Electronic supplementary materials *Mendeleev Commun.*, 2021, **31**, 867-868

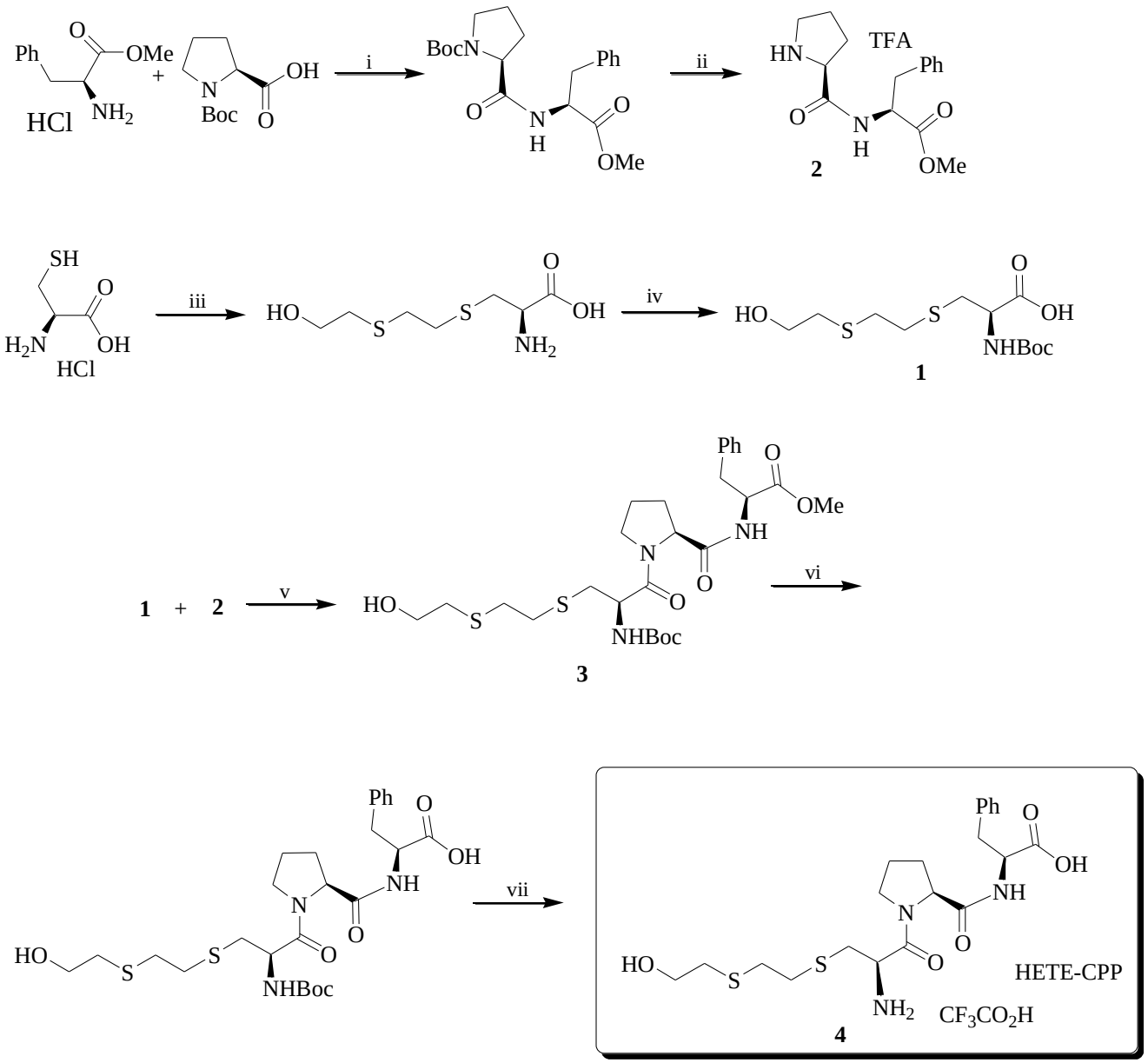
Synthesis of a tripeptide biomarker of exposure to sulfur mustard for support of OPCW biomedical proficiency tests

Victoria V. Abzianidze, Yulia V. Chelusnova, Georgy V. Karakashev, Dmitrii M. Kochura, Michael A. Leninskii, Alexander Yu. Gorbunov, Mariya D. Shachneva, Denis V. Krivorotov and Anton I. Ukolov

Table of contents:

I General remarks

^{1}H-NMR spectra were acquired on an AVANCE III 400 MHz NMR spectrometer (Bruker, Rheinstetten, Germany) in DMSO-d_6. Mass spectra data were acquired on a TSQ Quantum Access Max Mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). High-resolution mass spectra (HRMS) were acquired on a LTQ Orbitrap Velos spectrometer (Thermo Scientific) and on a Bruker MicrOTOF and on a Bruker Daltonics SolariX XR FT-ICR mass spectrometer. Optical rotations were acquired on an Optical Activity Polaar 3005 Polarimeter using a 2.5 cm cell with a Na 589 nm filter and the concentration of samples was denoted as *c*. Organic solvents used were dried by standard methods when necessary. Commercially available reagents were used without further purification. All reactions were monitored by TLC with silica gel coated plates (EMD/Merck KGaA, Darmstadt, Germany), with visualization by UV light and by charring with 0.1% ninhydrin in EtOH. Column chromatography was performed using Merck 60 Å 70–230 mesh silica gel.



Scheme S1 *Reagents and conditions:* i, HBTU, DIPEA, DMF, room temperature, 18 h; ii, TFA, CH_2Cl_2, room temperature, 18 h; iii, $Cl(CH_2)_2S(CH_2)_2OH$, MeONa, MeOH, room temperature, 2 h; iv, Boc_2O, $NaHCO_3$, 1,4-dioxane/water, room temperature, 18 h; v, HBTU, DIPEA, DMF, room temperature, 18 h; vi, LiOH, THF/MeOH/H_2O, room temperature, 18 h; vii, TFA, room temperature, 10 min.

II Synthesis

(2*R*)-2-(*tert*-Butoxycarbonylamino)-3-[2-(2-hydroxyethylsulfanyl)ethylsulfanyl]propanoic acid 1

Step 1. (2*R*)-2-Amino-3-[2-[(2-hydroxyethylsulfanyl)ethylsulfanyl]propanoic acid. Sodium metal (0.413 g, 17.9 mmol) was dissolved in a solution of L-cysteine hydrochloride (0.943 g, 5.98 mmol) in methanol (20 ml). 2-(2-Chloroethylsulfanyl)ethanol ('half sulfur mustard', 0.925 g, 6.58 mmol) was added dropwise, and the mixture was left for 2 h at room temperature. The solution was concentrated *in vacuo*. The crude intermediate product was used in the next step without further purification. MS revealed an $[M + H]^+$ ion with exact mass 226.06, corresponding to the molecular formula $C_7H_{15}NO_3S_2$.

Step 2. Sodium bicarbonate (1 g, 0.012 mol) was added to a 1,4-dioxane/water solution (45 ml, 1:2) containing intermediate product from the step 1 (5.98 mmol). The mixture was cooled to 0 °C, and di-*tert*-butyl dicarbonate (1.37 g, 6.3 mmol) was added. The resulting stirred mixture was allowed to warm to room temperature overnight, and then poured into H_2O (100 ml) and extracted with EtOAc (2×50 ml). The combined organic extracts were dried over $MgSO_4$ and concentrated *in vacuo*. The resulting white solid was used in the next step without further purification. MS revealed an $[M + Na]^+$ ion with exact mass 348.06, corresponding to the molecular formula $C_{12}H_{22}NNaO_5S_2$.

Methyl L-prolyl-L-phenylalaninate hydrotrifluoroacetate 2

Step 1. Methyl 1-(*tert*-butoxycarbonyl)-L-prolyl-L-phenylalaninate. Boc-L-Pro (1 g, 4.64 mmol) was dissolved in DMF (10 ml), and then HBTU (1.760 g, 4.64 mmol) and DIPEA (0.882 ml, 5.06 mmol) were added at 0 °C. After 10 min, L-phenylalanine methyl ester hydrochloride (0.92 g, 4.22 mmol) was added together with DIPEA (0.882 ml, 5.06 mmol). The mixture was stirred for additional 10 min at 0 °C, allowed to warm to room temperature, and stirred overnight. The solvent was removed *in vacuo*, and the residue was suspended in EtOAc and washed with 5% citric acid (2×200 ml), $NaHCO_3$, and brine. The organic layers were combined, dried over $MgSO_4$, filtered, and concentrated *in vacuo* to leave a crude solid product. The product was purified by silica gel column chromatography (CH_2Cl_2/MeOH, 50:1) to obtain the pure title product (76%). R_f 0.75 (CH_2Cl_2/MeOH 20:1). MS revealed an $[M + H]^+$ ion with exact mass 377.19, corresponding to the molecular formula $C_{20}H_{28}N_2O_5$.

***Step 2*.** The intermediate product from step 1 was deprotected by treatment with TFA/CH_2Cl_2 (20 ml, 1:1) overnight to give product **2** (quantitative). MS revealed an $[M + H]^+$ ion with exact mass 277.14, corresponding to the molecular formula $C_{15}H_{20}N_2O_3$.

Methyl *N*-(*tert*-butoxycarbonyl)-*S*-[2-(2-hydroxyethylthio)ethyl]-L-cysteinyl-L-prolyl-L-phenylalaninate 3. Compound **1** (1 g, 0.00307 mol) was dissolved in DMF (10 ml), and then HBTU (1.165 g, 0.00307 mol) and DIPEA (0.583 ml, 0.00335 mol) were added at 0°C. After 10 min, compound **2** (1.089 g, 0.00279 mol) was added together with DIPEA (0.583 ml, 0.00335 mol). The mixture was stirred for additional 10 min at 0°C, allowed to warm to room temperature, and stirred overnight. The solvent was removed under reduced pressure, and the residue was suspended in EtOAc and washed with 5% citric acid (200 ml), $NaHCO_3$ and brine. The organic layers were combined, dried over $MgSO_4$, filtered, and concentrated *in vacuo* to leave crude solid product. The product was purified by silica gel column chromatography (CH_2Cl_2/MeOH, 30:1) to obtain pure title product **3** (70%). MS revealed an $[M + H]^+$ ion with exact mass 584.22, corresponding to the molecular formula $C_{27}H_{41}N_3O_7S_2$.

S-[2-(2-Hydroxyethylthio)ethyl]-L-cysteinyl-L-prolyl-L-phenylalanine hydrotrifluoroacetate 4 (HETE–CPP)

Step 1. *N*-(*tert*-Butoxycarbonyl)-*S*-[2-(2-hydroxyethylthio)ethyl]-L-cysteinyl-L-prolyl-L-phenylalanine. Lithium hydroxide monohydrate (0.235 g, 5.6 mmol) was added to a solution of compound **3** (1.09 g, 1.87 mmol) in THF/MeOH/H_2O (50 ml, 3:l:l), and the mixture was stirred at room temperature overnight. The solution was acidified to pH 2 with saturated aqueous citric acid, and the mixture was extracted with EtOAc (3x20 ml). The organic layer was dried over $MgSO_4$ and concentrated under vacuum to give the intermediate product (80%). The resulting highly hygroscopic white solid was used in the next step without further purification. HRMS revealed an $[M+H]^+$ ion with exact mass 570.2297, corresponding to the molecular formula $C_{26}H_{40}N_3O_7S_2$.

Step 2. A solution of product from step 1 (0.5 g, 0.878 mmol) in trifluoroacetic acid (5 ml) was stirred at room temperature for 10 min. After removing the solvent by rotary evaporation, the residue was purified by reversed phase HPLC (eluent acetonitrile–0.1% TFA, 45:55) to give 60 mg (15%) of yellow oil. HRMS revealed an $[M+H]^+$ ion with exact mass 470.17771, corresponding to the molecular formula $C_{21}H_{31}N_3O_5S_2$. ^{1}H NMR (400 MHz, DMSO-d_6), δ, ppm: 12.78 (br s, 1H), 8.50-8.07 (m, 2H), 7.33-7.19 (m, 5H), 4.80 (br s, 1H), 4.57-4.28 (m, 3H), 3.77-3.67 (m, 1H), 3.58-3.46 (m, 3H), 3.15-2.89 (m, 3H), 2.86-2.58 (m, 8H), 2.33 (m, 1H), 2.14-2.01 (m, 1H), 1.99-1.71 (m, 3H).

III Analysis of plasma

The quantitative and qualitative analysis of the plasma samples for *S*-hydroxyethylthioethyl tripeptide adduct [S-HETE]-Cys-Pro-Phe from protein adducts (proteinase K digestion and SPE) was performed by LC-HRMS/MS (+ESI).

Sample preparation for plasma [S-HETE]-Cys-Pro-Phe analysis by LC-HRMS/MS (+ESI)

Acetone (0.6 ml) was added to plasma (100 µl). The mixture was centrifuged for 3 min at 1500 rpm. The supernatant was removed, and the precipitate was dried for 15 min at room temperature. 100 µl of 0.1 % formic acid in deionized water, 800 µl of 50 mM ammonium bicarbonate (pH 7.8), and 200 µl of 10 mg/ml proteinase K (from Tritirachium album, Sigma P8044-5G) in 50 mM ammonium bicarbonate were added to the dry precipitate. The resulting solution was shaken for 5 min and then heated at 50 °C for 90 min. Oasis HLB (60 mg/3 ml) cartridge was preconditioned with MeOH (1 ml) and deionized water (1 ml) on a vacuum manifold. The solution was loaded onto Oasis HLB, the cartridge was washed with deionized water (1.0 ml), and the target analyte was eluted with acetonitrile (1 ml). The eluate was dried over nitrogen at 60 °C and reconstituted in 50 µl of 0.1 % formic acid in deionized water.

IV Characterization and spectra charts

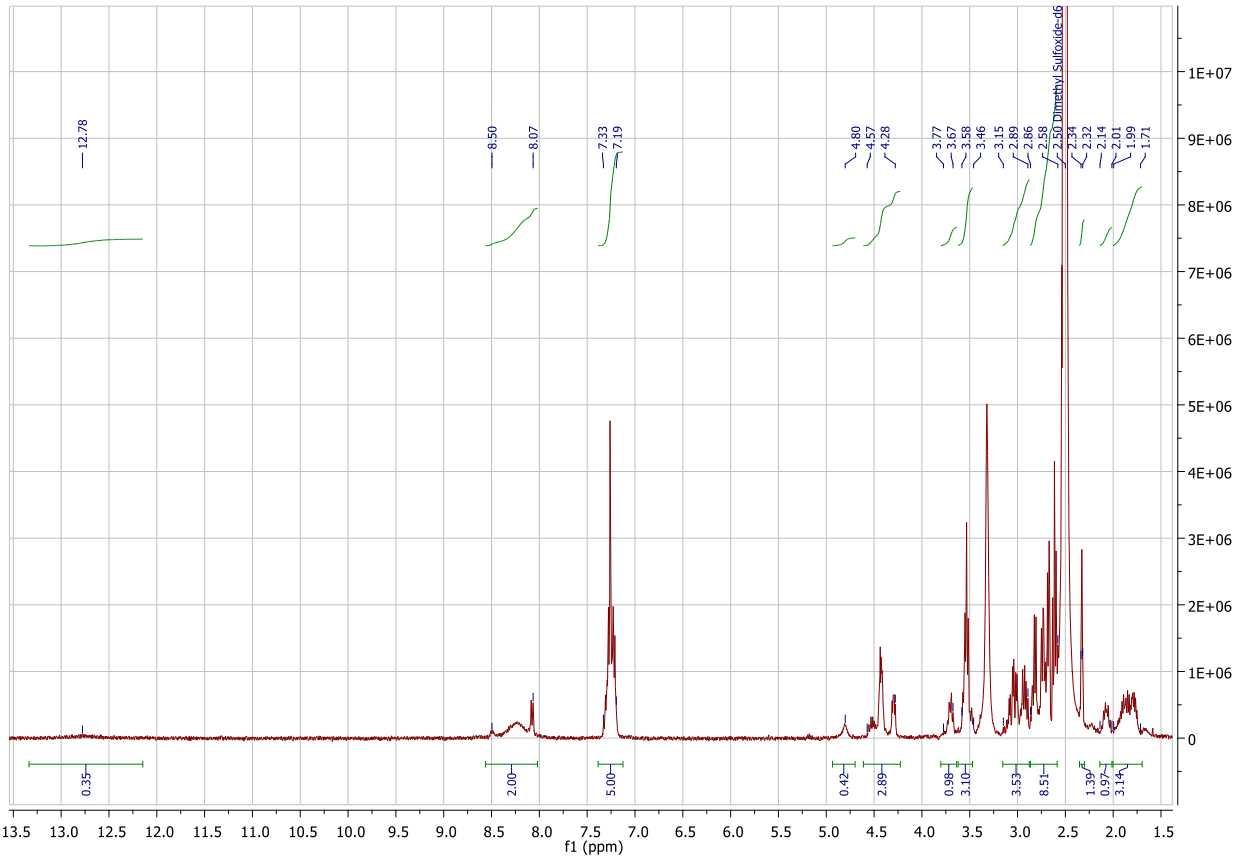


Figure S1 ^{1}H NMR spectrum of HETE-CPP **4** in DMSO-d_6

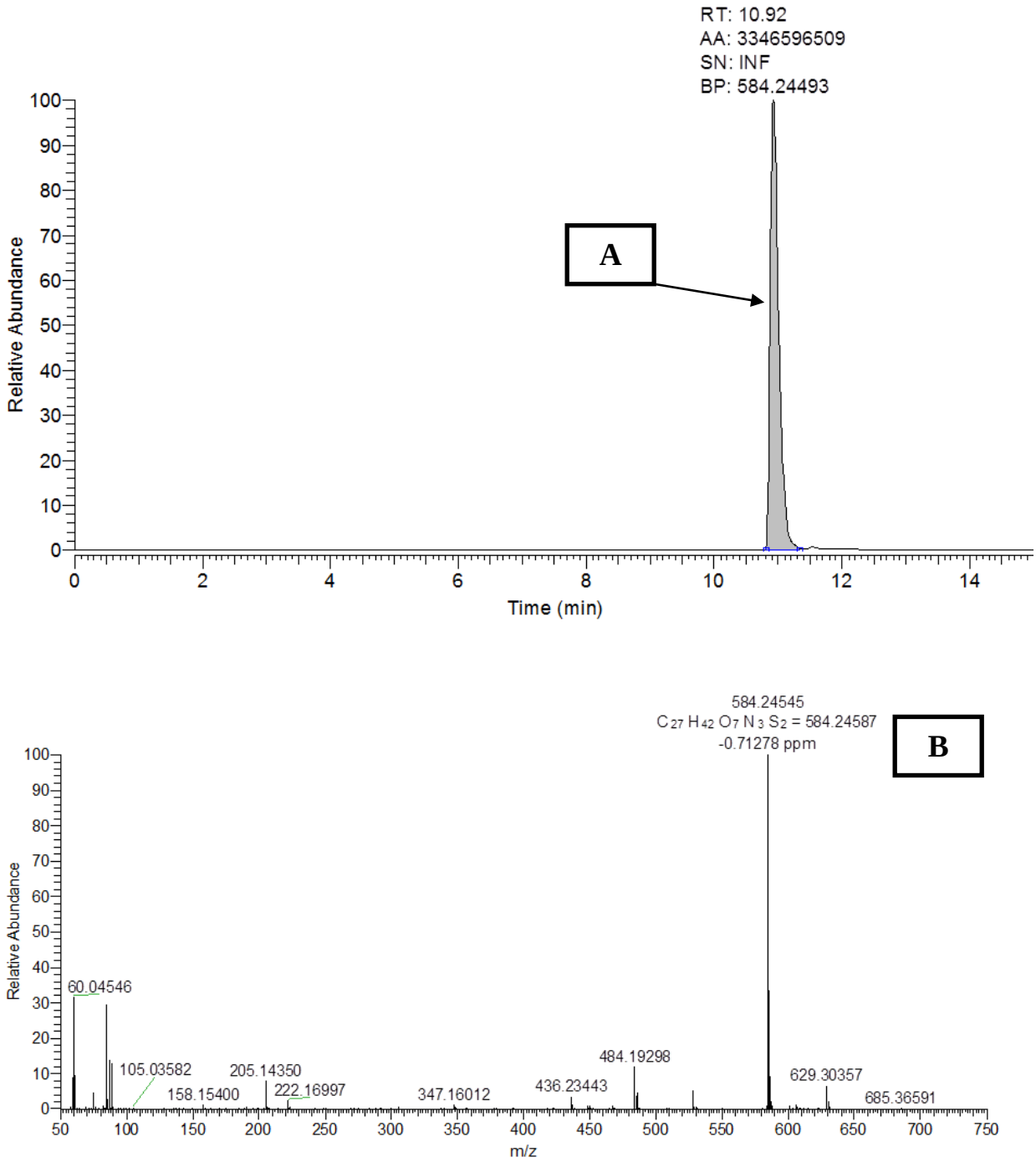


Figure S2 Intermediate product **3** (584.2459 m/z) HRMS data: the base peak intensity chromatogram obtained from ESI-MS analysis (A) and mass-spectrum (B) obtained from ESI-MS analysis.

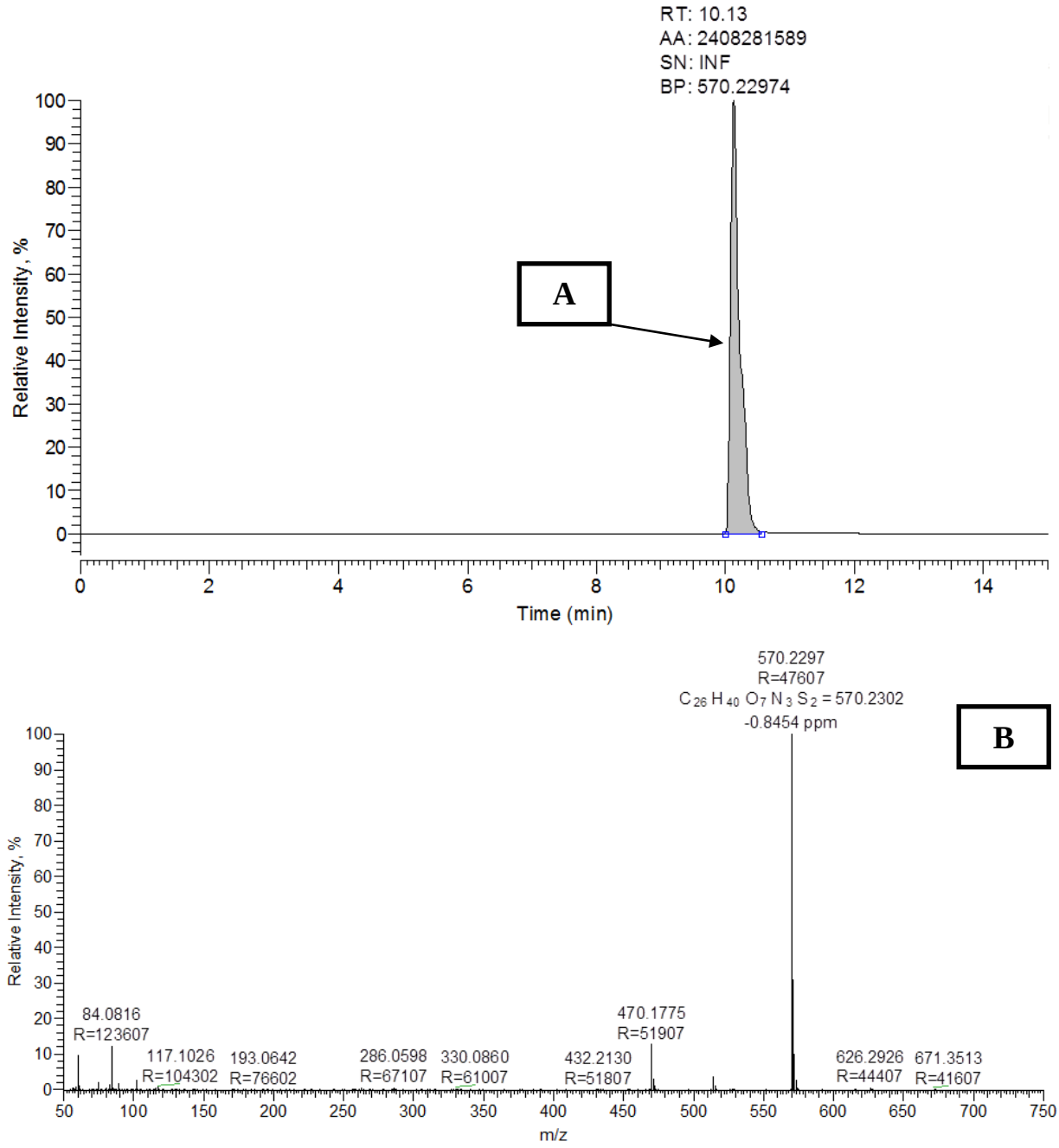


Figure S3 Intermediate product from step vi, Scheme S1, (570.2302 m/z) HRMS data: the base peak intensity chromatogram obtained from ESI-MS analysis (A) and mass-spectrum (B) obtained from ESI-MS analysis.

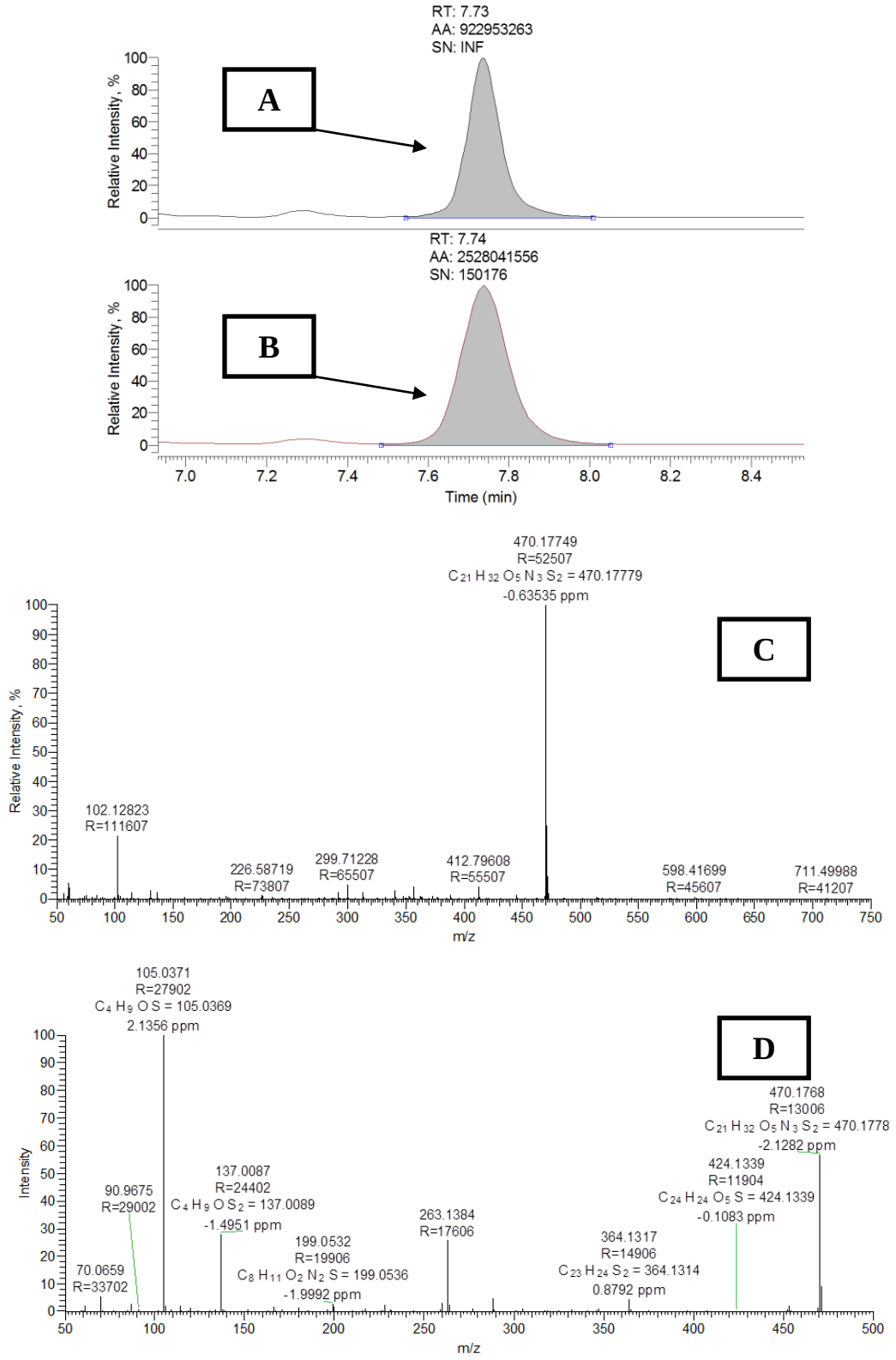


Figure S4 A - the base peak intensity chromatogram of HETE–CPP **4** with 470.1778 m/z obtained from ESI-MS analysis; B - the base peak intensity chromatogram of characteristic product-ions 105.03686 and 137.00893 obtained by higher energy collision dissociation (HCD) of HETE–CPP **4**; C – mass-spectrum of HETE–CPP obtained from ESI-MS analysis; D - mass-spectrum of HETE–CPP **4** obtained by higher energy collision dissociation (HCD).