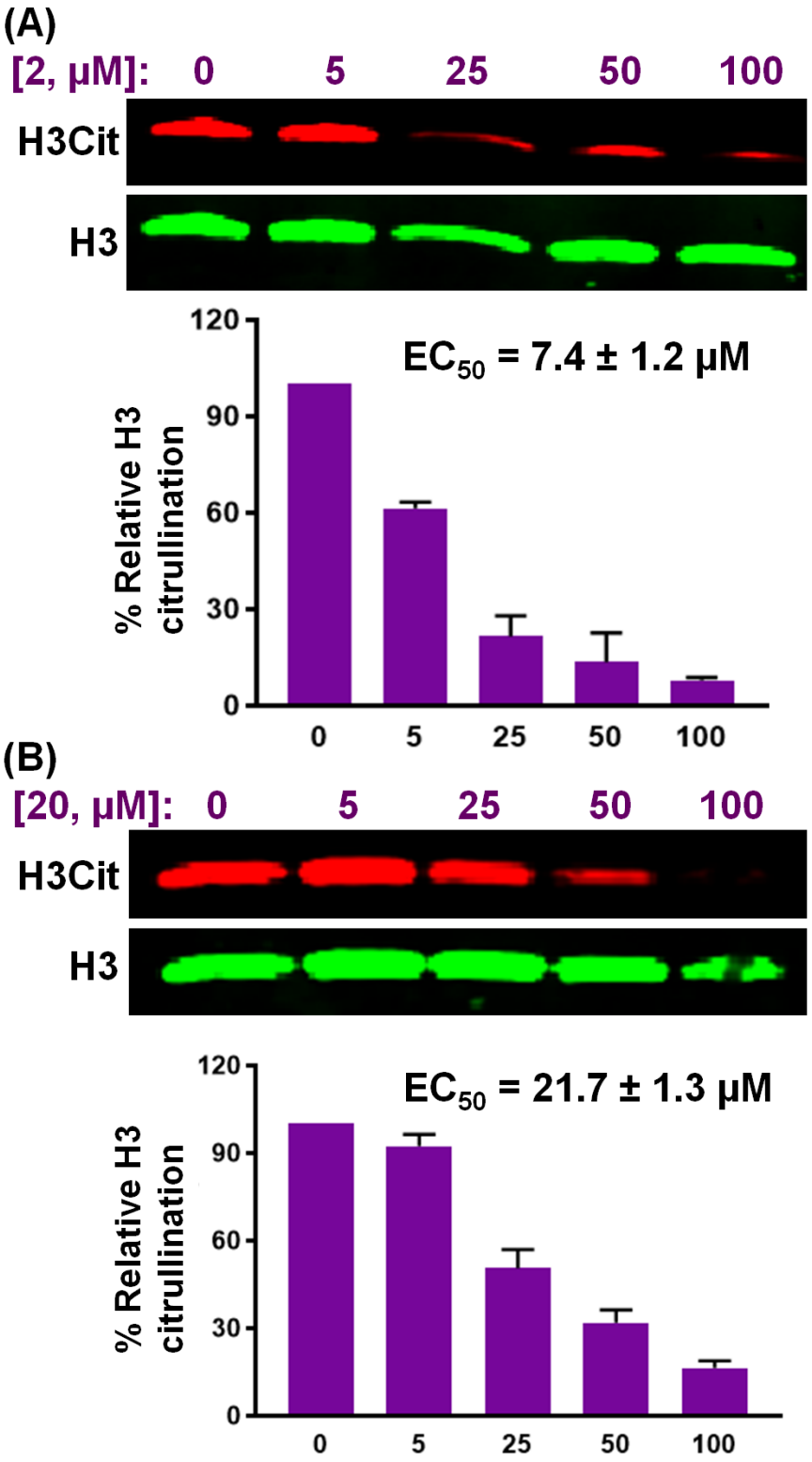
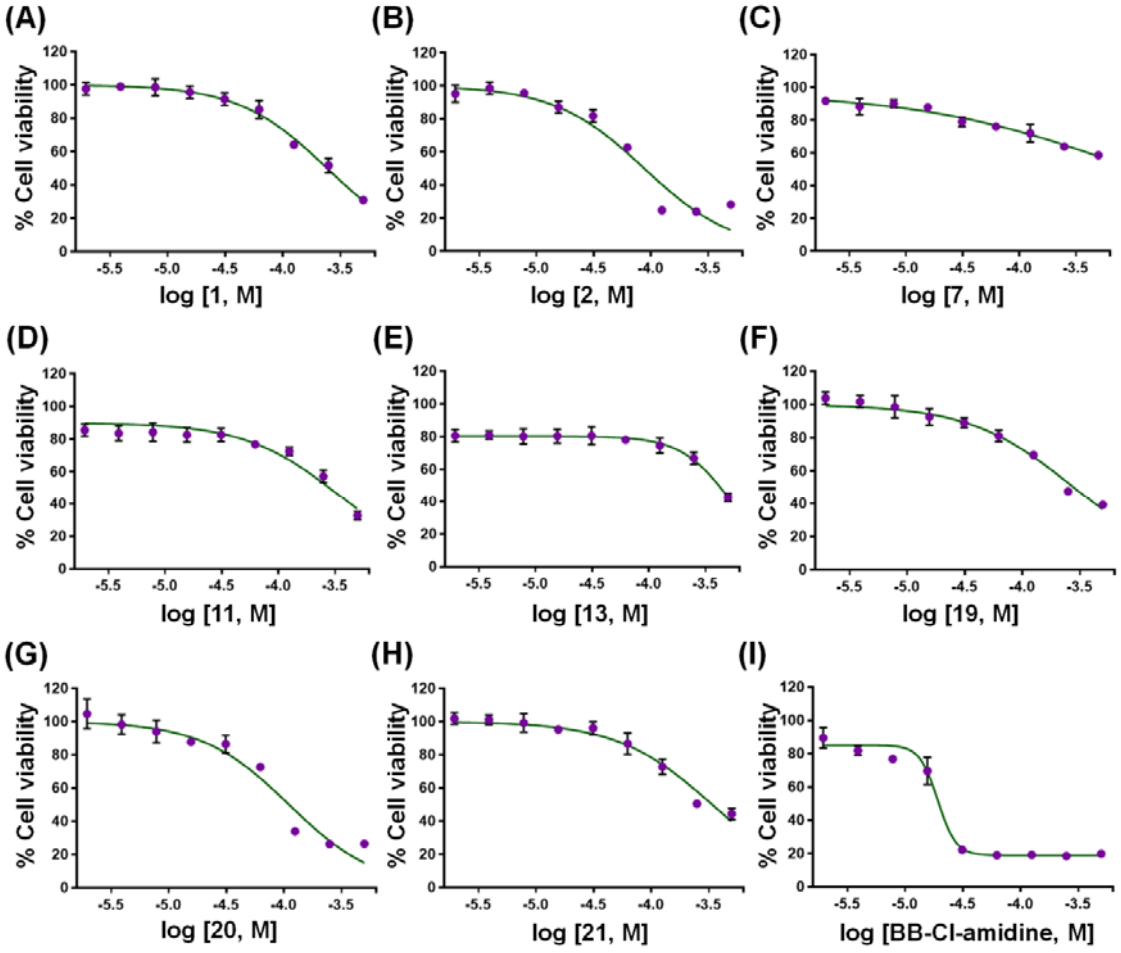


Figure S8. Inhibition of histone H3 citrullination in HEK293TPAD1 cells by compounds **2** (A) and **20** (B). I refers to the inhibitor concentrations in μM. Quantification of each band yielded the H3Cit/H3 ratio, from which % relative H3 citrullination was calculated. Cells were treated with various concentrations of inhibitor in the presence of 1 mM $CaCl_2$ and 5 μM ionomycin (calcium ionophore) in 1X HBS (pH 7.4) at 37 ^{0}C for 3 h.



(J)

Compounds	Cytotoxicity EC_{50} (μM)	Efficacy Index (Cytotoxicity : Potency)
1	243 ± 1.0	16.8
2	88 ± 1.1	11.9
7	>500	NA†
11	351 ± 3.4	11.0
13	>500	NA†
19	280 ± 1.1	9.6
20	111 ± 1.1	5.1
21	338 ± 1.1	NA‡
BB-Cl-amidine	19.2 ± 1.1	NA‡

NA: Not applicable. †Uncertain EC_{50}. ‡Potancy was not determined

Figure S9. Inhibition of cell-proliferation of HEK293TPAD1 cells by compounds **1** (A), **2** (B), **7** (C), **11** (D), **13** (E), **19** (F), **20** (G), **21** (H) and BB-Cl-amidine (I). The EC_{50} values are listed in the table (J). Potency refers to the efficiency of these compounds to inhibit histone H3 citrullination in HEK293TPAD1 cells. Comparison of the EC_{50} values for cytotoxicity and potency (efficacy index) indicates that these compounds inhibit histone H3 citrullination in HEK293TPAD1 cells without causing significant toxicity.

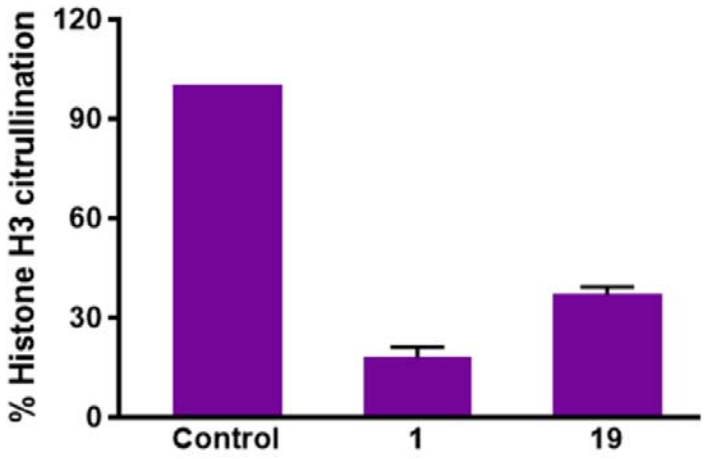


Figure S10. Quantification of inhibition of histone H3 citrullination in mice embryos by **1** and **19**. Mouse embryos (4-cell) cultured with DMSO (control), **1** or **19** (100 μM) were incubated with anti-histone H3Cit2/8/17 and DAPI as immunofluorescence staining agents for citrullinated histone H3 and DNA, respectively. Both **1** and **19** significantly inhibit histone H3 citrullination.

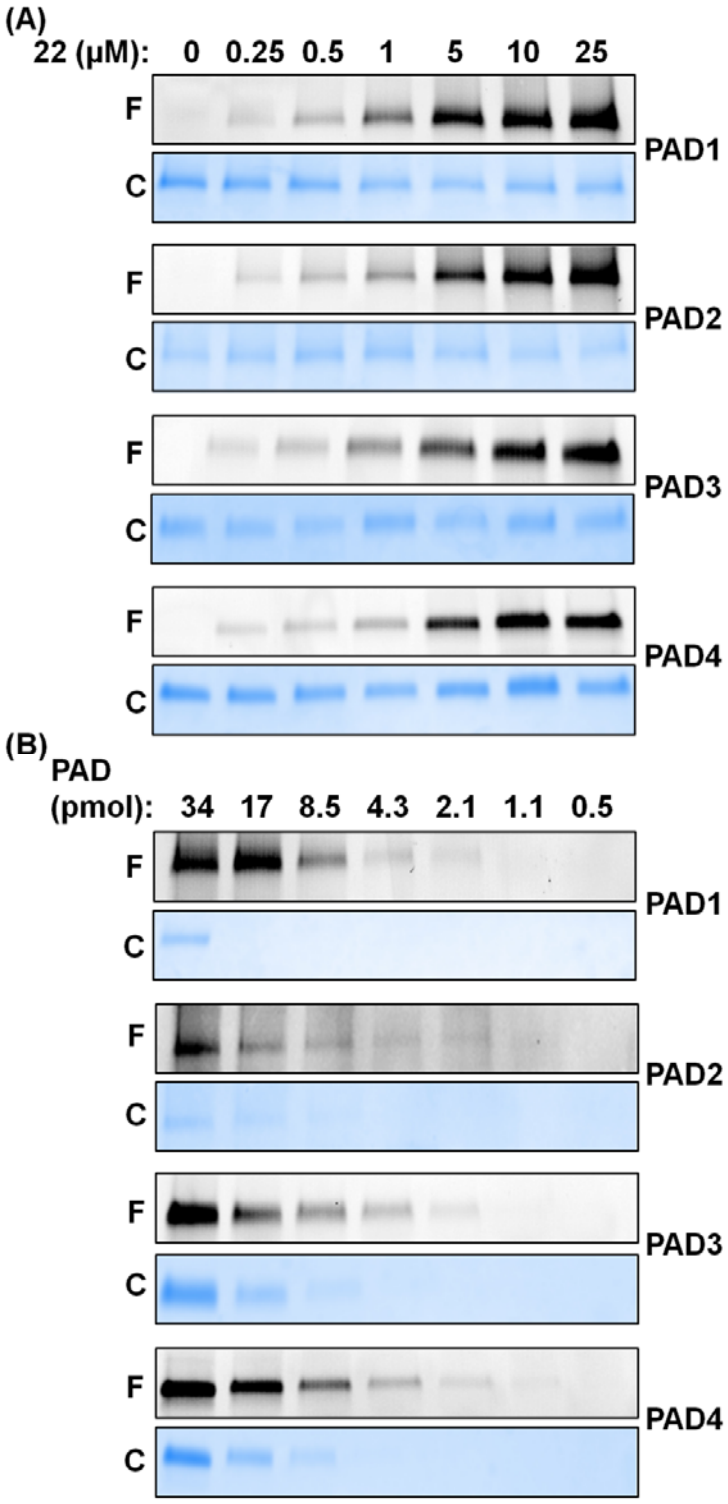


Figure S11. (A) Dose-dependent in vitro labelling of PADs by compound **22**. (B) Limit of detection of PADs by compound **22**. 1.4 μM of PAD and 10 μM of **22** were used for the dose-dependence and limit of detection studies, respectively. F and C refer to fluorograph and Coomassie stain, respectively.

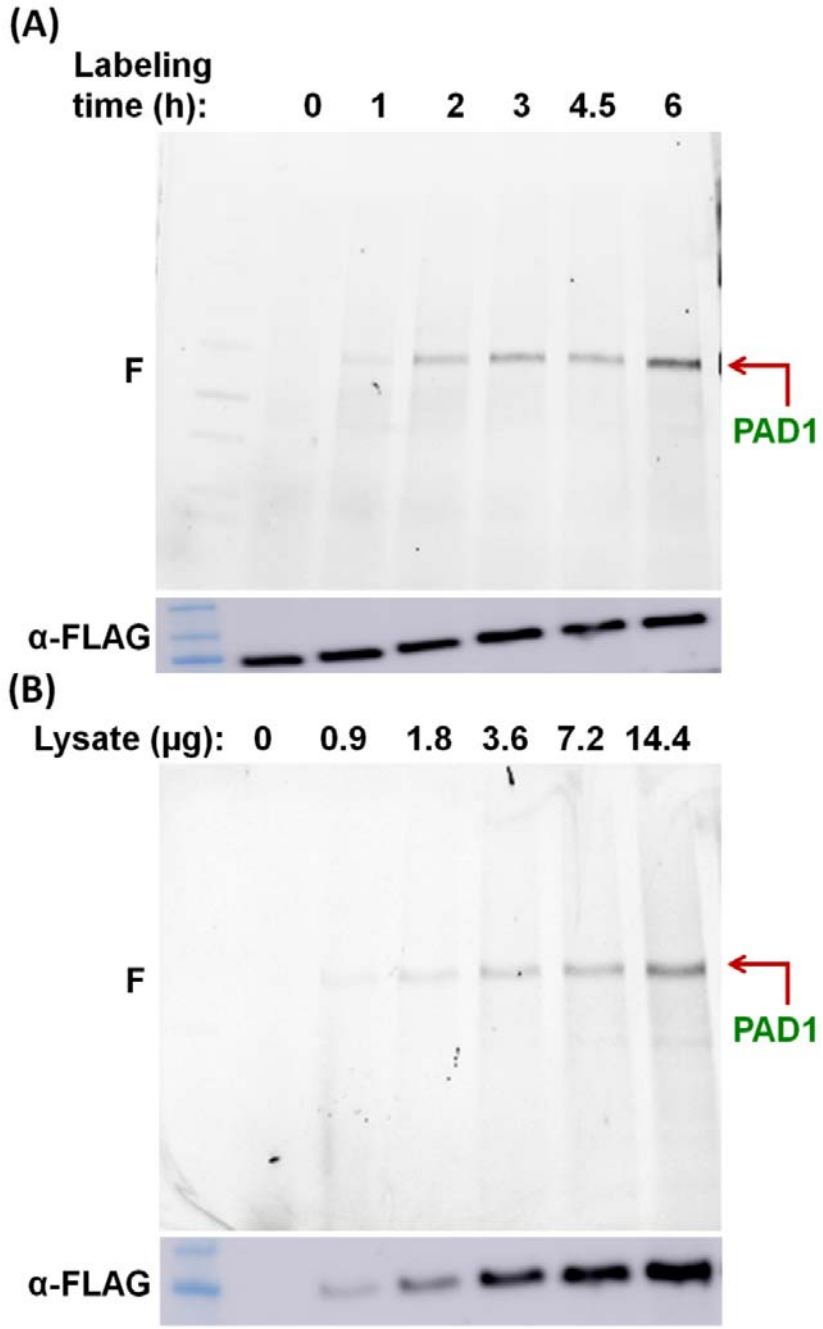


Figure S12. (A) Time-dependent labelling of PAD1 in HEK293TPAD1 cells by **21**. HEK293TPAD1 cells were treated with **21** (5 µM) in serum-free DMEM for various times. After the treatment, cells were harvested and flash-frozen in liquid nitrogen. Cell lysates were treated with TAMRA-N_3, TCEP, TBTA and $CuSO_4$ for 2 h at room temperature. Time-dependent labelling of PAD1 by **21** was recorded on a Typhoon scanner. (B) Limit of detection of PAD1 in HEK293TPAD1 cells by **21**. For this study, cells were treated with 5 µM of **21** in serum-free DMEM for 3 h and decreasing amounts (indicated by total protein content) of lysates were clicked with TAMRA-N_3 in the presence of TCEP, TBTA and freshly prepared $CuSO_4$. PAD1 content in HEK293TPAD1 cells was found to be 0.06% of the total protein and taking this into account, limit of detection of PAD1 in HEK293TPAD1 cells by **21** was calculated to be 72 pmol. F stands for fluorograph.

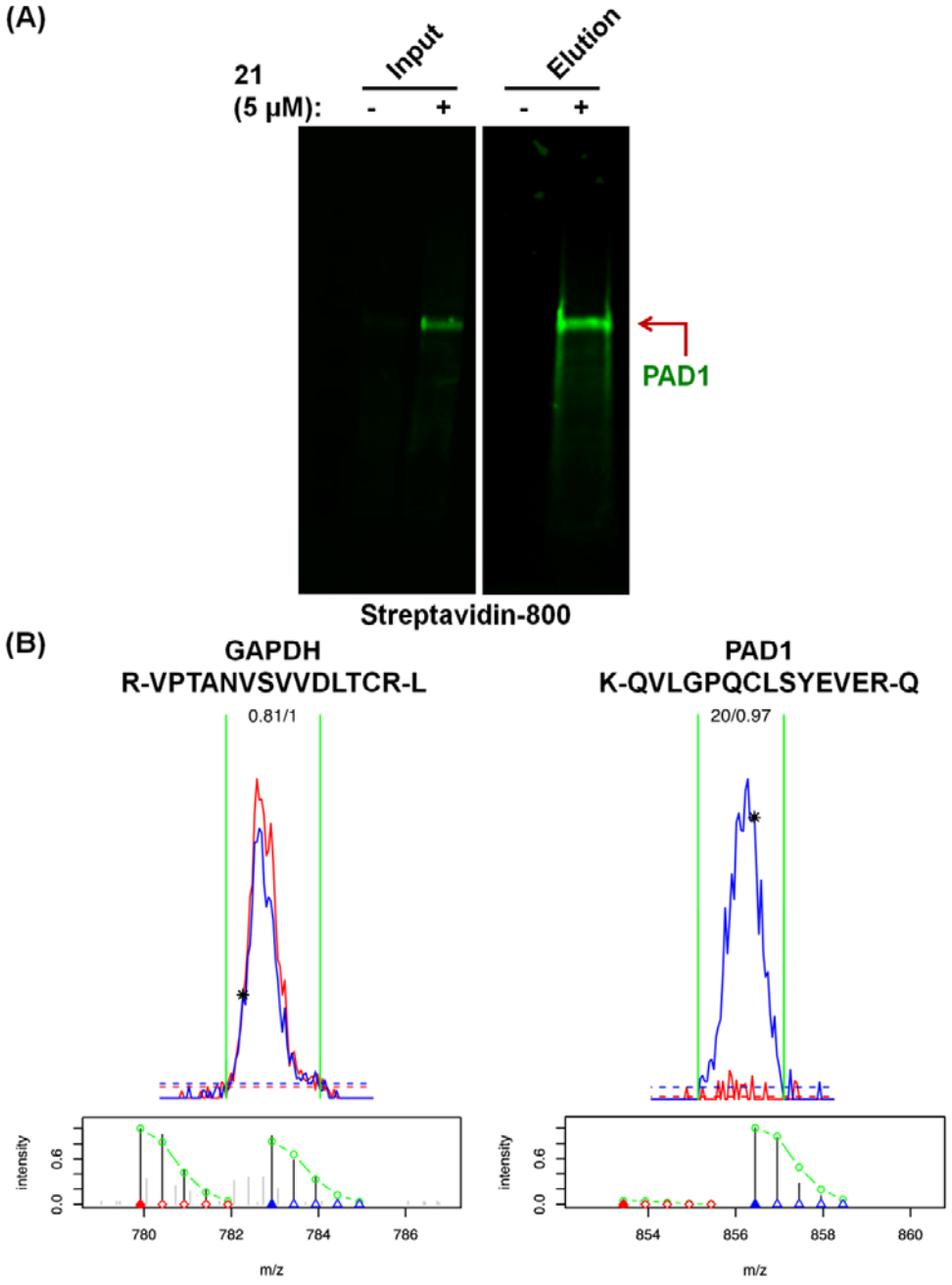


Figure S13. (A) Pull-down of PAD1 in HEK293TPAD1 cells on Streptavidin-agarose beads using **21** and Biotin-N_3. HEK293TPAD1 cells were treated with 5 μM **21** in the presence of 1 mM $CaCl_2$ and 5 μM ionomycin at 37 ^{0}C for 3 h, and the labelled proteins were clicked with Biotin-N_3 in the presence of TCEP, TBTA and $CuSO_4$. Biotinylated proteins were then selectively captured on Streptavidin-agarose beads and were eluted by heating the beads in a mixture of 4 M urea, 15 mM biotin and 2.5X SDS-PAGE loading dye at 95 ^{0}C for 30 min. The input and elution were analyzed by western blot using streptavidin-800. (B) Representative isotopic comparison of peptides from GAPDH and PAD1 obtained by trypsin digestion of the pulled-down proteins from HEK293TPAD1 cells by **21**. **21**-treated and control samples were labelled with heavy (blue) and light (red) formaldehyde, respectively. GAPDH shows a ratio close to 1 (i.e. no change), while PAD1 shows a strong enrichment with a ratio of ~20.

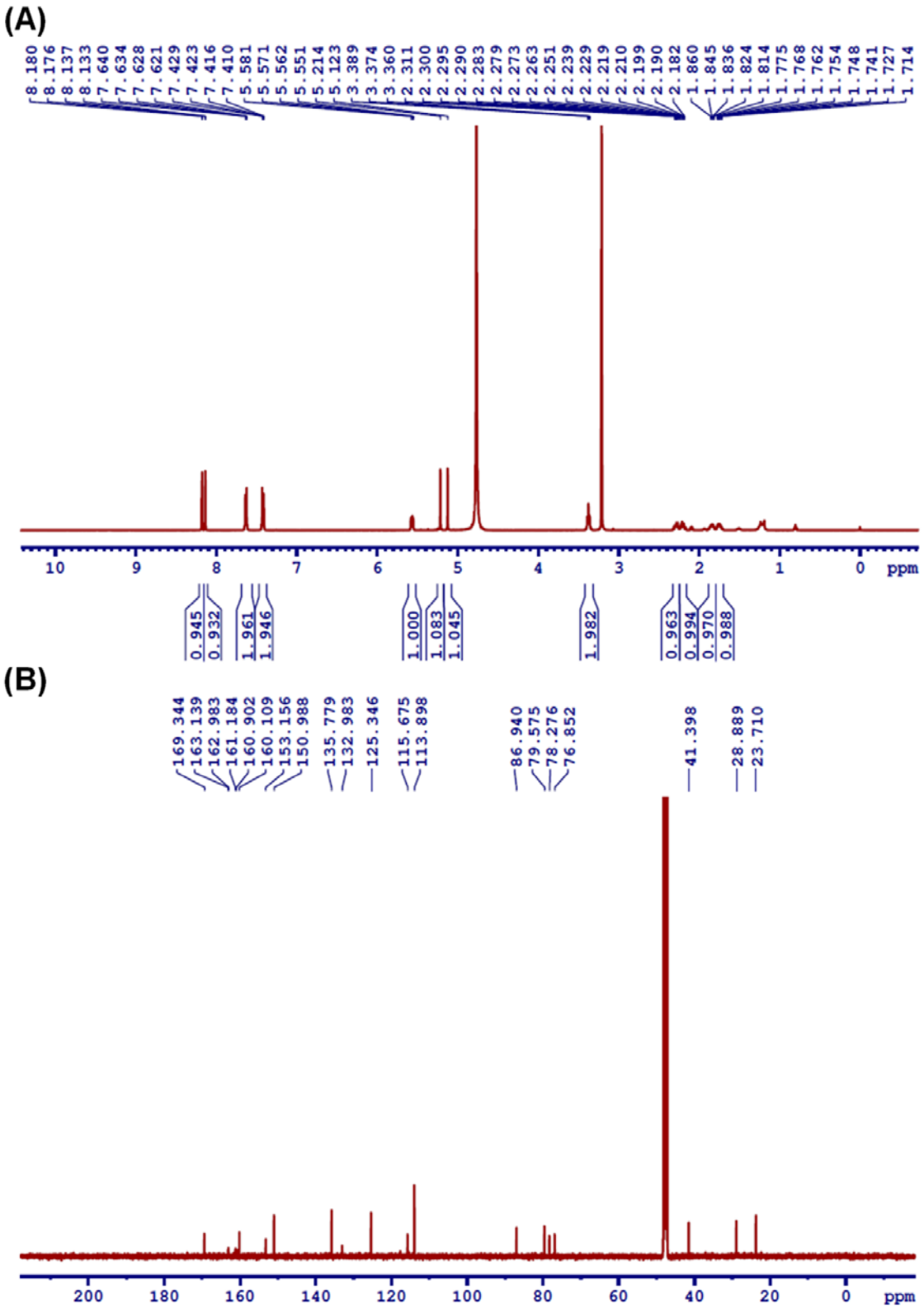


Figure S14. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **1** in CD_3OD.

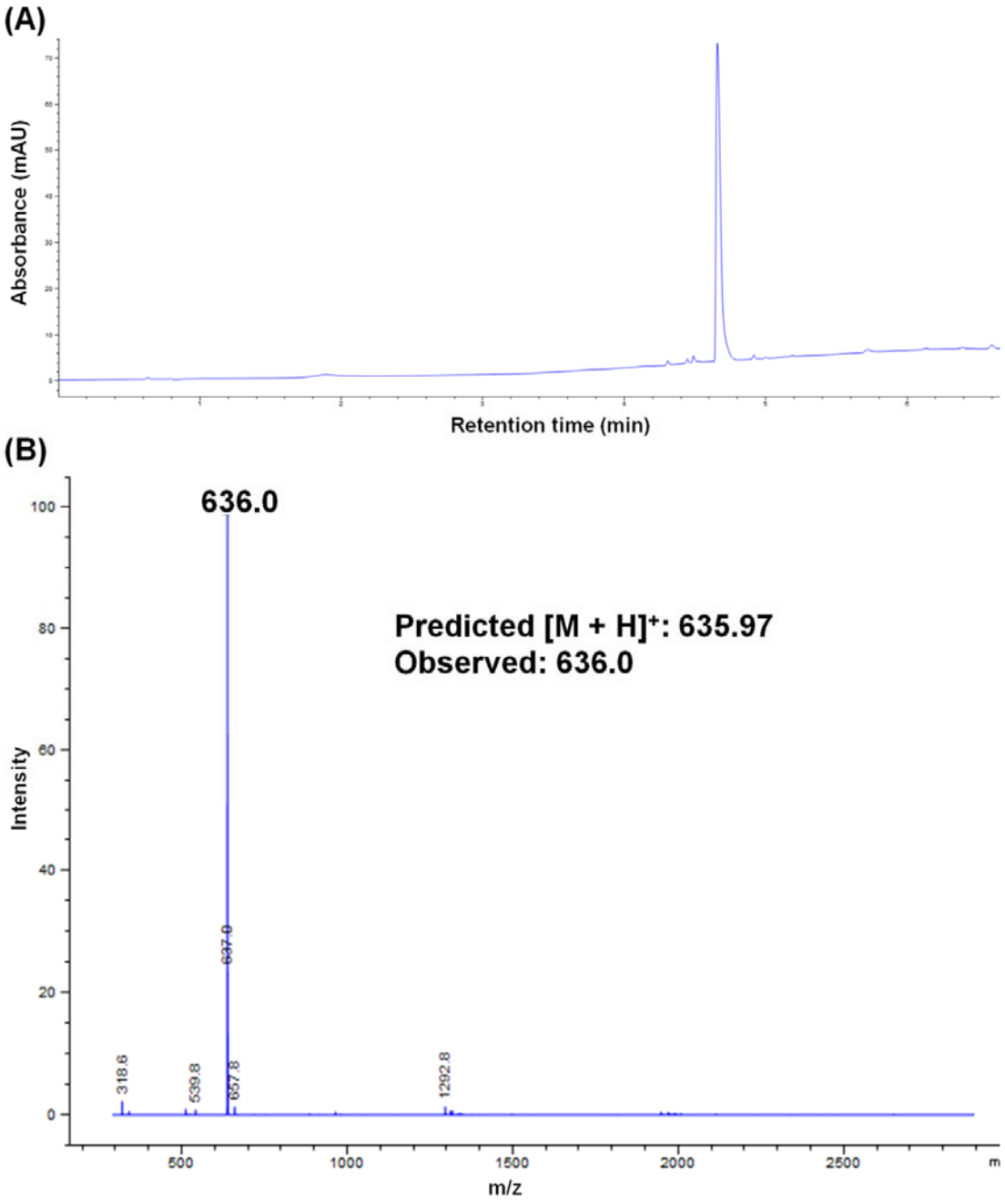


Figure S15. HPLC trace (A) and ESI-Mass spectra (B) of compound **1**.

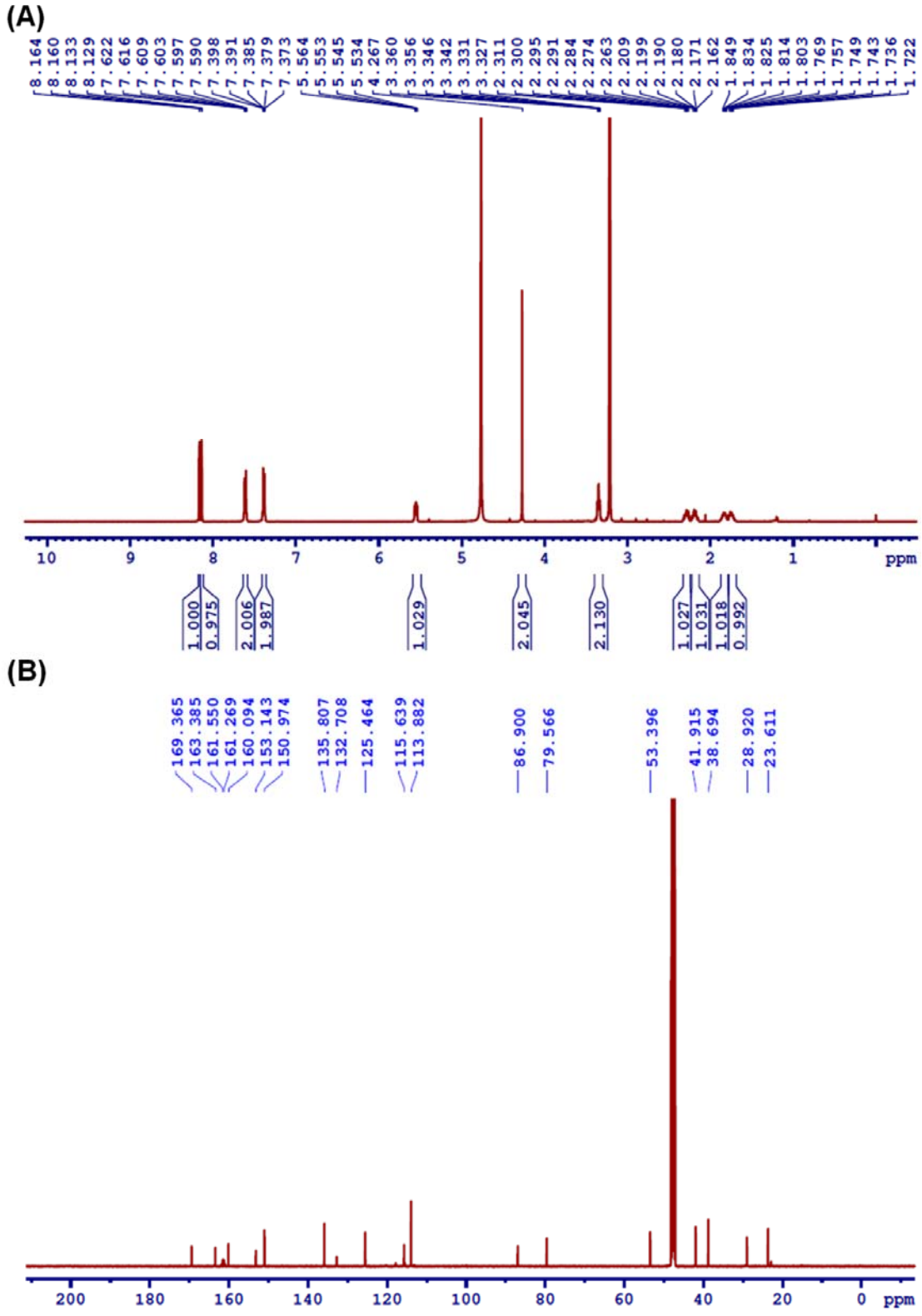


Figure S16. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **2** in CD_3OD.

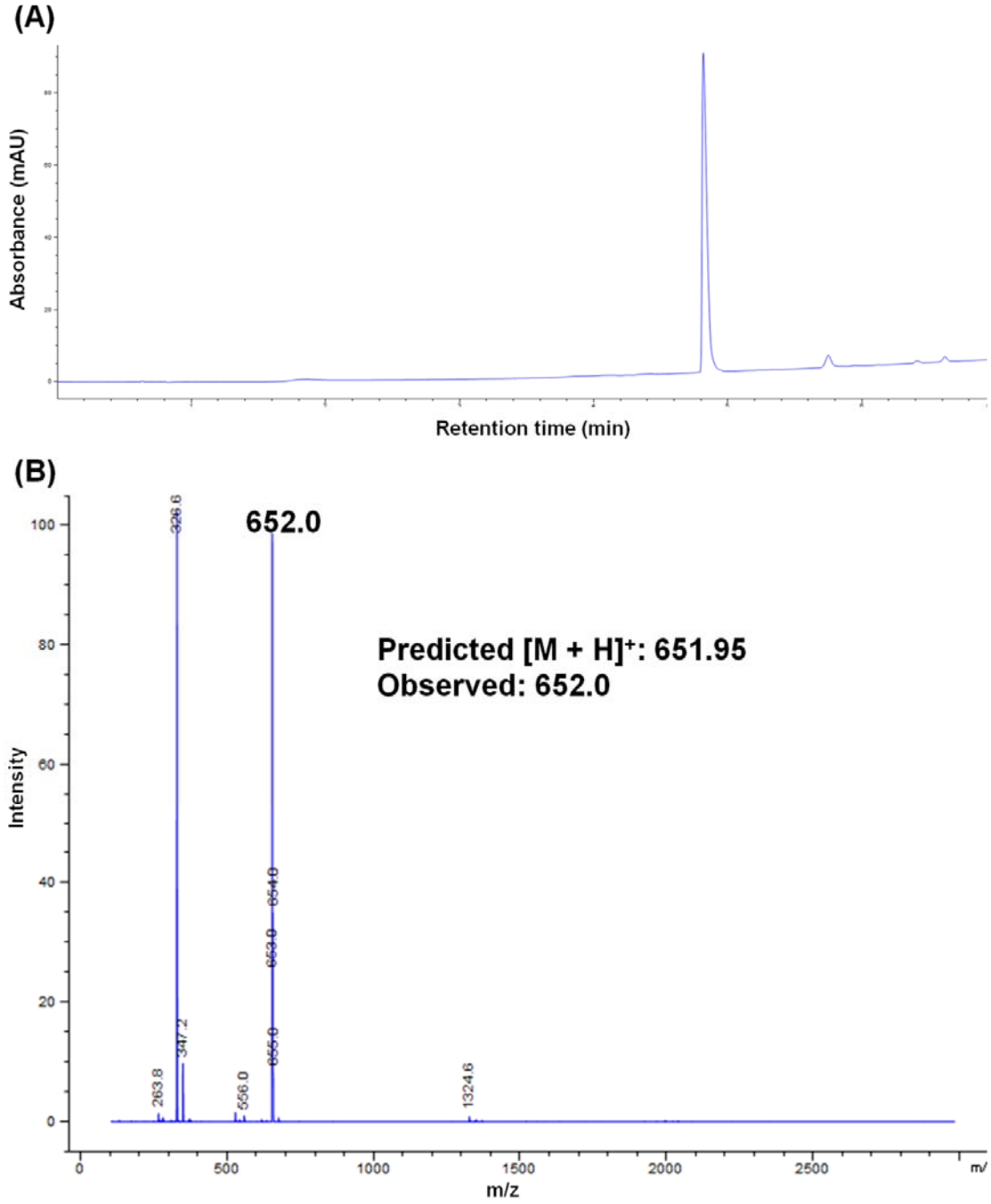


Figure S17. HPLC trace (A) and ESI-Mass spectra (B) of compound **2**.

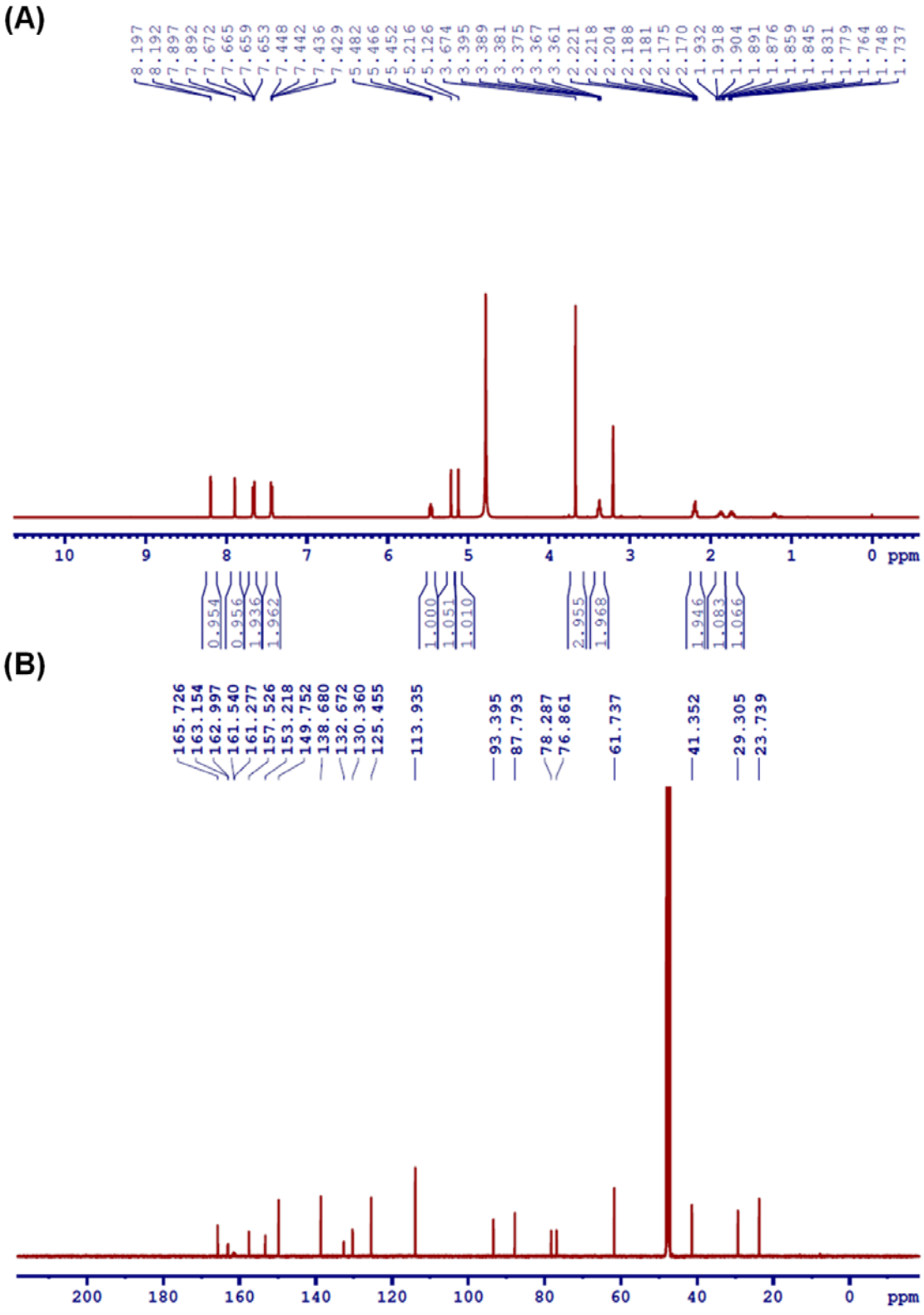


Figure S18. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **3** in CD_3OD.

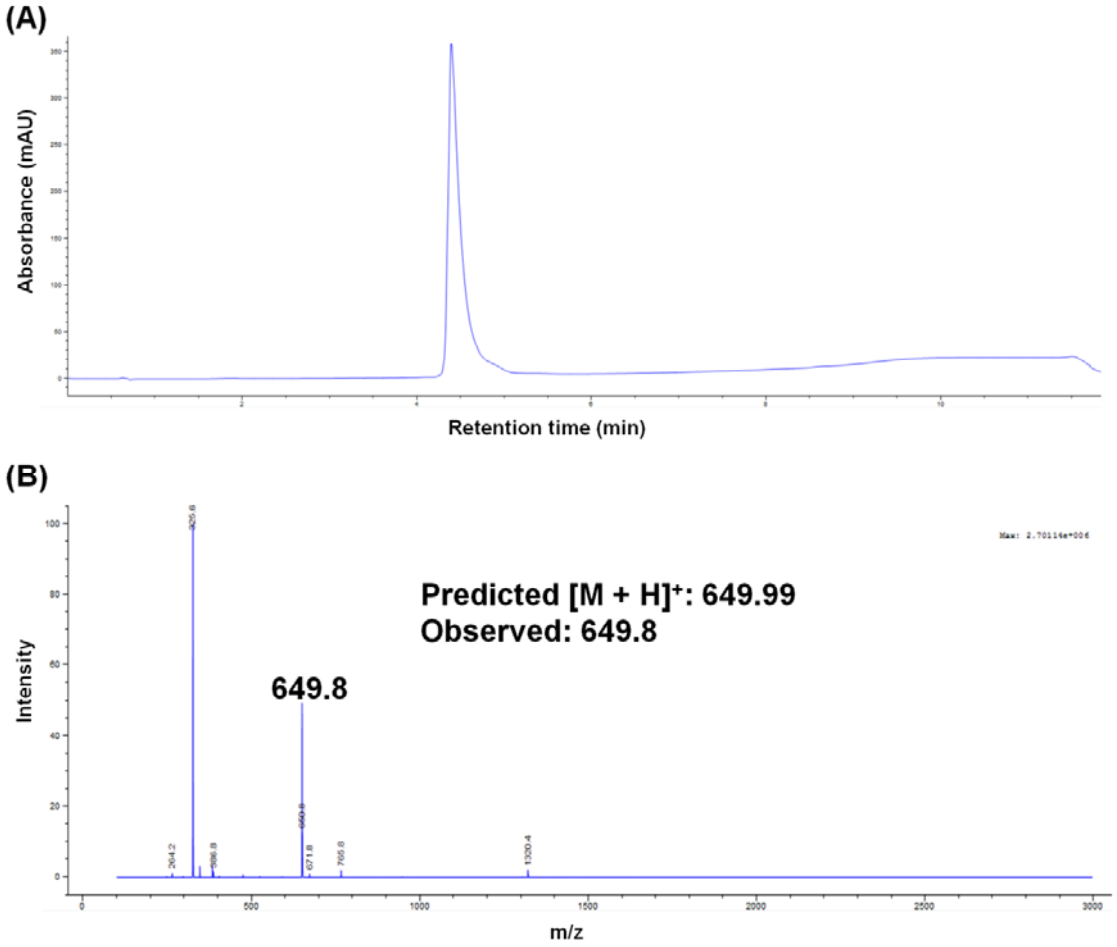


Figure S19. HPLC trace (A) and ESI-Mass spectra (B) of compound **3**.

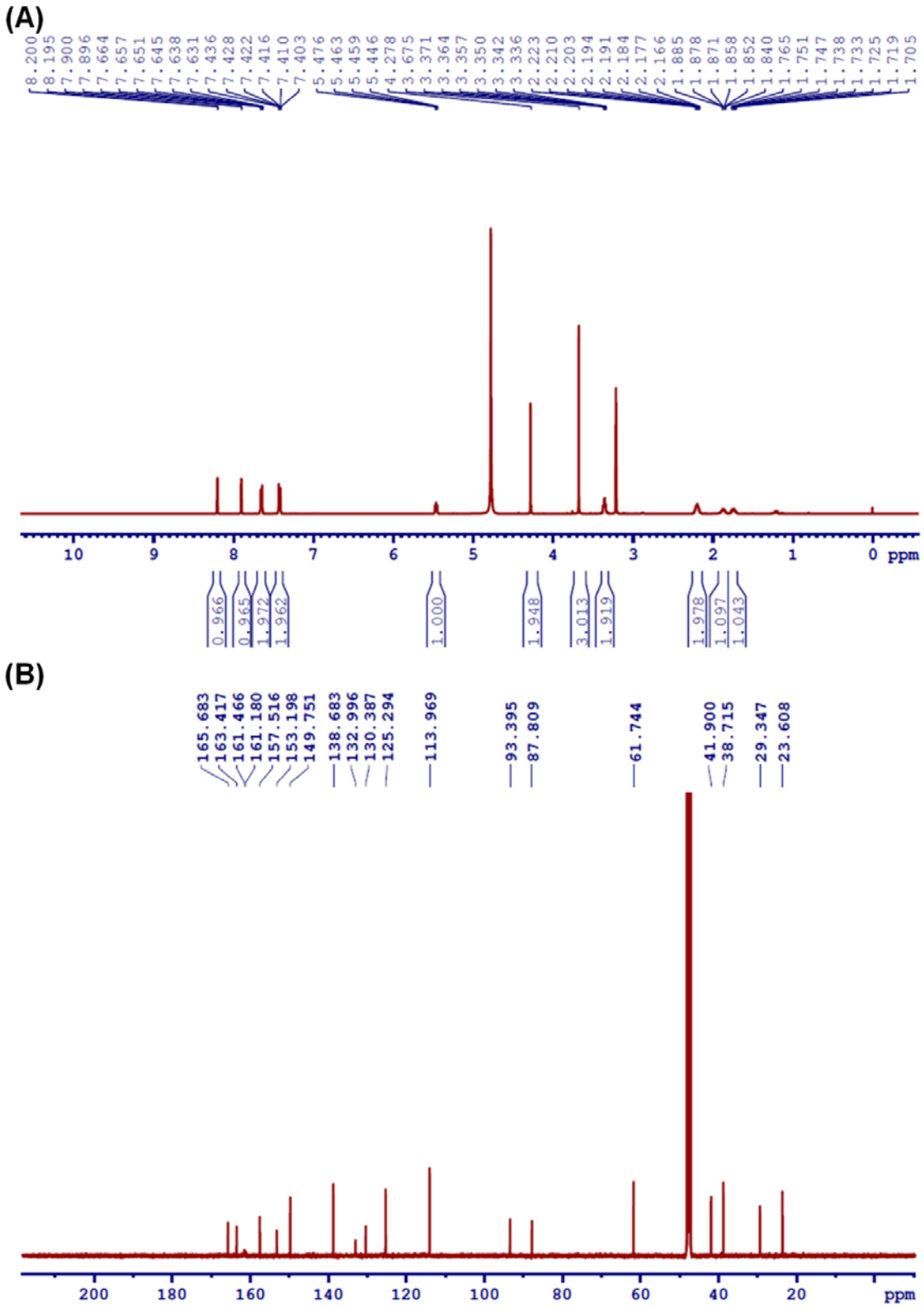


Figure S20. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **4** in CD_3OD.

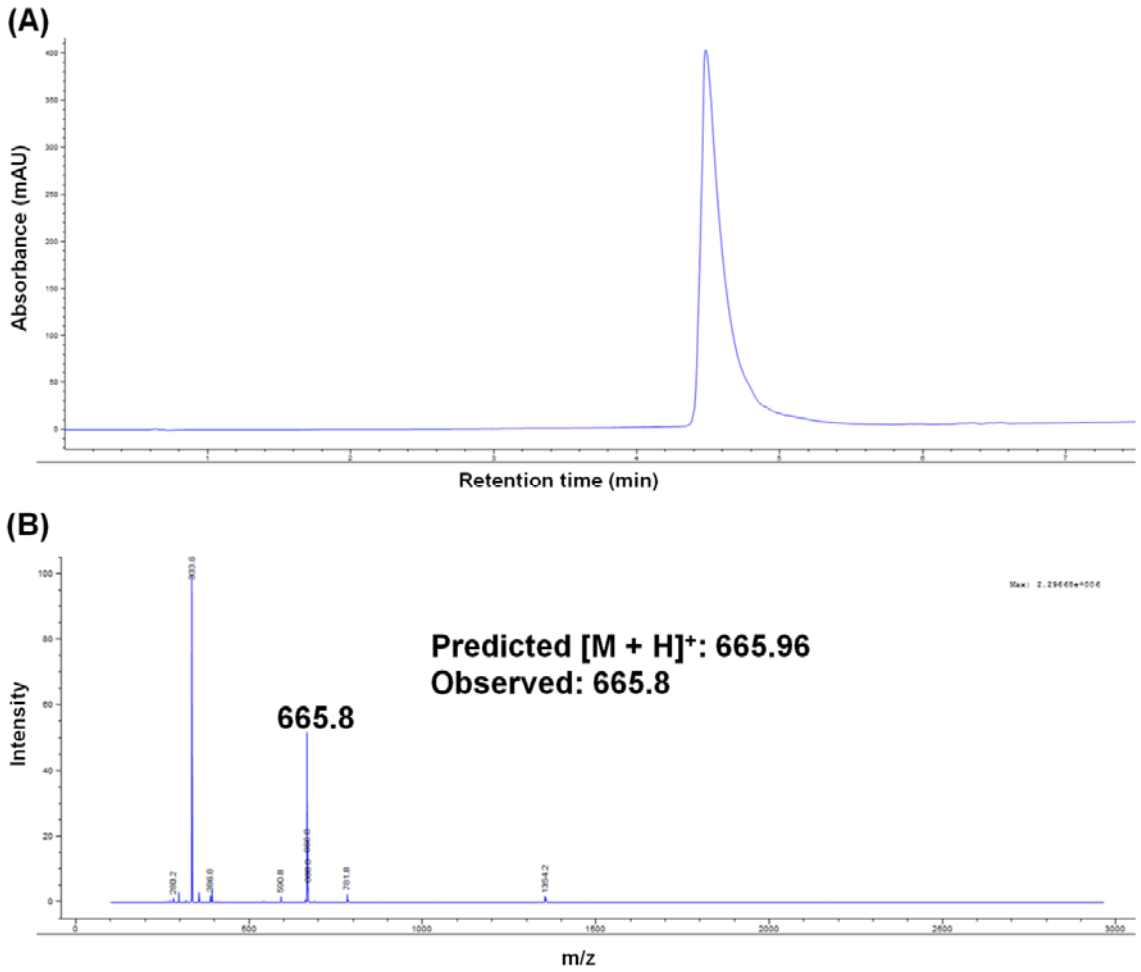


Figure S21. HPLC trace (A) and ESI-Mass spectra (B) of compound **4**.

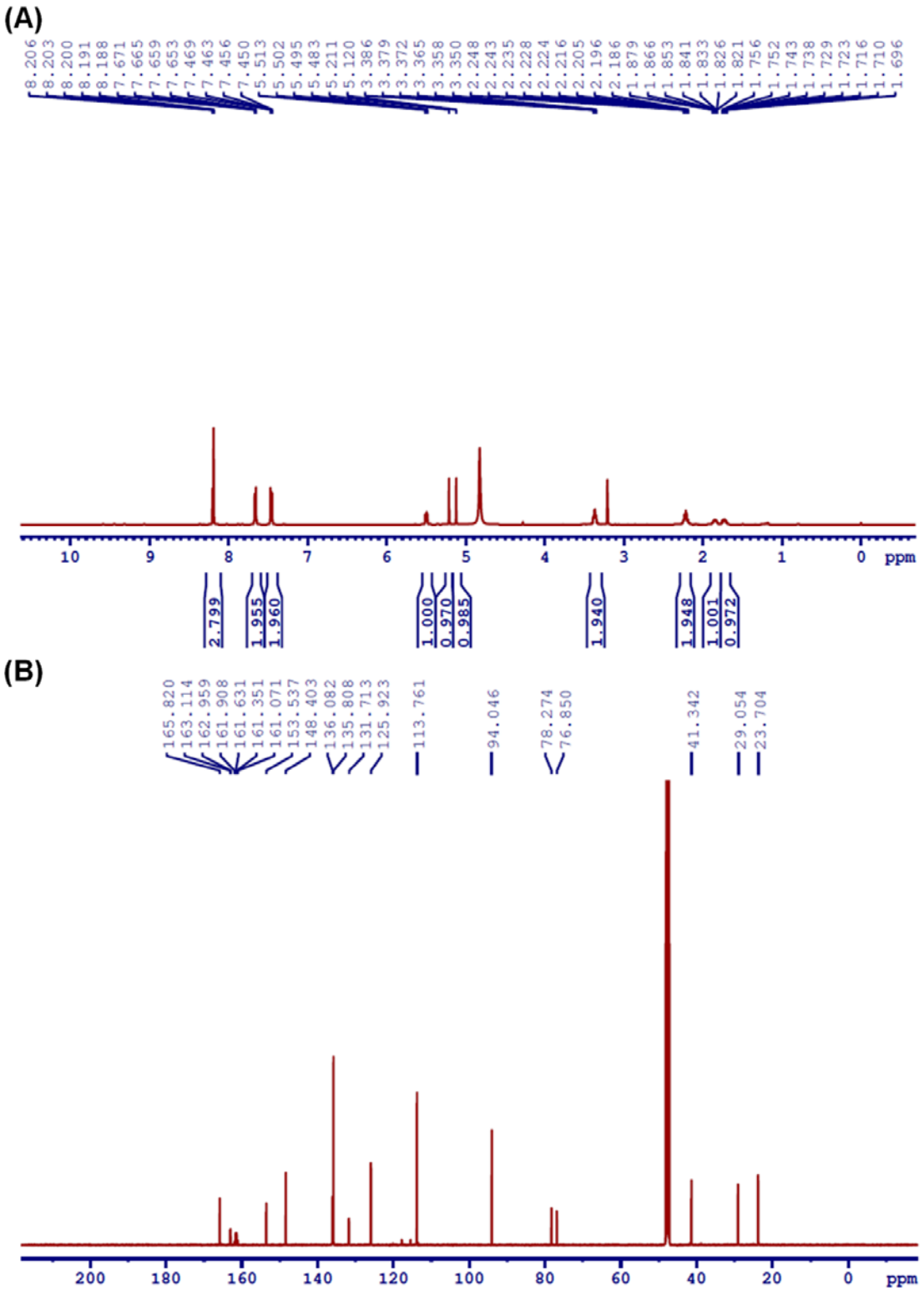


Figure S22. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **5** in CD_3OD.

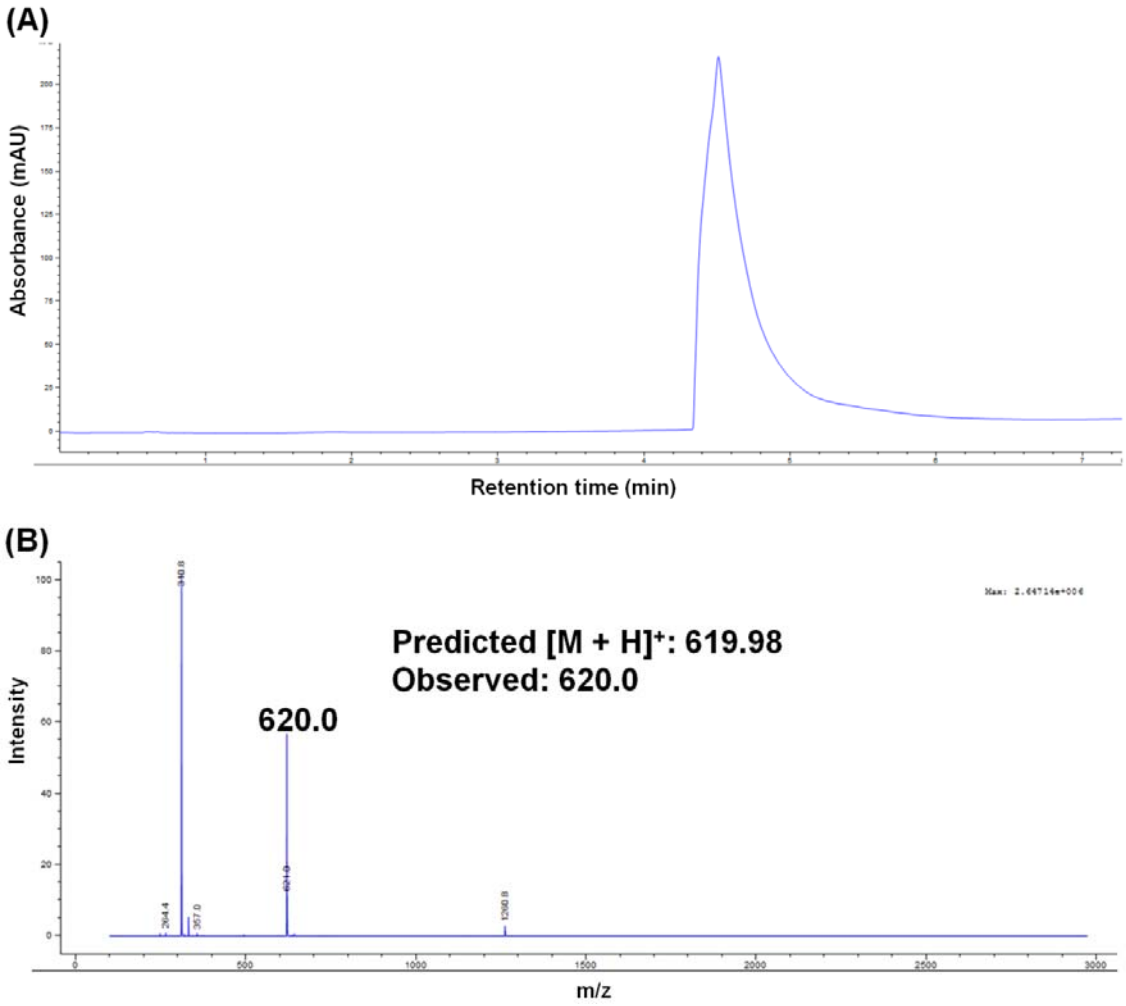


Figure S23. HPLC trace (A) and ESI-Mass spectra (B) of compound **5**.

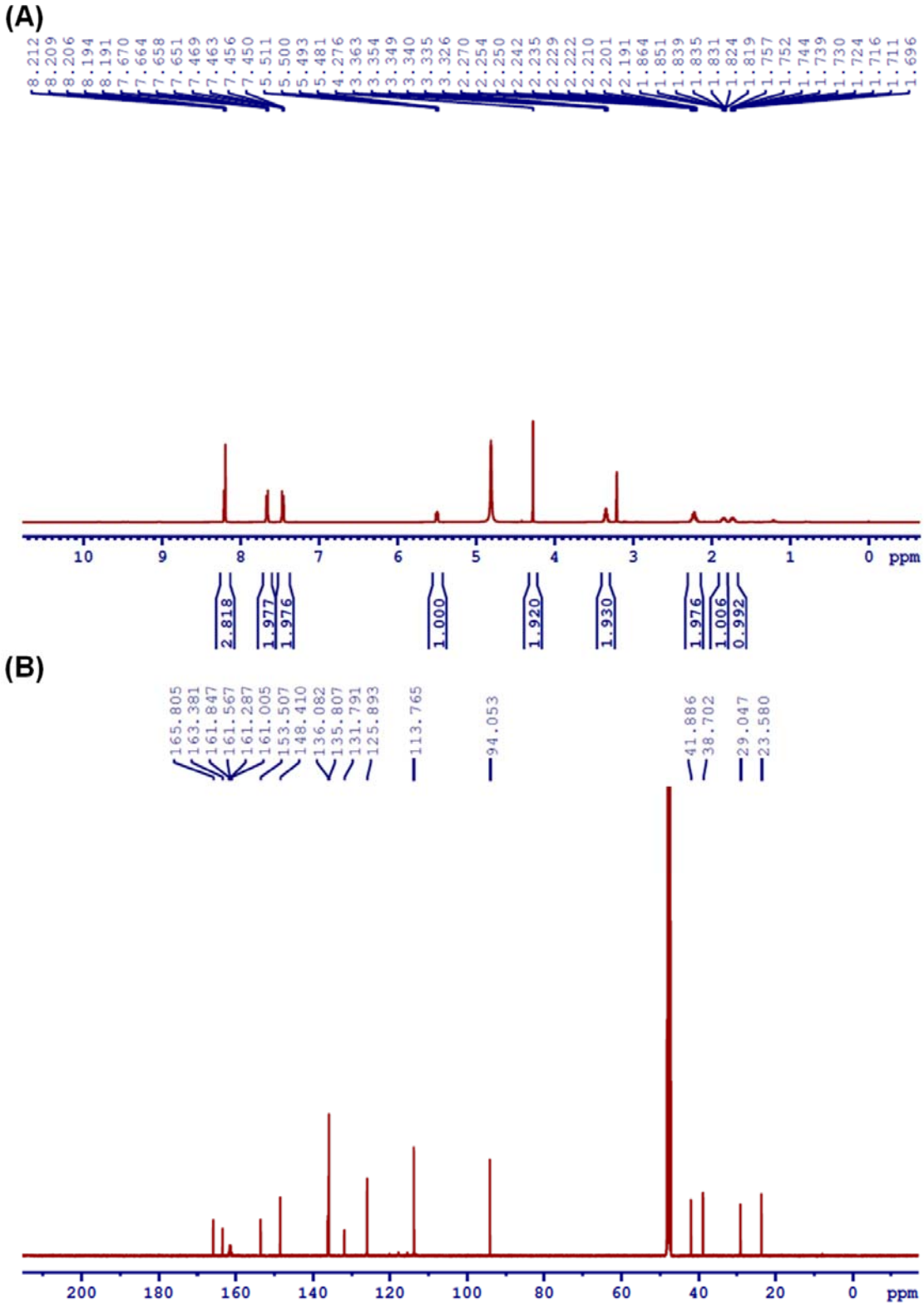


Figure S24. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **6** in CD_3OD.

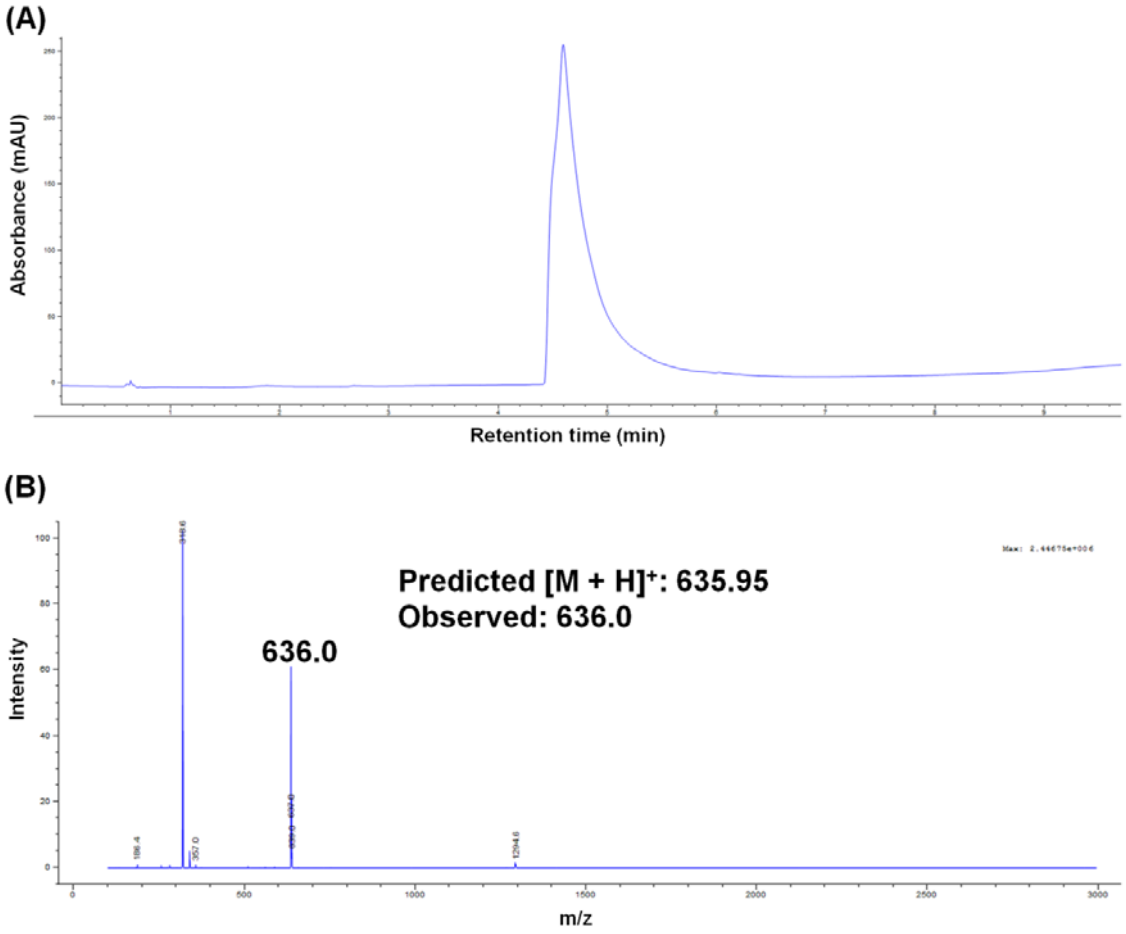


Figure S25. HPLC trace (A) and ESI-Mass spectra (B) of compound **6**.

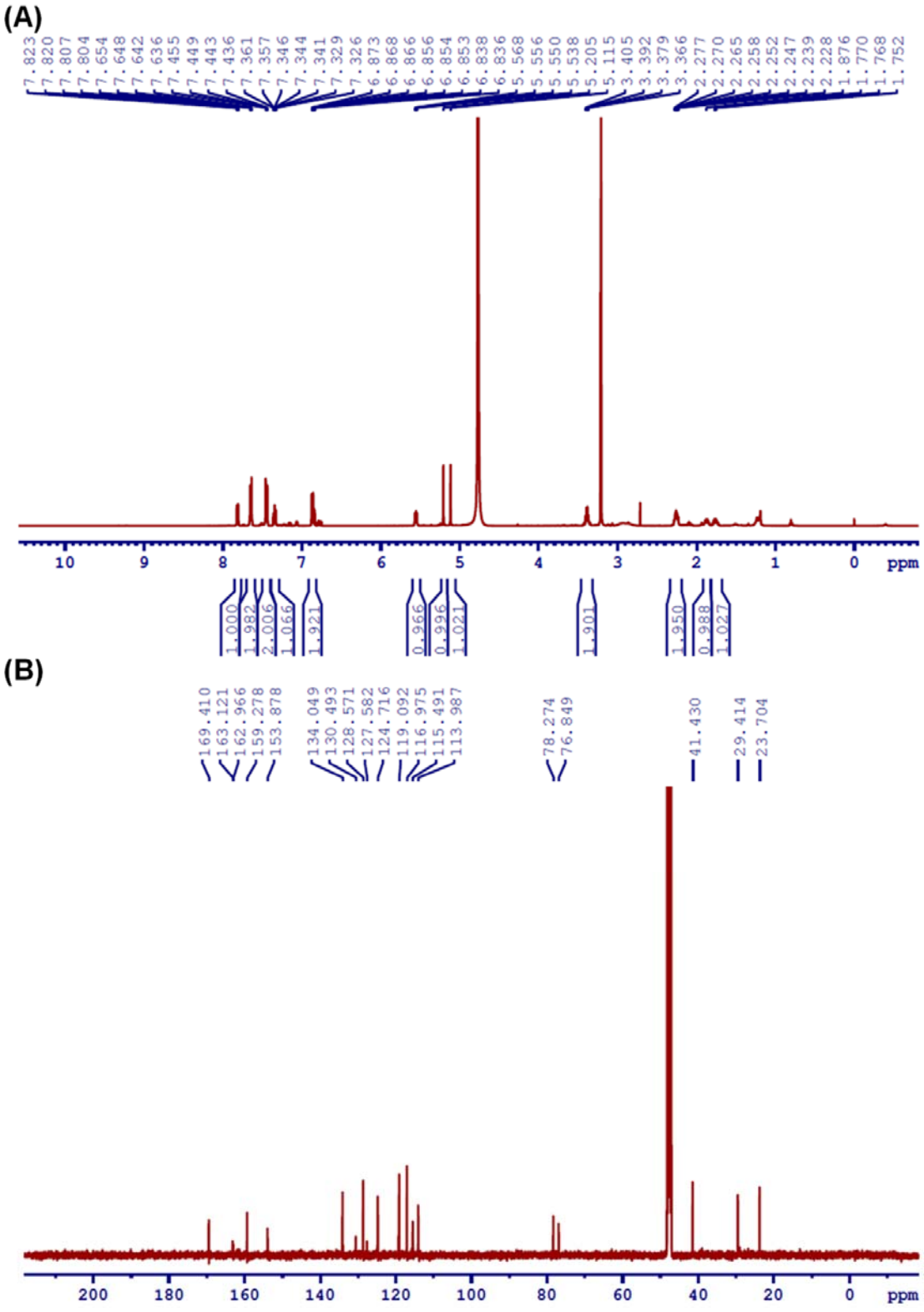


Figure S26. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **7** in CD_3OD.

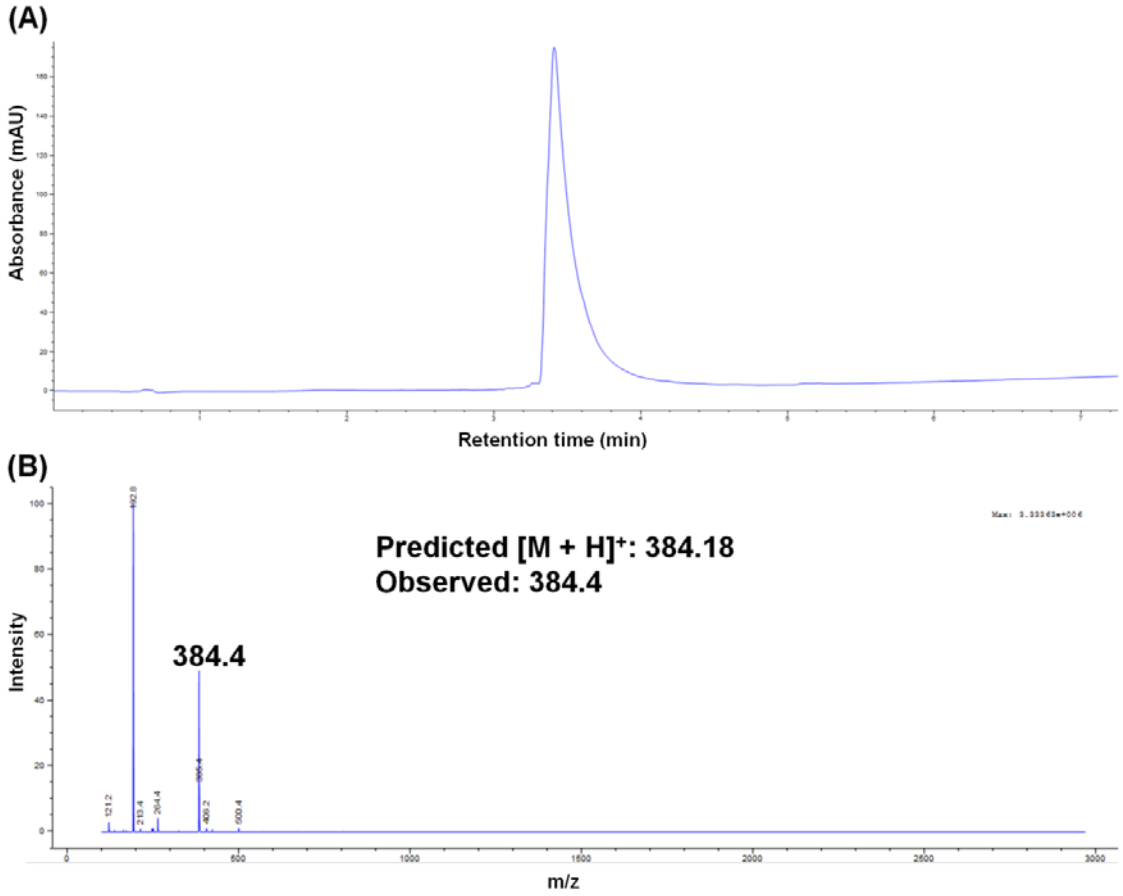


Figure S27. HPLC trace (A) and ESI-Mass spectra (B) of compound **7**.

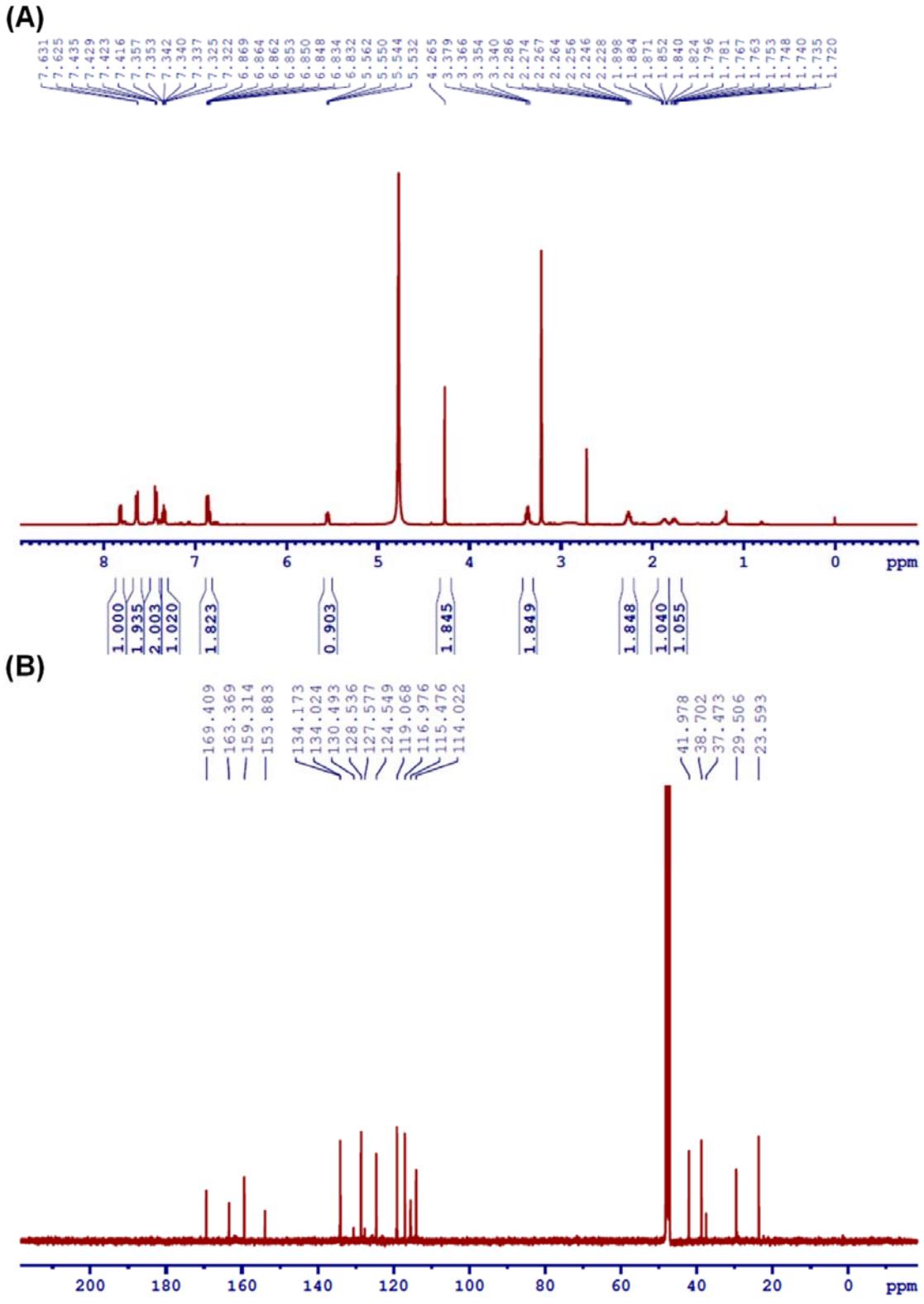


Figure S28. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **8** in CD_3OD.

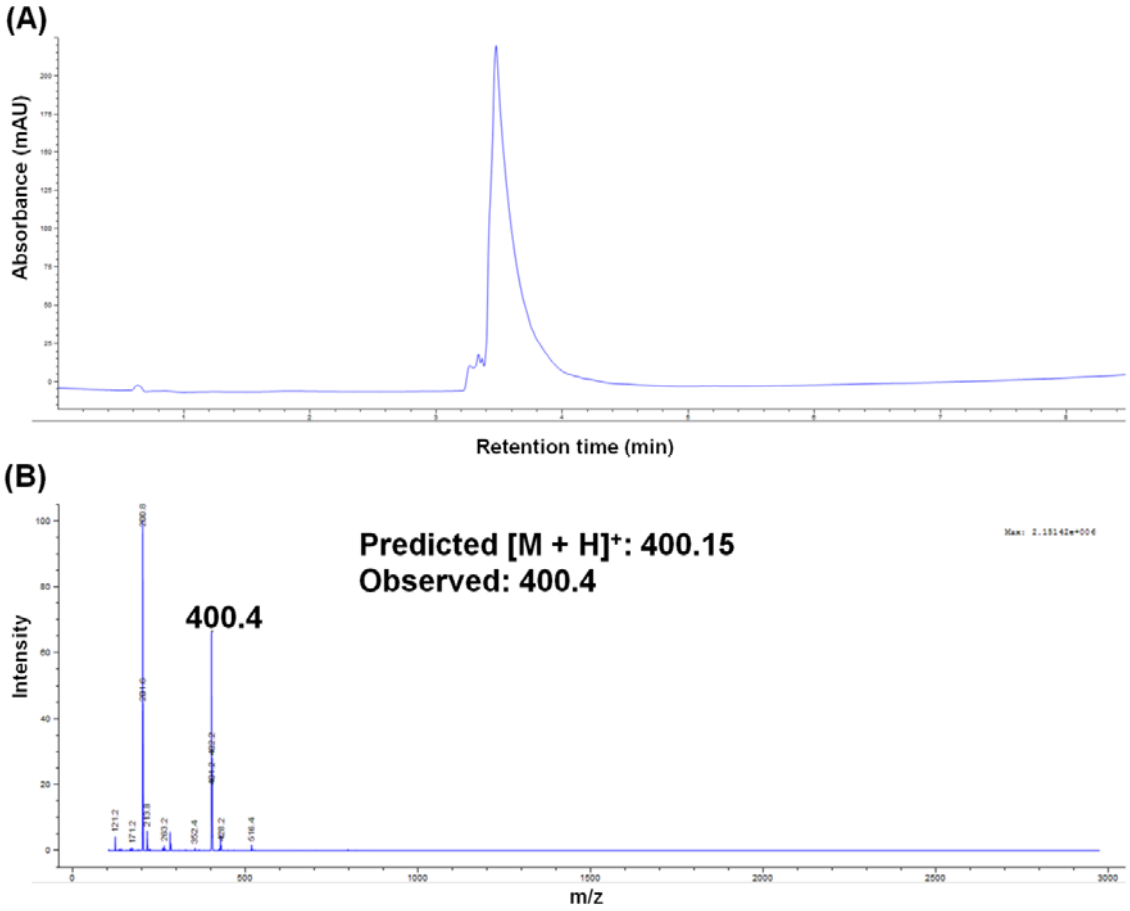


Figure S29. HPLC trace (A) and ESI-Mass spectra (B) of compound **8**.

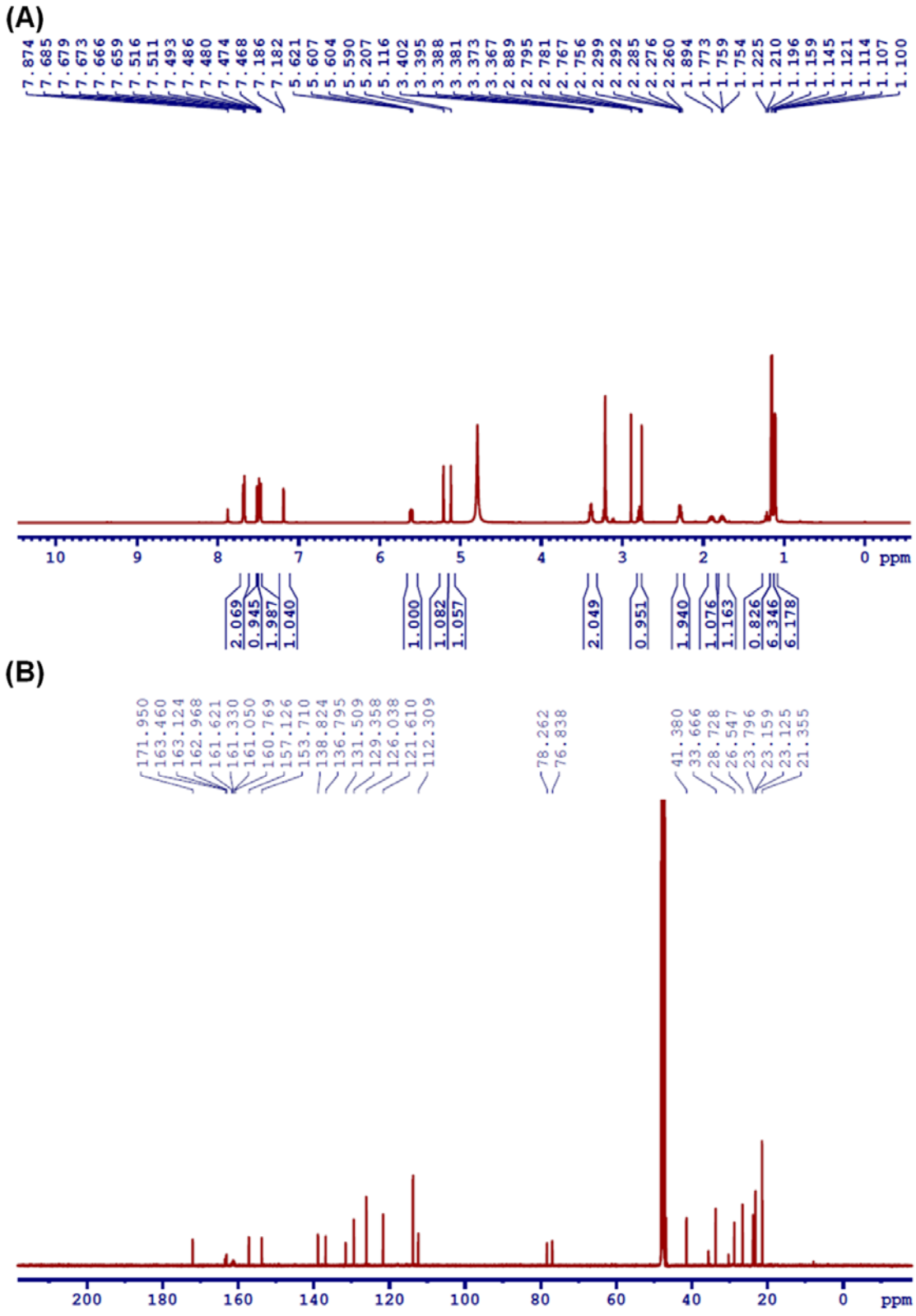


Figure S30. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **9** in CD_3OD.

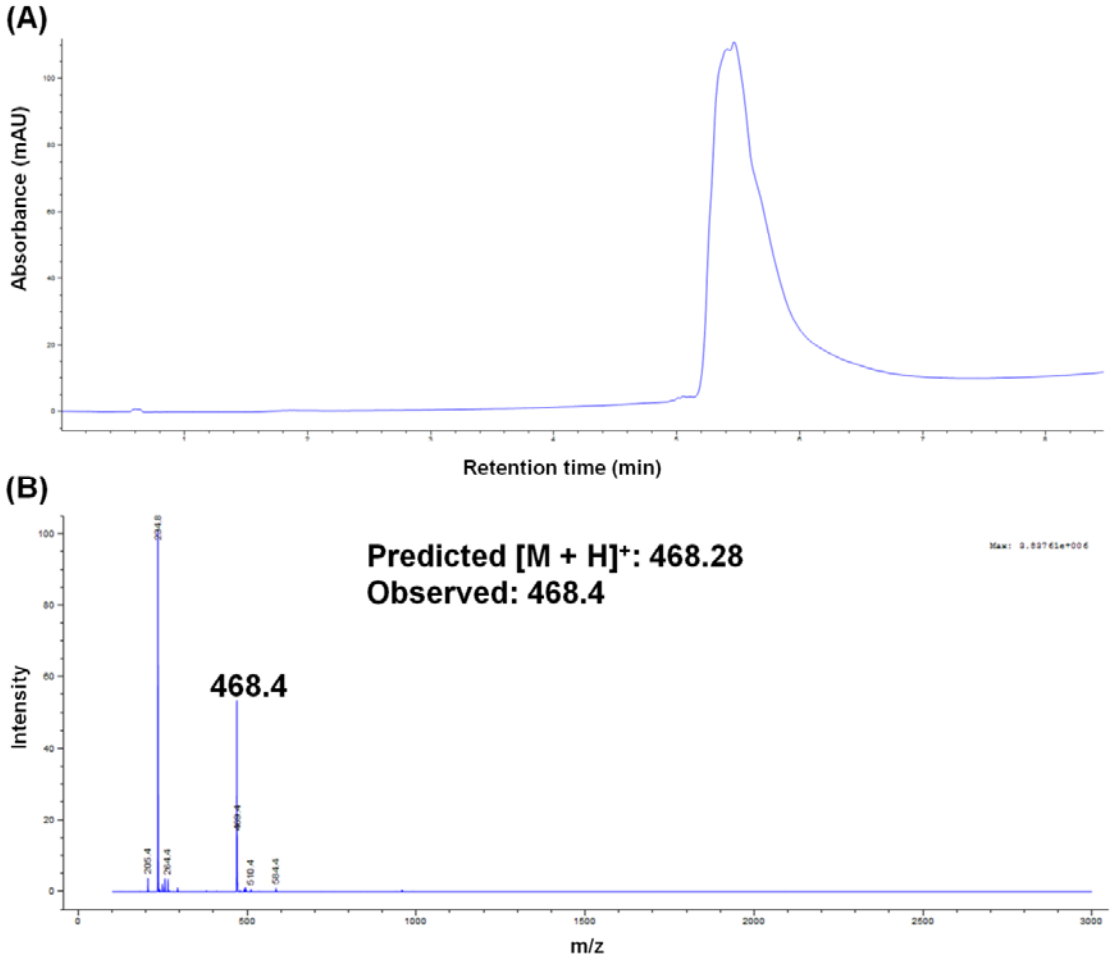


Figure S31. HPLC trace (A) and ESI-Mass spectra (B) of compound **9**.

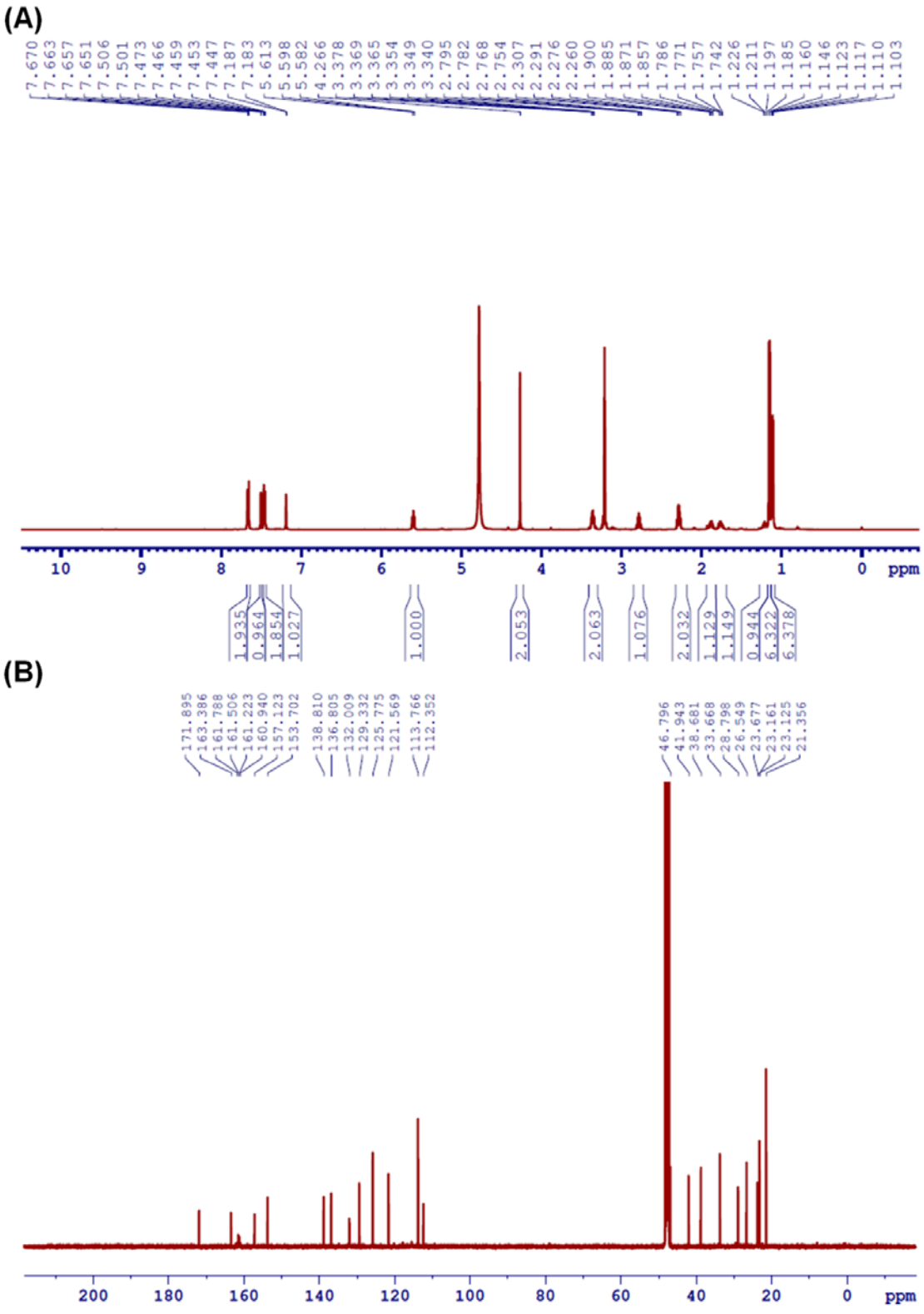


Figure S32. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **10** in CD_3OD.

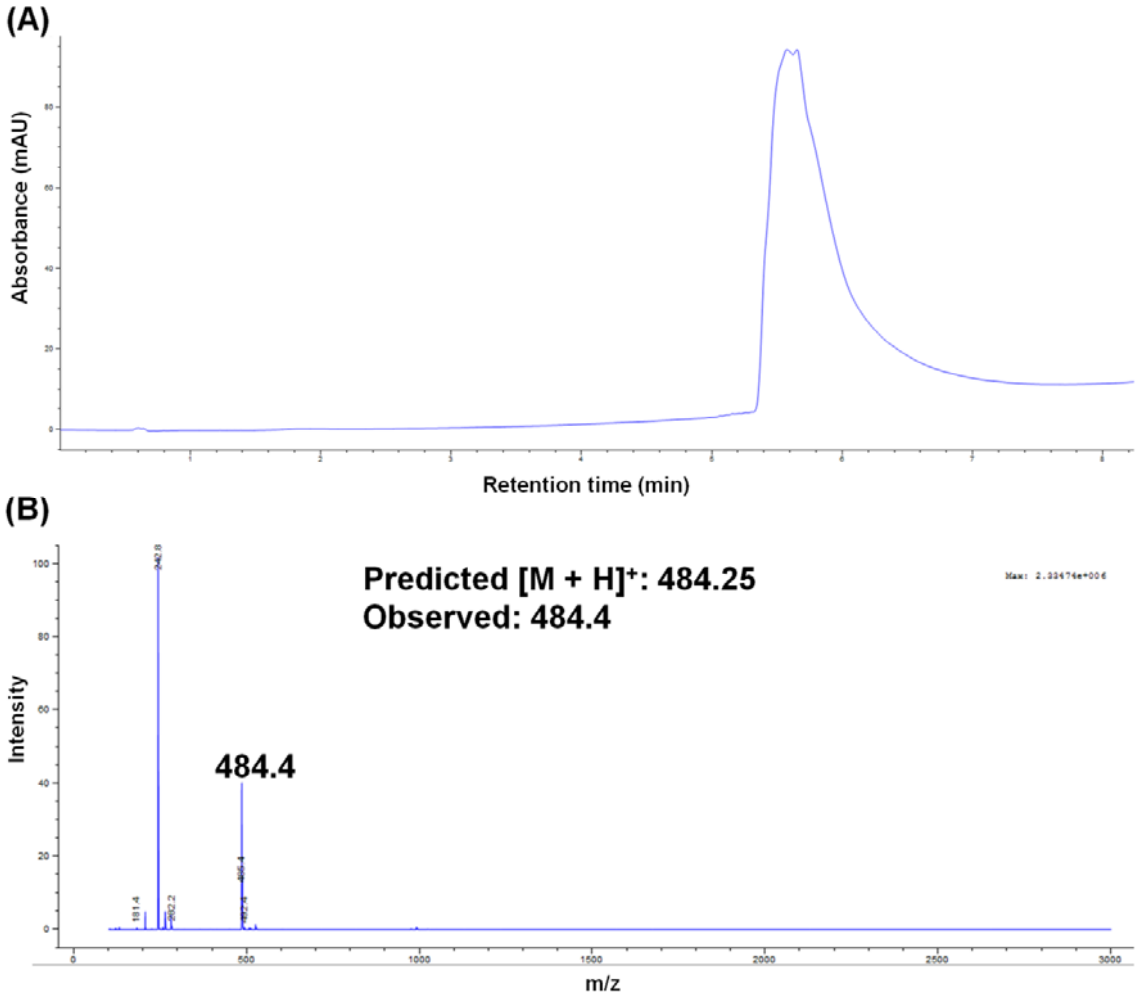


Figure S33. HPLC trace (A) and ESI-Mass spectra (B) of compound **10**.

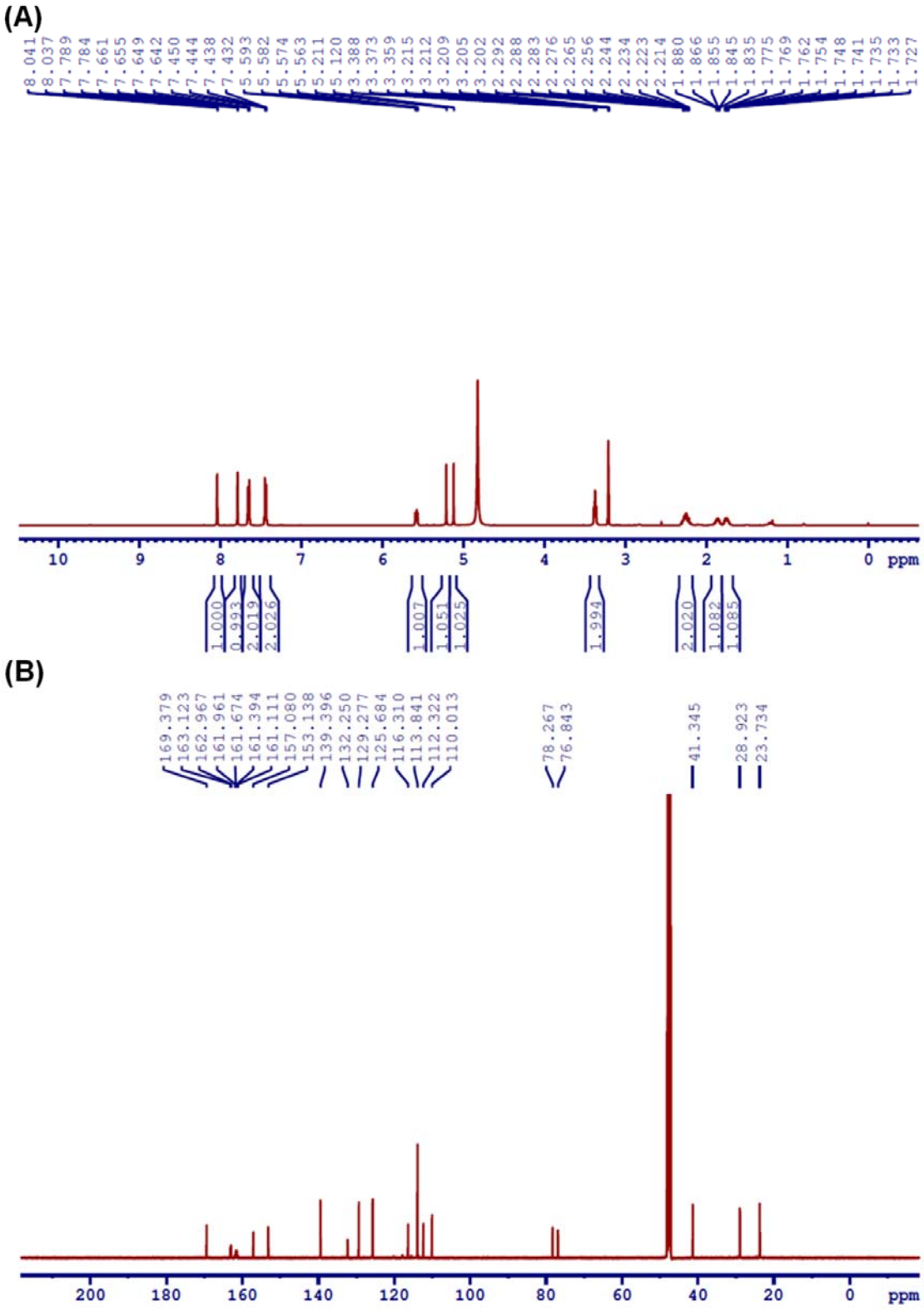


Figure S34. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **11** in CD_3OD.

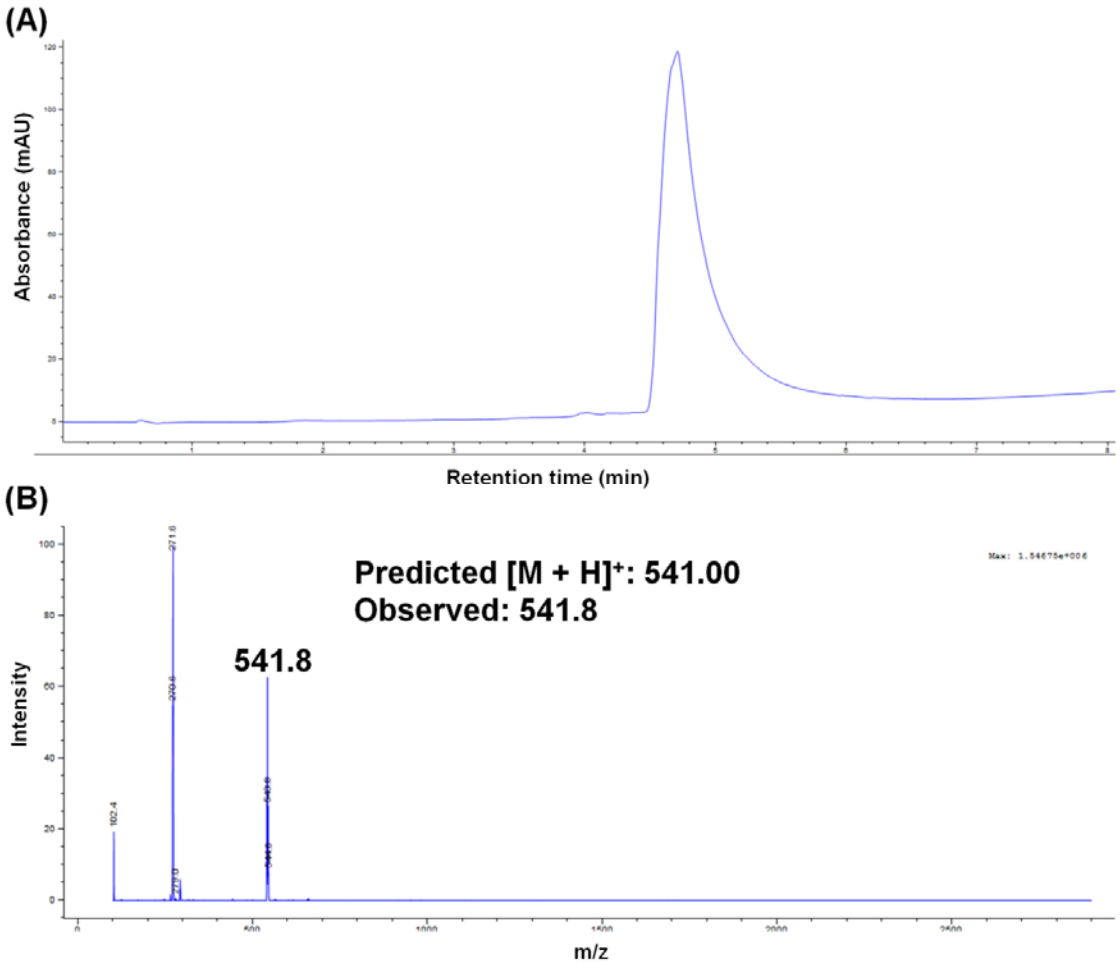


Figure S35. HPLC trace (A) and ESI-Mass spectra (B) of compound **11**.

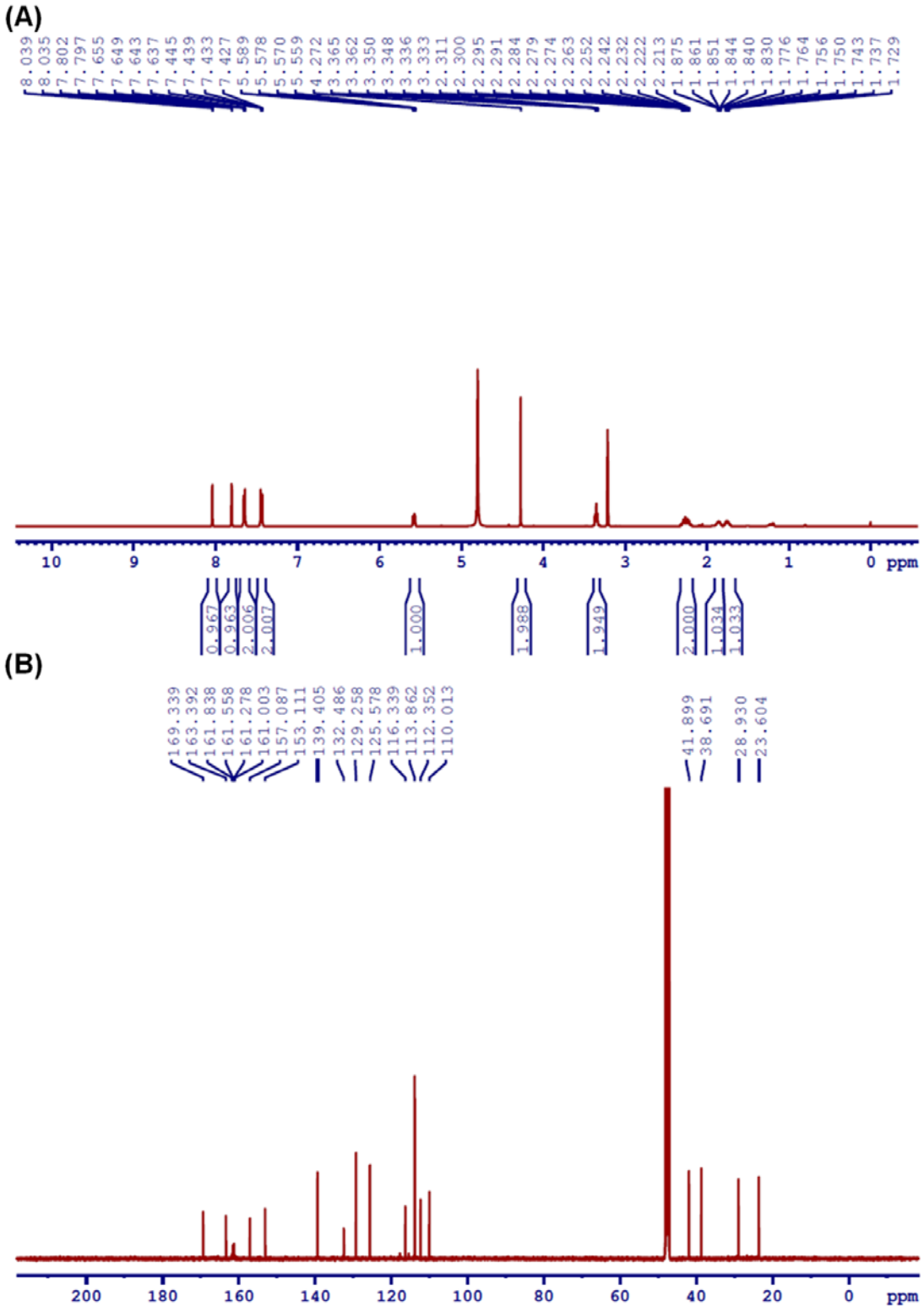


Figure S36. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **12** in CD_3OD.

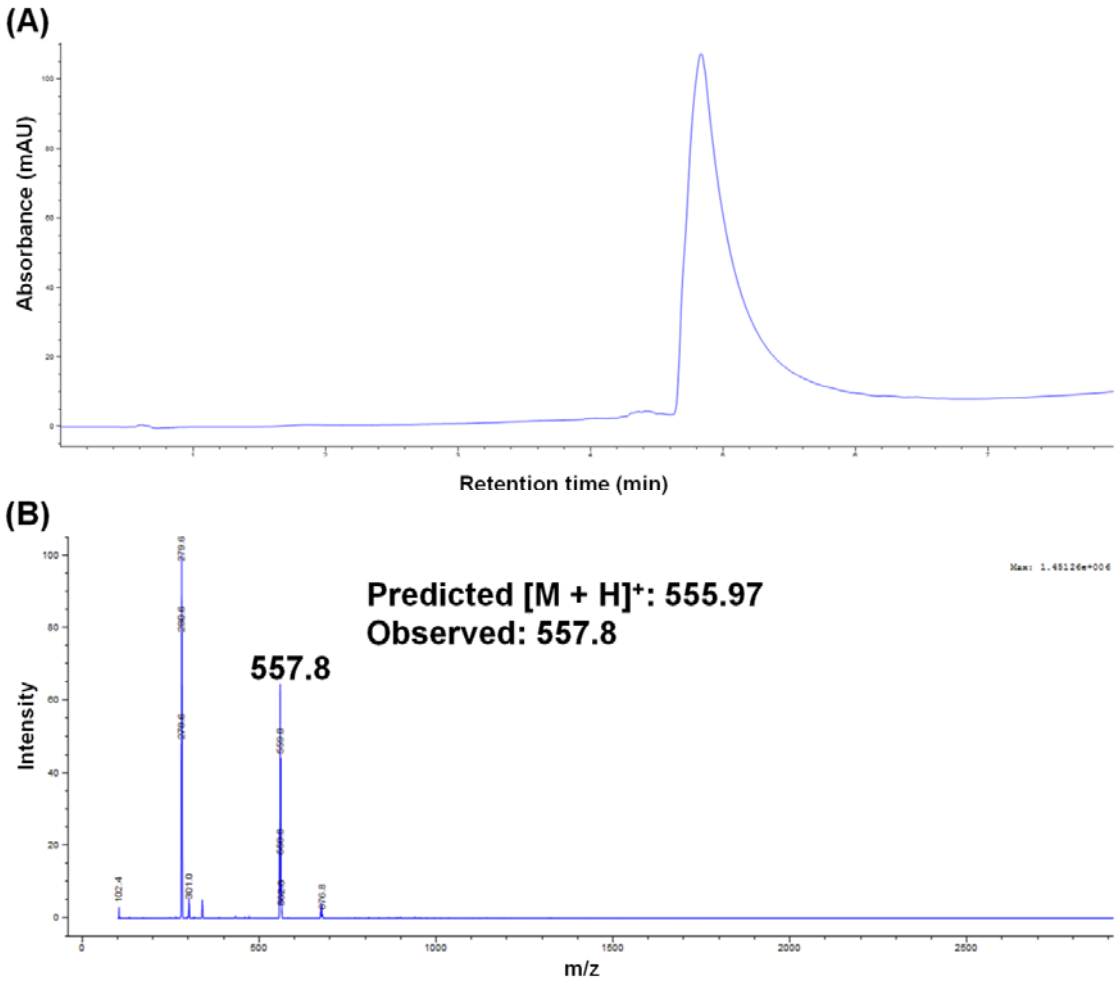


Figure S37. HPLC trace (A) and ESI-Mass spectra (B) of compound **12**.

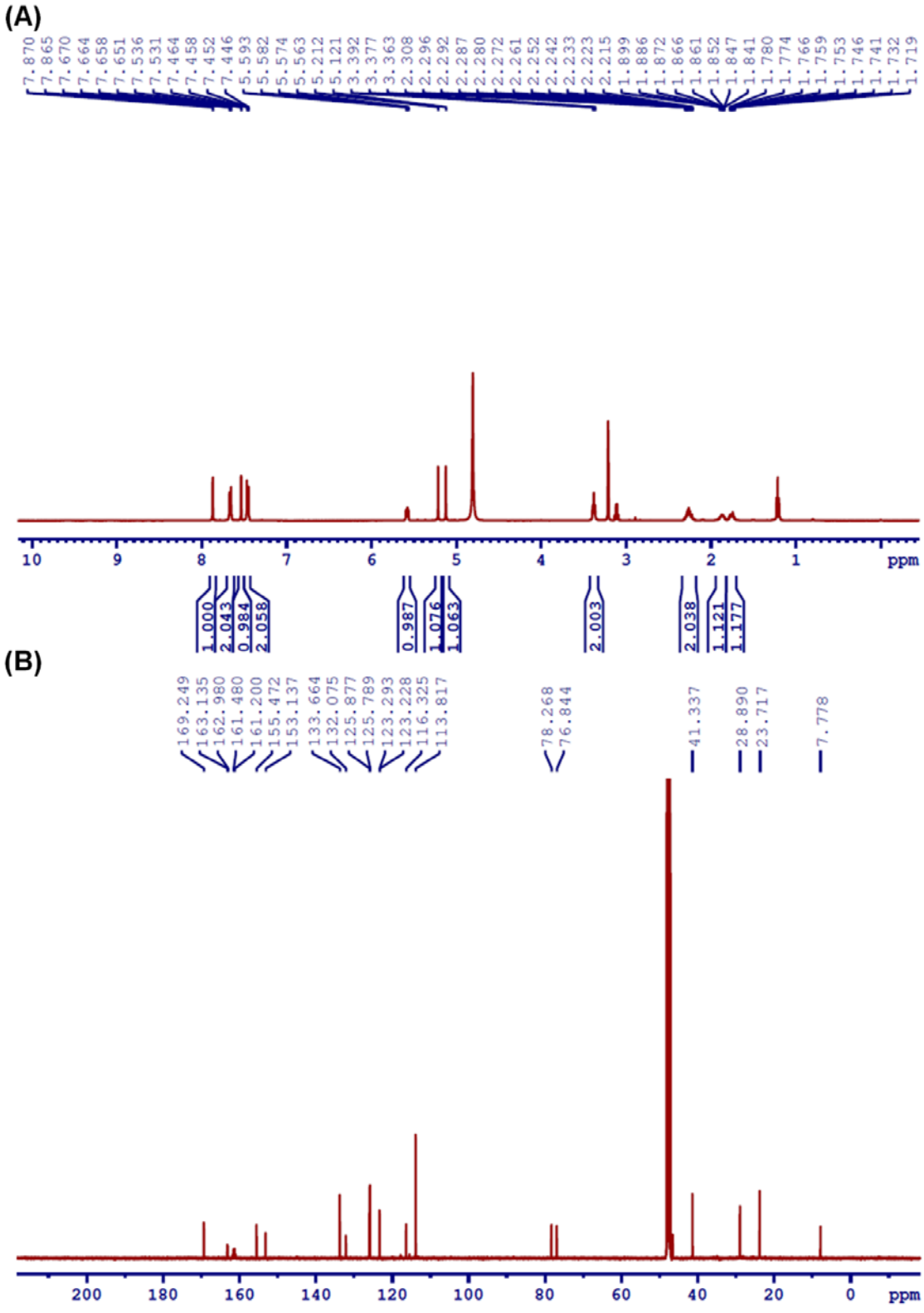


Figure S38. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **13** in CD_3OD.

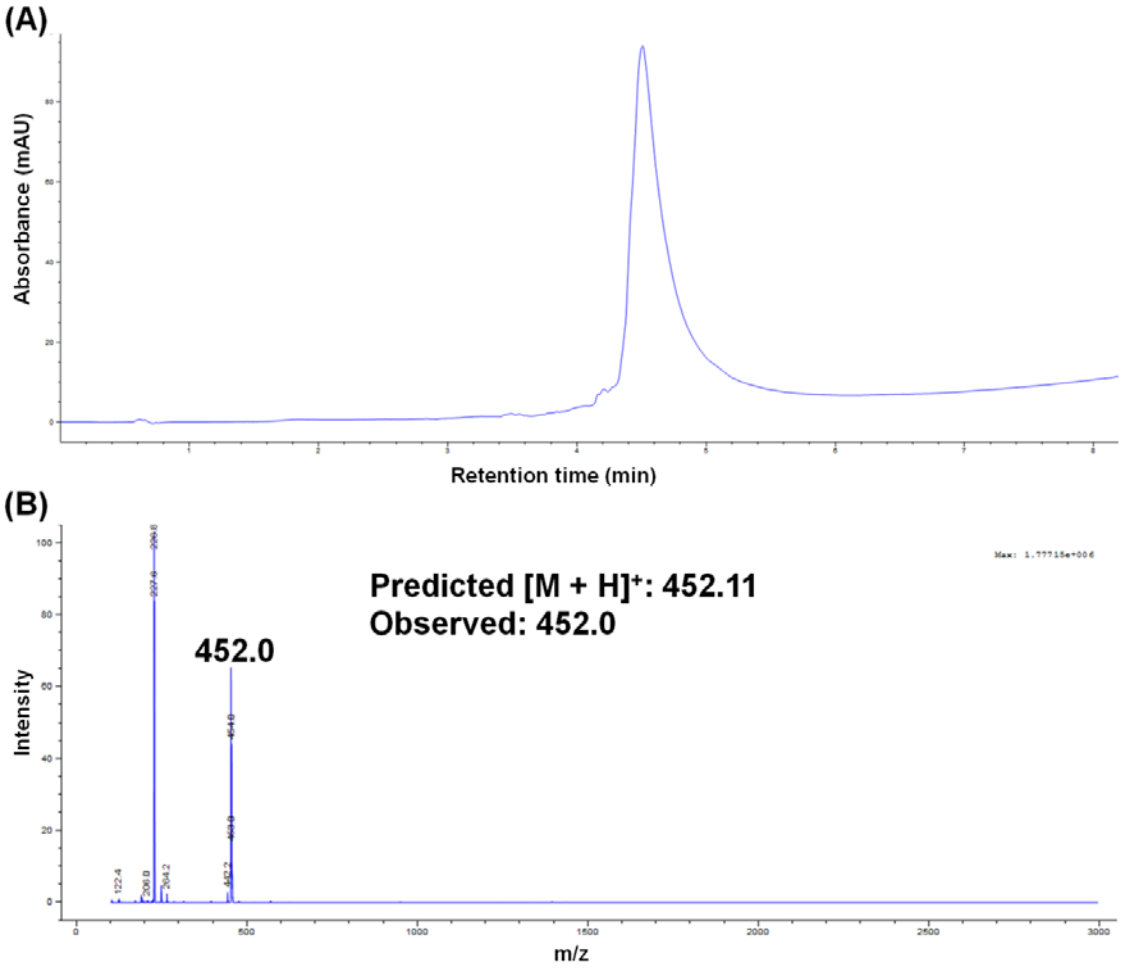


Figure S39. HPLC trace (A) and ESI-Mass spectra (B) of compound **13**.

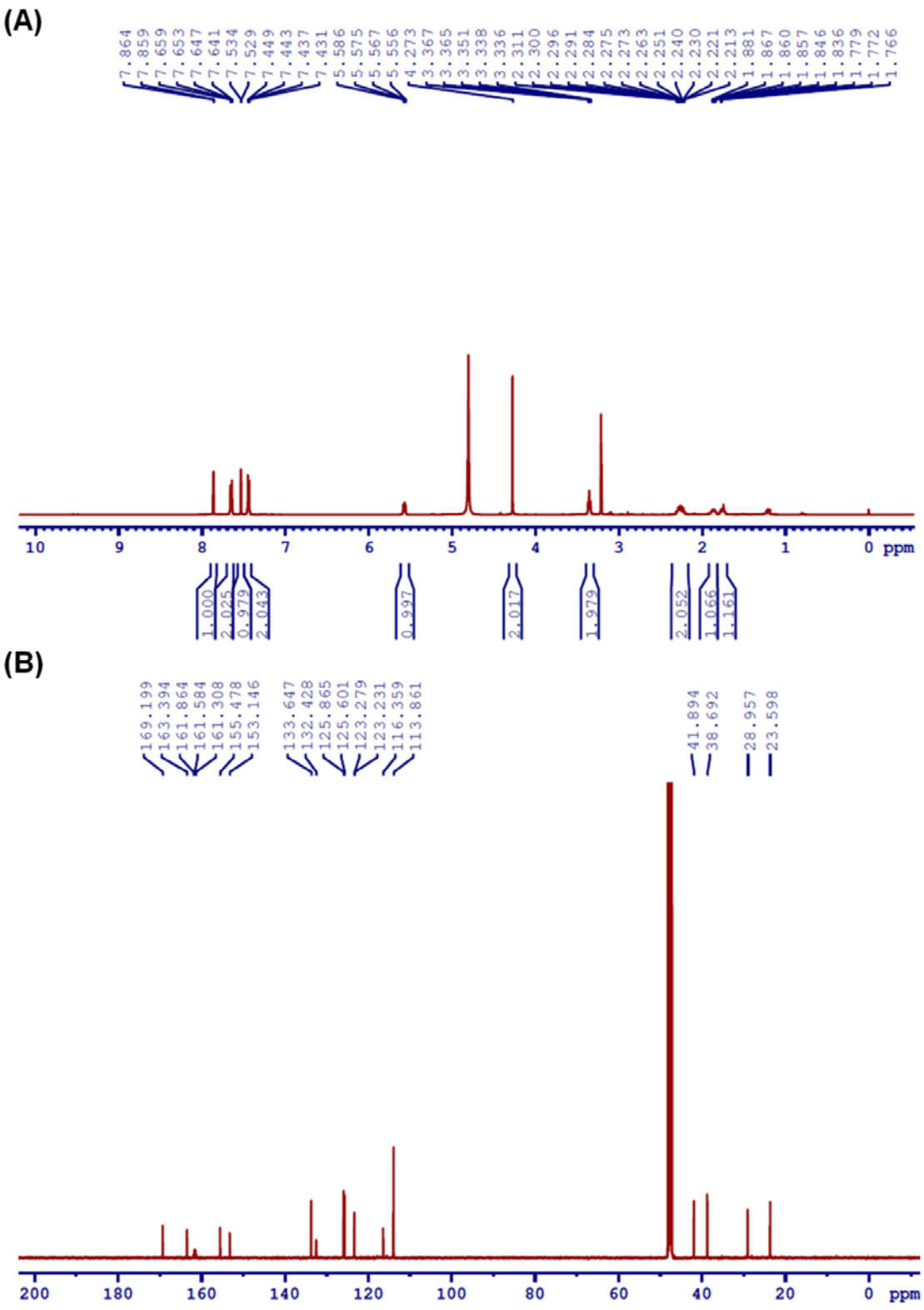


Figure S40. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **14** in CD_3OD.

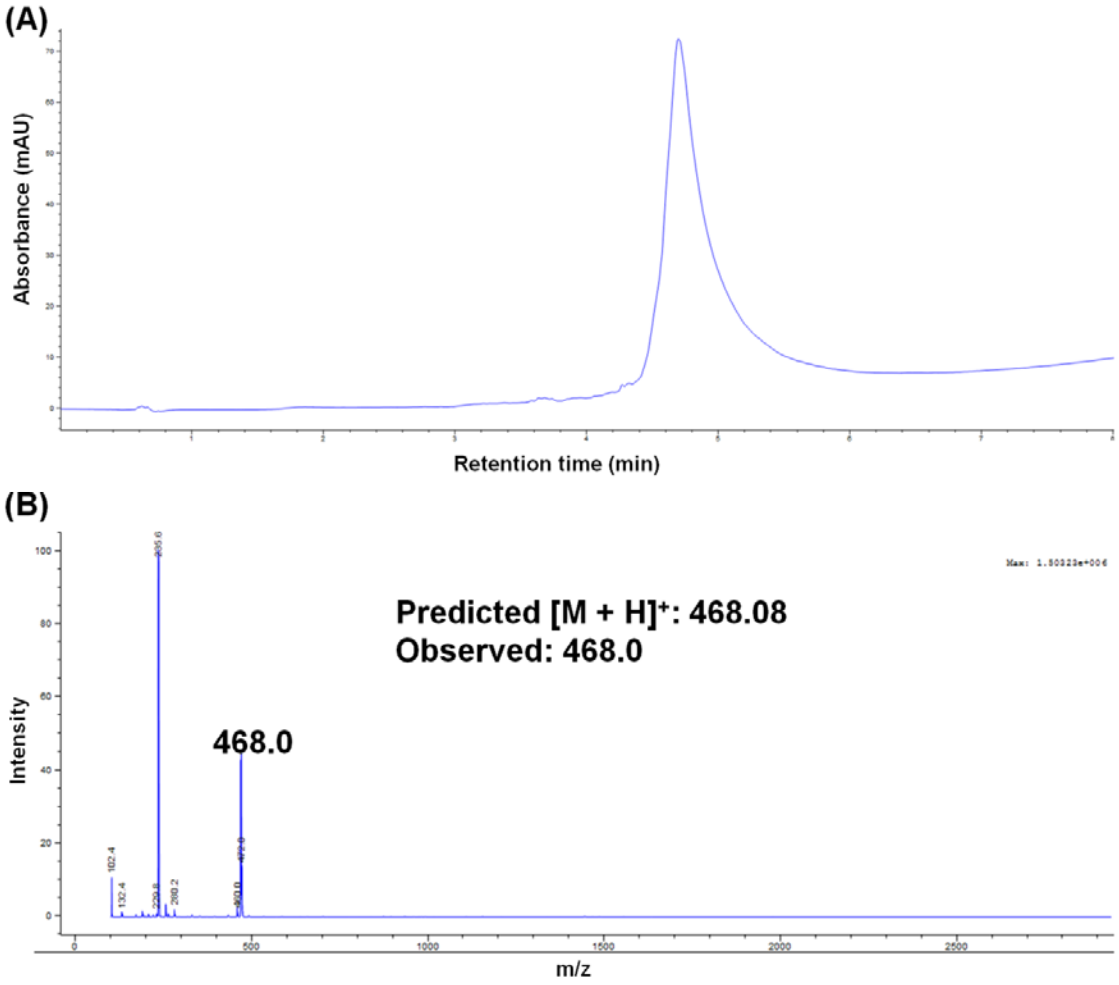


Figure S41. HPLC trace (A) and ESI-Mass spectra (B) of compound **14**.

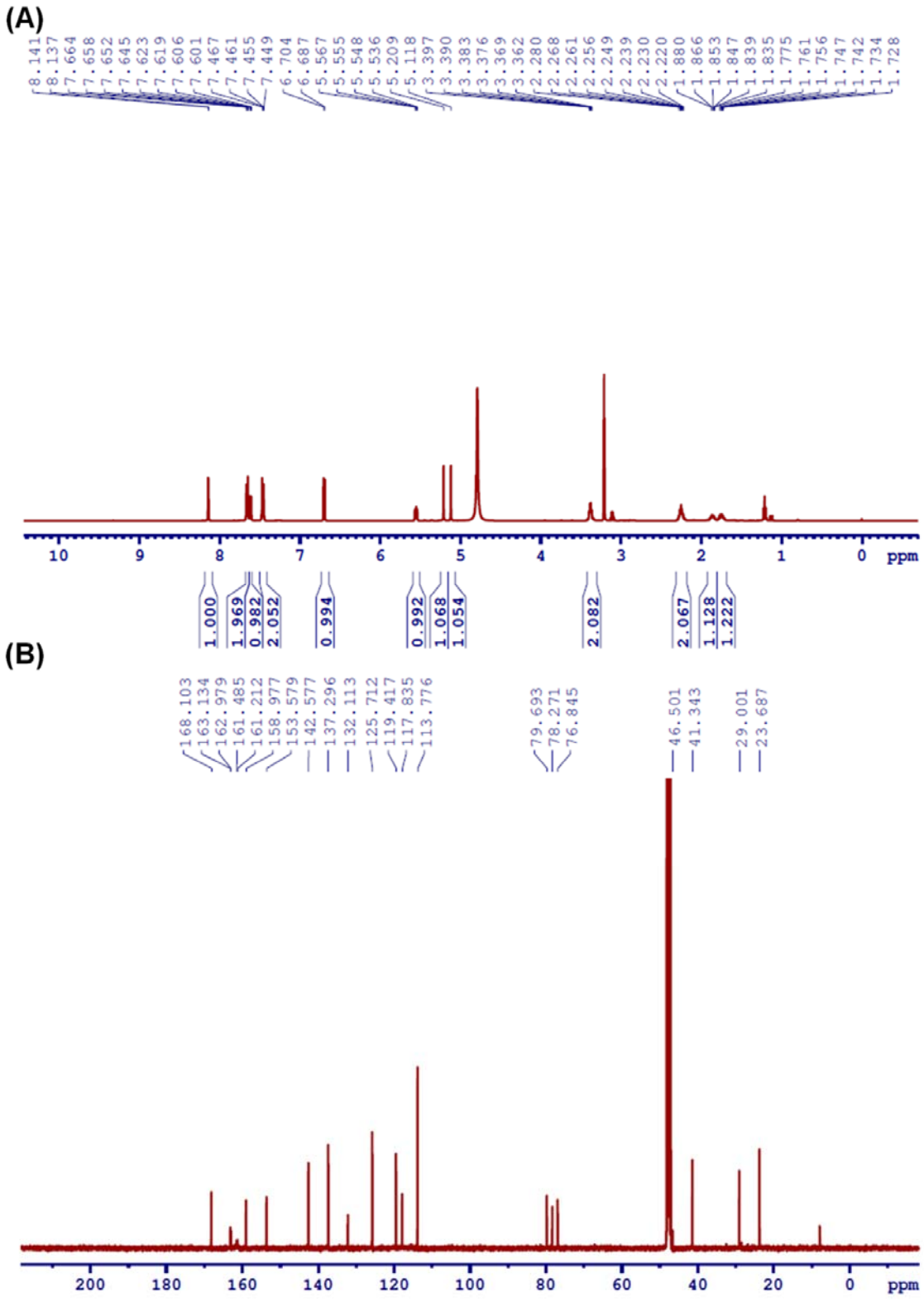


Figure S42. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **15** in CD_3OD.

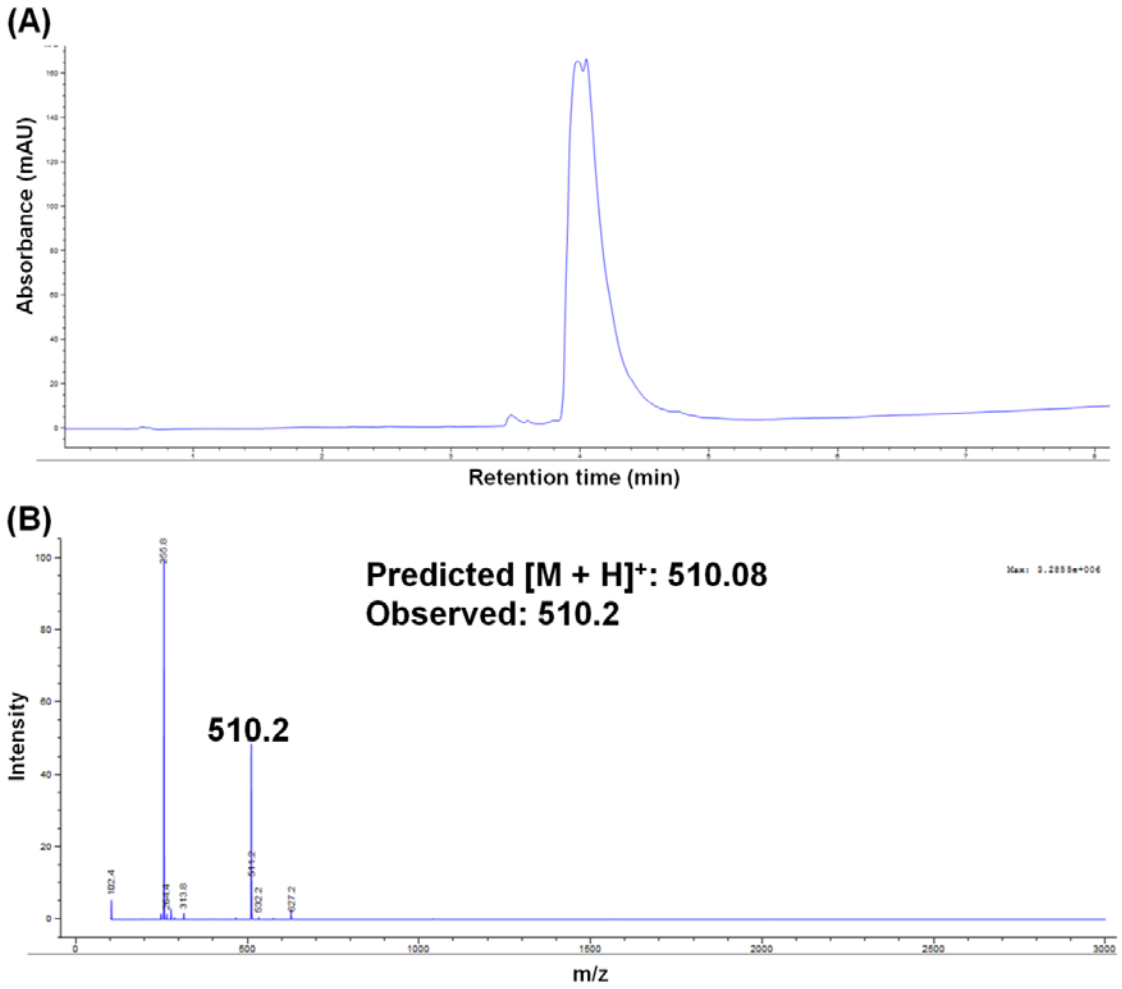


Figure S43. HPLC trace (A) and ESI-Mass spectra (B) of compound **15**.

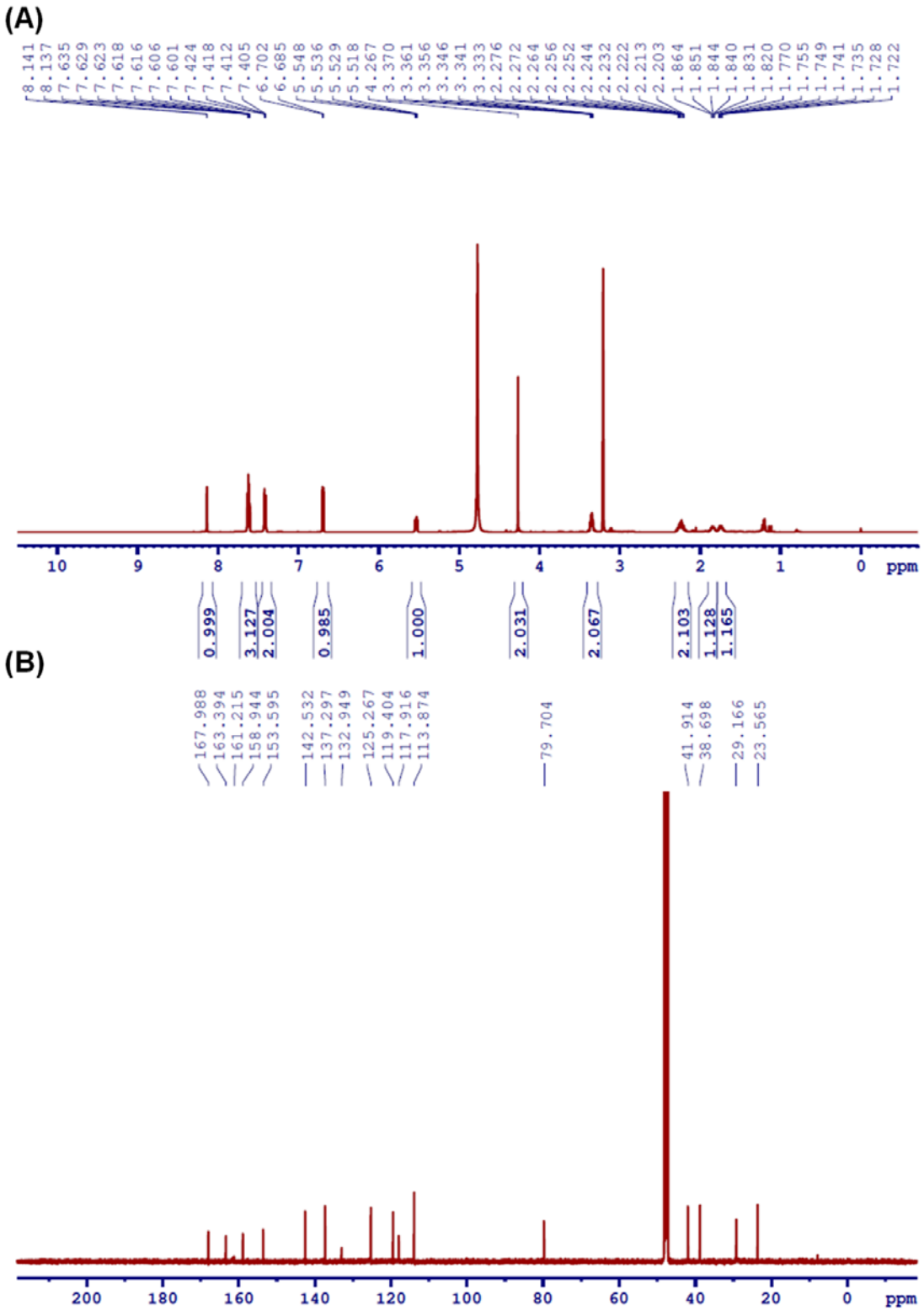


Figure S44. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **16** in CD_3OD.

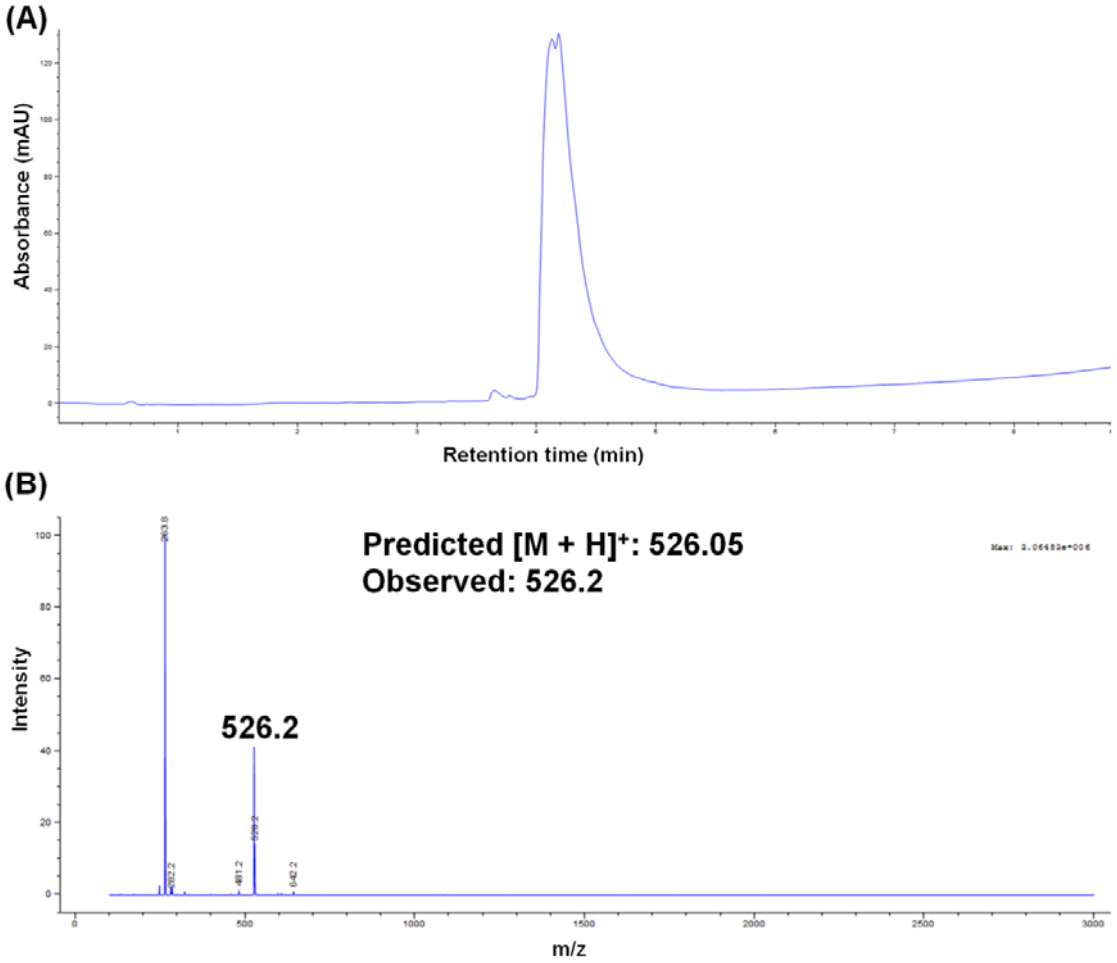


Figure S45. HPLC trace (A) and ESI-Mass spectra (B) of compound **16**.

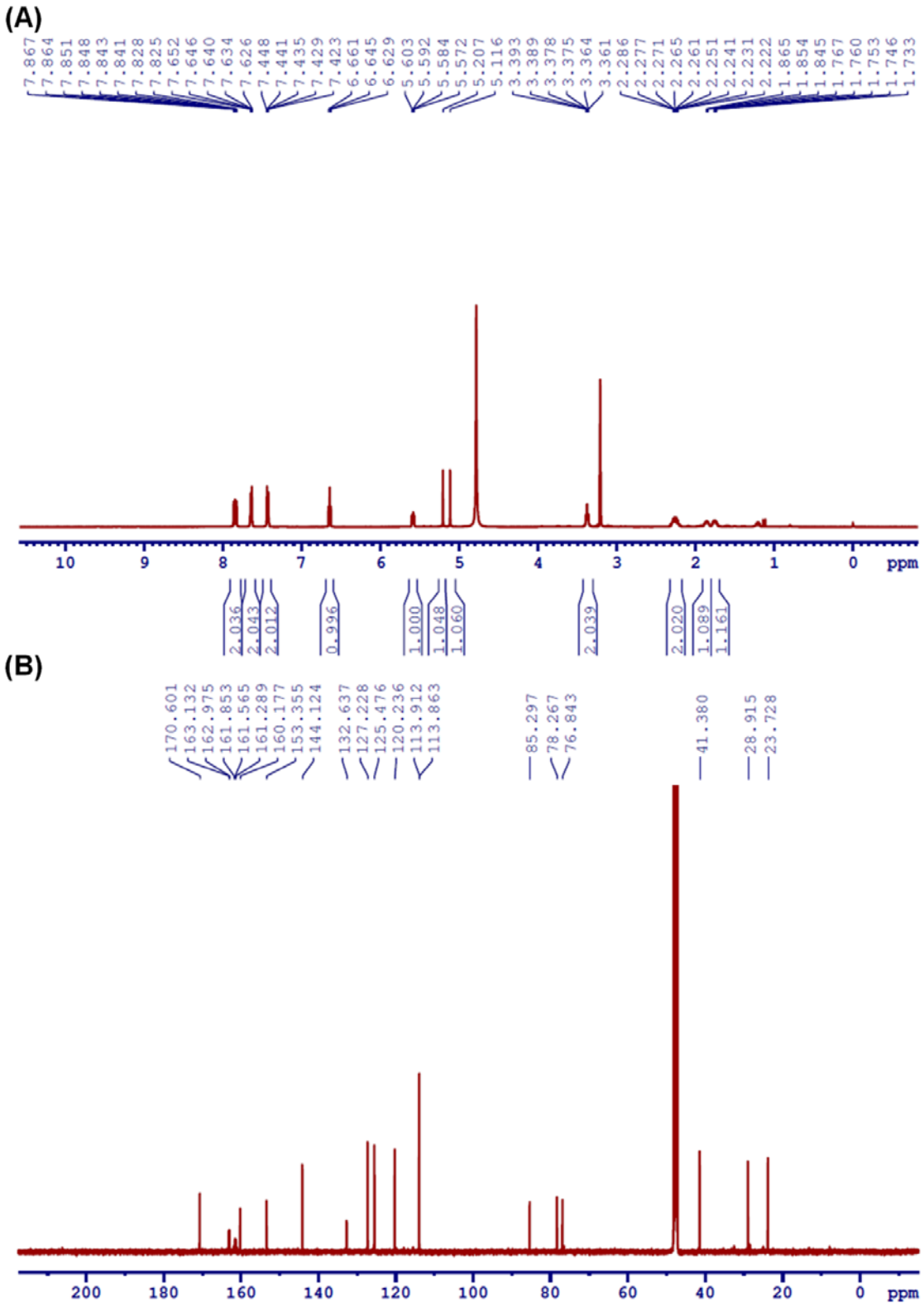


Figure S46. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **17** in CD_3OD.

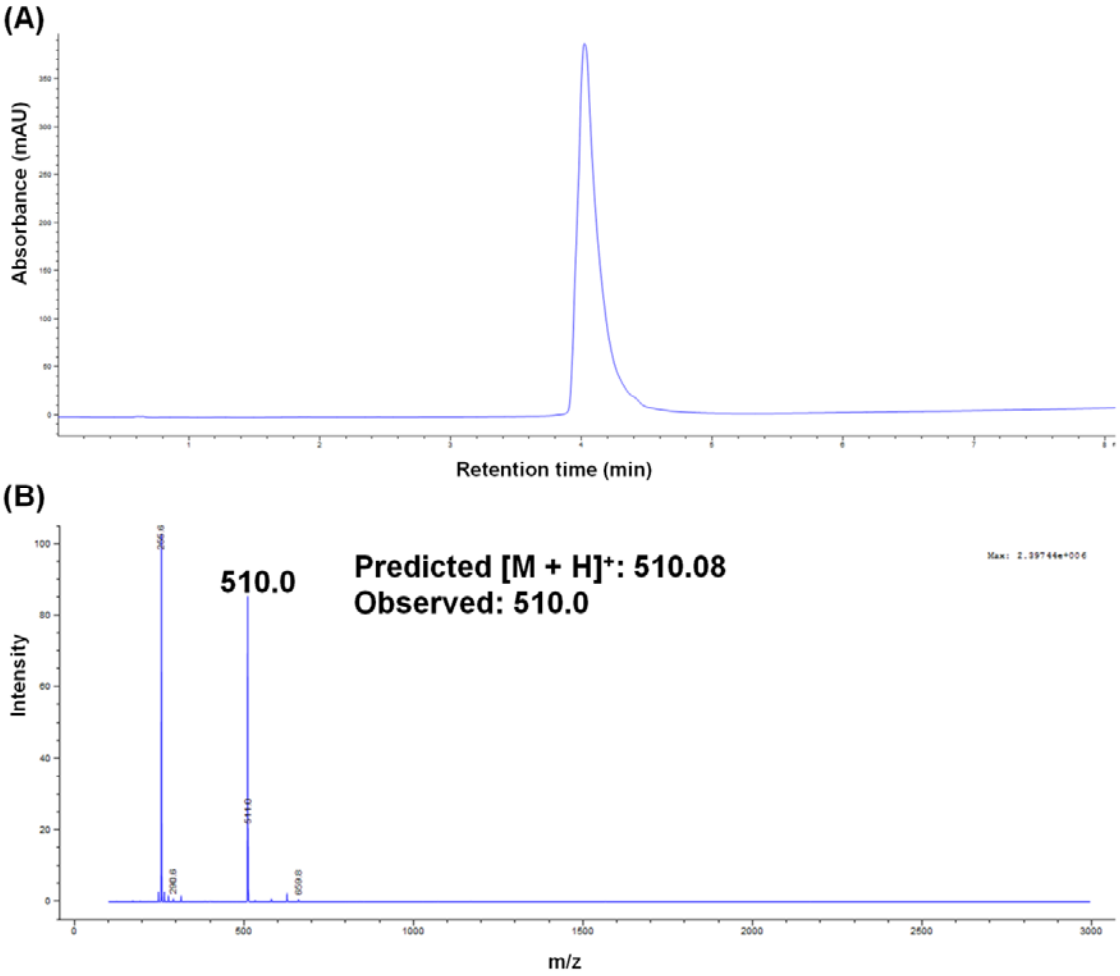


Figure S47. HPLC trace (A) and ESI-Mass spectra (B) of compound **17**.

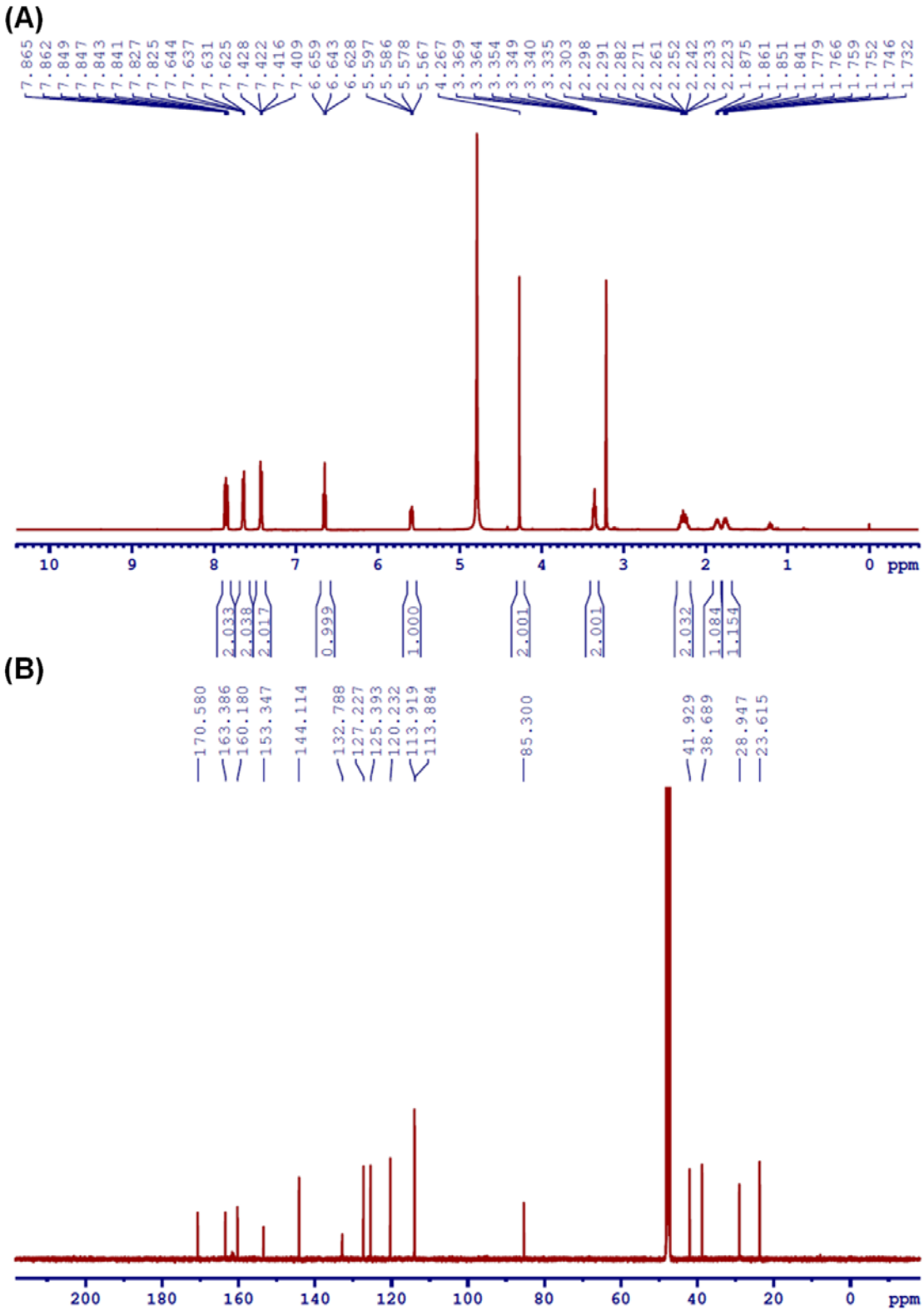


Figure S48. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **18** in CD_3OD.

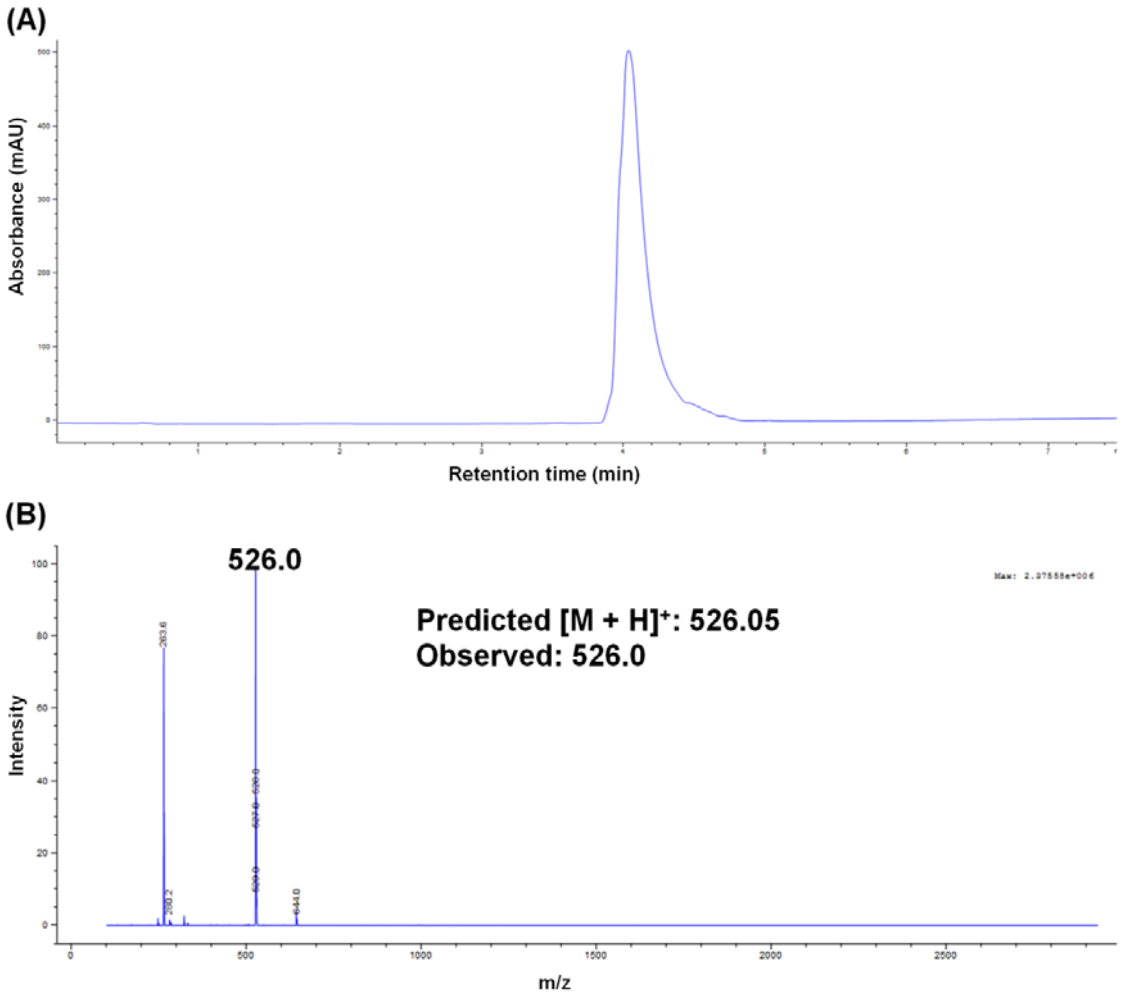


Figure S49. HPLC trace (A) and ESI-Mass spectra (B) of compound **18**.

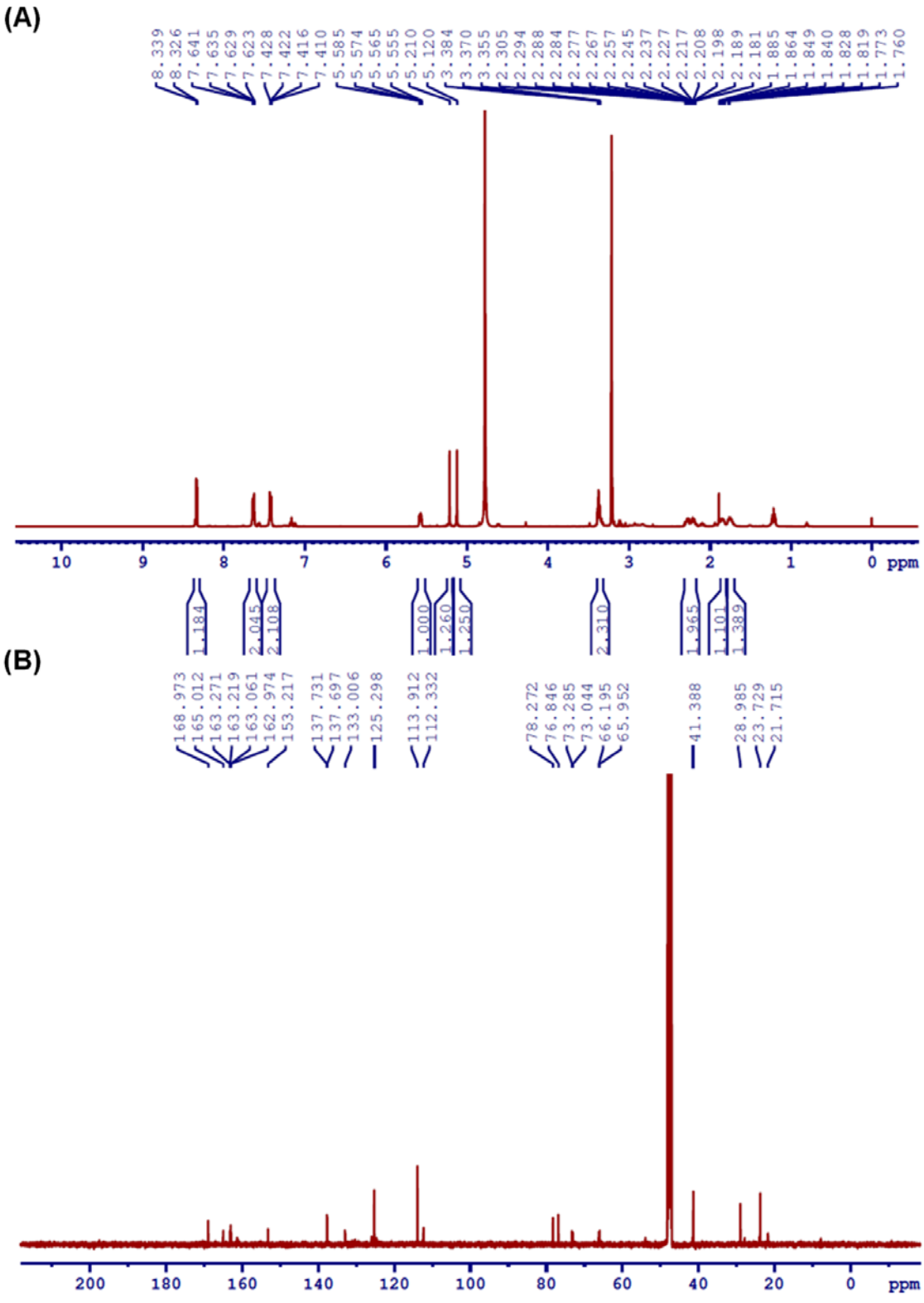


Figure S50. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **19** in CD_3OD.

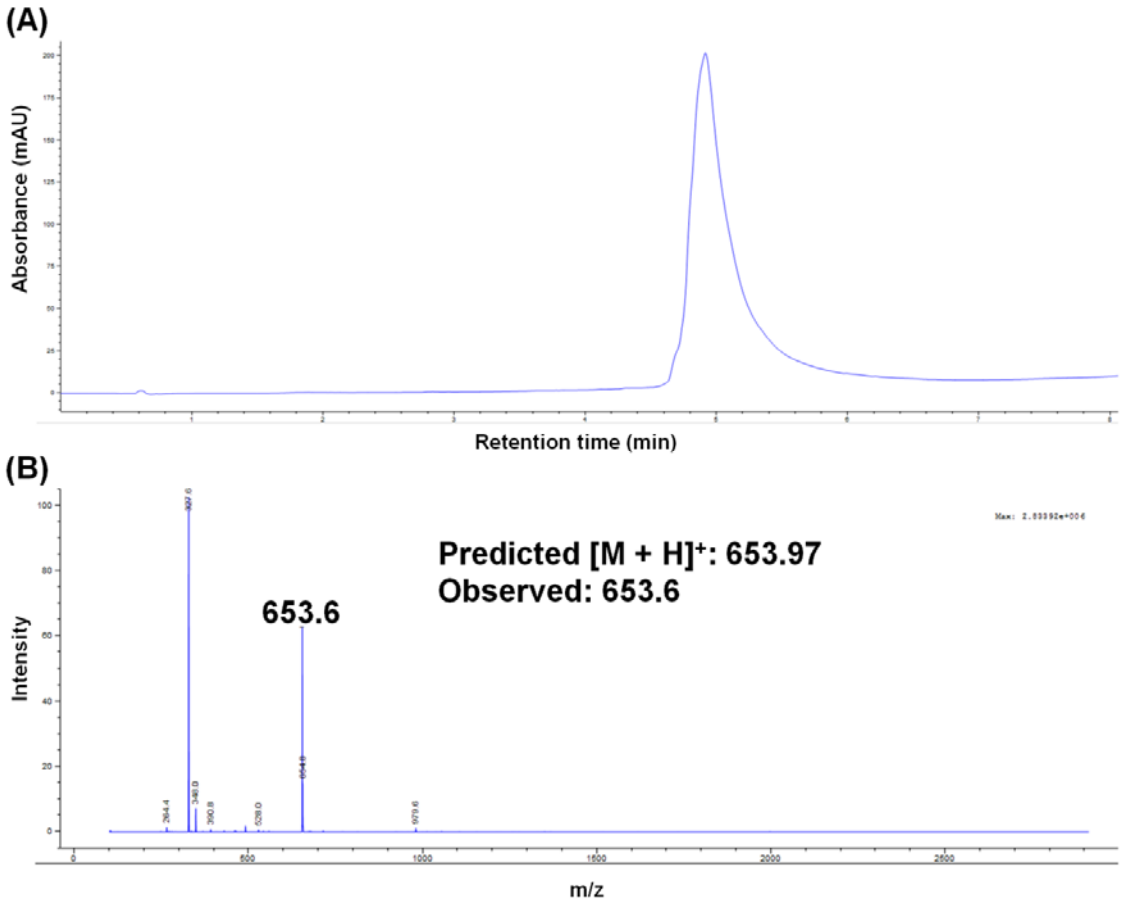


Figure S51. HPLC trace (A) and ESI-Mass spectra (B) of compound **19**.

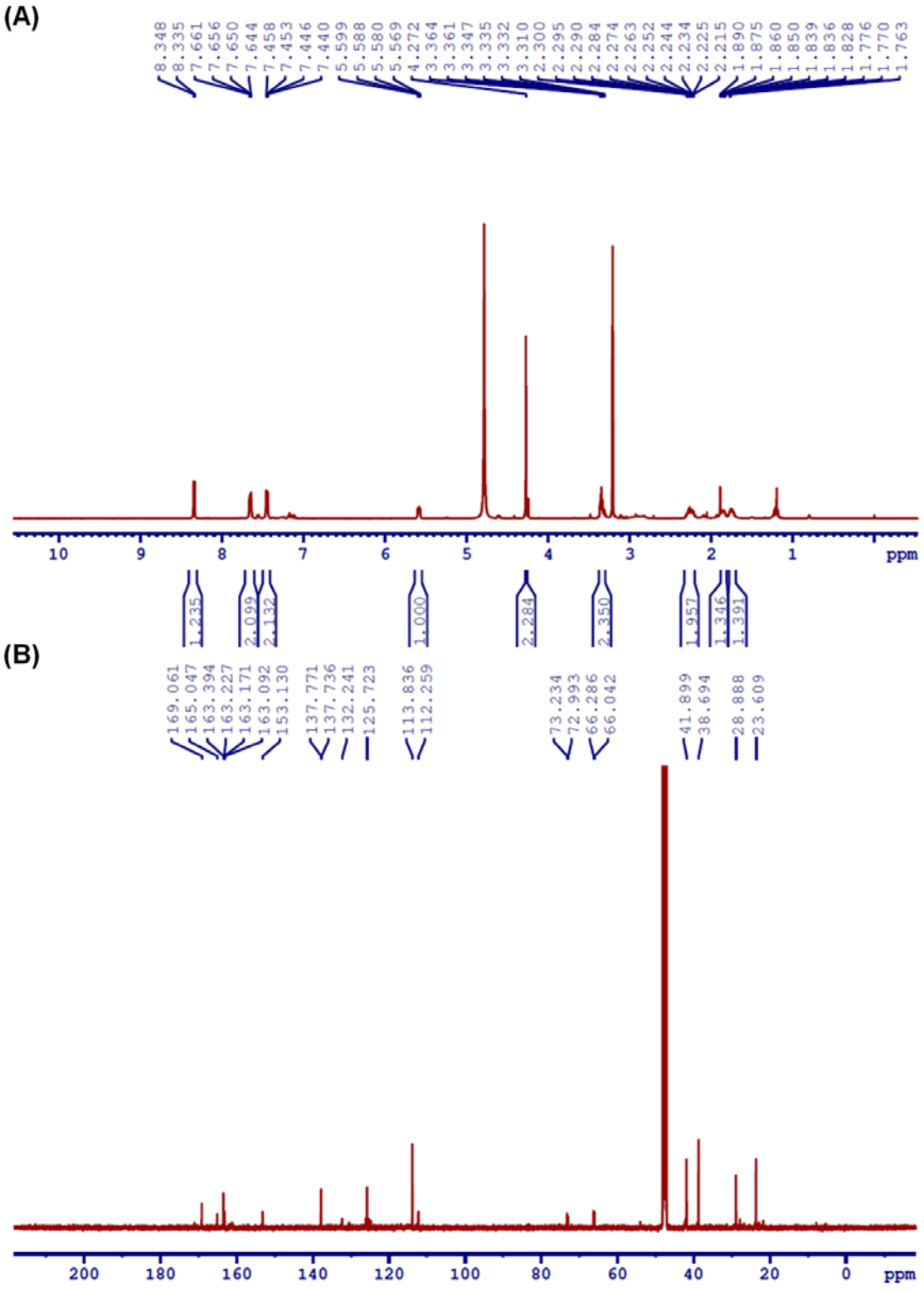


Figure S52. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **20** in CD_3OD.

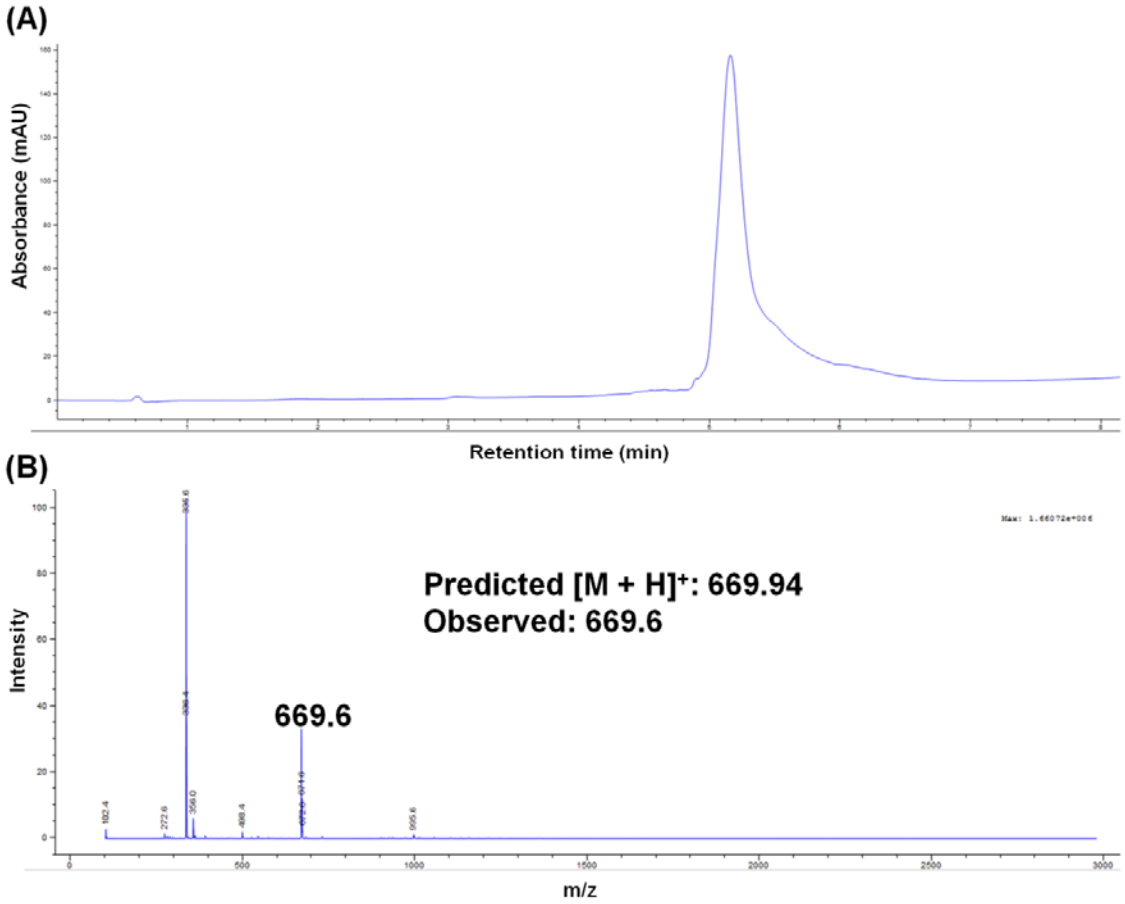


Figure S53. HPLC trace (A) and ESI-Mass spectra (B) of compound **20**.

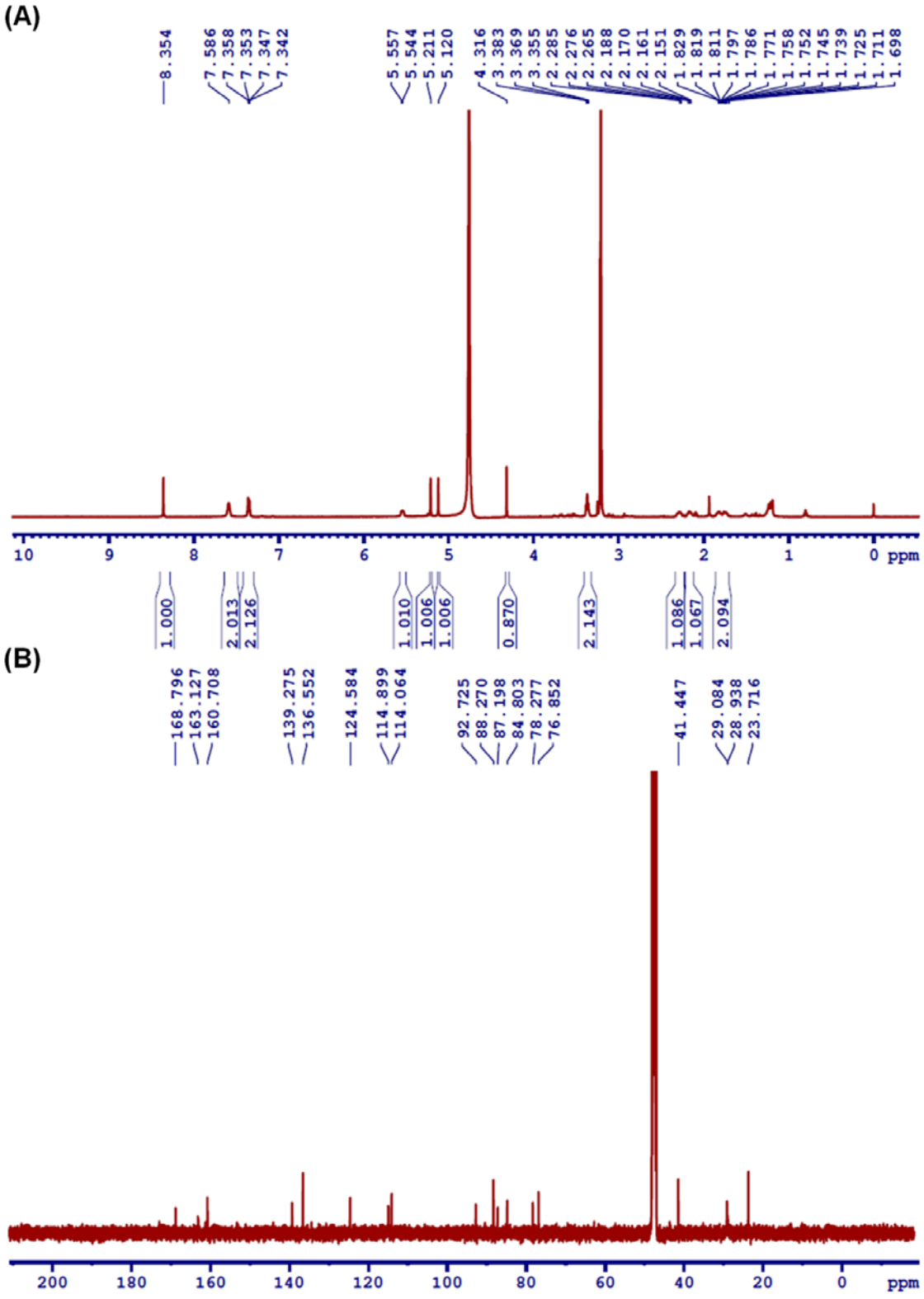


Figure S54. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **21** in CD_3OD.

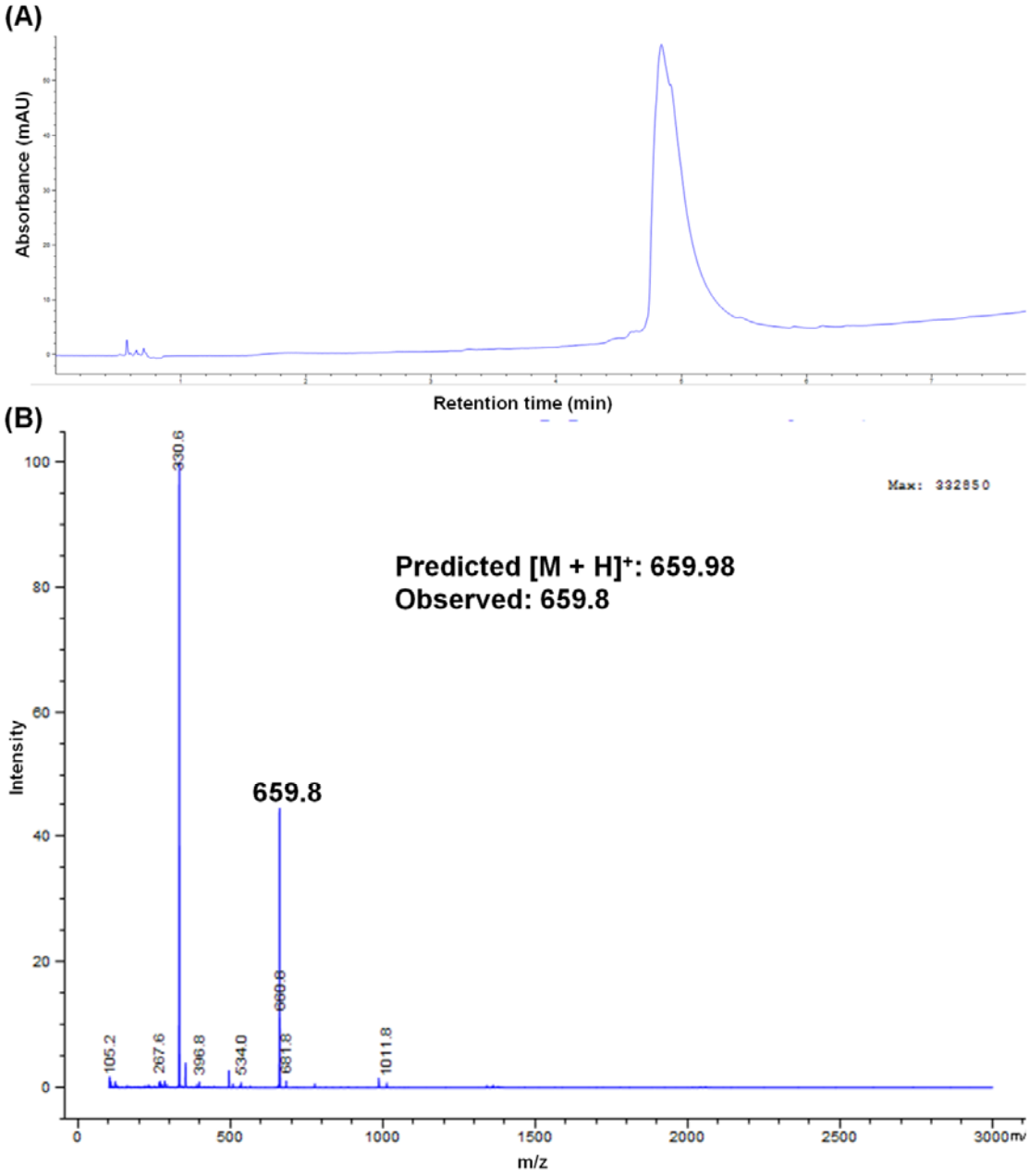


Figure S55. HPLC trace (A) and ESI-Mass spectra (B) of compound **21**.

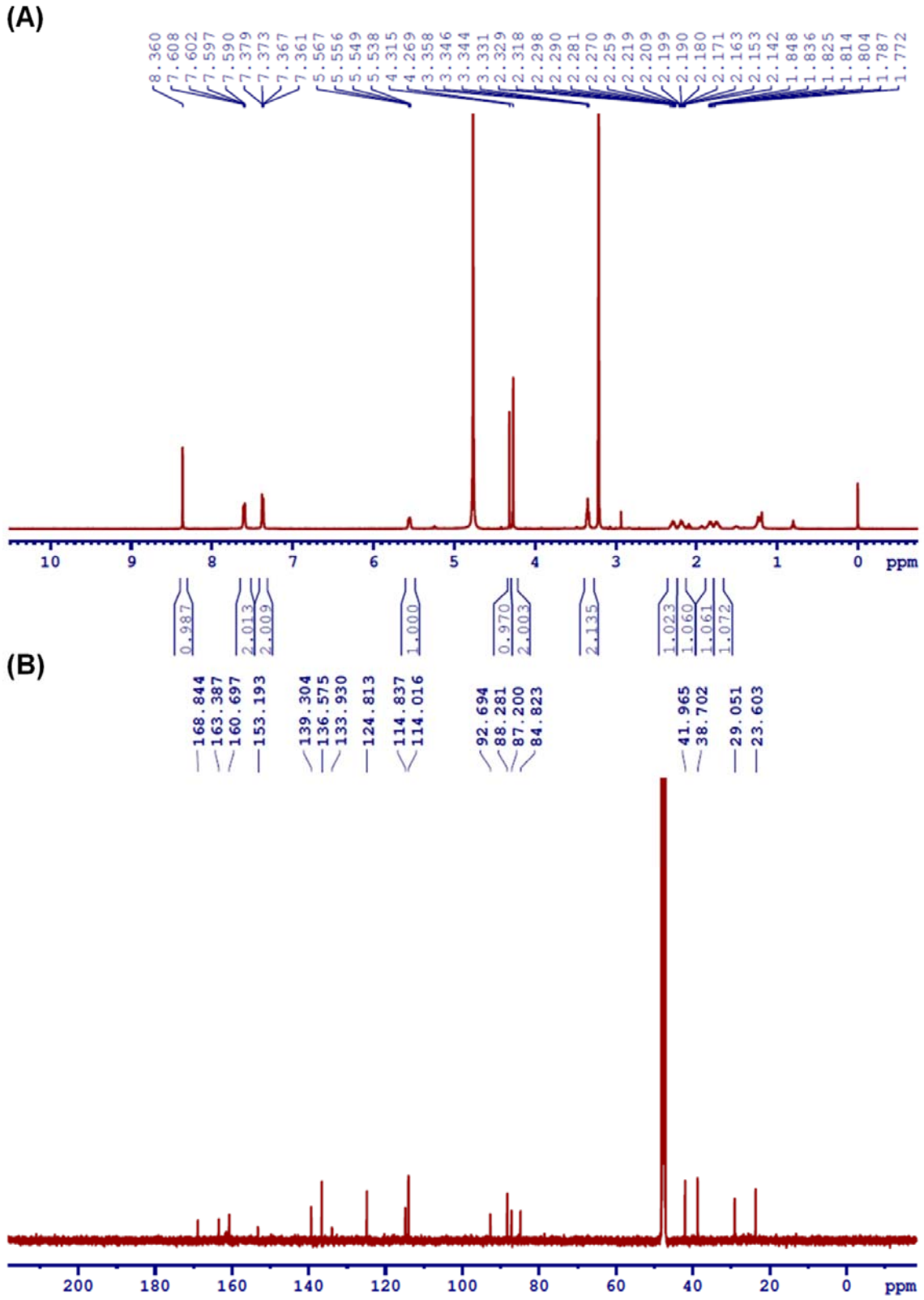


Figure S56. ^{1}H (A) and ^{13}C (B) NMR spectra of compound **22** in CD_3OD.

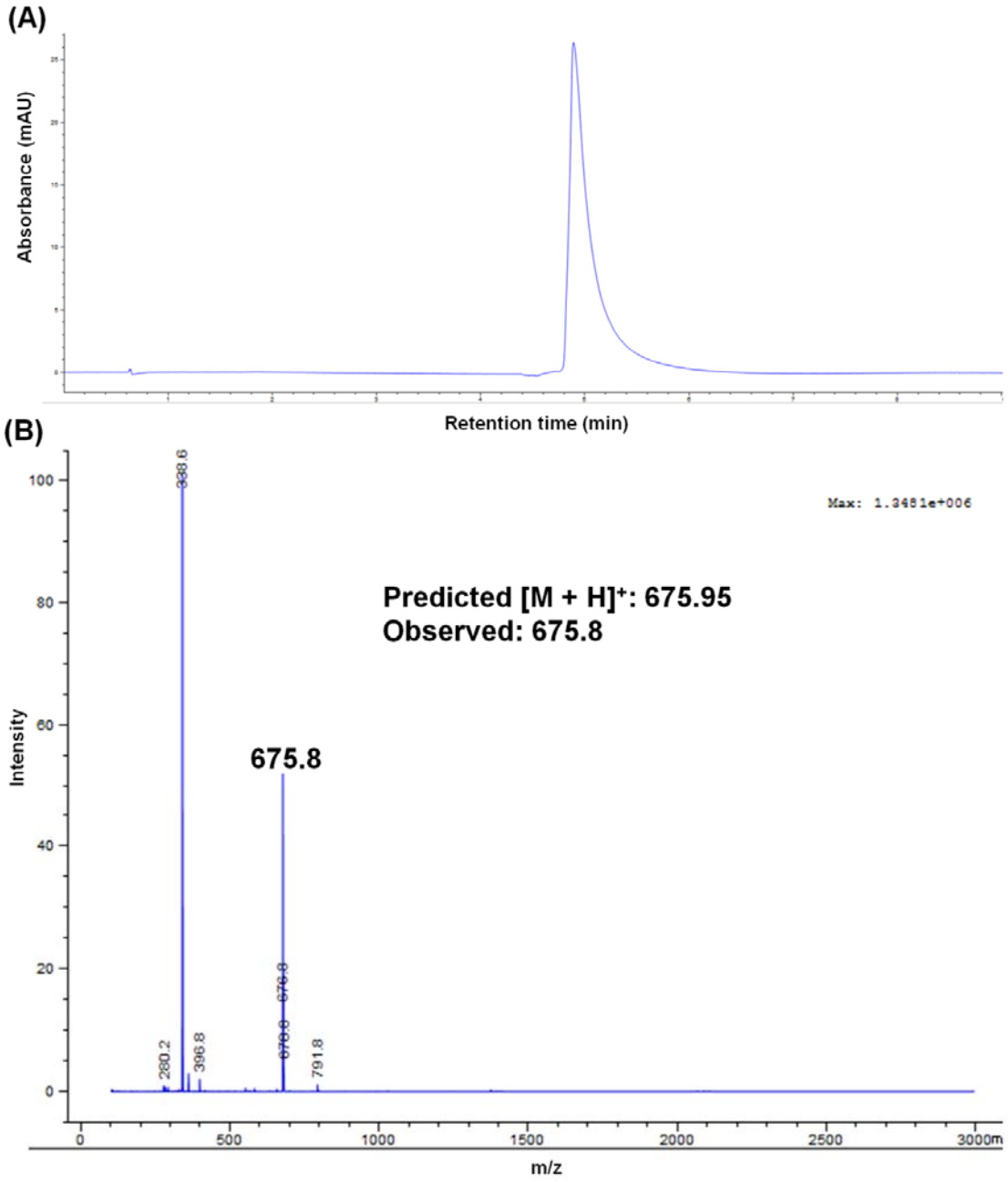


Figure S57. HPLC trace (A) and ESI-Mass spectra (B) of compound **22**.

References

[1] a) C. J. Dreyton, B. Knuckley, J. E. Jones, D. M. Lewallen, P. R. Thompson, *Biochemistry* **2014**, *53*, 4426-4433; b) B. Knuckley, C. P. Causey, J. E. Jones, M. Bhatia, C. J. Dreyton, T. C. Osborne, H. Takahara, P. R. Thompson, *Biochemistry* **2010**, *49*, 4852-4863.

[2] J. S. Knight, V. Subramanian, A. A. O'Dell, S. Yalavarthi, W. Zhao, C. K. Smith, J. B. Hodgin, P. R. Thompson, M. J. Kaplan, *Ann. Rheum. Dis.* **2015**, *74*, 2199-2206.

[3] a) B. Knuckley, C. P. Causey, P. J. Pellechia, P. F. Cook, P. R. Thompson, *Chembiochem* **2010**, *11*, 161-165; b) Y. Luo, K. Arita, M. Bhatia, B. Knuckley, Y. H. Lee, M. R. Stallcup, M. Sato, P. R. Thompson, *Biochemistry* **2006**, *45*, 11727-11736.

[4] a) P. L. Kearney, M. Bhatia, N. G. Jones, L. Yuan, M. C. Glascock, K. L. Catchings, M. Yamada, P. R. Thompson, *Biochemistry* **2005**, *44*, 10570-10582; b) M. Knipp, M. Vasak, *Anal. Biochem.* **2000**, *286*, 257-264.

[5] S. Mondal, S. S. Parelkar, M. Nagar, P. R. Thompson, *ACS Chem. Biol.* **2018**, *13*, 1057-1065.

[6] E. Weerapana, A. E. Speers, B. F. Cravatt, *Nat. Protoc.* **2007**, *2*, 1414-1425.

[7] J. K. Eng, A. L. McCormack, J. R. Yates, *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976-989.

[8] D. L. Tabb, W. H. McDonald, J. R. Yates, 3rd, *J. Proteome Res.* **2002**, *1*, 21-26.

[9] E. Weerapana, C. Wang, G. M. Simon, F. Richter, S. Khare, M. B. Dillon, D. A. Bachovchin, K. Mowen, D. Baker, B. F. Cravatt, *Nature* **2010**, *468*, 790-795.