Supplementary file

Antiproliferative potential of *Physalis peruviana*-derived magnolin against pancreatic cancer: A comprehensive *in vitro* and in silico study

Ahmed M. Sayed,[a] Seham S. El-Hawary,[b] Usama Ramadan Abdelmohsen[c,d] and Mosad A. Ghareeb[*e]

[a]Department of Pharmacognosy, Faculty of Pharmacy, Nahda University, Beni-Suef 62513, Egypt

[b]Pharmacognosy Department, Faculty of Pharmacy, Cairo University, Cairo, Egypt

[c]Department of Pharmacognosy, Faculty of Pharmacy, Minia University, Minia 61519, Egypt

[d]Department of Pharmacognosy, Faculty of Pharmacy, Deraya University, New Minia 61111, Egypt

[e]Medicinal Chemistry Department, Theodor Bilharz Research Institute, Kornaish El-Nile, Warrak El-Hadar, Imbaba (P.O. 30), Giza 12411, Egypt

Abstract

Physalis peruviana L., is a common edible fruit in Egypt and other regional countries. In the present study, we investigated its crude extract as a potential source of antiproliferative secondary metabolites. Upon bioactivity guided-solvent fractionation, ethyl acetate extract showed preferential activity toward the human pancreatic cancer cell line (PANC-1) with an IC_{50} value of 5.23 ± 0.2 µg/mL. The subsequent HR-LCMS-guided and biological activity-guided isolation revealed magnolin as a potent preferential antiproliferative agent against PANC-1 with an IC_{50} of 0.51 ± 0.46 µM that was comparable to that of the positive control Doxorubocin (IC_{50} of 0.17 ± 0.15 µM). Moreover, magnolin showed much less cytotoxicity in comparison with the positive control Doxorubicin (6.96% and 30.48% growth inhibition, respectively at 5 µg/mL) towards normal human cells (i.e. dermal fibroblasts; HDFa). Furthermore, magnolin was able to induce a concentration-dependent suppression of the formation of PANC-1 colonies, where the treatment of the tumor cells with 25 nM, 50 nM, and 100 nM of the compound resulted in a 36%, 57, and 78% reduction, respectively in the PANC-1 colony formation. Additionally, magnolin was observed to limit PANC-1 tumor cell migration in tumor cells wound healing assay indicating a substantial anti-migratory effect against PANC-1 cell lines. A subsequent, *in silico*-based study of this compound structure putatively suggested matrix metalloproteinase-3 (MMP3) as the molecular target that mediate these observed effects on PANC-1 cells. Absolute binding free energy estimation ($\Delta G_{binding}$) and 100 ns long molecular dynamics simulation (MDS) experiments indicated that magnolin structure has a good affinity towards the MMP3's active site and can achieve significantly stable binding inside it. Accordingly, upon experimental validation, magnolin was found to inhibit the catalytic activity of MMP3 in a dose-dependent manner with a nanomolar IC_{50} value of 185 nm $\pm$ 4.86 and K_i of 112 nm $\pm$ 6.31. In conclusion, our results clearly revealed that magnolin derived from *P. peruviana* is an interesting antiproliferative and antimetastatic agent against PANC-1 cells with potent inhibitory activity against MMP3. Further *in vivo* evaluation will be of great interest in the future.

Keywords: *Physalis peruviana*, magnolin, PANC-1, antimigitatory, in silico, MMP3 inhibitor.
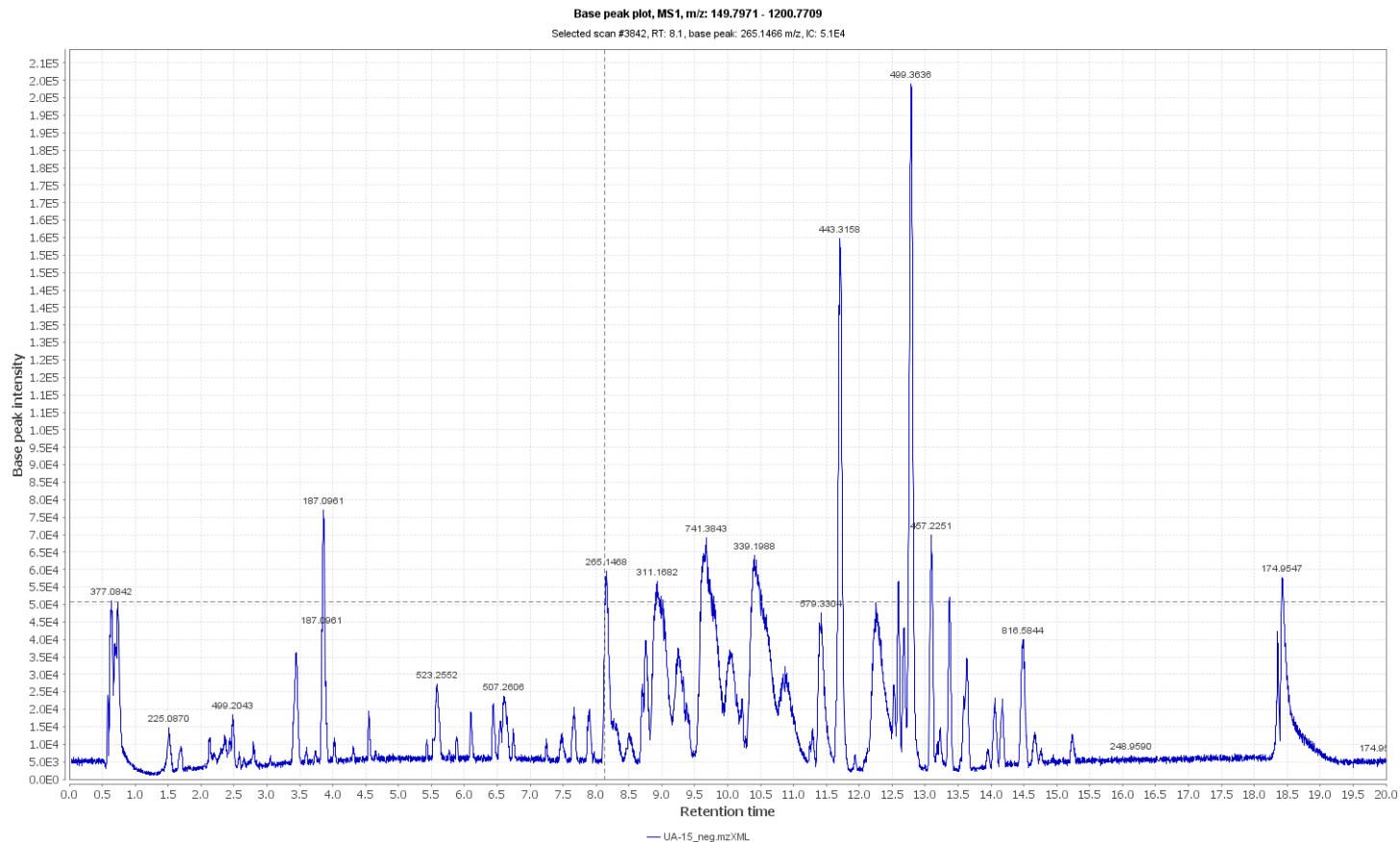
Base peak plot, MS1, m/z: 149.7971 - 1200.7709

Selected scan #3842, RT: 8.1, base peak: 265.1466 m/z, IC: 5.1E4

Fig. S1 LC-HRESIMS chromatogram of the dereplicated metabolites of *P. peruviana* (positive mode).
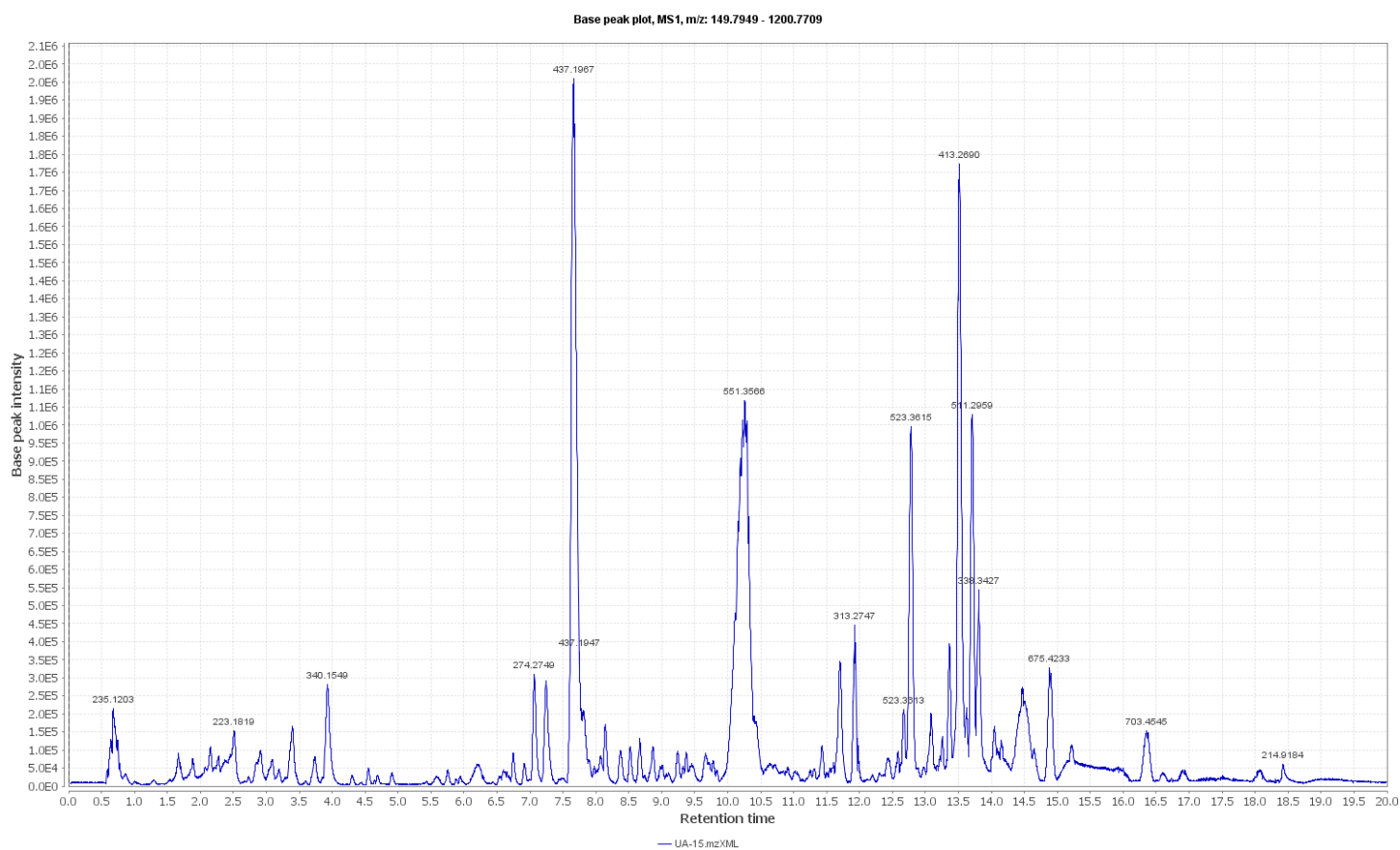
Base peak plot, MS1, m/z: 149.7949 - 1200.7709

Fig. S2 LC-HRESIMS chromatogram of the dereplicated metabolites of *P. peruviana* (negative mode).
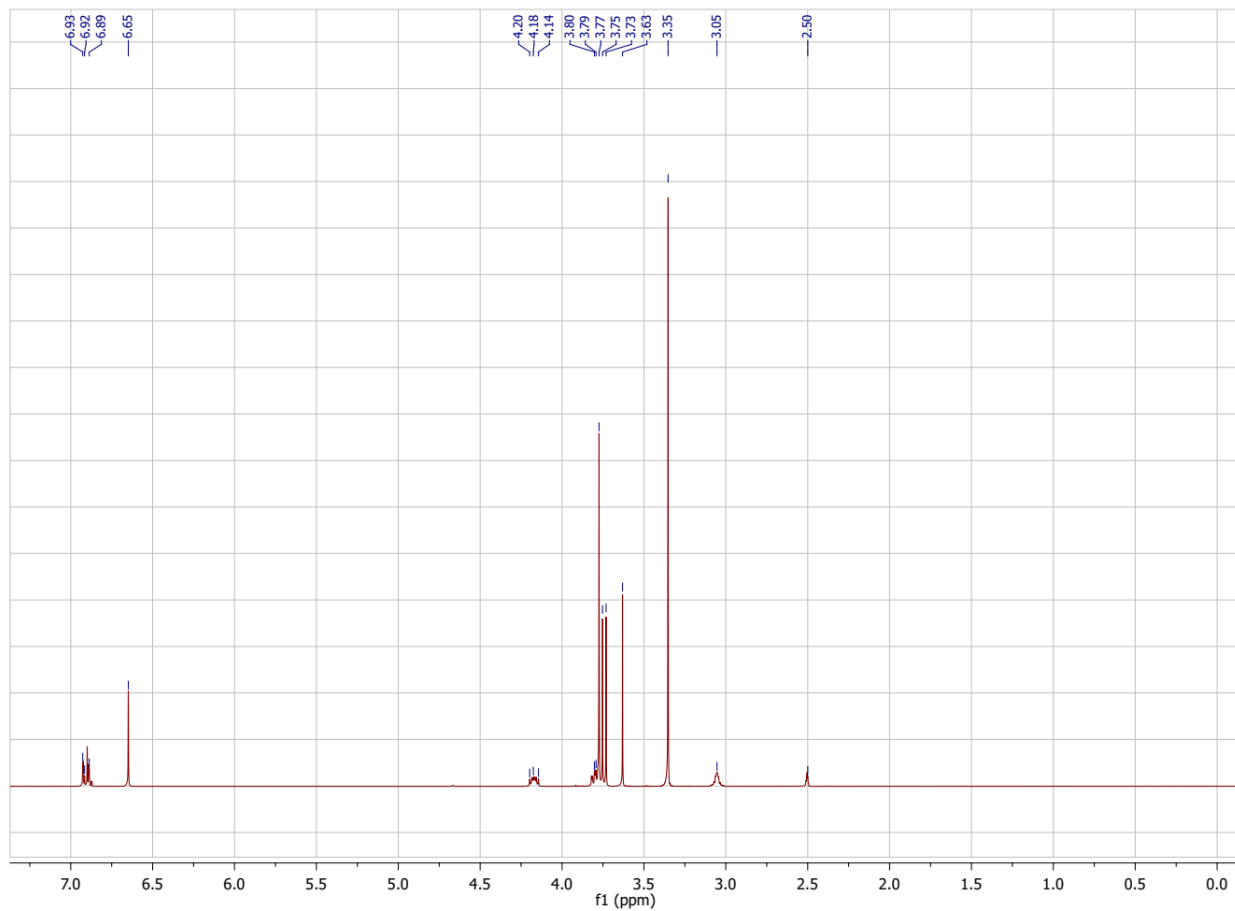
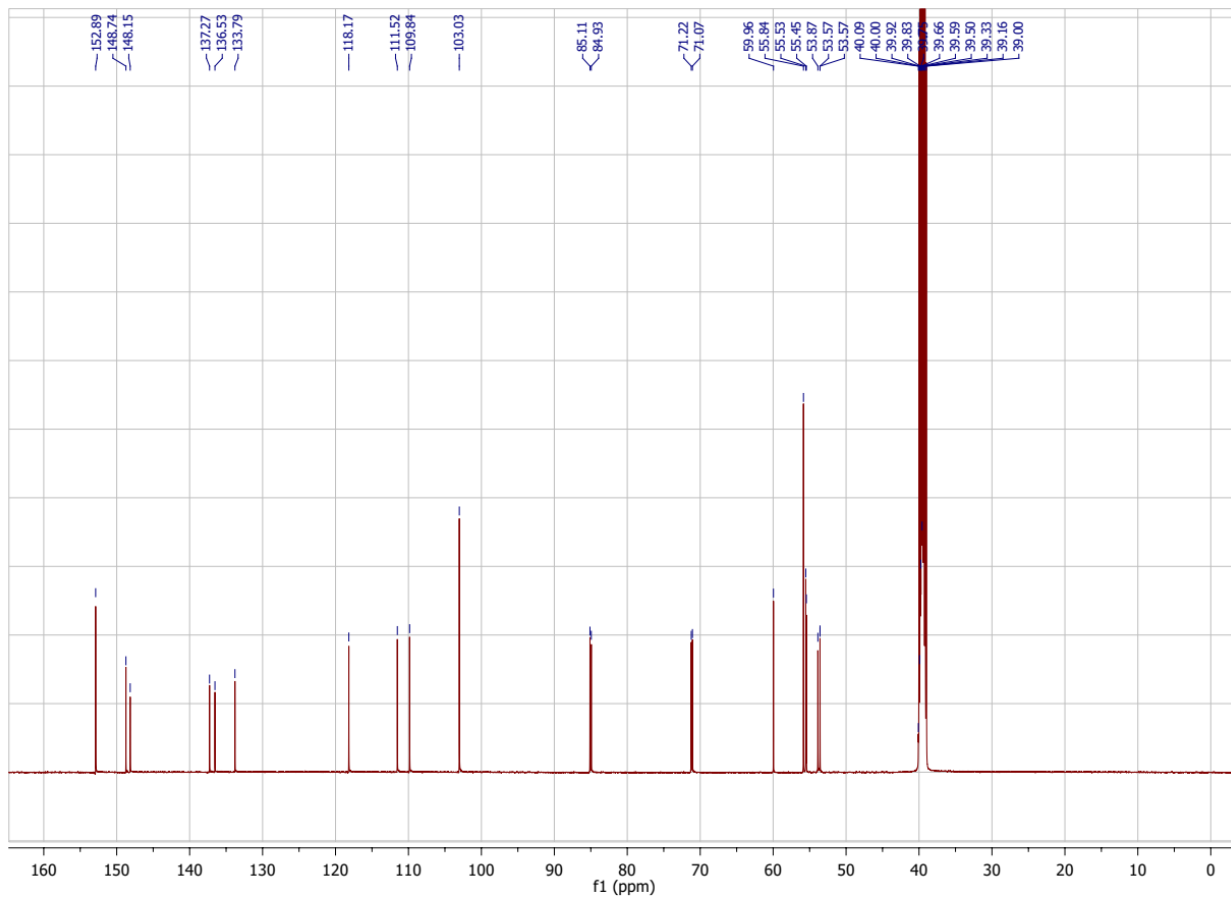**Fig. S3** ¹H-NMR spectrum of magnolin in DMSO-*d6*, 400 MHz.

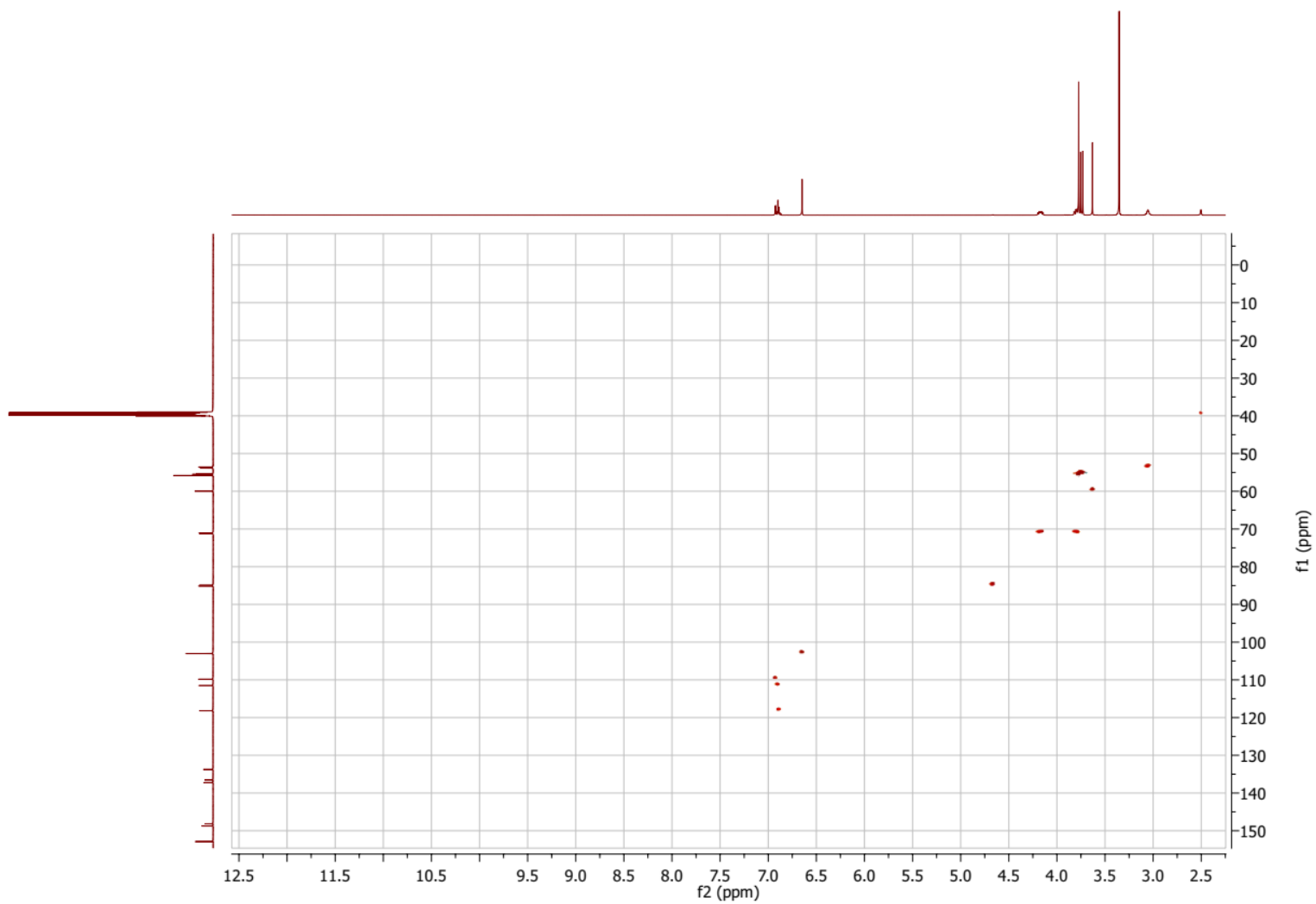**Fig. S4** ^{13}C-NMR spectrum of magnolin in DMSO-*d6*, 100 MHz.

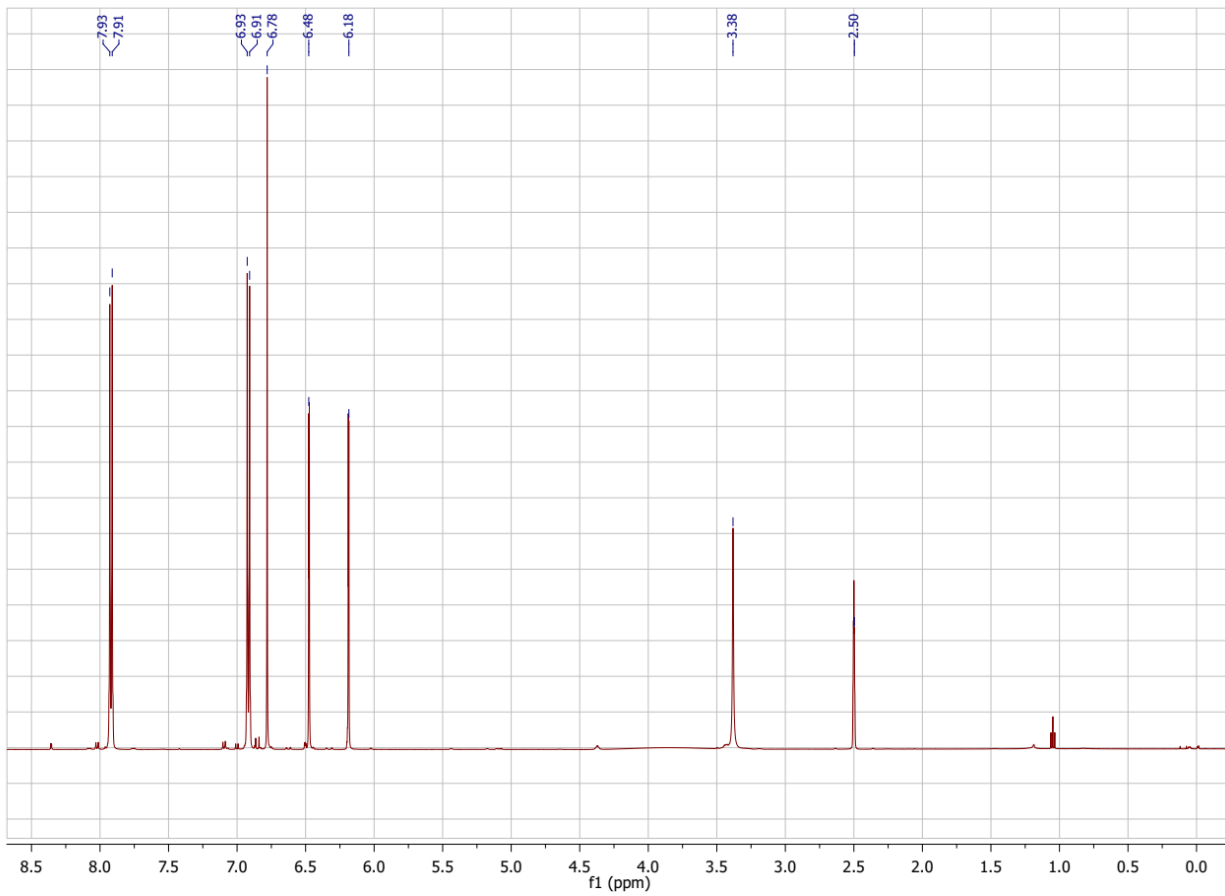**Fig. S5** ^{1}H-^{13}C-HSQC spectrum of magnolin in DMSO-*d6*.

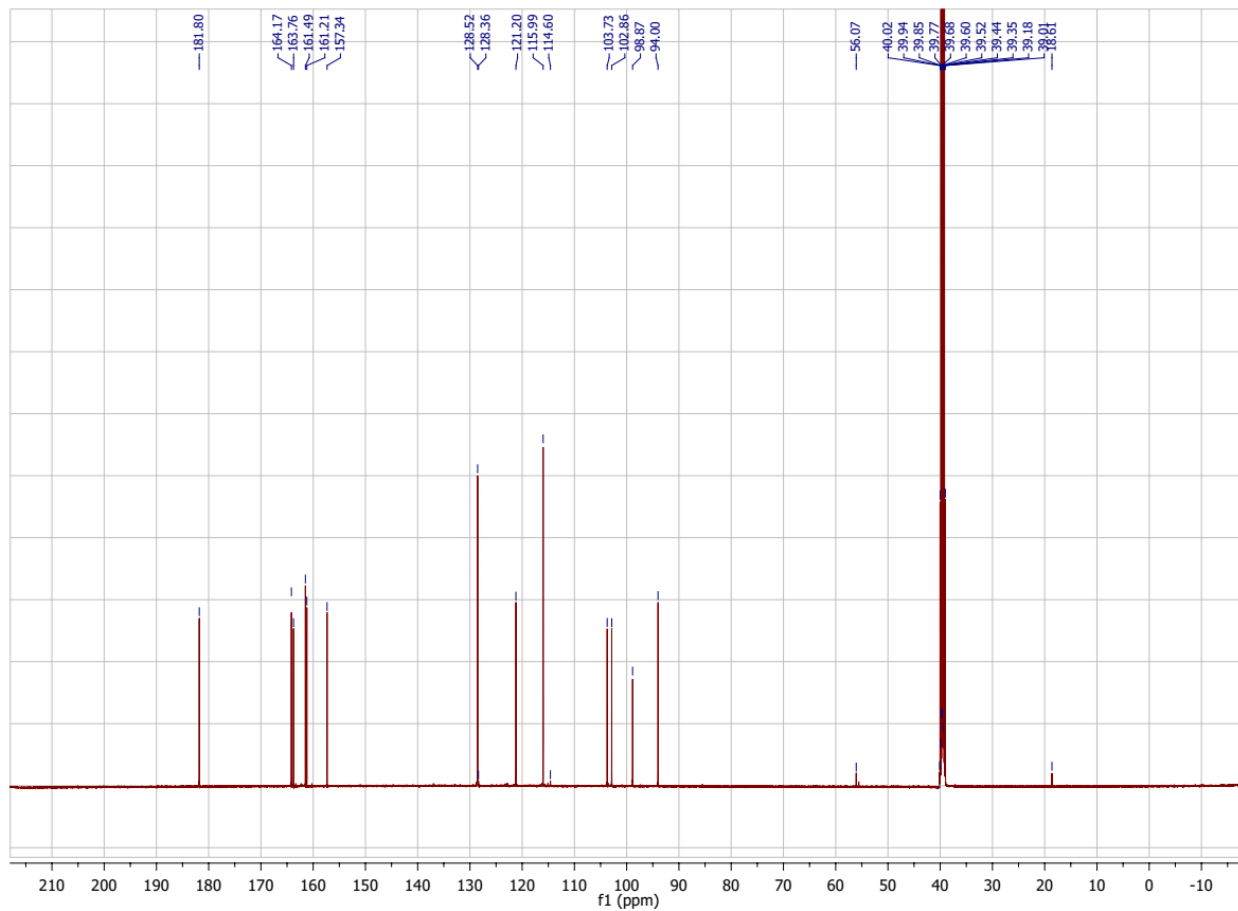**Fig. S6** ^{1}H-NMR spectrum of apigenin in DMSO-*d6*, 400 MHz.

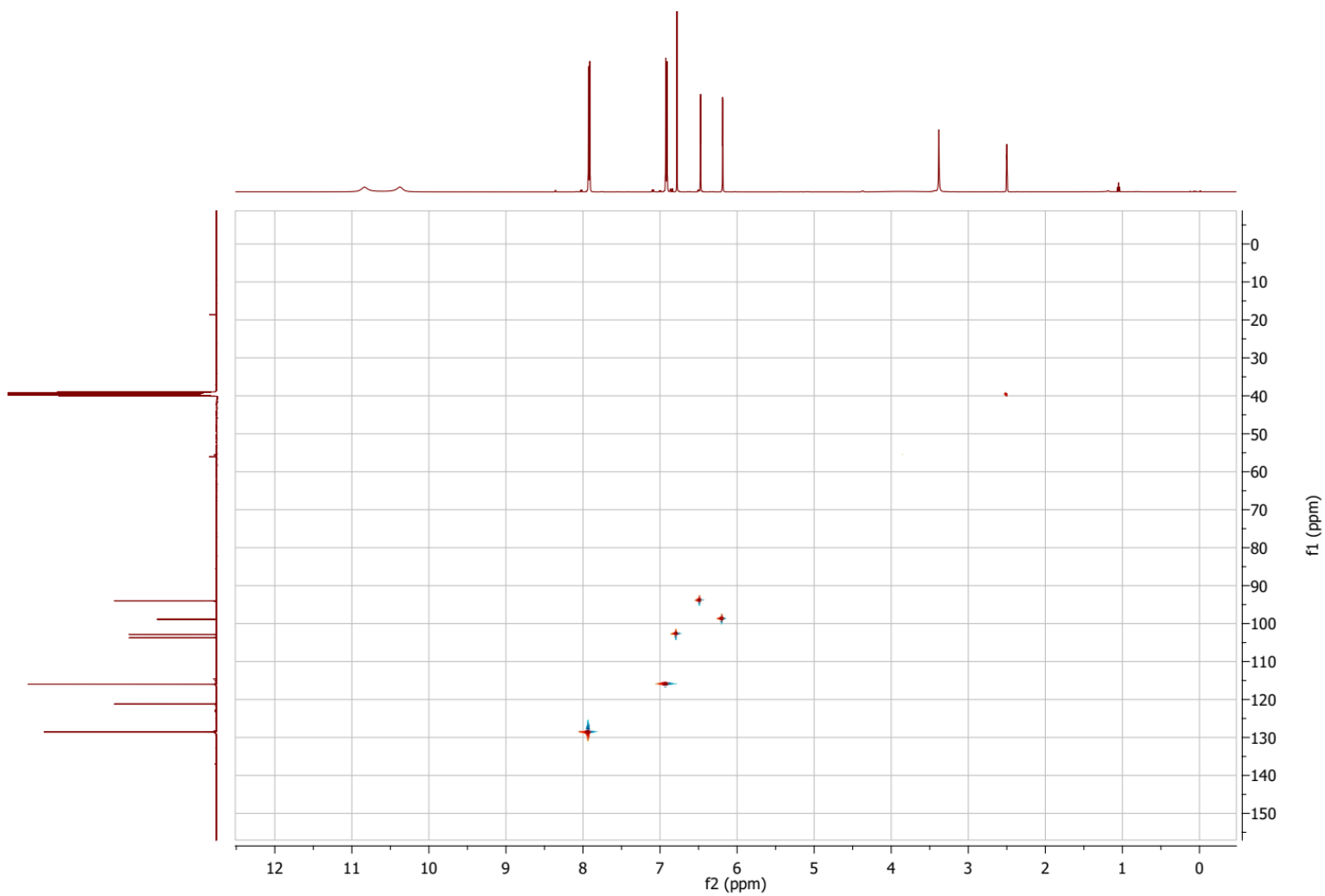**Fig. S7** ^{13}C-NMR spectrum of apigenin in DMSO-*d6*, 100 MHz.

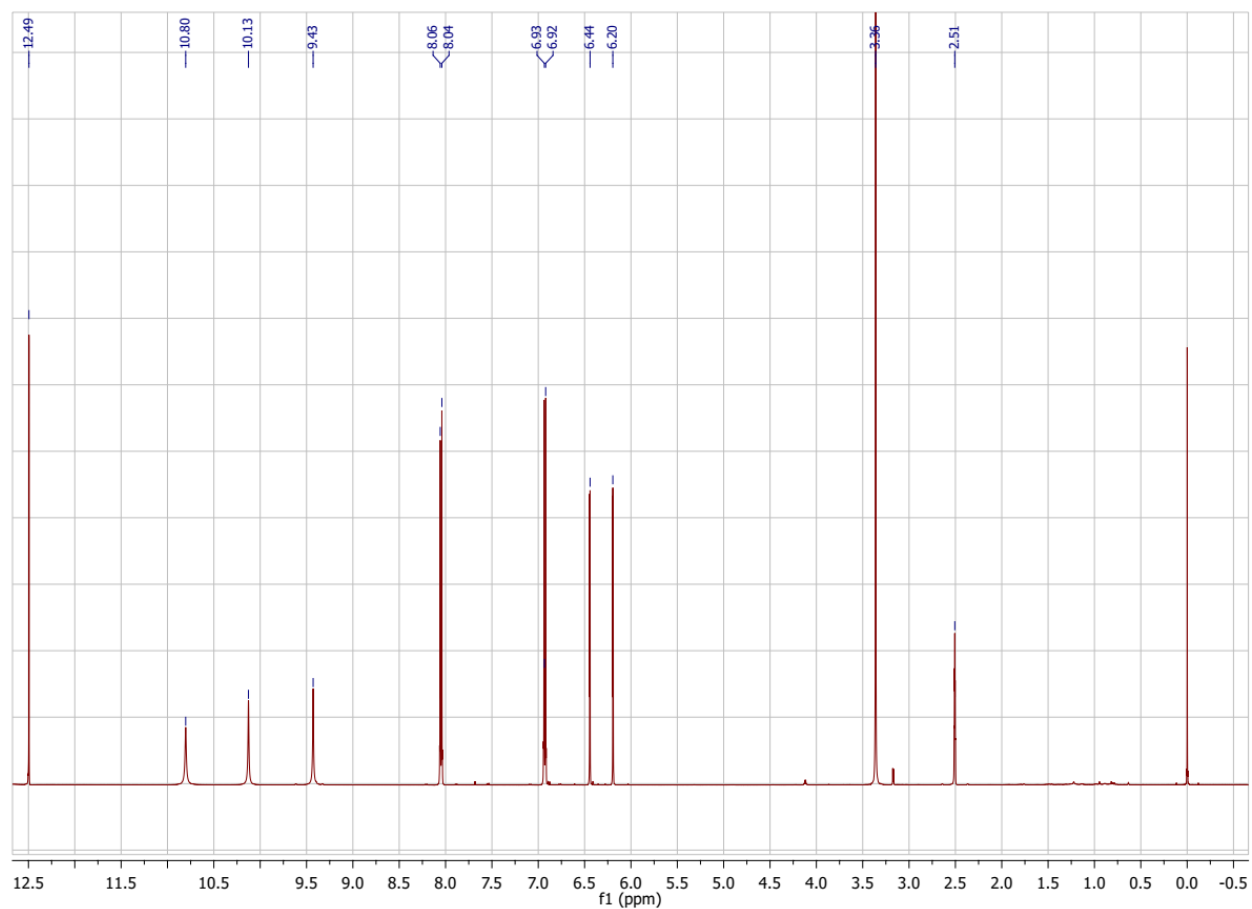**Fig. S8** ^{1}H-^{13}C-HSQC spectrum of apigenin in DMSO-*d6*.

Fig. S9 ^{1}H-NMR spectrum of kaempferol in DMSO-*d6*, 400 MHz.
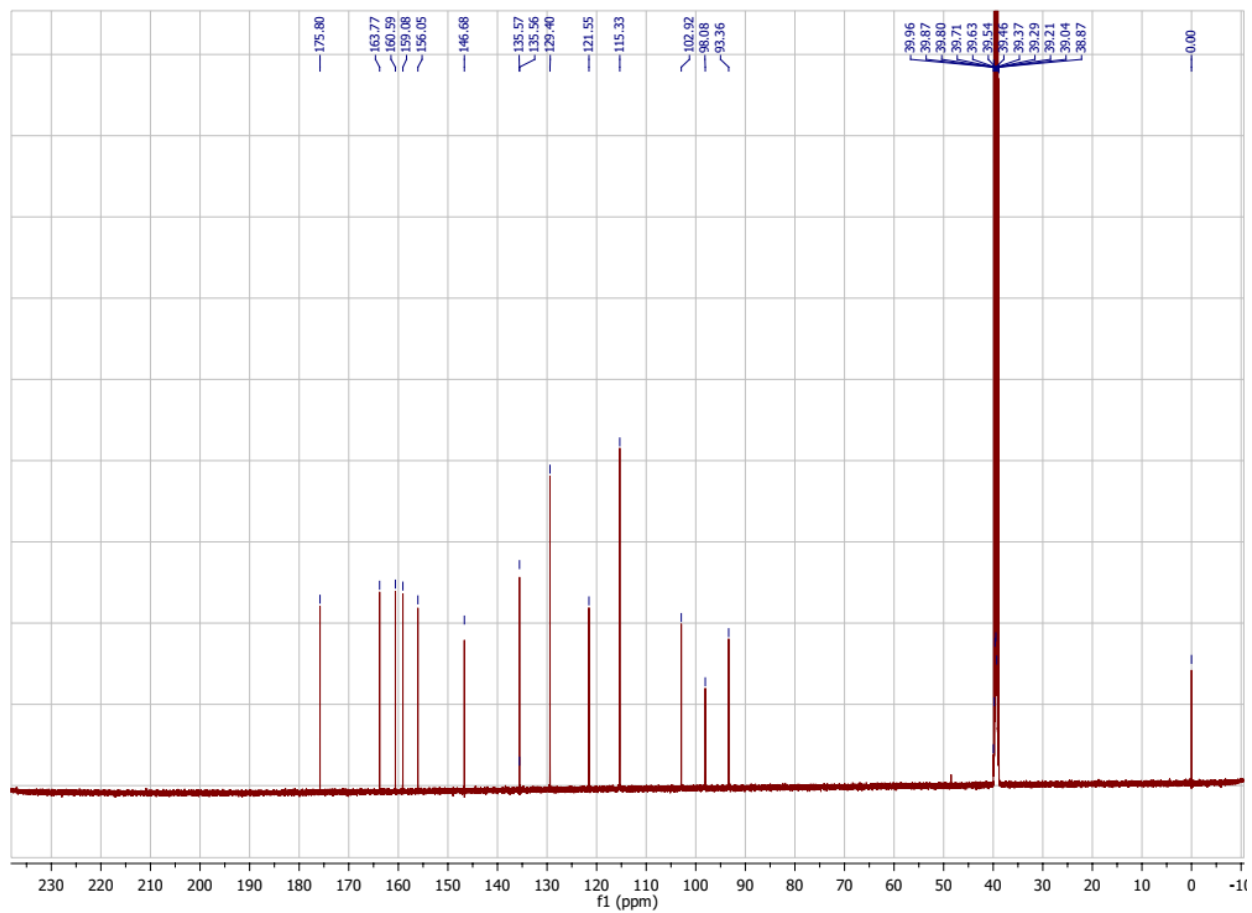
Fig. S10 ¹³C-NMR spectrum of kaempferol in DMSO-*d6*, 100 MHz.

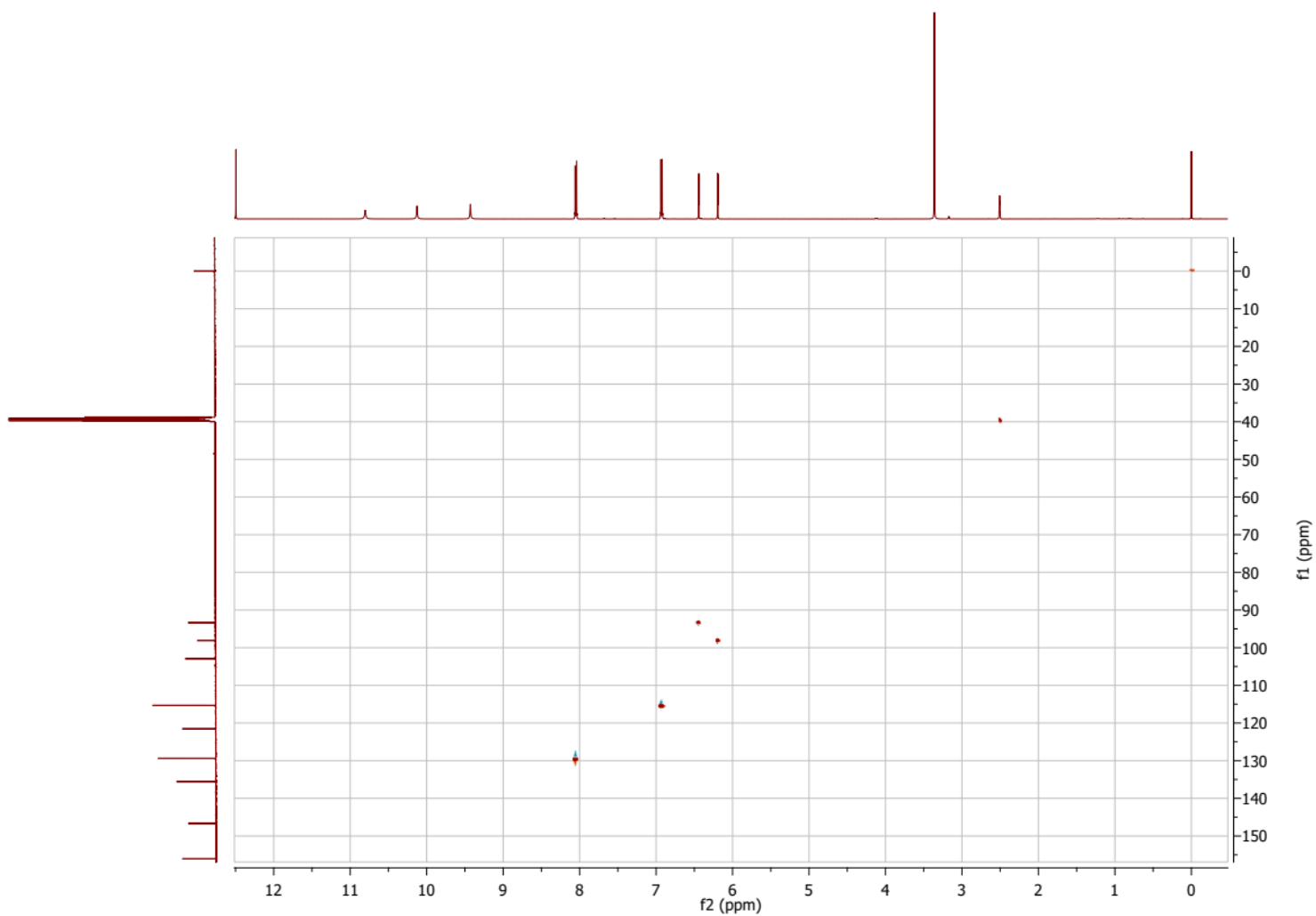**Fig. S11** ^{1}H-^{13}C-HSQC spectrum of kaempferol in DMSO-*d6*.

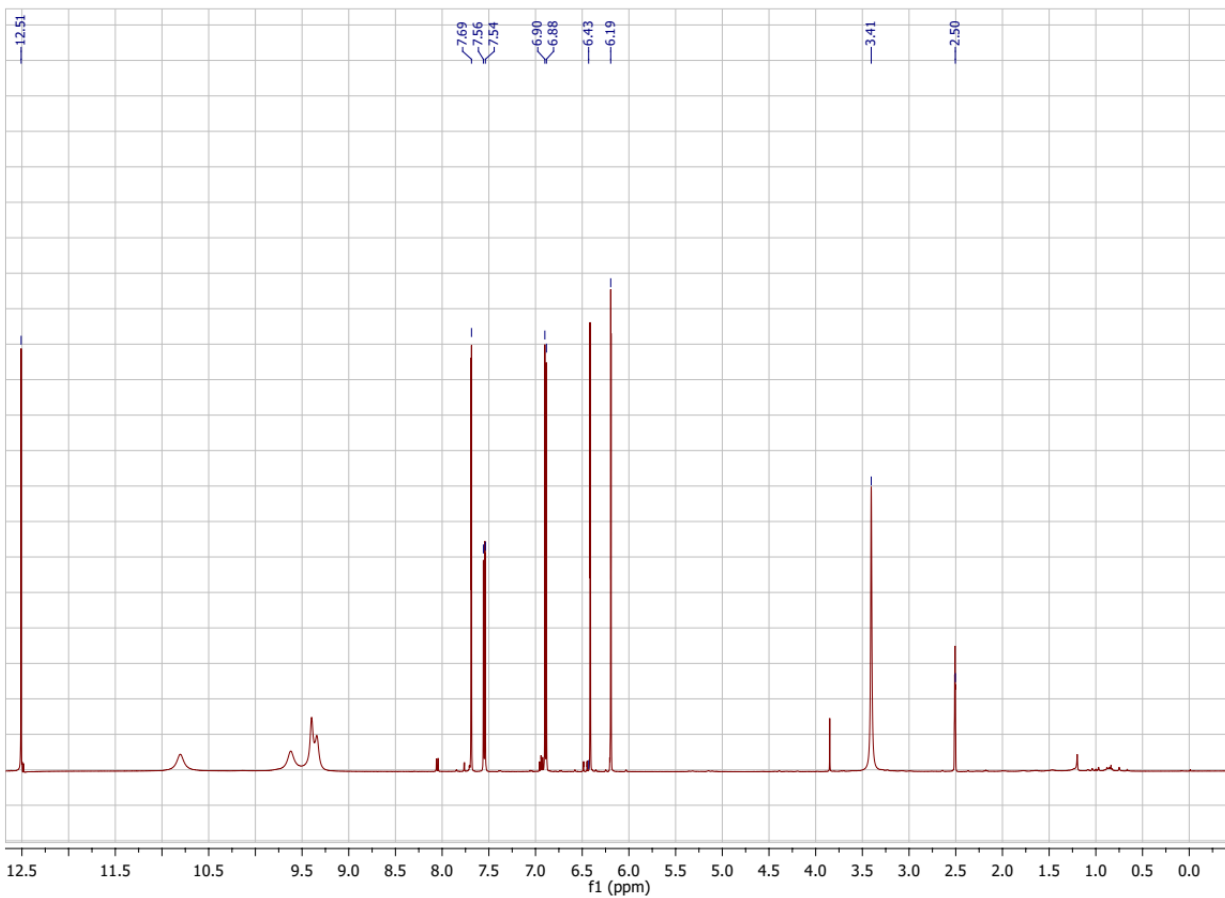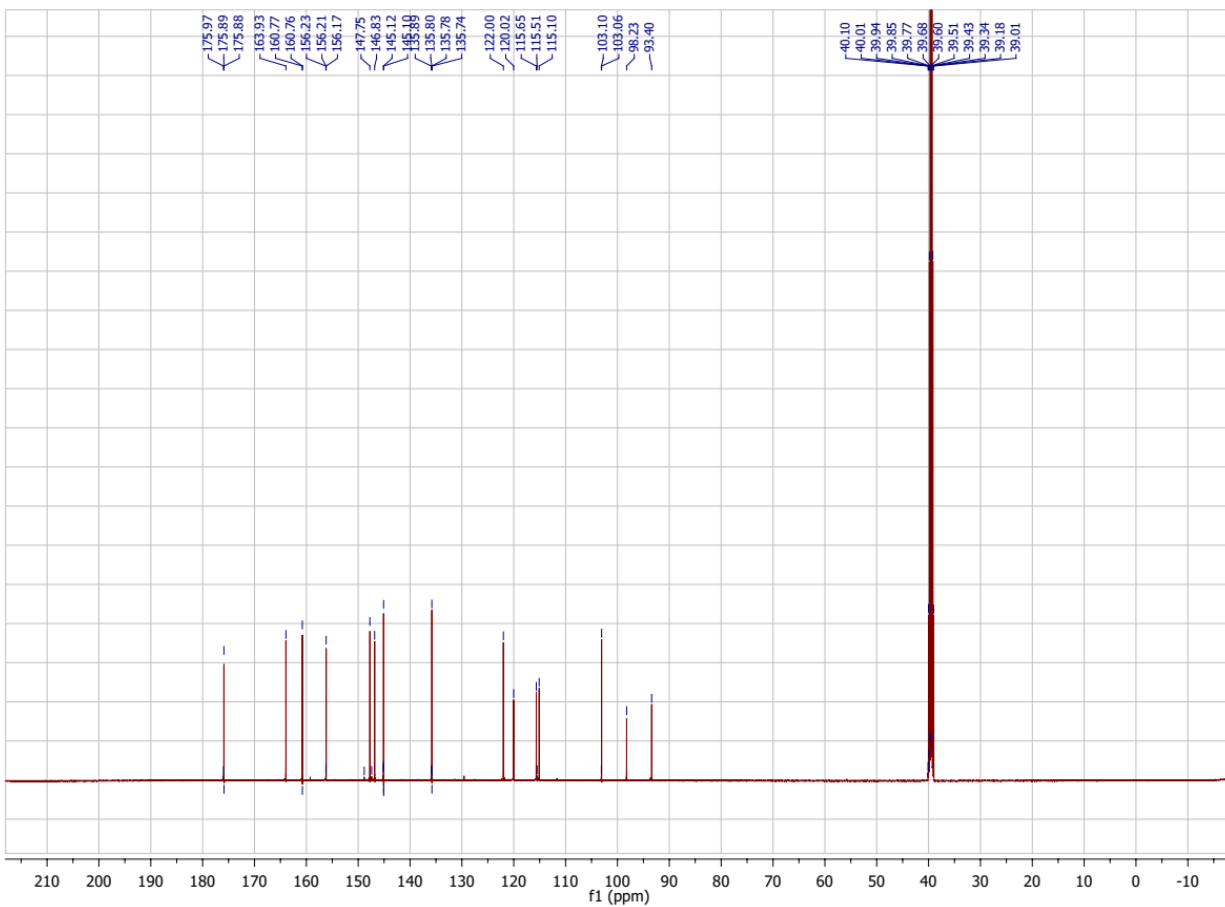**Fig. S12** ¹H-NMR spectrum of quercetin in DMSO-*d6*, 400 MHz.

Fig. S13 ^{13}C-NMR spectrum of quercetin in DMSO-*d6*, 100 MHz.
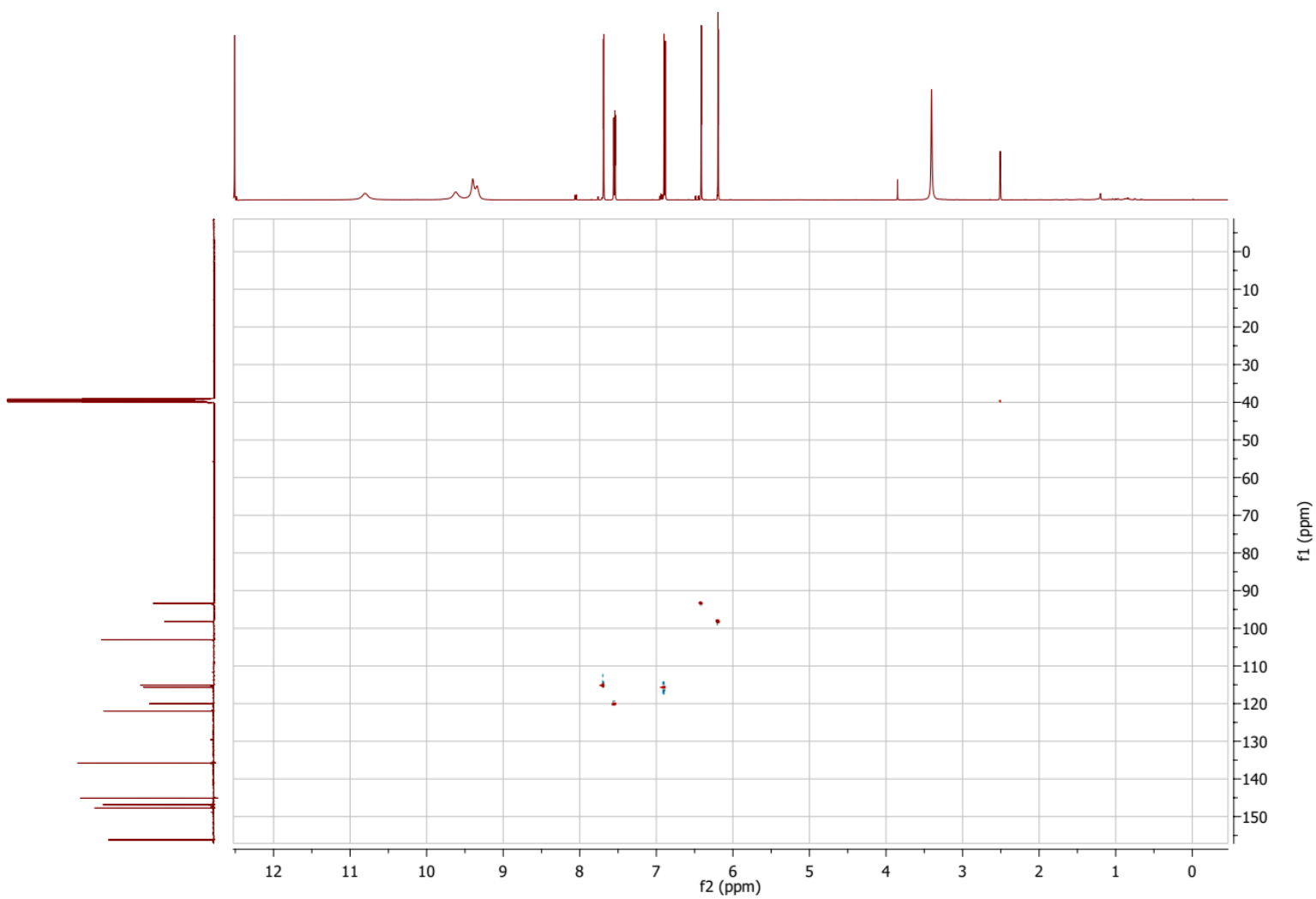
Fig. S14 ^{1}H-^{13}C-HSQC spectrum of quercetin in DMSO-*d6*.
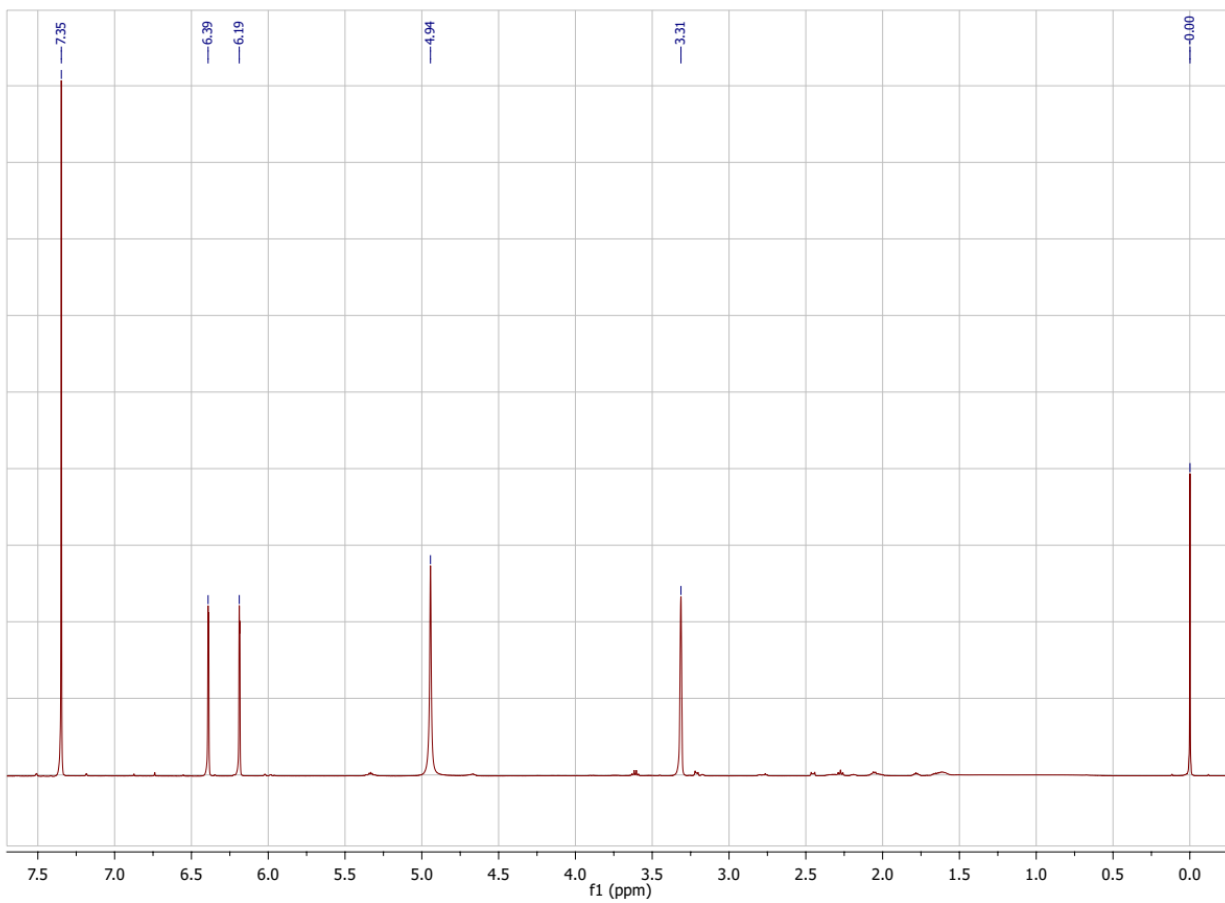
Fig. S15 ¹H-NMR spectrum of myricetin in CD₃COD, 400 MHz.
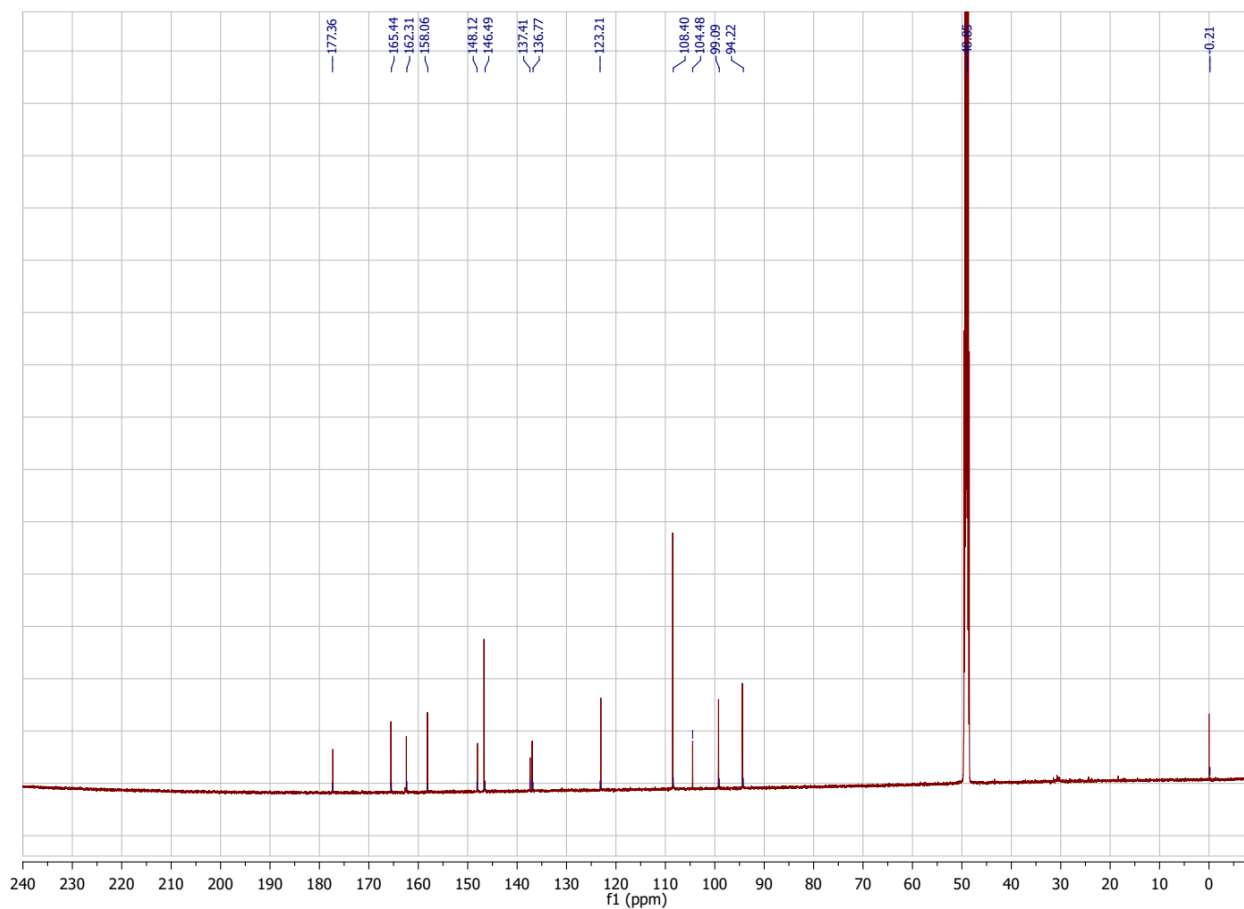
Fig. S16 ^{13}C-NMR spectrum of myricetin in CD$_3$COD, 100 MHz.
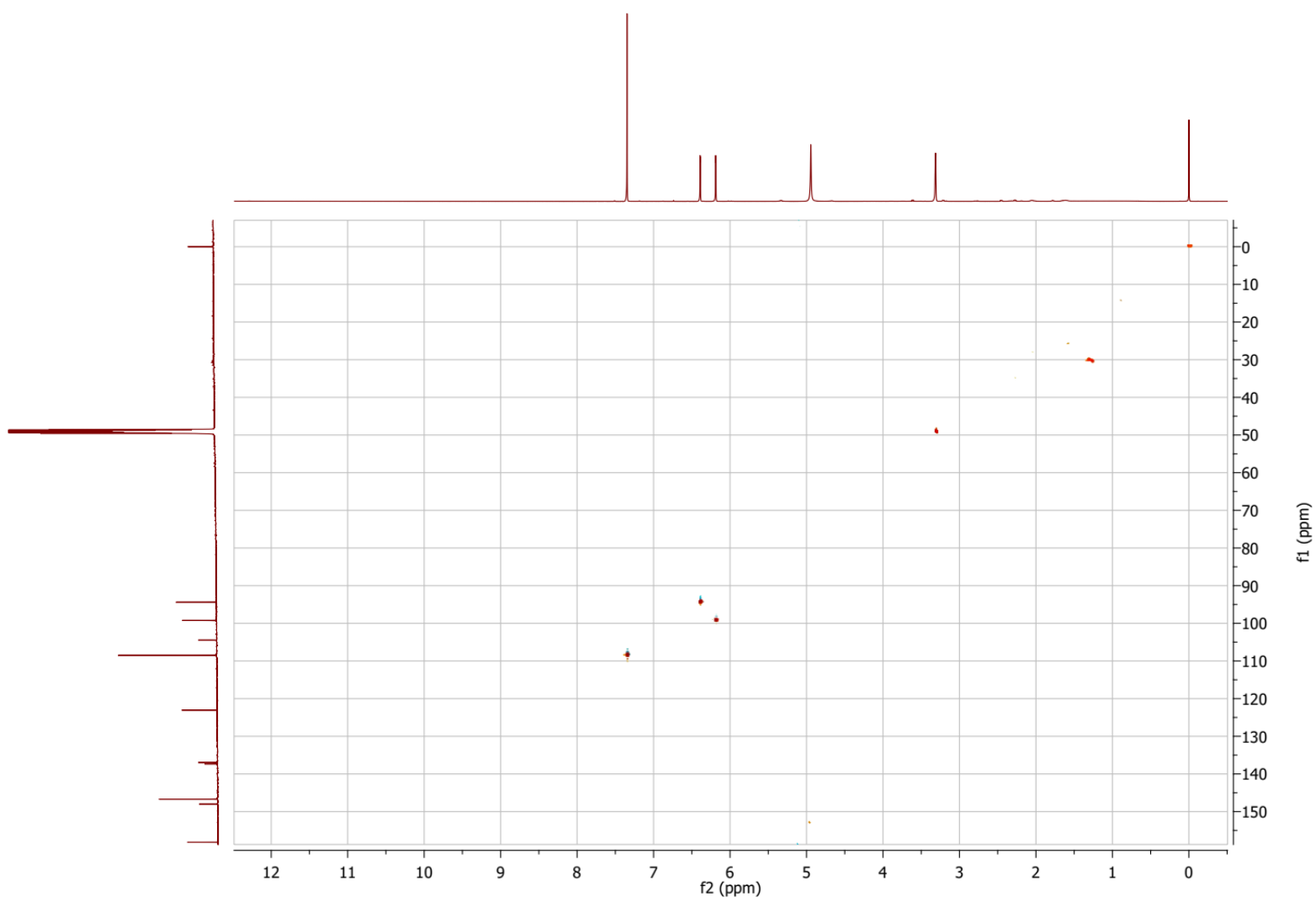
Fig. S17 ^{1}H-^{13}C-HSQC spectrum of myricetin in CD$_3$COD.
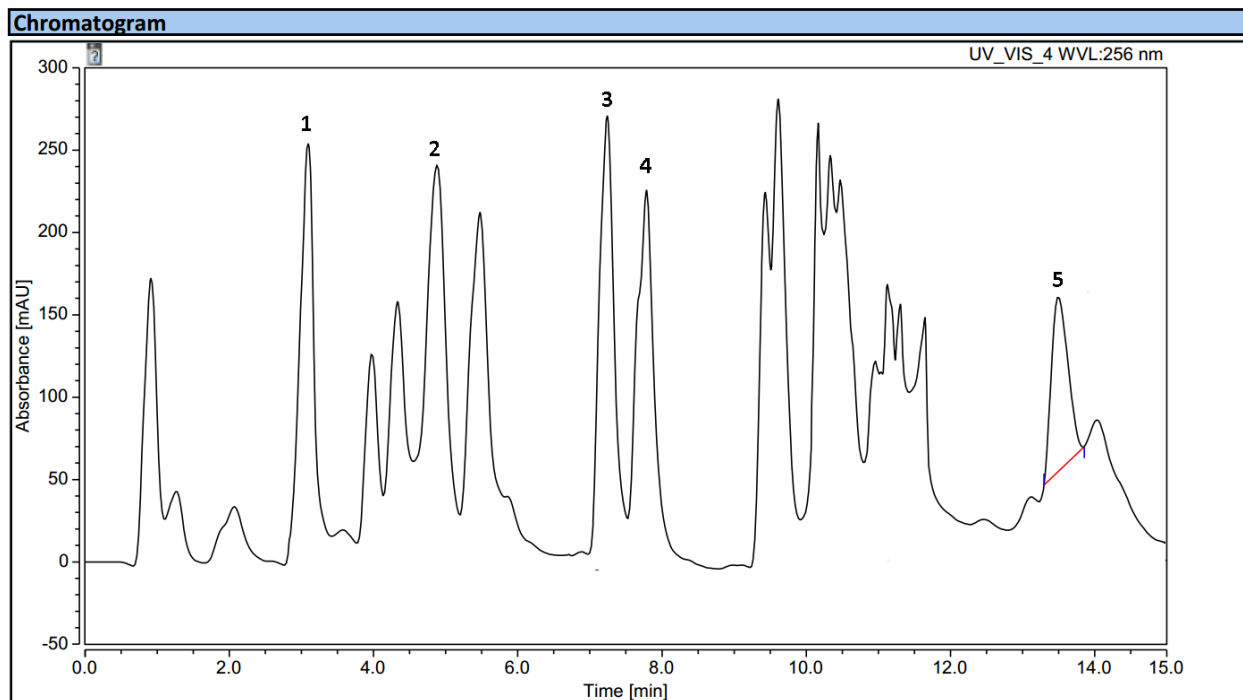
Fig. S18. UPLC chromatogram of the PEE showing the 5 isolated compounds. 1= myrecetin; 2= quercetin; 3= kaempferol; 4= apigenin; 5= magnolin.
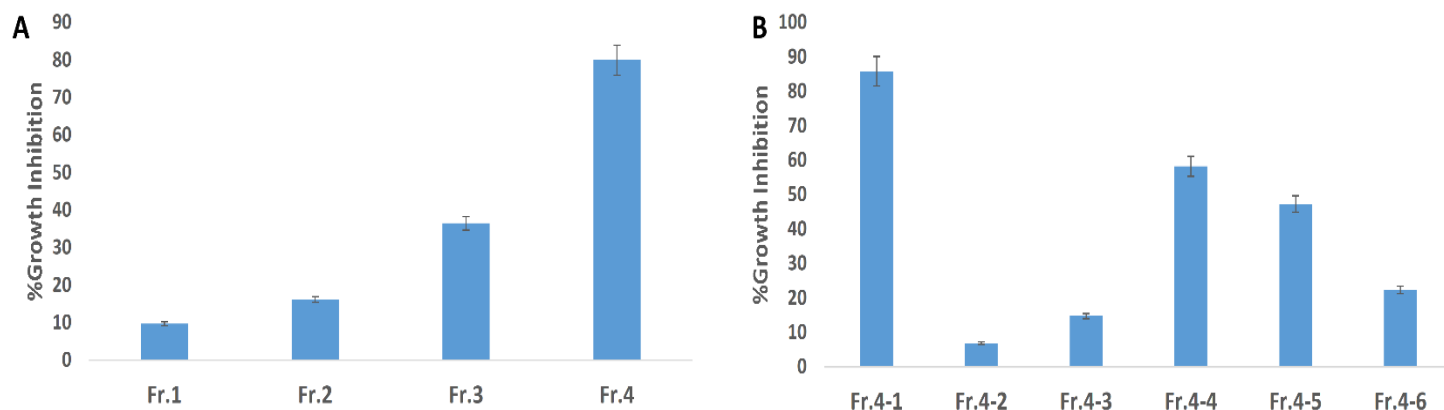
Fig. S19. Growth inhibition of PEE's fractions and sub-fractions (A and B, respectively) at concentration of 5 µg/mL against a set of human pancreatic cancer (PANC-1) cell lines using MTT assay.

Table S1. Concentration of the isolated compounds in PEE extract expressed as % *w/w*.

Compound	Concentration (%, *w/w*)
Myrecetin	0.38 ± 0.044
Quercetin	1.03 ± 0.056
Kaempferol	0.68 ± 0.033
Apigenin	0.53 ± 0.025
Magnolin	1.14 ± 0.062

Methods

LC-HRMS Chemical Profiling

The crude ethanolic extract was prepared at 1 mg/mL for mass spectrometry analysis. The recovered ethanolic extract was subjected to metabolic analysis using LC-HRESIMS according to previously reported protocol.[1] An Acquity Ultra Performance Liquid Chromatography system connected to a Synapt G2 HDMS quadrupole time-of-flight hybrid mass spectrometer (Waters, Milford, USA) was used. Positive and negative ESI ionization modes were utilized to carry out the high-resolution mass spectrometry coupled with a spray voltage at 4.5 kV, the capillary temperature at 320 °C, and mass range from *m/z* 150-1500. The MS dataset was processed, and data were extracted using MZmine 2.20 based on the established parameters (Elmaidomy et al., 2020). Mass ion peaks were detected and accompanied by chromatogram builder and chromatogram deconvolution. The local minimum search algorithm was addressed, and isotopes were also distinguished via the isotopic peaks of grouper. Missing peaks were displayed using the gap-filling peak finder. An adduct search along with a complex search was carried out. The processed data set was next subjected to molecular formula prediction and peak identification. The positive and negative ionization mode data sets from the respective extract were dereplicated against the DNP (Dictionary of Natural Products) databases.

UPLC- assisted quantification of the isolated compounds

For flavonoids' quantification in the extracts, the corresponding authentic standards (compounds 1-5) were dissolved in 50% aqueous acetonitrile to prepare stock solutions (1 mg mL^{-1}). Thereafter, they were diluted to obtain a series of working solutions that were kept at −4 °C and

were then brought to room temperature just before use. Subsequently, each dried extract was dissolved in 50% aqueous acetonitrile in a 10 mL volumetric flask. After that, 1 mL was filtered through 0.22 μm membrane filter before introduction to the LC system. The same mobile phase system was used (i.e. the mobile phase used for LC-HRMS analysis) over 10 min at a flow rate of 1 mL min^{-1} (Figure S18). A calibration curve was then obtained for each authentic flavonoid by plotting its peak areas vs. its different concentrations, and then these calibration curves were used to calculate the quantity of each compound in the plant extracts. The experiments were performed in triplicate.

Molecular dynamics Simulation

Desmond v. 2.2 software was used for performing MDS experiments.[2-4] This software applies the OPLS-2005 force field. Protein systems were built using the System Builder option, where the protein structure was checked for any missing hydrogen's, the protonation states of the amino acid residues were set (pH = 7.4), and the co-crystalized water molecules were removed. Thereafter, the whole structure was embedded in an orthorhombic box of TIP3P water together with 0.15 M Na$^+$ and Cl$^-$ ions in 20 Å solvent buffer. Afterward, the prepared systems were energy minimized and equilibrated for 10 ns. For protein-ligand complexes, the top-scoring poses were used as starting points for simulation. Desmond software automatically parameterizes inputted ligands during the system building step according to the OPLS force field. For simulations performed by NAMD,[5,6] the protein structures were built and optimized by using the QwikMD toolkit of the VMD software. The parameters and topologies of the compounds were calculated using the Charmm27 force field with the online software Ligand Reader and Modeler (http://www.charmm-gui. org/?doc=input/ligandrm, accessed on 3 May 2022).[7] Afterward, the generated parameters and topology files were loaded to VMD to readily read the protein–ligand complexes without errors and then conduct the simulation step.

Absolute Binding Free Energy Calculation

Binding free energy calculations (ΔG) were performed using the free energy perturbation (FEP) method [5]. This method was described in detail in the recent article by Kim and coworkers [6]. Briefly, this method calculates the binding free energy ΔGbinding according to the following equation: $\Delta G_{\text{binding}} = \Delta G_{\text{Complex}} - \Delta G_{\text{Ligand}}$. The value of each ΔG is estimated from a separate simulation using NAMD software. All input files required for simulation by NAMD can be

prepared by using the online website Charmm-GUI (https://charmm-gui.org/?doc=input/afes.abinding). Subsequently, we can use these files in NAMD to produce the required simulations using the FEP calculation function in NAMD. The equilibration (5 ns long) was achieved in the NPT ensemble at 300 K and 1 atm (1.01325 bar) with Langevin piston pressure (for "Complex" and "Ligand") in the presence of the TIP3P water model. Then, 10 ns FEP simulations were performed for each compound, and the last 5 ns of the free energy values was measured for the final free energy values.[6] Finally, the generated trajectories were visualized and analyzed using VMD software. It worth noting that Ngo and co-workers in their recent benchmarking study found that the FEP method of determination of ΔG was the most accurate method in predicting M[Pro] inhibitors.[7]

References

1 H. A. Alhadrami, A. M. Sayed, A. O. El-Gendy, Y. I. Shamikh, Y. Gaber, W. Bakeer, N. H. Sheirf, E. Z. Attia, G. M. Shaban, B. A. Khalifa, C. J. Ngwa, G. Pradel, M. E. Rateb, H. M. Hassan, D. H. M. Alkhalifah, U. R. Abdelmohsen and W. N. Hozzein, A metabolomic approach to target antimalarial metabolites in the *Artemisia annua* fungal endophytes. *Sci. Rep.,* 2021, **11**, 1-11.

2. K .J. Bowers, D. E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan and D. E. Shaw, Scalable algorithms for molecular dynamics simulations on commodity clusters. In Proceedings of the SC'06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, Tampa, FL, USA, 11–17 November 2006; IEEE: New York, NY, USA, 2006; p. 43.

3 S. Release, 3: Desmond Molecular Dynamics System, DE Shaw Research, New York, NY, 2017; Maestro-Desmond Interoperability Tools, Schrödinger: New York, NY, USA, 2017.

4 J. C. Phillips, R. Braun, W. Wang, J. Gumbart; E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten Scalable molecular dynamics with NAMD. *J. Comput. Chem*., 2005, **26**, 1781-1802.

5 L. L. C. Schrodinger, Maestro, Version 9.0; Schrodinger LLC: New York, NY, USA, 2009.

6 S. Kim, H. Oshima, H. Zhang, N. R. Kern, S. Re, J. Lee, B. Rous, Y. Sugita, W. Jiang and W. Im, CHARMM-GUI free energy calculator for absolute and relative ligand solvation and binding free energy simulations. *J. Chem. Theory Comput*., 2020, **16**, 7207-7218.

7 S. T. Ngo, N. M. Tam, P. M. Quan and T. H. Nguyen, Benchmark of Popular Free Energy Approaches Revealing the Inhibitors Binding to SARS-CoV-2 M^{pro}. *J. Chem. Inf. Model*., 2021, **61**, 2302-2312.