**Supporting Information**

# Genomics-Guided Discovery of Thailanstatins A, B and C as Pre-mRNA Splicing Inhibitors and Antiproliferative Agents from *Burkholderia thailandensis* MSMB43

Xiangyang Liu,[†,‡] Sreya Biswas,[†] Michael G. Berg,[§] Christopher M. Antapli,[†] Feng Xie,[⊥] Qi Wang,[⊥] Man-Cheng Tang,[¶] Gong-Li Tang,[¶] Lixin Zhang,[⊥] Gideon Dreyfuss,[§] and Yi-Qiang Cheng[*,†,‡]

[†]Department of Biological Sciences, and Department of Chemistry and Biochemistry, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201, United States
[‡]UNT System College of Pharmacy, University of North Texas Health Science Center, Fort Worth, Texas 76107, United States
[§]Howard Hughes Medical Institute, Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, United States
[⊥]CAS Key Laboratory of Pathogenic Microbiology & Immunology, Institute of Microbiology, Chinese Academy of Sciences (CAS), Beijing 100190, China
[¶]State Key Laboratory of Bio-organic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China

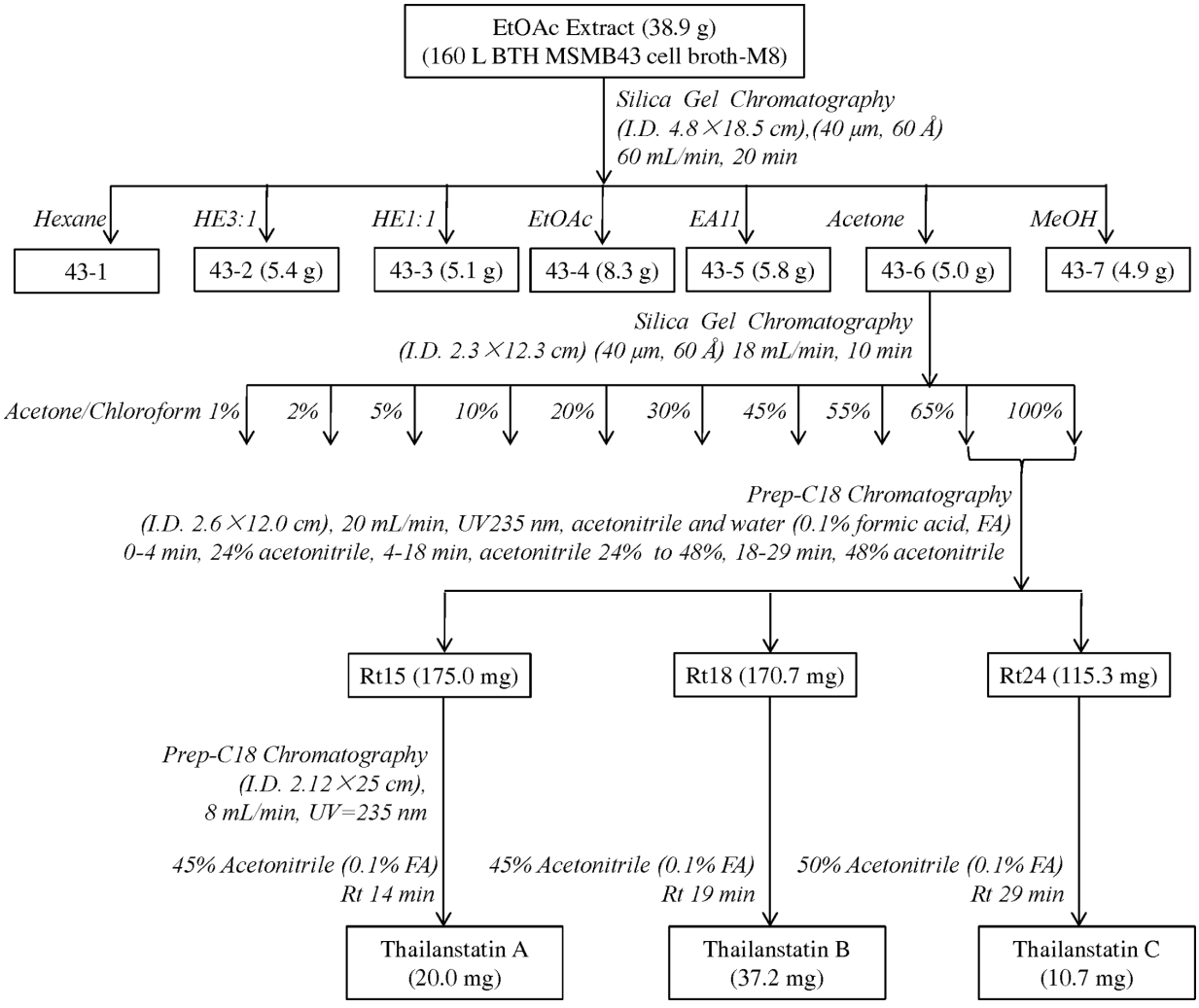
*Corresponding author e-mail: ycheng@uwm.edu

**Table of Content** **Page**

## Materials, Methods and Results

**Isolation and Structural Characterization of Thailanstatins A, B and C.** *B. thailandensis* MSMB43 strain was routinely activated from a permanent glycerol stock on an LB agar plate containing 50 $\mu$g mL$^{-1}$ of apramycin (Am$^{50}$) and grew at 37 °C for 24 h as a starting plate. Seed culture was prepared by growing the bacterium in a small volume of LB medium containing Am$^{50}$ at 30 °C for 24 h in a rotary shaker. For bacterial fermentation in flask, seed culture was inoculated at 4% (v/v) into 20 2-L flasks, each containing 500 mL of sterilized M8 medium (5 g L$^{-1}$ glucose, 5 g L$^{-1}$ peptone, 3 g L$^{-1}$ NaCl, 1.2 g L$^{-1}$ $Na_2HPO_4$ and 0.5 g L$^{-1}$ $KH_2PO_4$; pH 7.0), and the fermentation was proceeded at 30 °C for 72 h in a rotary shaker (200 rpm). For fermentation in fermentor, seed culture was inoculated at 1.67% (v/v) into two 20-L fermentors (BioFlo IV, New Brunswick Scientific Co.), each containing 12 L of sterilized M8 medium. The pH of bacterial culture was automatically maintained by the fermentor with 1 N HCl or 1 N NaOH, and the fermentation was proceeded at 30 °C for 72 h, with agitation at 200 rpm and air flow rate of 4 L/min. The repeated fermentation with 20 flasks and two fermentors led to an accumulation of 160 L of fermentation broth. Fermentation broth was extracted three times with ethyl acetate (3:2, *v*/*v*), and the extracts were combined and concentrated to dryness on a rotary evaporator at 35 °C.

Isolation and purification steps of thailanstatins are summarized in the following scheme; LC-MS analysis was routinely performed to track fractions that contain the target compounds during the process:



The crude extract was subjected to two steps of silica gel chromatography and one step of ODS C18 chromatography on an YFLC AI-580 flash chromatography system (Yamazen, Japan). In the first step, approximately each 5 g of crude extract was resuspended in 25 mL of ethyl acetate, mixed with 30 g of silica gel and packed into an injection column (3.0 x 12.5 cm, Cat. No. W830 silica gel, Yamazen), which was mounted atop a silica gel Universal Column (4.8 x 18.5 cm, 200 g silica gel, 40 μm, 60 Å). The column system was sequentially eluted with 1.2 L of each of the following solvents: hexane, hexane:ethyl acetate (3:1, *v*/*v*), hexane:ethyl acetate (1:1, *v*/*v*), ethyl acetate, ethyl acetate:acetone (1:1, *v*/*v*), acetone, and methanol, all at a flow rate of 60 mL $min^{-1}$. The acetone fraction containing thailanstatins was concentrated to dryness at 35 °C. In the second step, the above resulting extract was resuspended in 15 mL of acetone, mixed with 15 g of silica gel and packed into an injection column (2.0 x 6.5 cm, 14 g silica gel), which was mounted atop a silica gel Universal Column (2.6 x 12.0 cm, 40 g silica gel, 40 μm, 60 Å). The column system was sequentially eluted with 1%, 2%, 5%, 10%, 20%, 30%, 45%, 55%, 65% of acetone mixed with chloroform, and then with 100% acetone. The 65% and 100% acetone fractions containing thailanstatins were again concentrated and further fractionated on a Hi-FLASH ODS C18 injection column (2.0 x 6.5 cm, 50 μm, 120 Å) mounted atop of a Universal ODS C18 column (2.6 x 12.0 cm, 37 g, 50 μm, 120 Å) pre-equilibrated with 24% acetonitrile (in water added with 0.1% formic acid, FA). The column system was first eluted with 24% acetonitrile (+ 0.1% FA) for 4 min, then with a linear gradient from 24% to 48% acetronitrile (+ 0.1% FA) in 14 min, and then isocratically in 48% acetronitrile (+ 0.1% FA) for 10 min. Fractions containing thailanstatins A, B and C were eluted at 15 min, 18 min and 24 min, respectively. The final purification of thailanstatins was performed individually with a Varian ProStar HPLC system (210 binary pump and 330 photodiode array detector) equipped with an Agilent Prep-C18 column (21.2 × 250 mm, 10 μm). An isocratic elution scheme was performed with a flow rate of 8 mL $min^{-1}$ and UV monitored at 235 nm. Thailanstatins A and B were eluted by 45% acetonitrile (+ 0.1% FA) at 14 min and 19 min, respectively; thailanstatin A was eluted by 50% acetonitrile (+ 0.1% FA) at 29 min. Those individual elutes were dried by lyophilization which yielded 20.0 mg, 37.2 mg and 10.7 mg of white dry powder of thailanstatins A, B and C, respectively. Compounds were stored in sealed dark tubes in a –20 °C freezer.

Thailanstatin B, the most abundant compound of the three, was obtained as a white powder and its molecular formula $C_{28}H_{42}NClO_9$ was deduced from high-resolution electrospray-ionization mass spectrum (HR-ESI-MS) ($[M + H]^+$ *m*/*z*: found 572.2637, calc. 572.2621) and NMR data (Table S2, Table S4, Figure S15 through Figure S24). In HR-ESI-MS spectrum, each peak is accompanied with approximately one-third intensity of the quasi-molecular peak with two more daltons, which is the characteristic of the presence of a chlorine atom. The molecular formula was corroborated by $^{1}H$, $^{13}C$ and $^{15}N$ HSQC NMR (NH at 125.3 ppm) data. The UV spectrum of thailanstatin B in acetonitrile exhibited maximal absorption at approximately 235 nm (Figure S13). The IR spectra indicated the presence of two hydroxyl groups (3,428 $cm^{-1}$ and 3,424 $cm^{-1}$), one carboxyl group (3,054 $cm^{-1}$), one carbonyl group (1,716 $cm^{-1}$), one amide carbonyl group (1,667 $cm^{-1}$) and one conjugated diene linkage (1,633 $cm^{-1}$) (Figure S14). $^{13}C$ and DEPT135 spectra revealed the presence of five methyl groups, five methylene groups and 13 methine group, as well as the two quaternary carbons and three carbonyl carbons. The key ADEQUATE, COSY and HMBC data were shown in Figure S35A. Two spin systems C17-C1-C2 and C4-C5-C6-C7 were resolved in the ADEQUATE and $^{1}H$–$^{1}H$ COSY spectra are joined into the first tetrahydropyran ring with a chloromethyl group at C3 position and a carboxyl group at C17 position. Another three spin systems C9-C10, C12-C21 and C11-C12-C13-C14-C15-C16 were confirmed by $^{1}H$–$^{1}H$ COSY and ADEQUATE spectra to form another tetrahydropyran ring. These two tetrahydropyran rings are linked through the conjugated diene C9-C8-C7-C6 with a C8 substituted methyl group based on the HMBC correlations between C20, C7, C8 and C9. It is also observed that the second tetrahydropyran ring is connected with the spin system C2'-C3'-C4'-C5' through a C2' O-NH group in the HMBC spectrum. The final correlation from the C4' position of the C2'-C3'-C4'-C5' system to the C1''-C2'' spin system yielded the 28 skeleton of thailanstatin B. When the $^{1}H$ and $^{13}C$ NMR data were compared to those of FR901464 (Figure S36 and Figure S37), the spectra showed significant similarities except for the

absence of the C17 methyl group ($\delta_C$ 29.1 ppm) at the C1 position and two more chemical shift ($\delta_C$ 37.9 ppm and $\delta_C$ 175.4 ppm) attached to C1 atom. Additionally, the methylene group with chemical shift at 49.9 ppm in $^{13}C$ spectrum and 3.56 ppm in $^{1}H$ spectrum indicated that this $CH_2$ is connected to a chlorine atom. According to the molecular formula and the already known -NH group, there should be three hydroxyl groups in the structure, which were not observable in $^{1}H$ NMR spectrum but indicated in IR spectrum.

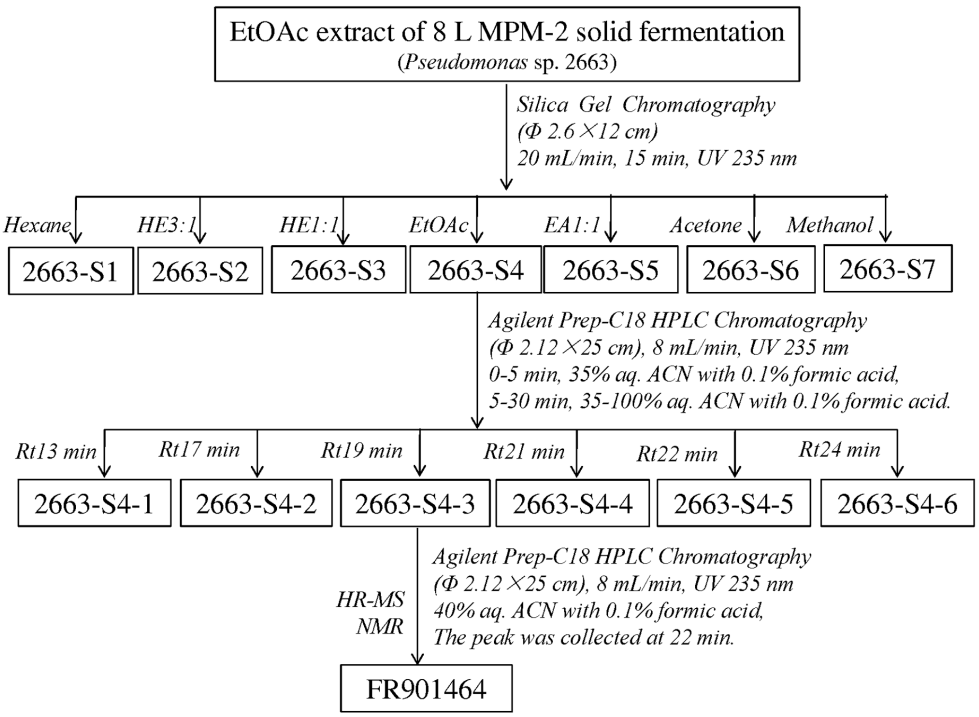
Thailanstatin A was obtained as a white powder and has a molecular formula $C_{28}H_{41}NO_9$ as established by HR-TOF-MS ($[M + H]^+$ *m*/*z*: found 536.2864, calc. 536.2854) and NMR data (Table S1, Table S4, Figure S1 through Figure S12). This formula is one hydrogen and chlorine atoms less than that of thailanstatin B. Empirical chemical shift data and $^{1}H$-$^{13}C$ HMBC experiments indicated the presence of three ester carbonyl carbons (173.8, 170.3, 164.9 ppm), one carboxylic carbonyl carbon (174 ppm), five olefin (143.6, 138.0, 129.4, 123.0, 122.4 ppm), two quaternary carbons (134.5, 57.1 ppm), six alcohol methylene carbons (81.1, 76.2, 75.9, 70.1, 68.6, 68.6 ppm), one methylene carbon (29.1 ppm) and five methyl carbons (21.0, 19.8, 17.4, 14.7, 12.3 ppm). The similarities of $^{1}H$ and $^{13}C$ NMR data between thailanstatins A and B suggest that thailanstatin A is an analogue of thailanstatin B. The HMBC signals between C3 and C19 of thailanstatin A suggested the presence of the epoxide moiety at C3 atom, which is similar to the C3 epoxide moiety ($\delta_C$ 48.1 ppm) for FR901464.

Thailanstatin C was obtained as a white powder with a molecular formula of $C_{30}H_{46}NClO_9$ deduced from HR-ESI-MS ($[M + H]^+$ *m*/*z*: found 600.2966, calc. 600.2933) and NMR data (Table S3, Table S4, Figure 25 through Figure S34). The molecular composition of thailanstatin C indicates a mass of 28 more than that of thailanstatin B. The $^{1}H$-NMR data of thailanstatin C displays a high degree similarity to that of thailanstatin B, except for the additional observance of two more methyl groups (overlapped at $\delta_H$ 1.16 ppm $\delta_C$ 18.9 ppm). The structure of thailanstatin C differs from that of thailanstatin B at the 1" atom by replacement of a methyl group with an isopropyl group according to the HMBC correlations.

The relative configuration of thailanstatin B, the most abundant compound in the thailanstatin family, was determined using NMR experiments and the configurations of the remaining family members thailanstatins A and C were assumed to be the same as that of thailanstatin B. The geometry of C2' and C6 double bond was proposed as *cis* (*Z*) based on the vicinal coupling constants $J_{H2'\text{-}H3'}$=11.6 Hz and $J_{H6\text{-}H7}$=16 Hz (Table S4). The *trans* (*E*) configuration of the double bonds at C8-C9 was indicated by the chemical shift of C20 at 12.8 ppm (less than 20 ppm)[1] and the observations of the NOESY correlation between H7 and H9 and between H6 and H20. The nine chiral carbons of thailanstatin B were assigned according to the related NOESY correlations, which were divided into two parts (Figure S35B). The first part is located at C1~C5 atoms in the first tetrahydropyran ring. The trans-diaxial orientations between 4-H and 5-H are supported by the observation of NOESY signals between 4-H and one of 2-H and that between 10-H and both of the two 20-$H_2$. The cross peaks between 5-H and 19-$H_2$ and between 5-H and 17-$H_2$ in NOESY spectrum suggested a 1, 3- diaxial-disposition of C19 and C17. The second part is located at C11~C15 atoms in the second tetrahydropyran ring. The observation of signals between 15-H and 13-$H_2$ and between 15-H and 11-H pointed out the 1, 3-diaxial orientation. The observation of strong NOESY correlations between NH and 21-$H_3$ suggested a 1, 3-diaxial interaction between NH and the methyl group 13-$CH_3$. So the stereo features of atoms C1, C3, C4, C5, C11, C12, C14, and C15 in thailanstatin B were *S*, *S*, *R*, *R*, *S*, *S*, *R*, and *R*, respectively. At the end, the stereo center of C4' was predicted to be *R* on the basis of the strong homology of *tstC* ACP and ketoreductase (KR) domain to those of *fr9C* ACP and KR domain (Figure S39).

**Isolation of FR901464 as a Reference Compound.** *Pseudomonas* sp. No. 2663 strain was routinely activated from a glycerol stock on LB agar at 30 ºC for 2 days as a starting plate. Several single colonies from the plate were inoculated into a 1-L Erlenmeyer flask containing 250 mL of LB medium and incubated at 30 ºC for 24 h in a rotary shaker (150 rpm). This seed culture was evenly distributed into 40 2-L flasks, each containing 200 mL of sterilized MPM-2 agar medium [2.6 g $L^{-1}$ glucose, 2.3 g $L^{-1}$ yeast extract, 2.4 g $L^{-1}$ $Na_2HPO_4$, 0.5 g $L^{-1}$ $(NH_4)_2SO_4$, 0.06 g $L^{-1}$ $MgSO_4 \cdot 7H_2O$, and 15 g $L^{-1}$ agar; pH 7.0]. Semi-solid fermentation was carried out at room temperature (25 ºC) for 3 days without shaking.

Bacterial cells were killed by adding 200 mL of ethyl acetate to each flask, and the agar was sliced into approximately 1 $cm^3$ blocks and transferred into a 20-L glass container. Approximately 8 L of additional ethyl acetate was added into the container and the extraction was performed three times. The ethyl acetate extracts were combined and concentrated to dryness at 30 ºC on a rotary evaporator. Isolation and purification steps of FR901464 are summarized in the following scheme; LC-MS analysis was routinely performed to track fractions that contain the target compound during the process:




The crude extract (~2 g) was dissolved in $CH_2Cl_2$ and mixed with 20 g silica gel (Purasil Silica Gel, Whatman, 38-63 µm, 230-400 mesh, 60 Å pore size), then packed into two injection columns (2.0 I.D. × 7.5 cm). Each injection column was connected to a silica gel column (2.6 I.D. × 12 cm, same silica gel as that of the injection column) and applied on an AI-580 Yamazen Flash Chromatography System (Yamazen, Japan). Separation was accomplished by sequential isocratic elutions using hexane, hexane/ethyl acetate (3:1, *v*/*v*), hexane/ethyl acetate (1:1, *v*/*v*), ethyl acetate, ethyl acetate/acetone (1:1, *v*/*v*), acetone and methanol at a flow rate of 20 mL $min^{-1}$, each for 15 min. The elution was monitored at UV 235 nm. Fractions from each solvent elution were combined and concentrated to dryness on a rotary evaporator at 30 ºC.

The dry mass of ethyl acetate elution (184 mg) was re-dissolved in 1 ml of 35% acetonitrile solution and centrifuged at 10,000 g for 5 min. The supernatant was applied on a Varian ProStar HPLC system equipped with 210 Solvent Delivery Modules, a 330 Diode Array Detector and an Agilent Prep-C18 column (21.2 I.D. x 250 mm, 10 µm). The mobile phase was composed of water (+ 0.1% formic acid, FA) (buffer A) and acetonitrile (+ 0.1% FA) (buffer B). Once the sample was loaded, the column was first washed with 35% buffer B, and then eluted with a linear gradient from 35% buffer B to 100% buffer B in 25 min. The flow rate was 8.0 mL $min^{-1}$ and the elution was monitored at 235 nm. The fraction containing FR901464 was collected at 19 min. One more round of purification of FR901464 was performed with the same C18 column but eluted isocratically with 40% aqueous acetonitrile (+ 0.1% FA); the fraction containing FR901464 was collected at 22 min. The final elute was dried by lyophilization.

Collectively, 12 mg of FR901464 was obtained from 24 L of semi-solid fermentation. Compound was stored in a sealed dark tube in a –20 °C freezer.

**References**


1. Lange, G. L., and Lee, M. (1986) $^{13}C$ NMR determination of the configuration of methyl-substituted double bonds in medium- and large-ring terpenoids. *Magn. Res. Chem. 24*, 656-658.
2. Zhang, F., He, H. Y., Tang, M. C., Tang, Y. M., Zhou, Q., and Tang, G. L. (2011) Cloning and elucidation of the FR901464 gene cluster revealing a complex acyltransferase-less polyketide synthase using glycerate as starter units. *J. Am. Chem. Soc. 133*, 2452-2462.
3. Nakajima, H., Takase, S., Terano, H., and Tanaka, H. (1997) New antitumor substances, FR901463, FR901464 and FR901465. III. Structures of FR901463, FR901464 and FR901465. *J. Antibiot. 50*, 96-99.
4. Albert, B. J., Sivaramakrishnan, A., Naka, T., and Koide, K. (2006) Total synthesis of FR901464, an antitumor agent that regulates the transcription of oncogenes and tumor suppressor genes. *J. Am. Chem. Soc. 128*, 2792-2793.
5. Thompson, C. F., Jamison, T. F., and Jacobsen, E. N. (2001) FR901464: total synthesis, proof of structure, and evaluation of synthetic analogues. *J. Am. Chem. Soc. 123*, 9974-9983.

**Table S1.** 1D and 2D NMR Spectroscopic Assignments for Thailanstatin A ($CD_2Cl_2$, 298K).

| Pos. | $\delta_C$ (ppm) | HSQC | $\delta_H$ (mult, $J$ in Hz) | COSY | HMBC | NOESY |
|---|---|---|---|---|---|---|
| 1 | 68.6 | CH | 4.51, m | 2, 17 | 5, 19, 17 | 2, 17 |
| 2 | 38.1 | $CH_2$ | 2.60, dd (15, 5.0); 3.02 (15.0, 9.0) | 1 | 1, 17, 18 | |
| 3 | 57.1 | C | | | 5, 1, 19, 17 | |
| 4 | 70.1 | CH | 3.51, d (7.3) | 5 | 6, 5, 3, 19, 17 | 2, 5, 6, 17, 19 |
| 5 | 75.9 | CH | 4.27, t (6.3) | 6, 4 | 7, 6, 4, 1, 3 | 7, 6, 4, 19 |
| 6 | 123.0 | CH | 5.66, dd (6.0, 16.0) | 5 | 8, (9), (2'), 5, 4 | 7, 5, 4, 20 |
| 7 | 138.0 | CH | 6.37, d (16.0) | 6 | 8, 9, 5, 20 | 9, 6 |
| 8 | 134.5 | C | | | 7, 6, 20 | |
| 9 | 129.4 | CH | 5.48, t (7.0) | 10 | 7, 11, 10, 20 | 7, 11, 10, 21, 20 |
| 10 | 31.8 | $CH_2$ | 2.24, m; 2.38, m | 9, 11 | 8, 9, 11, 12 | 9, 11, 12, 21 |
| 11 | 81.1 | CH | 3.59, td (7.4, 2.5) | 10 | 9, 15, 13, 10, 16 | 9, 15, 10, 13, 12 |
| 12 | 29.1 | CH | 1.78 (overlap) | 21 | 14, 10 | 21 |
| 13 | 35.7 | $CH_2$ | 1.95, m | 12, 14 | 11, 15, 14, 22, 16 | |
| 14 | 47.0 | CH | 3.92, m | 13, NH | 12 | NH, 15, 13, 16 |
| 15 | 76.2 | CH | 3.74, qd (6.5, 2.0) | 16 | 11, 14, 16 | 11, 14, 13, 16 |
| 16 | 17.4 | $CH_3$ | 1.16, d (6.5) | 15 | 16, 14 | |
| 17 | 34.4 | $CH_2$ | 1.80, m; 2.12, d | 1 | 1, 3, 19, 2 | |
| 18 | 173.8 | C=O | | | 1, 19 | |
| 19 | 49.9 | $CH_2$ | 2.67, d (4.5); 2.99, d (4.6) | | 3 | 5 |
| 20 | 12.3 | $CH_3$ | 1.77, s | | 7, 8, 9 | |
| 21 | 14.7 | $CH_3$ | 1.01, d (7.3) | 12 | 11, 13, 12 | |
| 1' | 164.9 | C=O | | | 3', 2', NH | |
| 2' | 122.4 | CH | 5.84, dd (11.6, 1.0) | 3' | 1', 3', 4', 5' | 4', 11, 10, 12, 21 |
| 3' | 143.6 | CH | 5.95, dd (11.6, 8.0) | 2', 4' | 1', 2', 4', 5' | 2', 5' |
| 4' | 68.6 | CH | 6.33, m | 3', 5' | 1", 3', 2', 5' | 3', 5', 21 |
| 5' | 19.8 | $CH_3$ | 1.36, d (6.5) | 4' | 3', 7 | |
| 1" | 170.3 | C=O | | | 4', 2" | |
| 2" | 21.0 | $CH_3$ | 2.06, s | | 1" | |
| | | NH[a] | 6.69, d (8.7) | 14 | 1', 14, 15 | 13, 14, 16, 21, 2', 3' |

[a] Two hydroxy groups were not observable in NMR spectra.

**Table S2.** 1D and 2D NMR Spectroscopic Assignments for Thailanstatin B ($CDCl_3$, 298K).

| Pos. | $\delta_C$ (ppm) | HSQC | $\delta_H$ (mult, $J$ in Hz) | COSY | HMBC | ADEQUATE | NOESY |
|---|---|---|---|---|---|---|---|
| 1 | 68.5 | CH | 4.44, s | 17 | 17, 2 | 17, 2 | 2, 17, 19 |
| 2 | 34.6 | $CH_2$ | 2.13, 1.89 | 1 | | | |
| 3 | 72.2 | C | | | | | |
| 4 | 70.6 | CH | 3.52 (overlap) | | 6, 5, 4 | 3 | 5, 6, 7 |
| 5 | 71.3 | CH | 4.2, s | 4 | 7, 6 | | 4, 7, 17 |
| 6 | 124 | CH | 5.56, dd (16.0) | 5 | 8 | | 4, 20 |
| 7 | 138.5 | CH | 6.34, d (15.5) | 6 | 8, 9, 5, 20 | 8 | 4, 5, 9, 20 |
| 8 | 134.7 | C | | | | | |
| 9 | 128.9 | CH | 5.41 | 10 | 7, 20 | | |
| 10 | 31.7 | $CH_2$ | 2.36, 2.20 | | 8, 9, 11 | | 9, 20 |
| 11 | 81.4 | CH | 3.52 (overlap) | 10 | | 10, 12 | 9, 10, 12, 13, 14, 16, 2 |
| 12 | 28.9 | CH | 1.74 (overlap) | | | | |
| 13 | 35.7 | $CH_2$ | 1.98, 1.88 | | | 14, 12 | 21 |
| 14 | 47.2 | CH | 3.87 | | | 15, 13 | 11, 13, 15, 16, NH |
| 15 | 76.4 | CH | 3.68, d (5.0) | 16 | | 14, 16 | 11, 12, 13, 14, 16 |
| 16 | 17.6 | $CH_3$ | 1.14, d (5.3) | | 15, 14 | 15 | 11, 14, 15, NH |
| 17 | 37.9 | $CH_2$ | 3.18, t (11.0); 2.73, d (11.5) | 1 | | | 1, 5 |
| 18 | 175.4 | C | | | | | |
| 19 | 49.9 | $CH_2$ | 3.56 | | 3, 2 | | |
| 20 | 12.6 | $CH_3$ | 1.70, s | | 7, 8, 9 | 8 | |
| 21 | 14.7 | $CH_3$ | 0.96, d (6.2) | | 11, 13, 12 | 12 | 10, 9, 2' NH |
| 1' | 165.4 | C | | | | | |
| 2' | 122.4 | CH | 5.80, t (11.4) | | 1', 3', 4' | | 16, 21, NH |
| 3' | 143.8 | CH | 5.89, t (11.3) | | 1', 2', 4', 5' | 4' | 4', 5' |
| 4' | 68.9 | CH | 6.28, t (6.5) | 5' | 1", 3', 2', 5' | 3', 5' | 5' |
| 5' | 20 | $CH_3$ | 1.36, d (6.1) | | 3', 4' | 4' | |
| 1" | 170.6 | C | | | | | |
| 2" | 21.3 | $CH_3$ | 2.02, s | | 1" | 1" | |
| | | NH[a] | 6.74, d (7.0) | 14 | 1' | | 14, 16, 21, 4' |

[a] Three hydroxy groups were not observable in NMR spectra.

**Table S3.** 1D and 2D NMR Spectroscopic Assignments for Thailanstatin C ($CDCl_3$ 298K).

| Pos. | $\delta_C$ (ppm) | HSQC | $\delta_H$ (mult, $J$ in Hz) | COSY | HMBC | NOESY |
|---|---|---|---|---|---|---|
| 1 | 68.7 | CH | 5.91, m | 17, 2 | 3, 5 | 19, 17, 2 |
| 2 | 34.6 | $CH_2$ | 1.88, d (14.0); 2.14, d (14.0) | 1 | 3, 4 | 1, 19 |
| 3 | 72.2 | C | | | 7, 6, 4, 1 | |
| 4 | 70.5 | CH | 3.56 (overlap) | 5 | 3, 5, 19 | 6, 5, 1 |
| 5 | 71.4 | CH | 4.22, t (7.4) | 6, 4 | 7, 6, 3, 4 | 7, 6, 4, 17 |
| 6 | 123.7 | CH | 5.56, dd (7.0, 14.0) | 7, 5 | 8, 4 | 5, 4, 20 |
| 7 | 139.1 | CH | 6.38, d (15.0) | 6 | 8, 9, 5, 20 | 6, 9, 5 |
| 8 | 134.6 | C | | | 7, 6, 10, 20 | |
| 9 | 129.3 | CH | 5,46, t (8.5) | 10 | 9, 7, 10, 20 | 7, 11, 10, 12, 21 |
| 10 | 31.7 | $CH_2$ | 2.21, m; 2.35, m | 9, 11 | 8, 9, 11, 12 | 9, 11, 12, 20 |
| 11 | 81.3 | $CH_2$ | 3.54 (overlap) | 10 | 15, 10, 21 | 15, 10, 13, 12, 20 |
| 12 | 28.9 | CH | 1.77 (overlap) | 21 | | |
| 13 | 35.7 | $CH_2$ | 1.91, m; 2.01, m | 14, 12 | 15, 11, 14, 12, 21 | 14, 12 |
| 14 | 47.1 | CH | 3.92, m | NH, 13 | 12, 13 | NH, 15, 13, 14 |
| 15 | 76.4 | CH | 3.69, m | 16 | 11, 14, 16 | |
| 16 | 17.7 | $CH_3$ | 1.15 (overlap) | 15 | 15, 14 | |
| 17 | 37.9 | $CH_2$ | 2.72, d (12.0); 3.26, t (10.0) | 1 | 18, 1 | 5, 1, 2 |
| 18 | 175.5 | C=O | | | 17 | |
| 19 | 49.9 | $CH_2$ | 3.59 (overlap) | | 3, 4 | 2 |
| 20 | 12.6 | $CH_3$ | 1.74, s | | 7, 8, 9 | |
| 21 | 14.8 | $CH_3$ | 1.00, d (6.5) | 12 | 11, 13, 12 | |
| 1' | 165.2 | C=O | | | 3', 2' | |
| 2' | 122.1 | CH | 5.79, dd (12.0, 1.0) | 3' | 1', 3', 4' | NH, 4' |
| 3' | 144.5 | CH | 5.91, dd (12.0, 7.5) | 4' | 1', 2' | 4', 5' |
| 4' | 68.8 | CH | 6.30, m, | 7, 5' | 1", 3', 2', 5' | 3', 2', 5' |
| 5' | 20.0 | $CH_3$ | 1.39, d (6.0) | 4' | 3', 4' | |
| 1" | 176.6 | C=O | | | 4', 2", 3", 4" | |
| 2" | 34.1 | CH | 2.52, octet (6.0) | 3", 4" | 1", 3", 4" | 3", 4" |
| 3" | 18.9 | $CH_3$ | 1.16 (overlap) | 2" | 1", 2", 4" | |
| 4" | 18.9 | $CH_3$ | 1.16 (overlap) | 2" | 1", 2", 3" | |
| | | NH[a] | 6.54, d (7.0) | 14 | | |

[a] Three hydroxy groups were not observable in NMR spectra.

**Table S4.** Physico-Chemical Properties of Thailanstatins A, B and C.

| | Thailanstatin A | Thailanstatin B | Thailanstatin C |
|---|---|---|---|
| Appearance | White powder | White powder | White powder |
| Melting point | 80~82 °C | 98~104 °C | 89-97 °C |
| $[\alpha]_D^{23}$ | 4.0° (*c* 0.15, $CH_2Cl_2$) | -22.8° (*c* 0.14, $CH_2Cl_2$) | -10.0° (*c* 0.10, $CH_2Cl_2$) |
| Molecular formula | $C_{28}H_{41}NO_9$ | $C_{28}H_{42}NClO_9$ | $C_{30}H_{46}NClO_9$ |
| Molecular weight by HR-ESI-MS: | $[M + H]^+$ | $[M + H]^+$ | $[M + H]^+$ |
| Calculated | 536.2854 | 572.2621 | 600.2933 |
| Found | 536.2864 | 572.2637 | 600.2966 |
| UV $\lambda_{max}^{acetonitrile}$ nm(ε) | 236 (30,000) | 237 (35,435) | 237 (32,648) |
| Solubility | | | |
| Soluble | DMSO, acetonitrile, acetone, EtOAc, $CHCl_3$, $CH_2Cl_2$, $H_2O$ | DMSO, acetonitrile, acetone, EtOAc, $CHCl_3$, $CH_2Cl_2$, $H_2O$ | DMSO, acetonitrile, acetone, EtOAc, $CHCl_3$, $CH_2Cl_2$, $H_2O$ |
| Insoluble | Hexane | Hexane | Hexane |
| TLC (Rf value) System[a] | 0.76 | 0.86 | 0.89 |

[a]Plate: Silica gel 60 (Whatman), $CHCl_3$-acetone=2:3 (*v*/*v*), sprayed with 0.1% Bromocresol Green (BCG) in ethanol solution.

**Table S5.** $^{13}C$ NMR Assignments for FR901464.

| Position | Freshly purified natural FR901464: $\delta_C$ (ppm) | Previously reported natural FR901464[3]: $\delta_C$ (ppm) | Previously reported synthetic FR901464[4]: $\delta_C$ (ppm) | Previously reported synthetic FR901464[5]: $\delta_C$ (ppm) |
|---|---|---|---|---|
| 1 | 96.7 | 96.7 | 96.7 | 96.6 |
| 2 | 41.9 | 41.8 | 41.8 | 41.8 |
| 3 | 58.1 | 58.1 | 58.1 | 58.1 |
| 4 | 68.2 | 68.1 | 68.1 | 68.1 |
| 5 | 73.9 | 73.8 | 73.9 | 73.8 |
| 6 | 124.7 | 124.7 | 124.6 | 124.6 |
| 7 | 138.4 | 138.3 | 138.3 | 138.2 |
| 8 | 134.9 | 134.8 | 134.8 | 134.7 |
| 9 | 129.9 | 129.8 | 129.9 | 129.7 |
| 10 | 32.4 | 32.4 | 32.4 | 32.4 |
| 11 | 81.2 | 81.2 | 81.2 | 81.1 |
| 12 | 29.6 | 29.6 | 29.6 | 29.6 |
| 13 | 36.2 | 36.3 | 36.2 | 36.2 |
| 14 | 47.4 | 47.4 | 47.4 | 47.4 |
| 15 | 76.3 | 76.3 | 76.3 | 76.2 |
| 16 | 18.0 | 17.9 | 17.9 | 18 |
| 17 | 29.1 | 29.1 | 29.1 | 29.1 |
| 18 | 48.1 | 48 | 48.1 | 48.1 |
| 19 | 12.8 | 12.7 | 12.7 | 12.8 |
| 20 | 15.3 | 15.2 | 15.2 | 15.3 |
| 1’ | 165.0 | 165.0 | 165.0 | 164.8 |
| 2’ | 122.9 | 122.8 | 122.8 | 122.7 |
| 3’ | 143.9 | 143.9 | 143.9 | 143.7 |
| 4’ | 68.9 | 68.9 | 68.9 | 68.9 |
| 5’ | 20.2 | 20.2 | 20.1 | 20.2 |
| 1” | 170.7 | 170.6 | 170.6 | 170.4 |
| 2” | 21.4 | 21.4 | 21.4 | 21.4 |

**Table S6.** $^{1}$H NMR Assignments for FR901464.

| Position | Freshly purified natural FR901464: $\delta_H$ (mult, $J$ (Hz)) | Previously reported natural FR901464[3]: $\delta_H$ (mult, $J$ (Hz)) | Previously reported synthetic FR901464[4]: $\delta_H$ (mult, $J$ (Hz)) | Previously reported synthetic FR901464[5]: $\delta_H$ (mult, $J$ (Hz)) |
|---|---|---|---|---|
| 1-OH | 3.42, s | 3.38, s | 3.31, s | 3.40, s |
| 2axial 2equitorial | 2.33, d (14.0); 1.63, d (14.0) | 2.36, d (14.0); 1.66, d (14.0) | 2.34, d (14.4; 1.64, d (14.4) | 2.34, d (14.5); 1.64, d (14.0) |
| 4 | 3.57, dd (10.0, 10.0) | 3.58, dd (10.0, 10.0) | 3.57, app t (10.0) | 3.57, dd (10.0, 10.0) |
| 4-OH | 1.69, d (8.0) | 1.66, d (10.0) | 1.59, d (10.3) | 1.67, d (10.0) |
| 5 | 4.23, dd (10.0, 7.0) | 4.25, dd (10.0, 7.0) | 4.24, dd (9.3, 7.0) | 4.24, dd (9.0, 7.0) |
| 6 | 5.63, dd (7.0, 16.0) | 5.66, dd (7.0, 16.0) | 5.65, dd (15.7, 7.0) | 5.66, dd (7.0, 16.0) |
| 7 | 6.36, d (16.0) | 6.37, d (16.0) | 6.38, d (15.7) | 6.37, d (15.5) |
| 9 | 5.53, br t (7.0) | 5.53, br t (7.0) | 5.54, br t (7.0) | 5.53, br t (7.0) |
| 10 | 2.35, m; 2.21, m | 2.36, m; 2.24, m | 2.40−2.30, m<br>2.28−2.20, m | 2.35, m; 2.22, m |
| 11 | 3.53, m | 3.53, m | 3.57−3.50, m | 3.53, m |
| 12 | 1.77, m | 1.77, m | 1.77, m | 1.77, m |
| 13 | 1.92, m; 1.91, m | 1.94, m; 1.91, m | 1.95−1.93, m<br>1.93−1.91, m | 1.93, m; 1.92, m |
| 14 | 3.88, m | 3.90, m | 3.94−3.88, m | 3.90, m |
| 14-NH | 6.08, d (9.0) | 5.99, d (9.0) | 5.96, br d (8.9) | 6.00, d (9.0) |
| 15 | 3.65, qd (6.5, 2.0) | 3.66, qd (7.0, 2.0) | 3.66, qd (6.5, 2.2) | 3.65, qd (6.5, 2.5) |
| 16 | 1.11 (3H), d (6.6) | 1.11 (3H), d (7.0) | 1.11 (3H), d (6.4) | 1.11 (3H), d (6.5) |
| 17 | 1.41 (3H), s | 1.43 (3H), s | 1.43 (3H), s | 1.42 (3H), s |
| 18 | 3.05, d (4.6); 2.55, d (4.4) | 3.07, d (4.5); 2.55, d (4.5) | 3.06, d (4.5); 2.55, d (4.5) | 3.05, d (4.5); 2.55, d (4.5) |
| 19 | 1.77 (3H), s | 1.78 (3H), s | 1.78 (3H), br s | 1.78 (3H), s |
| 20 | 1.00 (3H), d (7.0) | 1.01 (3H), d (7.0) | 1.01 (3H), d (7.3) | 1.01 (3H), d (7.5 |
| 2’ | 5.71, dd (11.5, 1.0) | 5.71, dd (11.5, 1.0) | 5.71, dd (11.6, 1.3) | 5.71, dd (12.0, 1.5) |
| 3’ | 5.90, dd (11.6, 8.0) | 5.90, dd (11.5, 8.0) | 5.90, dd (11.6, 7.8) | 5.90, dd (11.6, 8.0) |
| 4’ | 6.25, m | 6.26, m | 6.26, m | 6.26, m |
| 5’ | 1.32 (3H), d (6.5) | 1.33 (3H), d (6.5) | 1.34 (3H), d (6.5) | 1.33 (3H), d (8.5) |

**Figure S1.** UV Spectrum of Thailanstatin A.

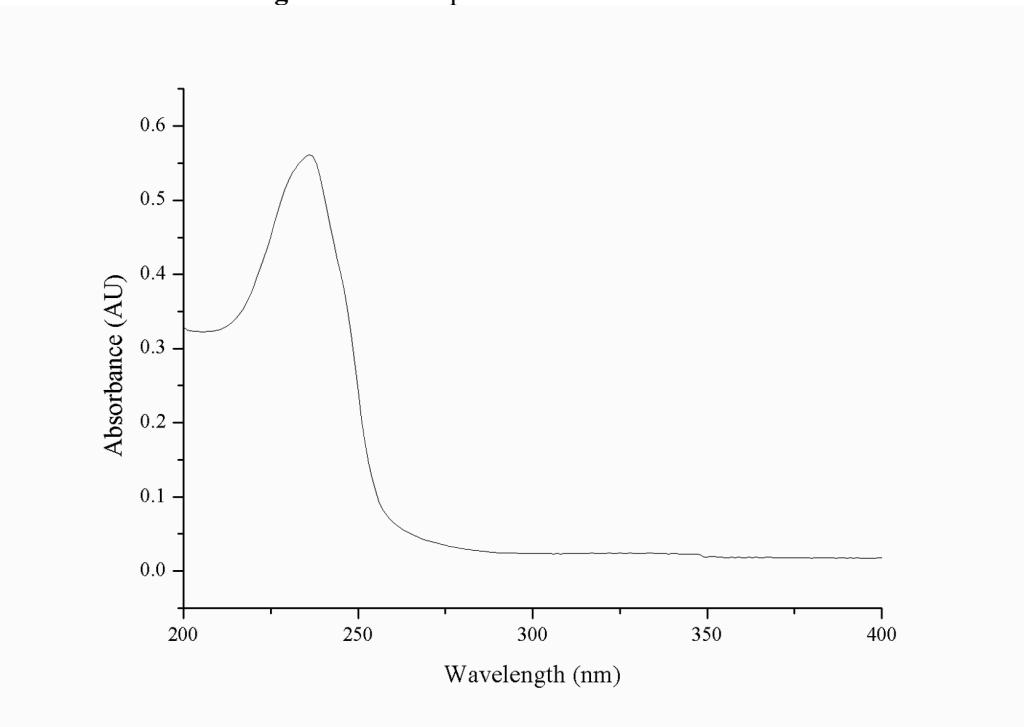



**Figure S2.** IR Spectrum of Thailanstatin A.

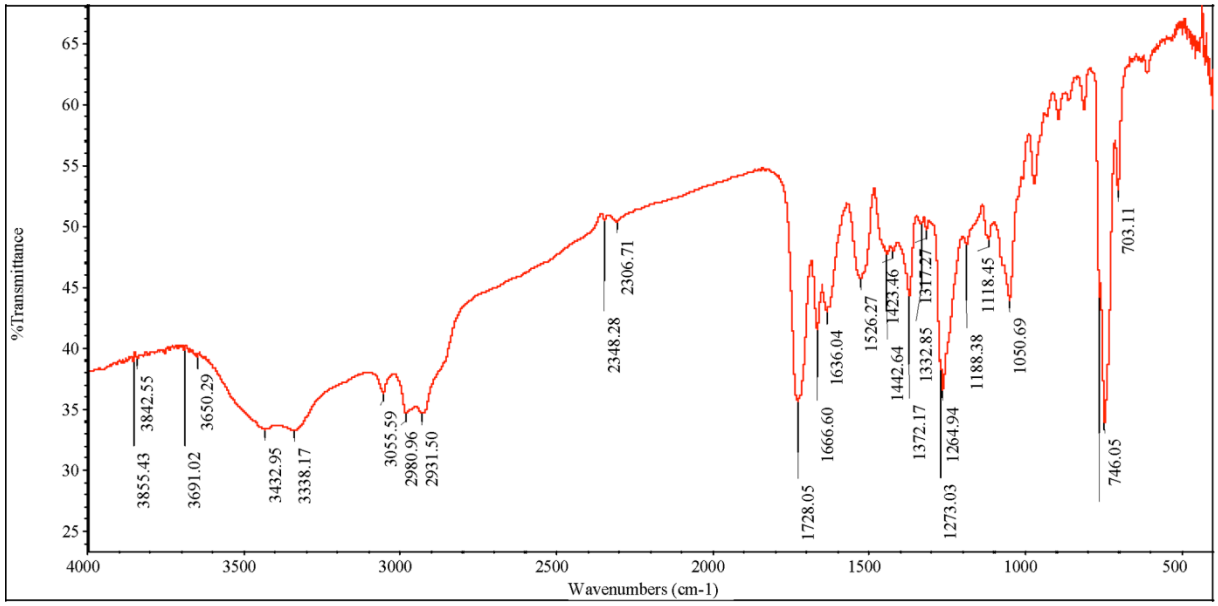
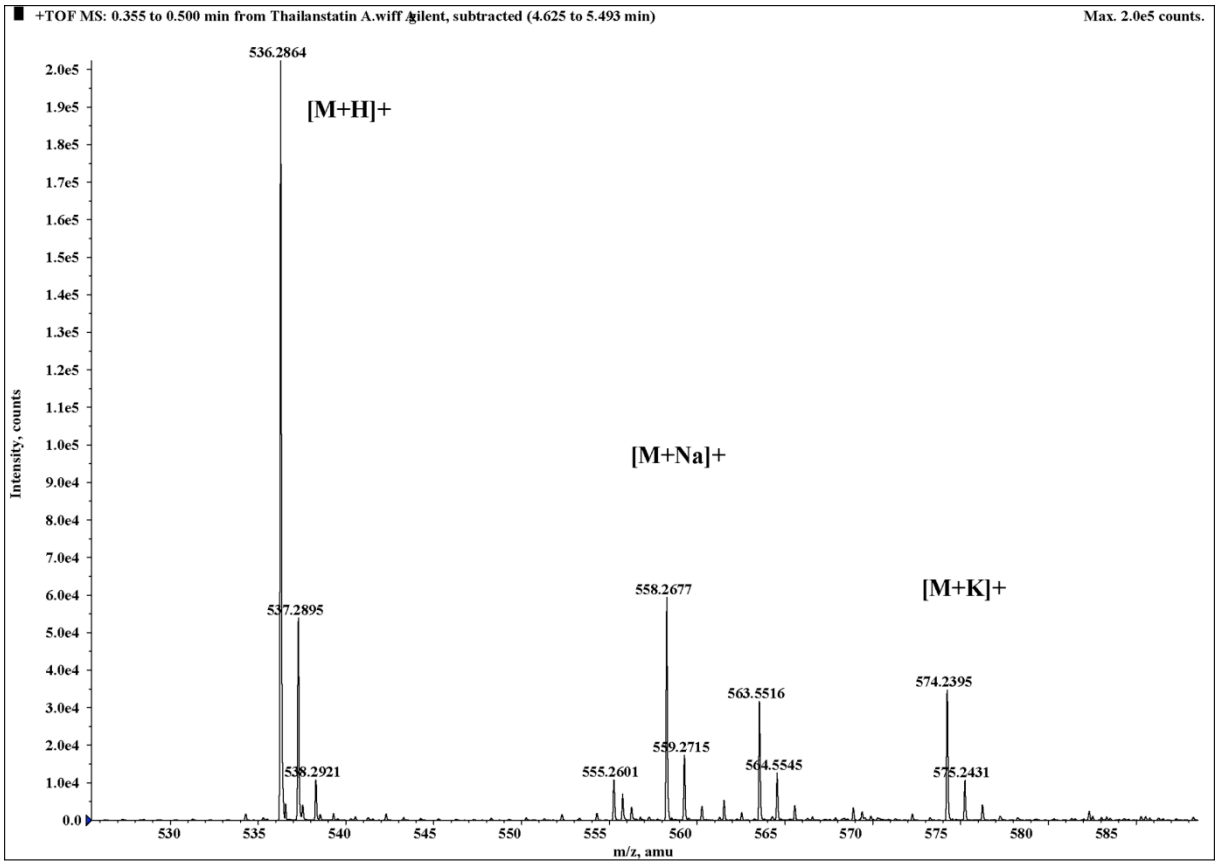


**Figure S3.** HR ESI-TOF MS Spectrum of Thailanstatin A.




**Thailanstatin A**

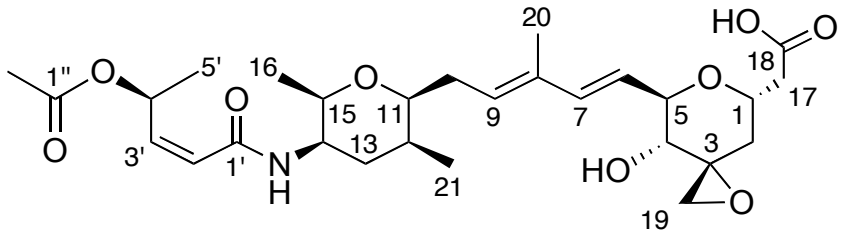


**Figure S4.** $^{1}H$ NMR Spectrum of Thailanstatin A ($CD_2Cl_2$, 298K, 500 MHz).

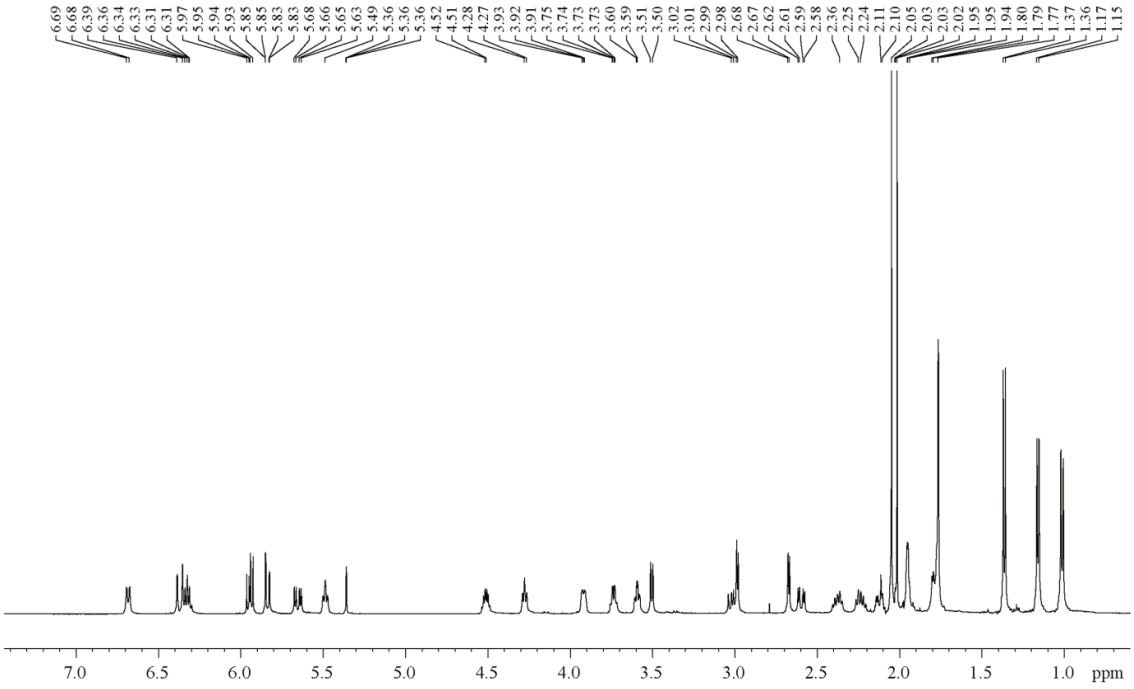



**Figure S5.** $^{13}C$ NMR Spectrum of Thailanstatin A ($CD_2Cl_2$, 298K, 125 MHz).

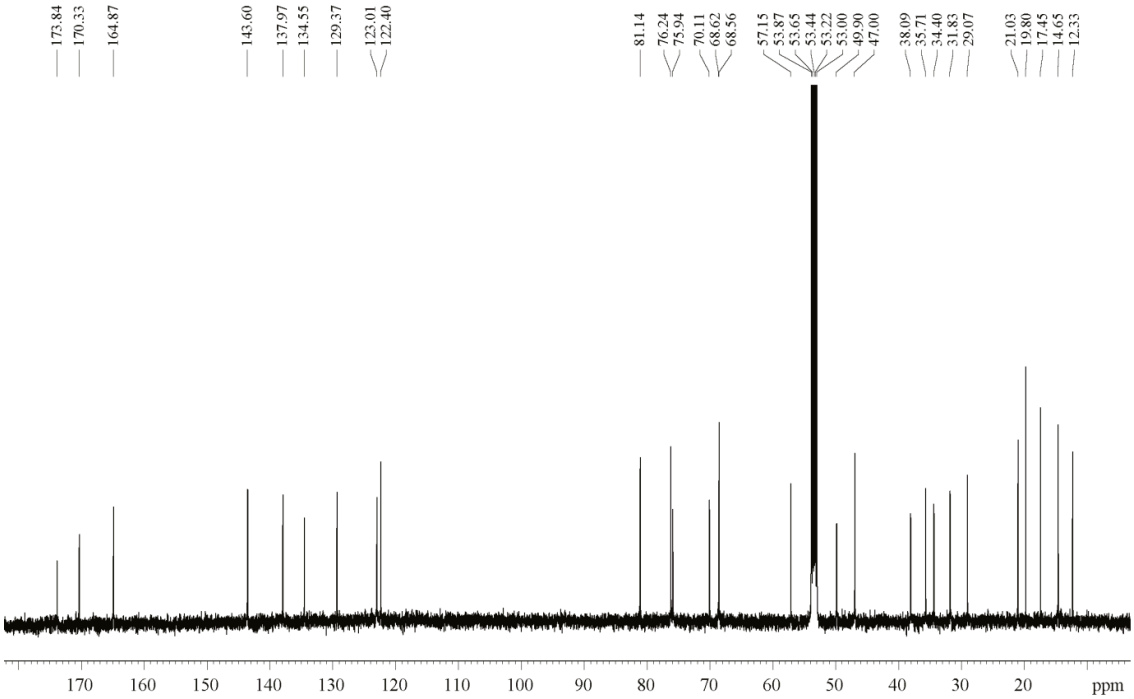


**Figure S6.** DEPT135 Spectrum of Thailanstatin A.

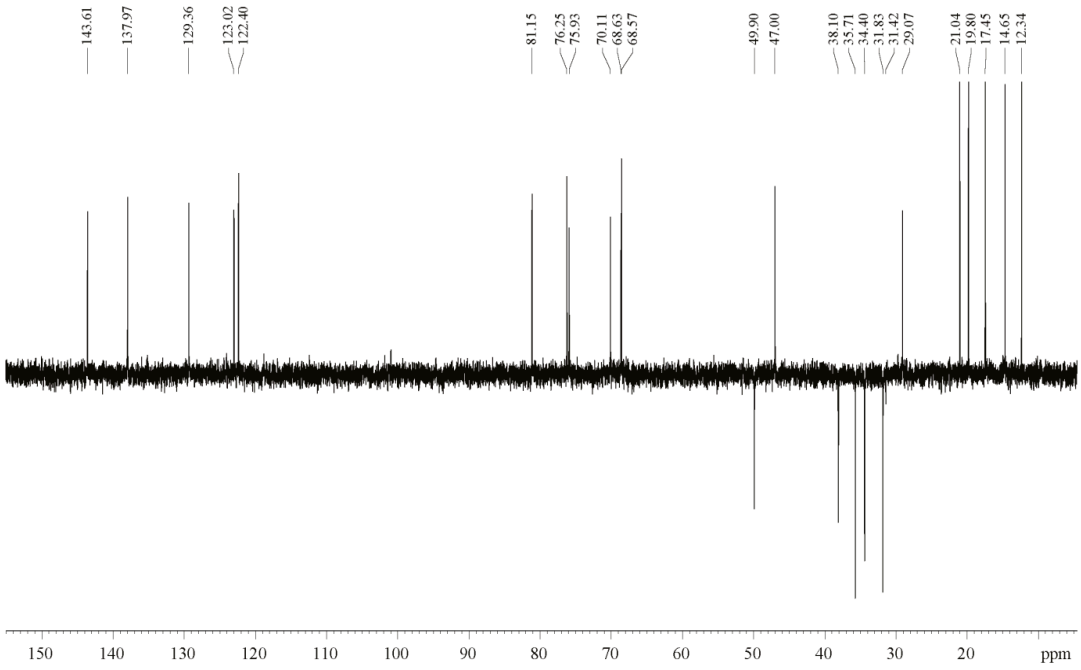



**Figure S7.** $^{1}H$-$^{13}C$ HSQC Spectrum of Thailanstatin A.

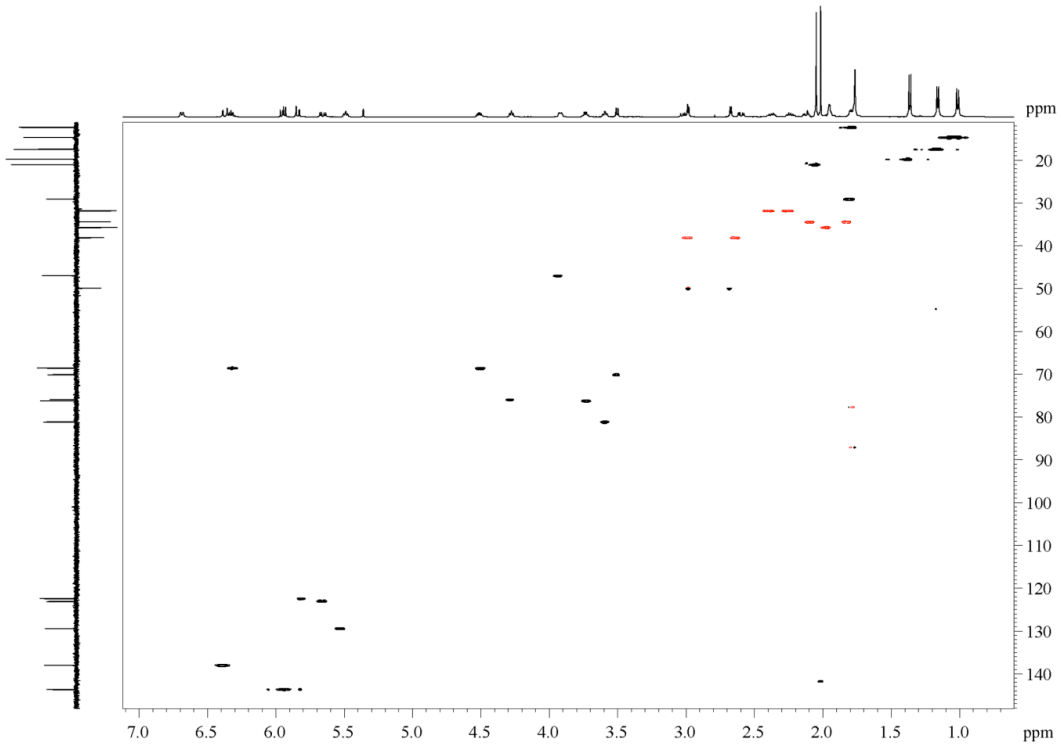


**Figure S8.** DQF COSY Spectrum of Thailanstatin A.

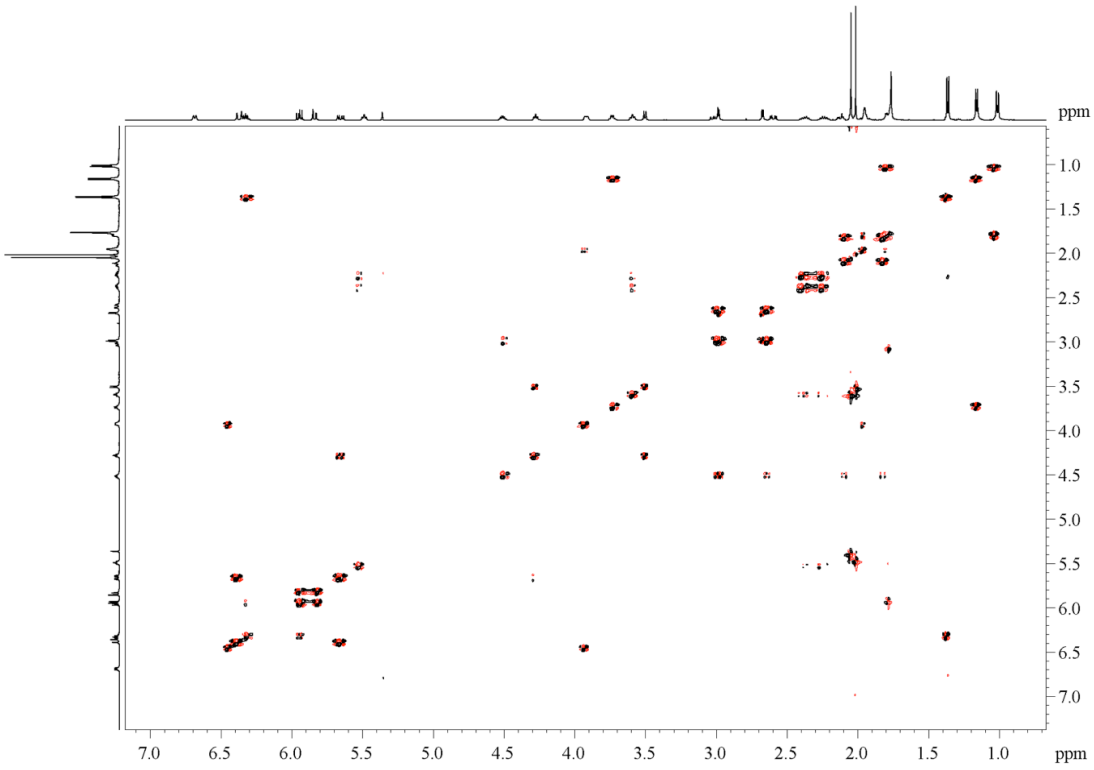



**Figure S9.** $^{1}H$-$^{13}C$ HMBC Spectrum of Thailanstatin A.

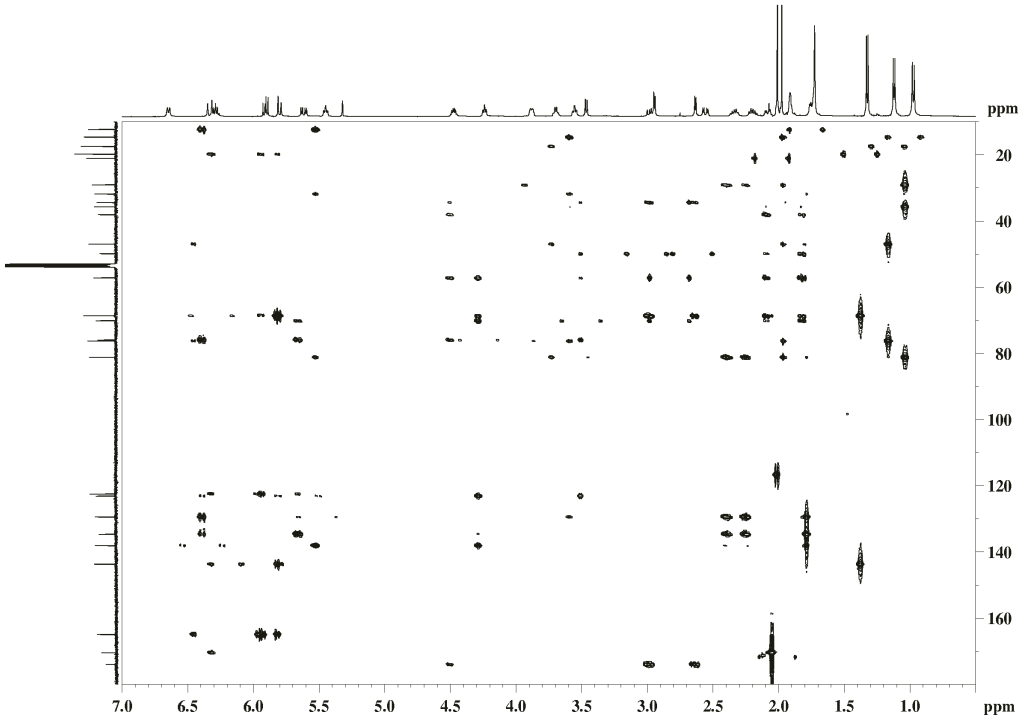


**Figure S10.** $^{1}$H-$^{1}$H NOESY Spectrum of Thailanstatin A.

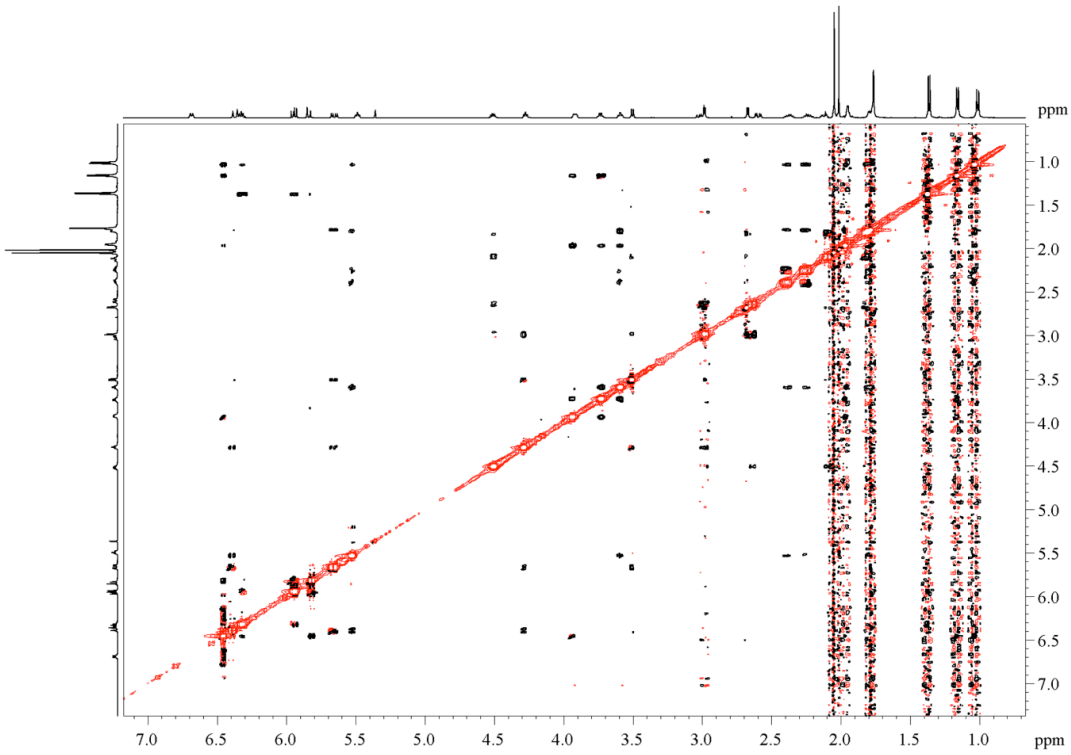



**Figure S11.** $^{1}$H-$^{1}$H ROESY Spectrum of Thailanstatin A.

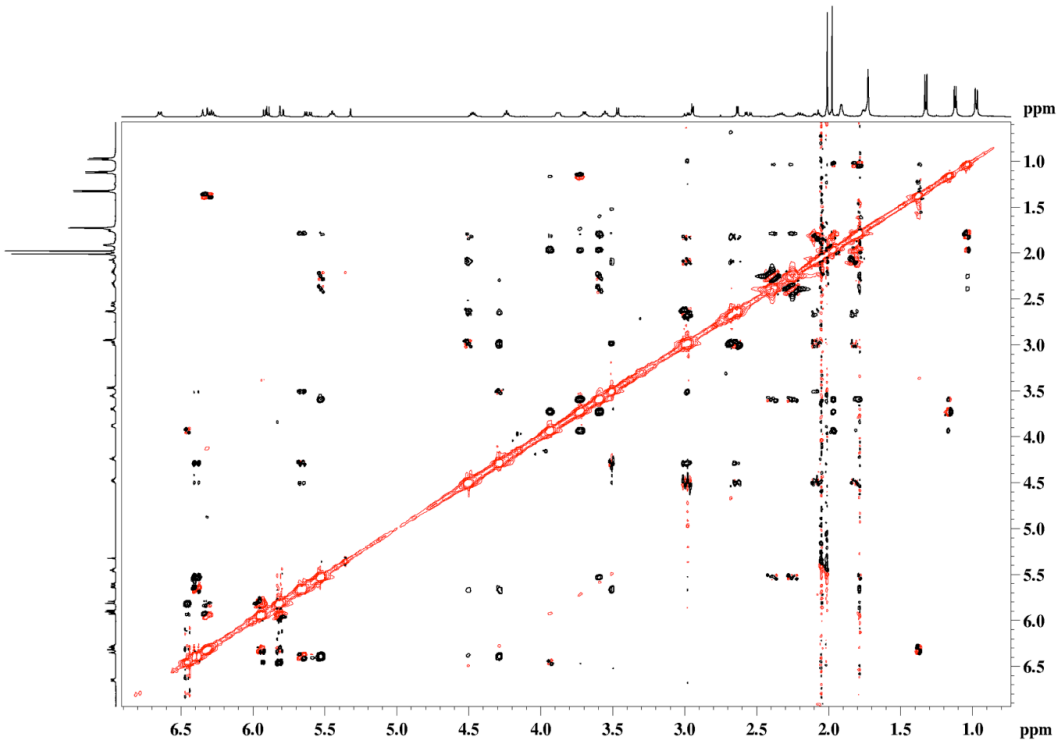


**Figure S12.** $^{1}H$-$^{15}N$ HSQC Spectrum of Thailanstatin A.

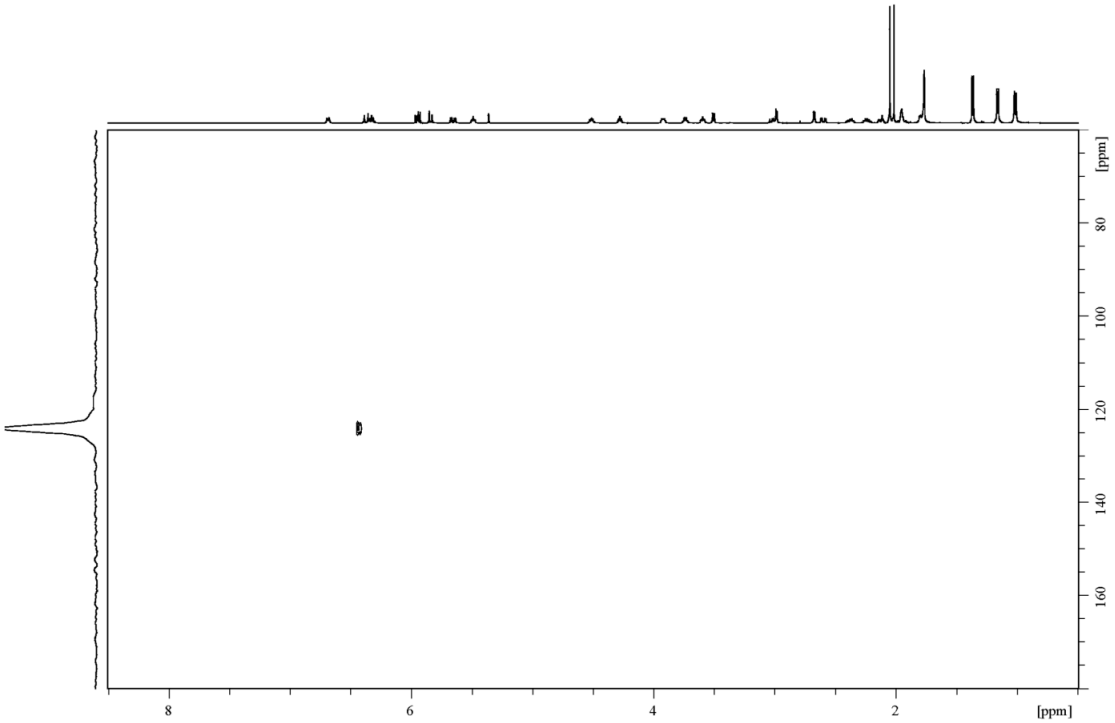


**Figure S13.** UV Spectrum of Thailanstatin B.

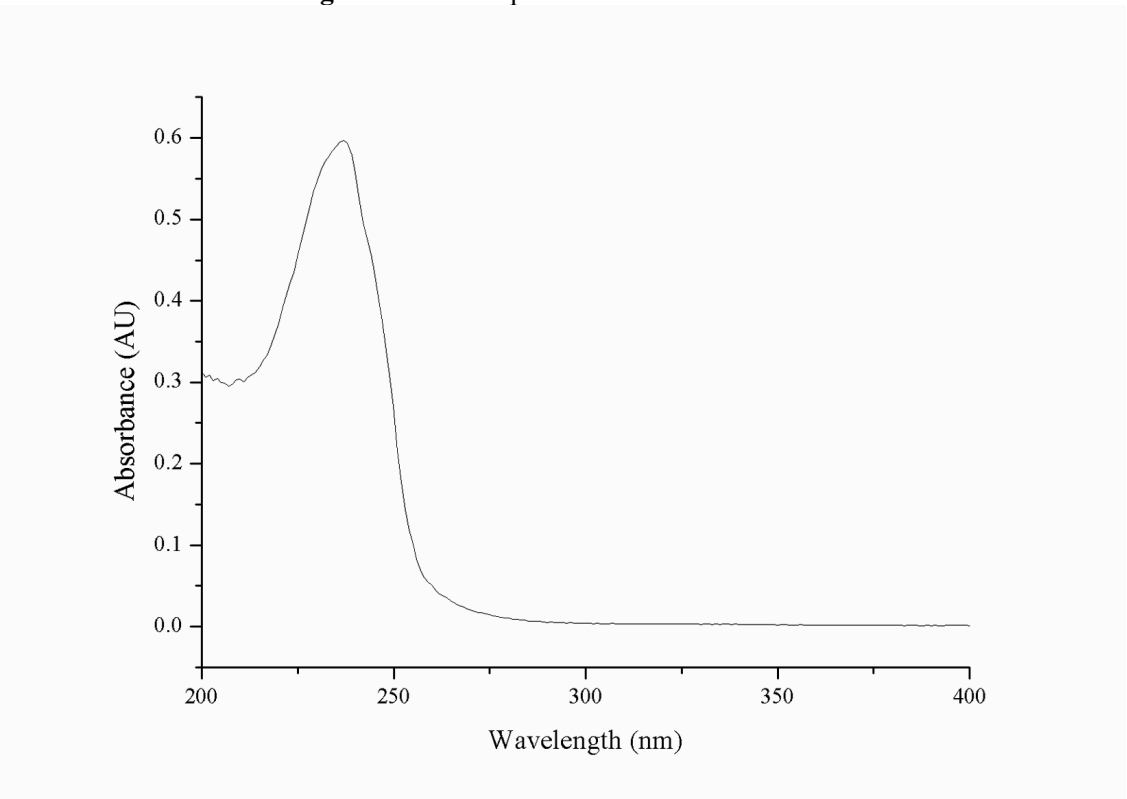



**Figure S14.** IR Spectrum of Thailanstatin B.

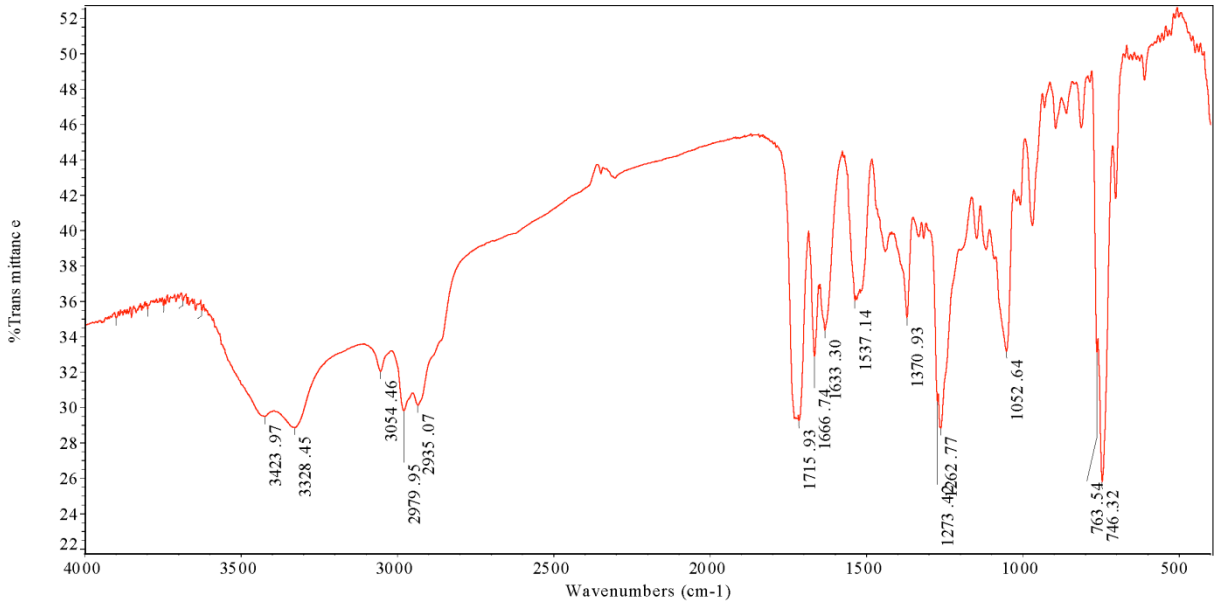
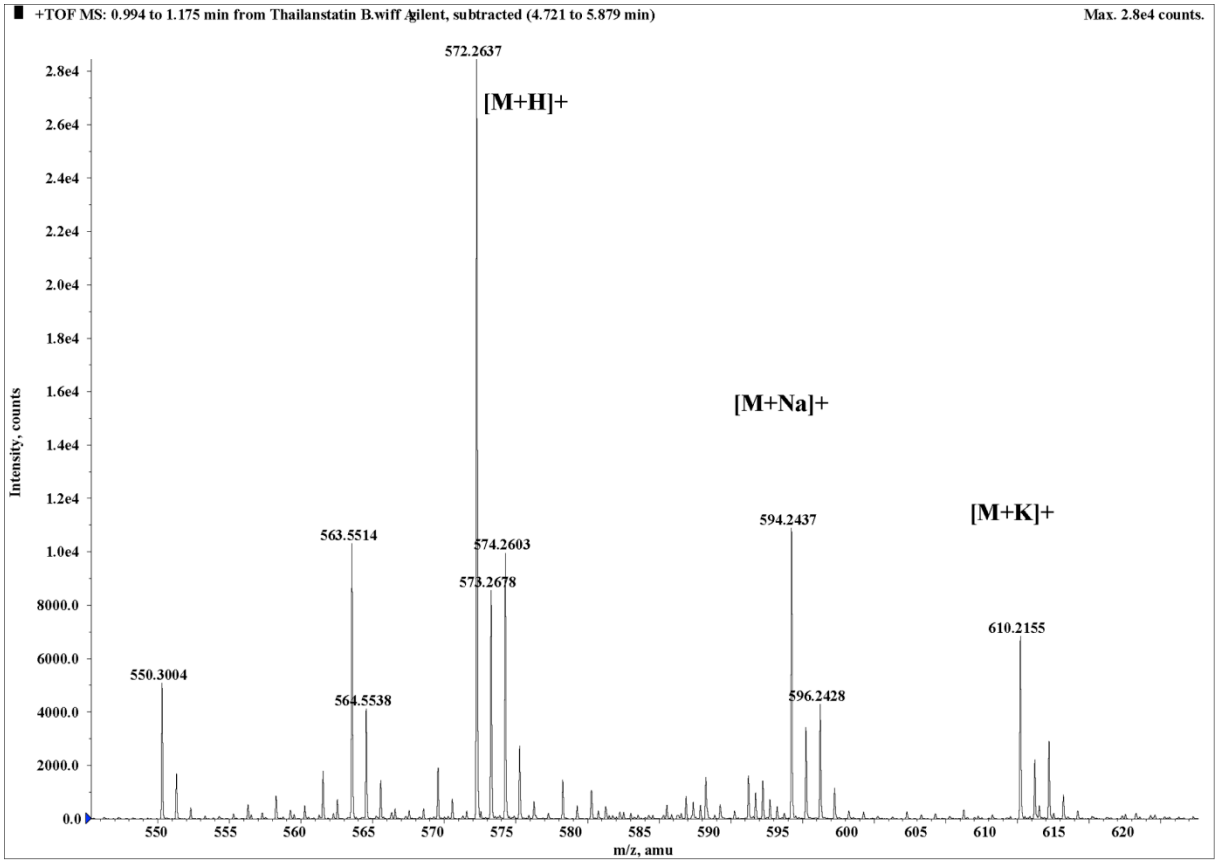


**Figure S15.** HR ESI-TOF MS Spectrum of Thailanstatin B.




**Thailanstatin B**

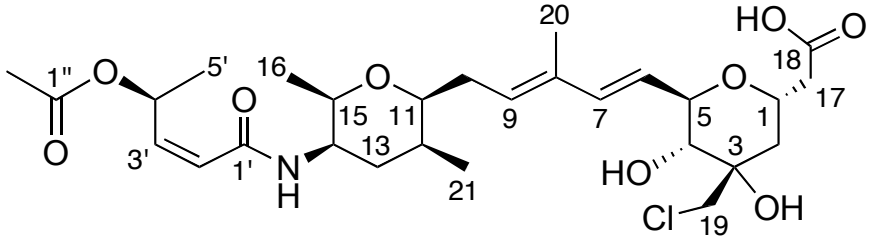


**Figure S16.** $^{1}$H NMR Spectrum of Thailanstatin B ($CDCl_3$, 298K, 500 HMz).

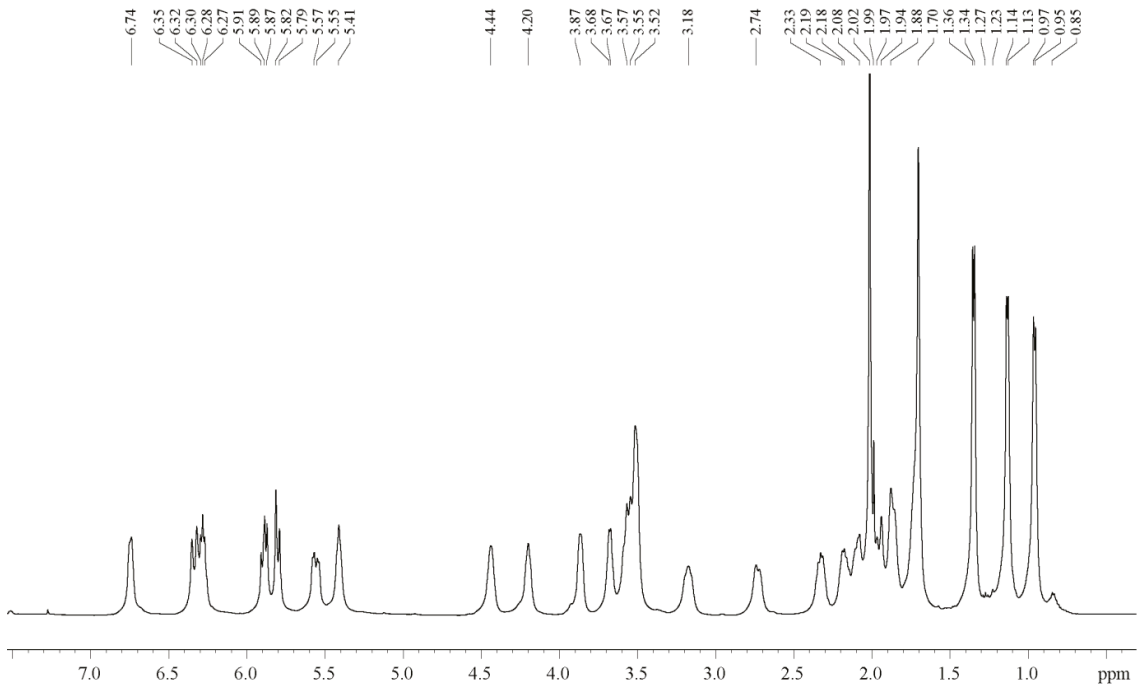



**Figure S17.** $^{13}$C NMR Spectrum of Thailanstatin B ($CDCl_3$, 298K, 125 HMz).

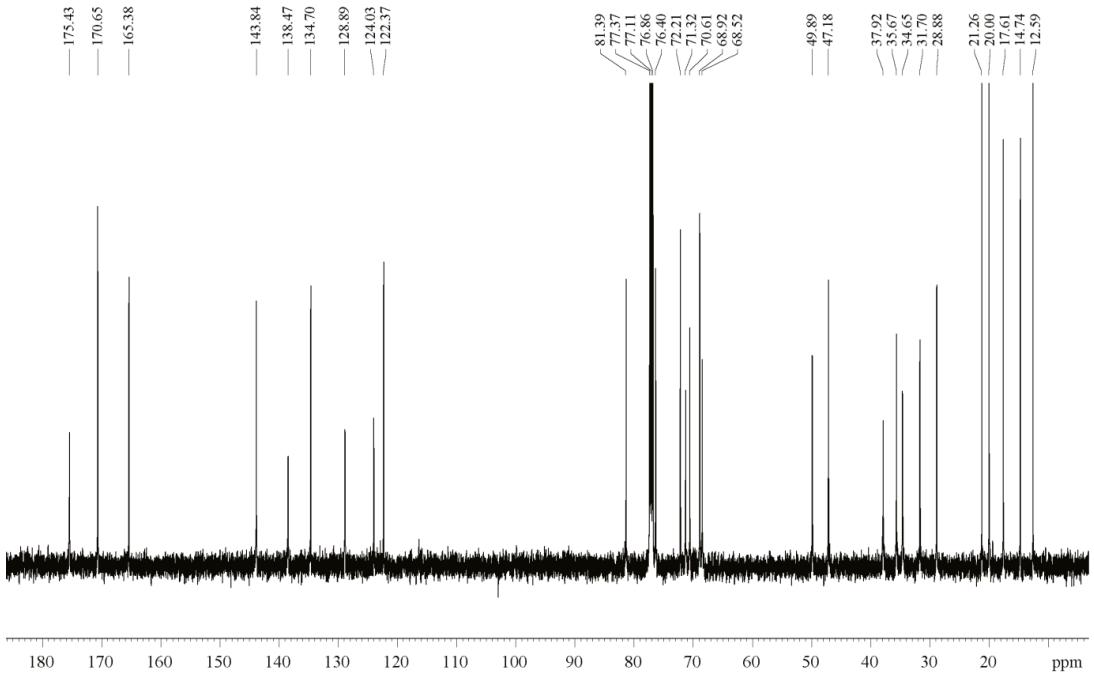


**Figure S18.** DEPT135 Spectrum of Thailanstatin B.

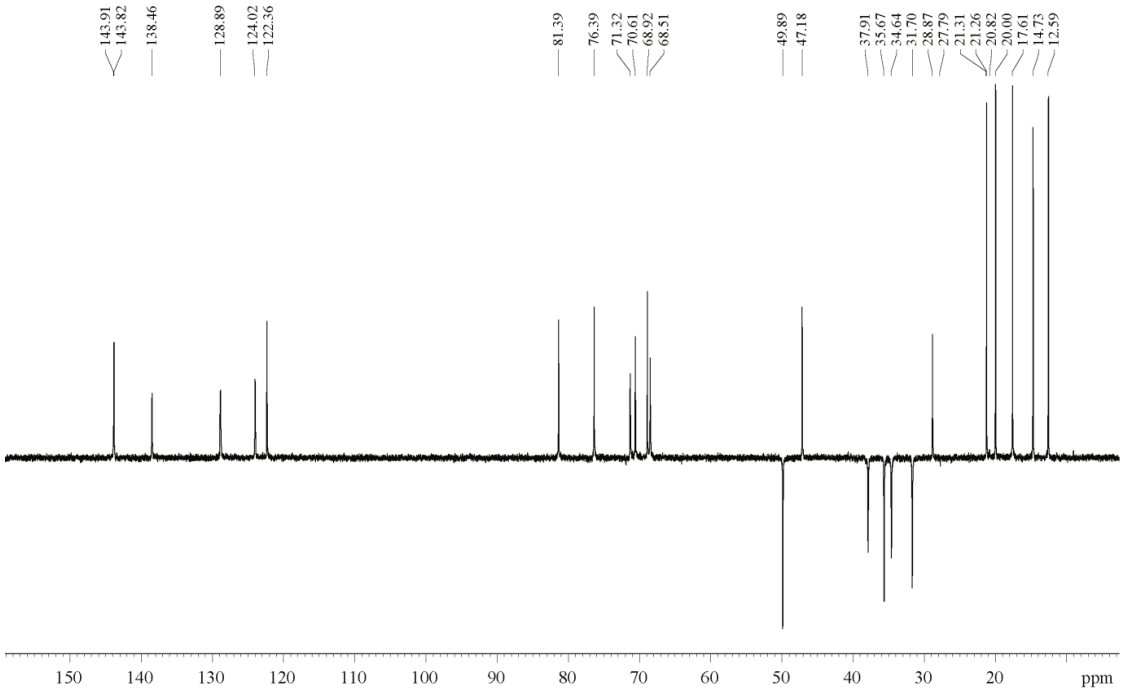



**Figure S19.** $^{1}H$-$^{13}C$ HSQC Spectrum of Thailanstatin B.

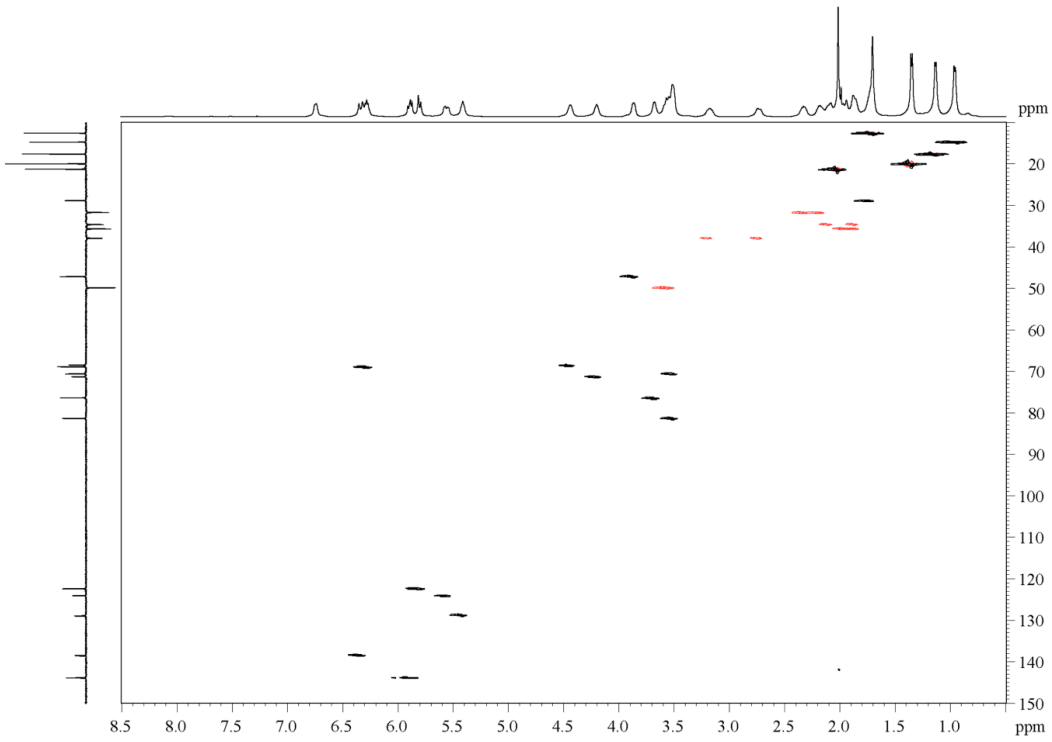


**Figure S20.** DQF COSY Spectrum of Thailanstatin B.

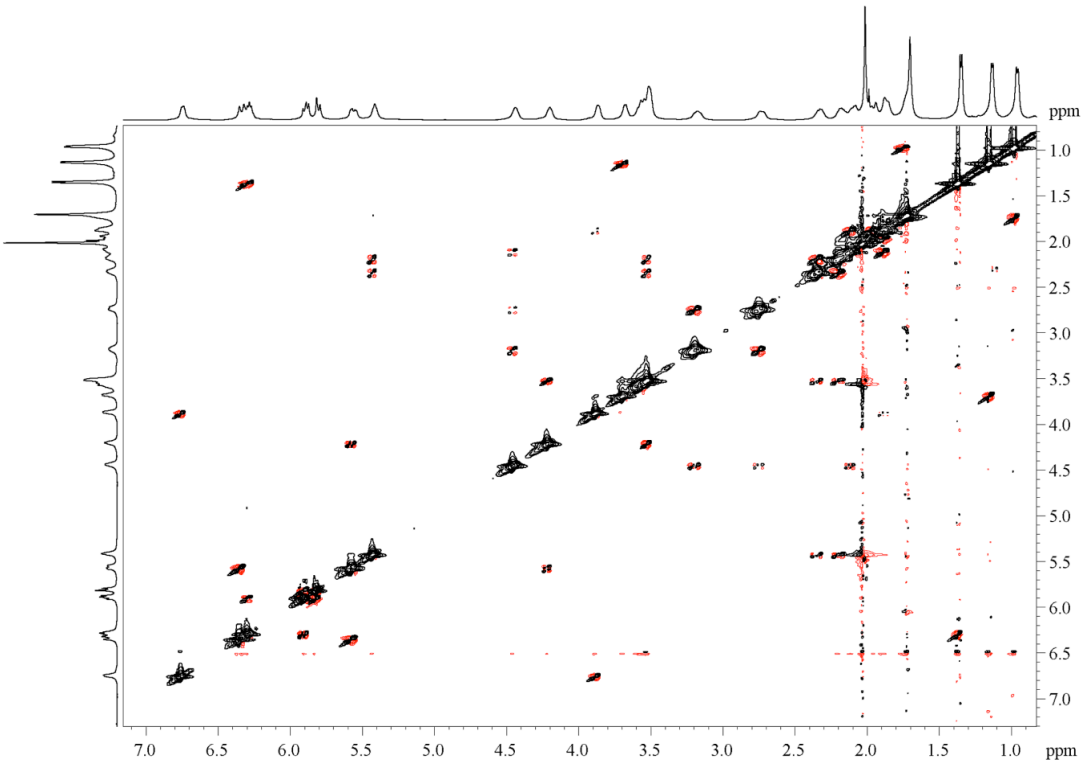



**Figure S21.** $^{1}H$-$^{13}C$ HMBC Spectrum of Thailanstatin B.

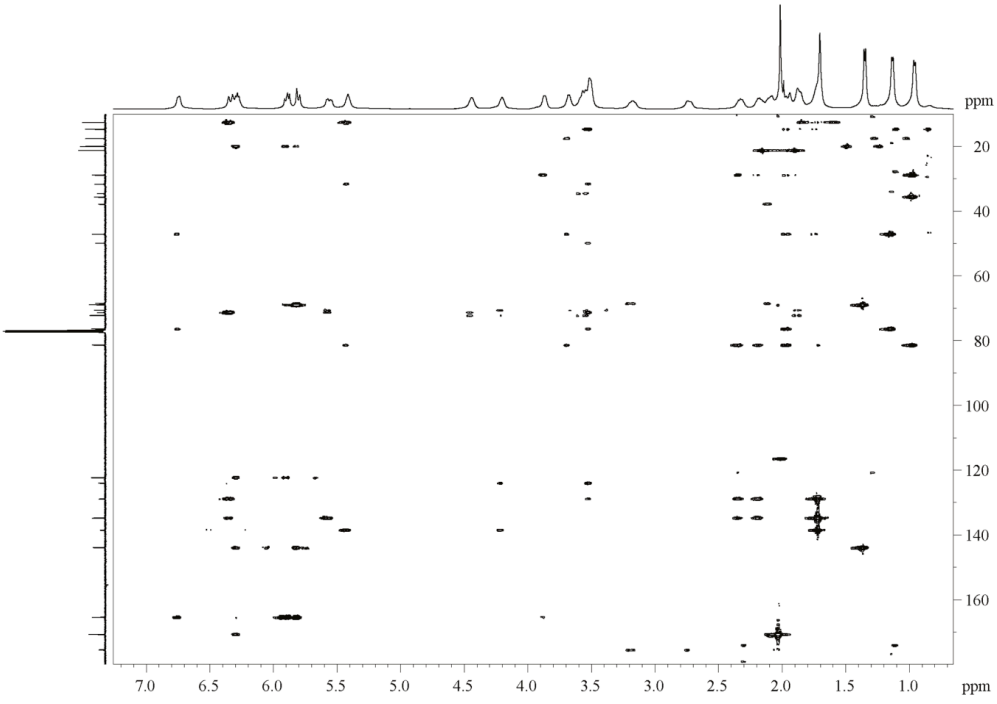


**Figure S22.** $^{1}$H-$^{1}$H NOESY Spectrum of Thailanstatin B.

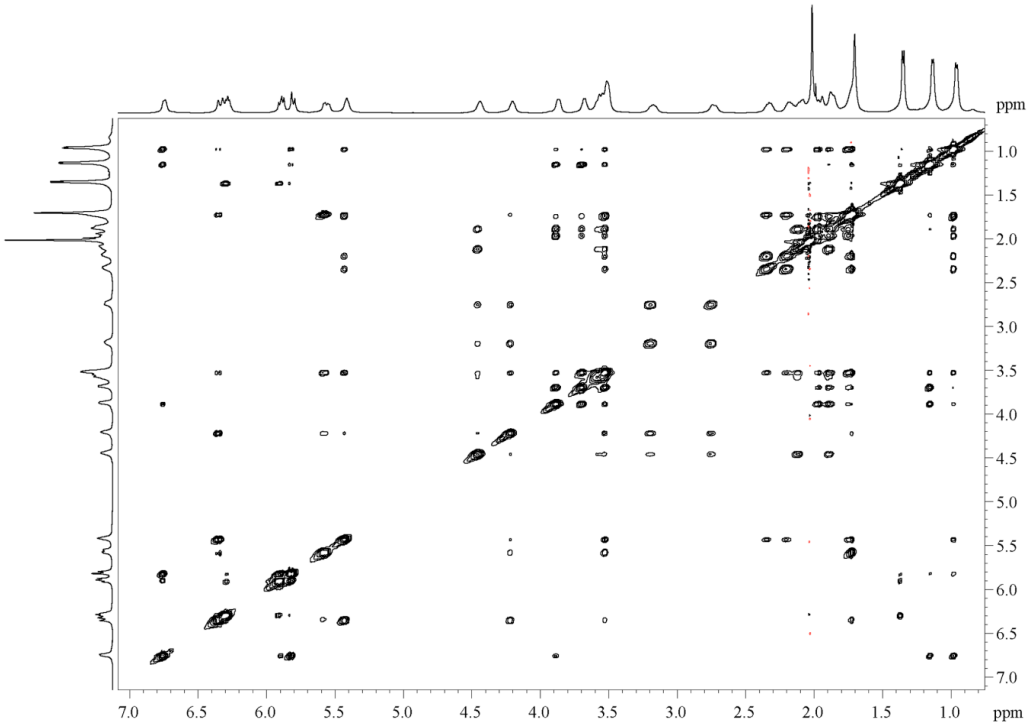



**Figure S23.** 1,1-ADEQUATE Spectrum of Thailanstatin B.

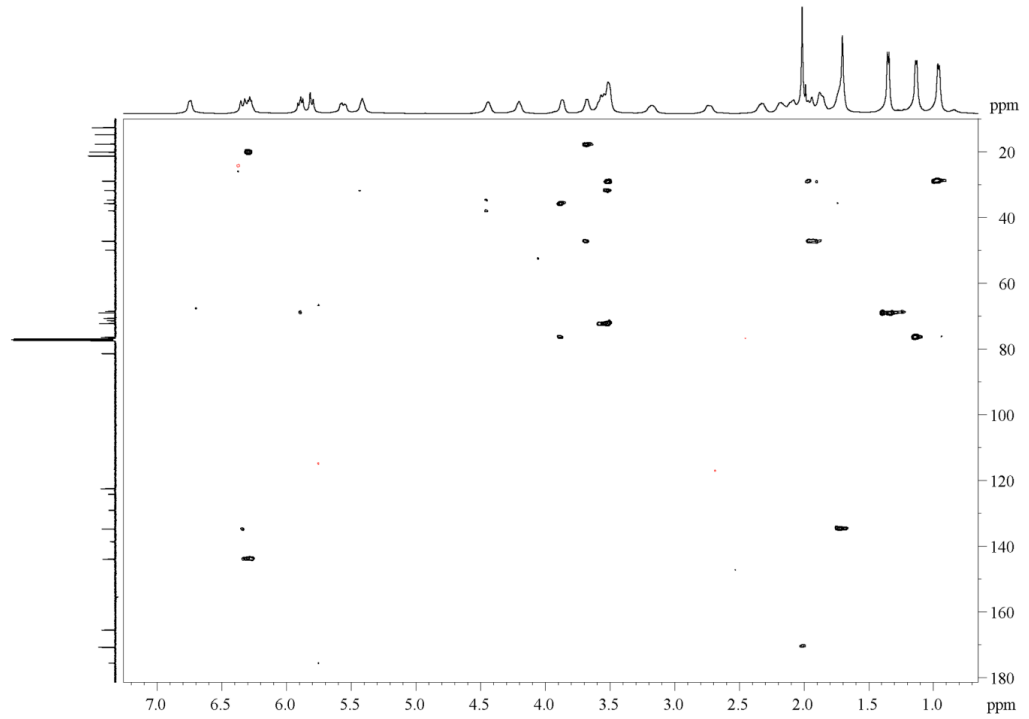


**Figure S24.** $^{1}H$-$^{15}N$ HSQC Spectrum of Thailanstatin B.

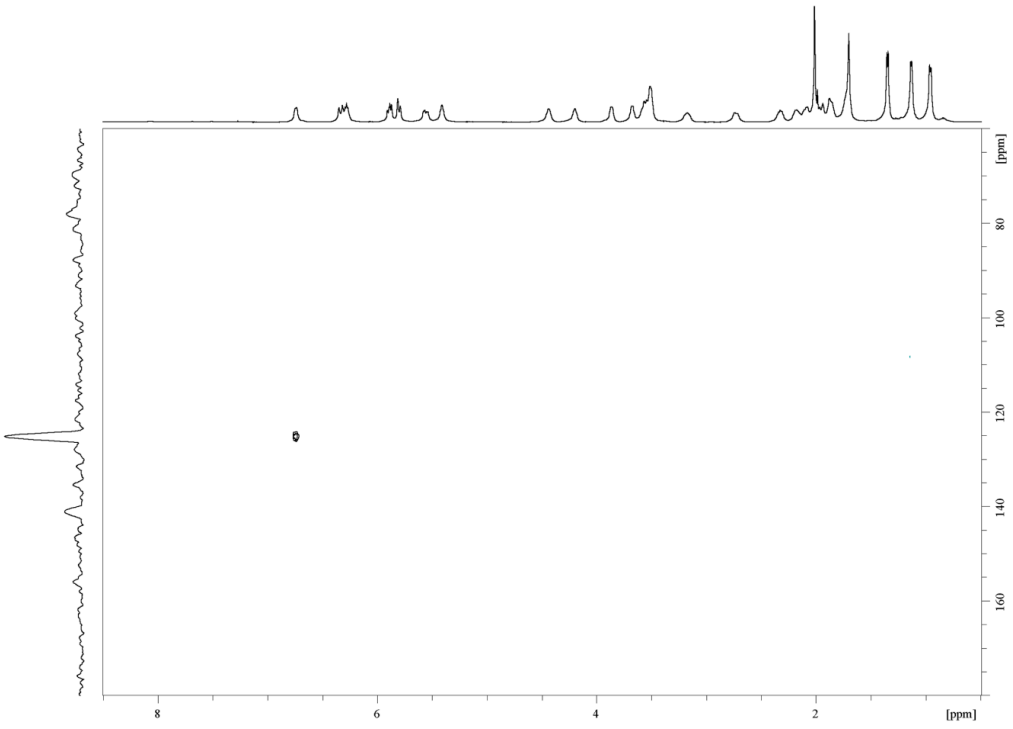


**Figure S25.** UV Spectrum of Thailanstatin C.

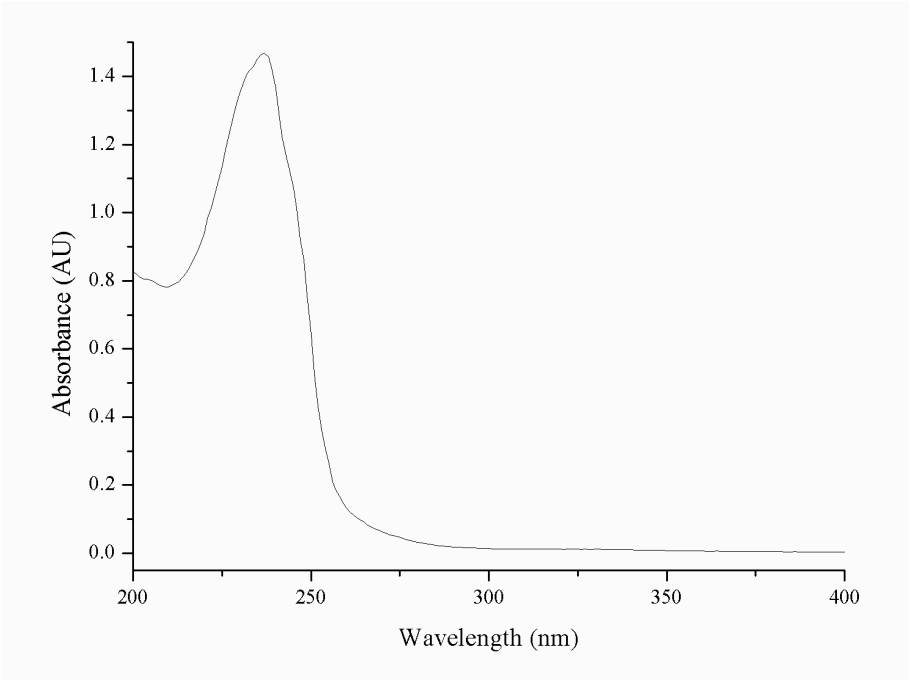



**Figure S26.** IR Spectrum of Thailanstatin C.

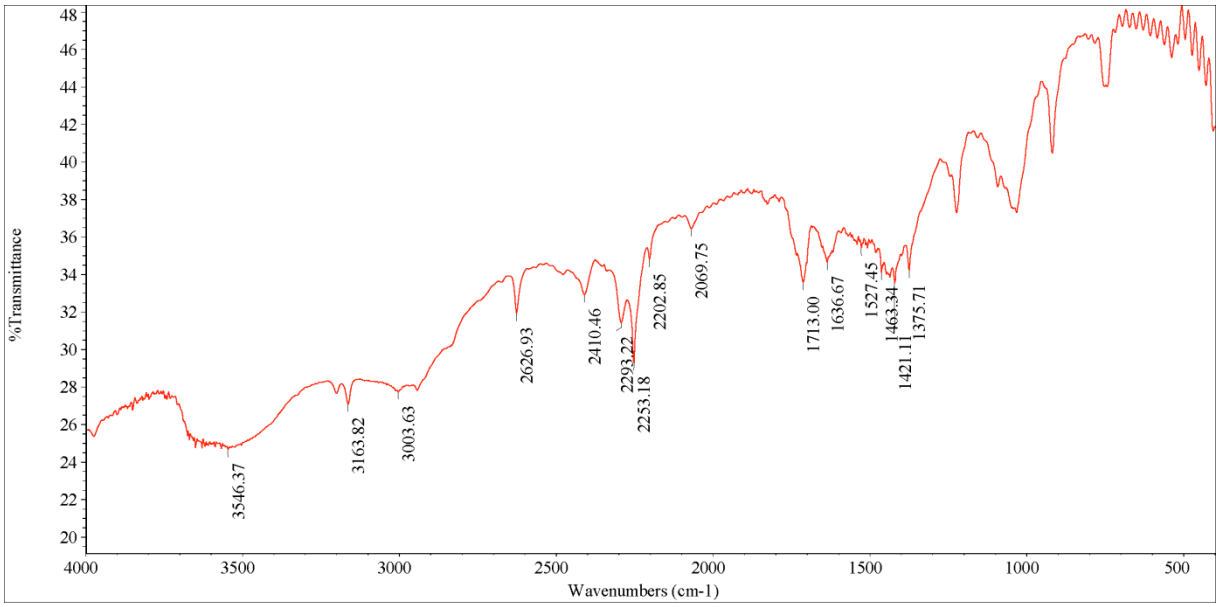
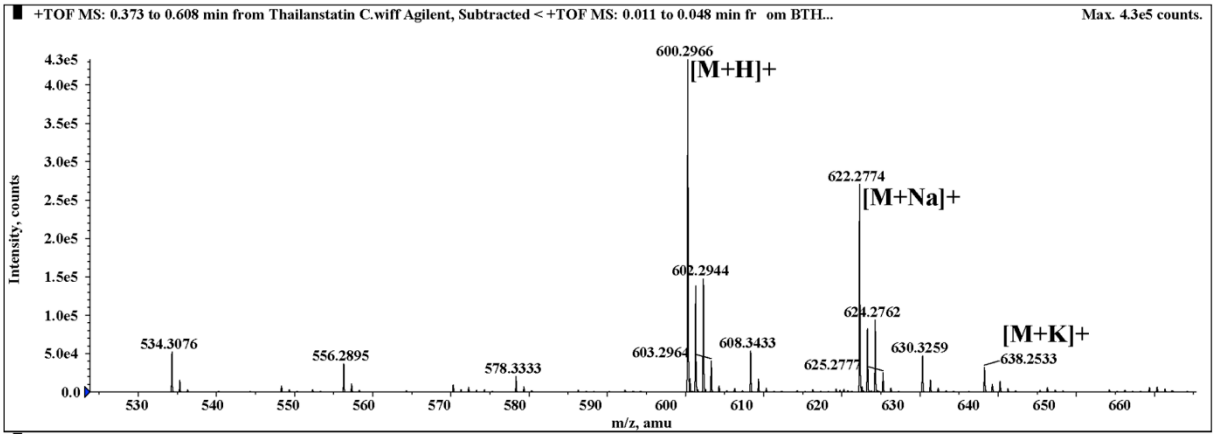


**Figure S27.** HR ESI-TOF MS Spectrum of Thailanstatin C.




**Thailanstatin C**

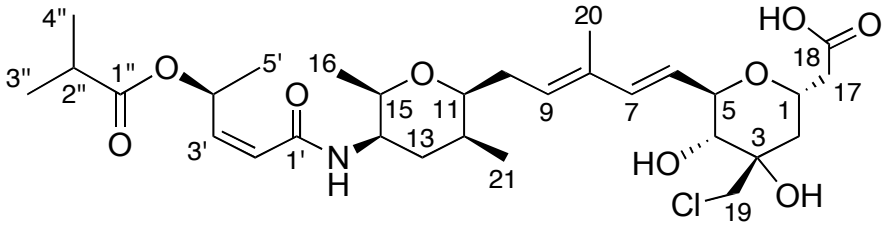


**Figure S28.** $^{1}$H NMR Spectrum of Thailanstatin C ($CDCl_3$, 298K, 500 HMz).

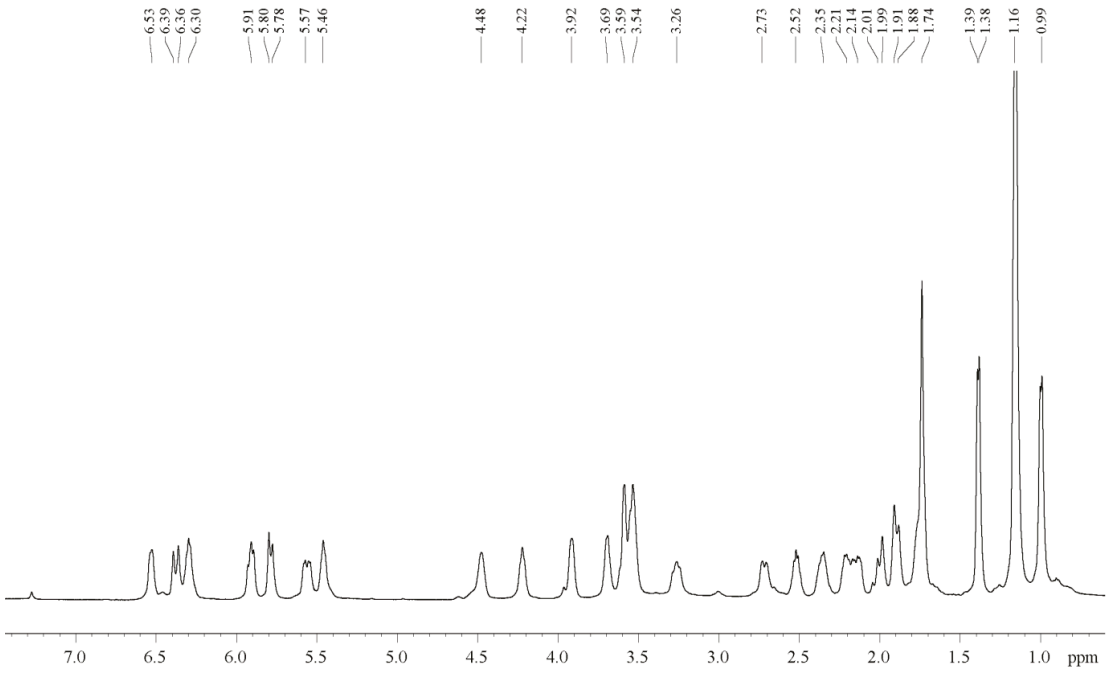



**Figure S29.** $^{13}$C NMR Spectrum of Thailanstatin C ($CDCl_3$, 298K, 125 HMz).

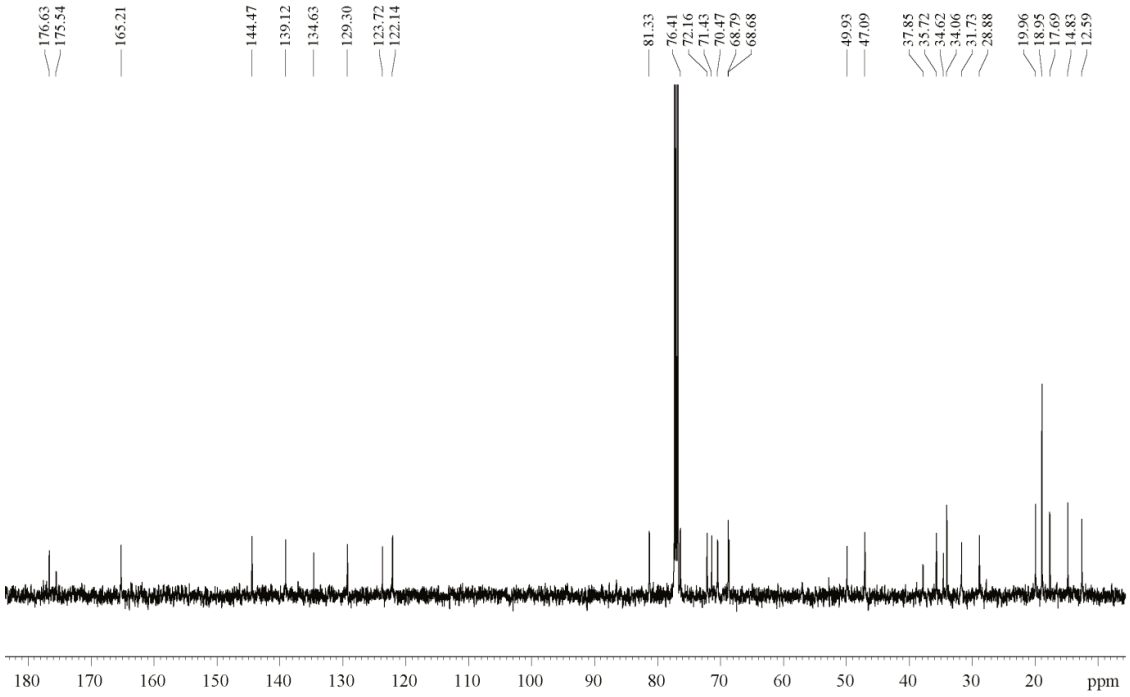


**Figure S30.** DEPT135 Spectrum of Thailanstatin C.

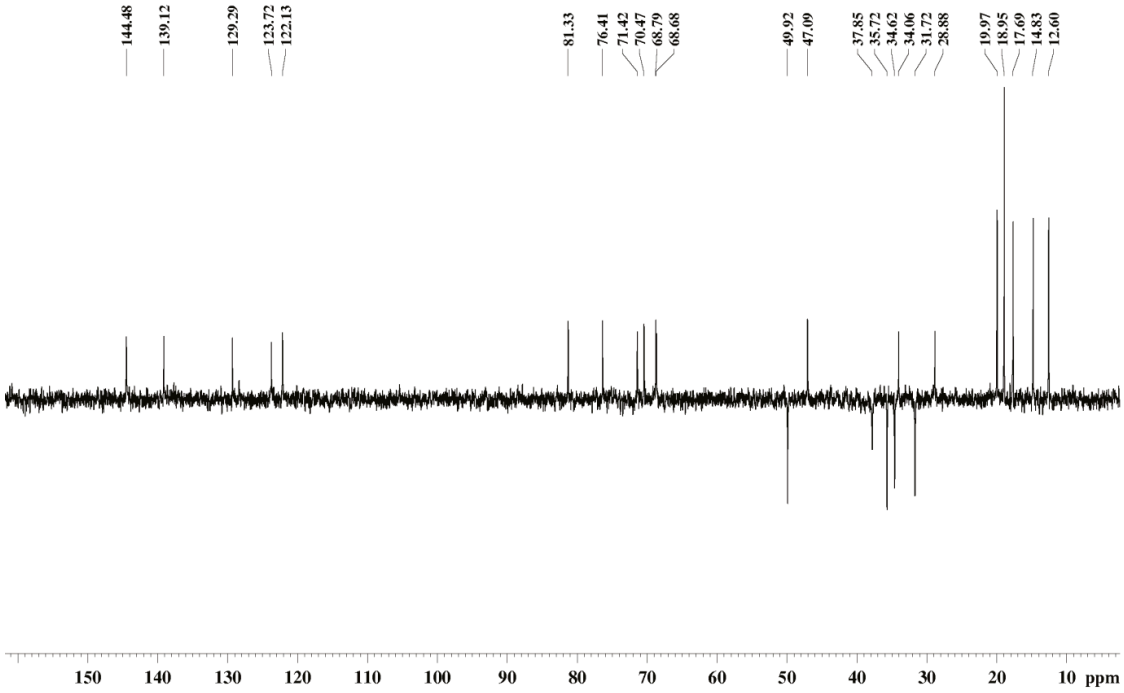



**Figure S31.** $^{1}H$-$^{13}C$ HSQC Spectrum of Thailanstatin C.

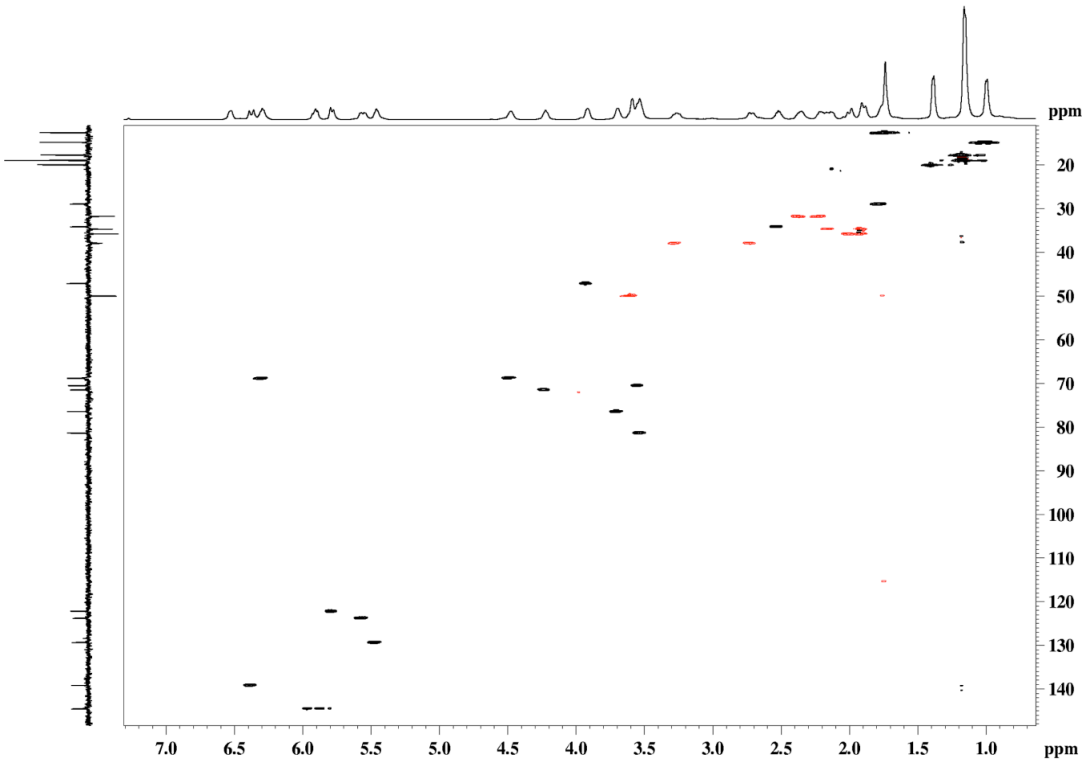


**Figure S32.** DQF COSY Spectrum of Thailanstatin C.

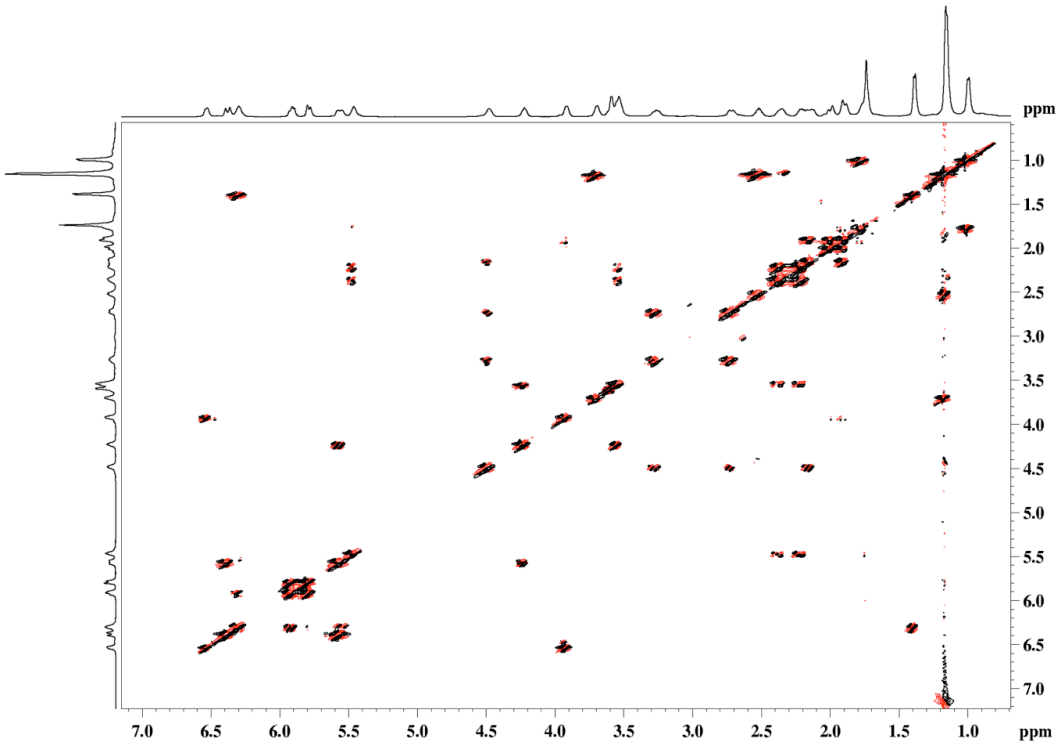



**Figure S33.** $^{1}H$-$^{13}C$ HMBC Spectrum of Thailanstatin C.

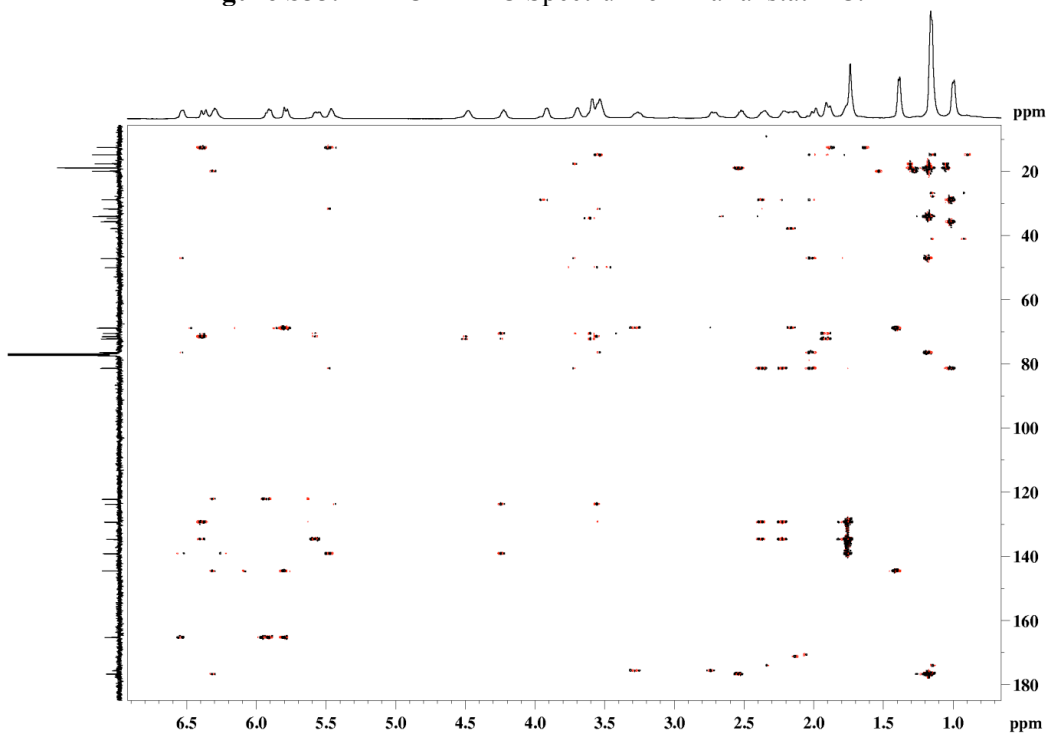


**Figure S34.** $^{1}$H-$^{1}$H NOESY Spectrum of Thailanstatin C.

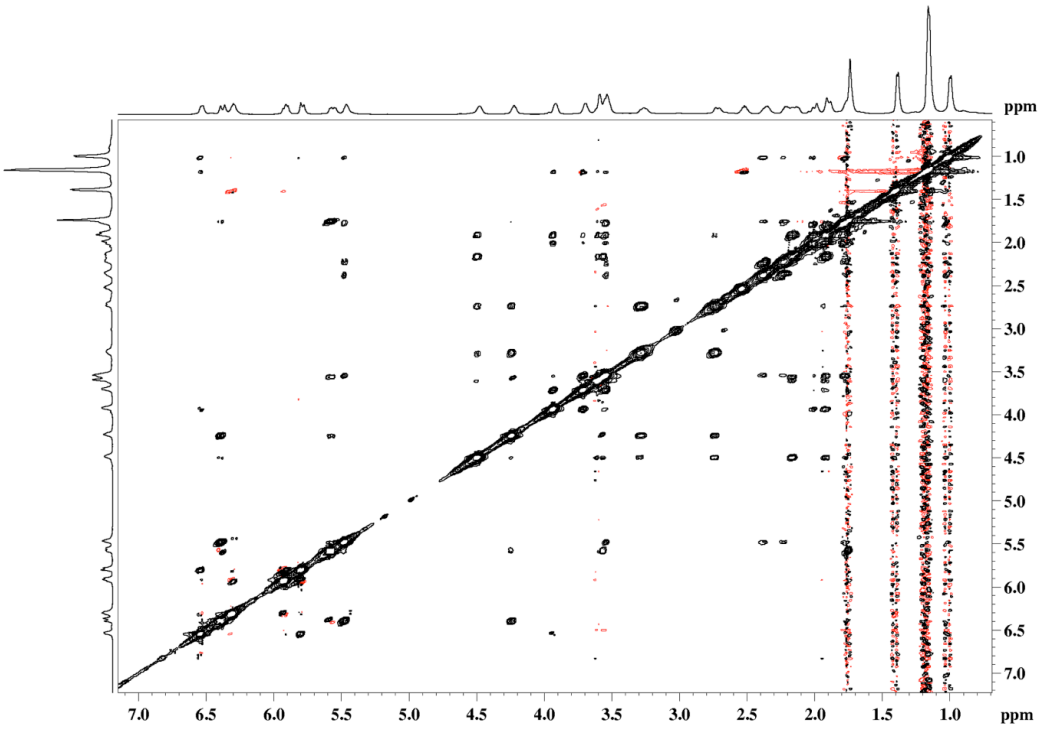
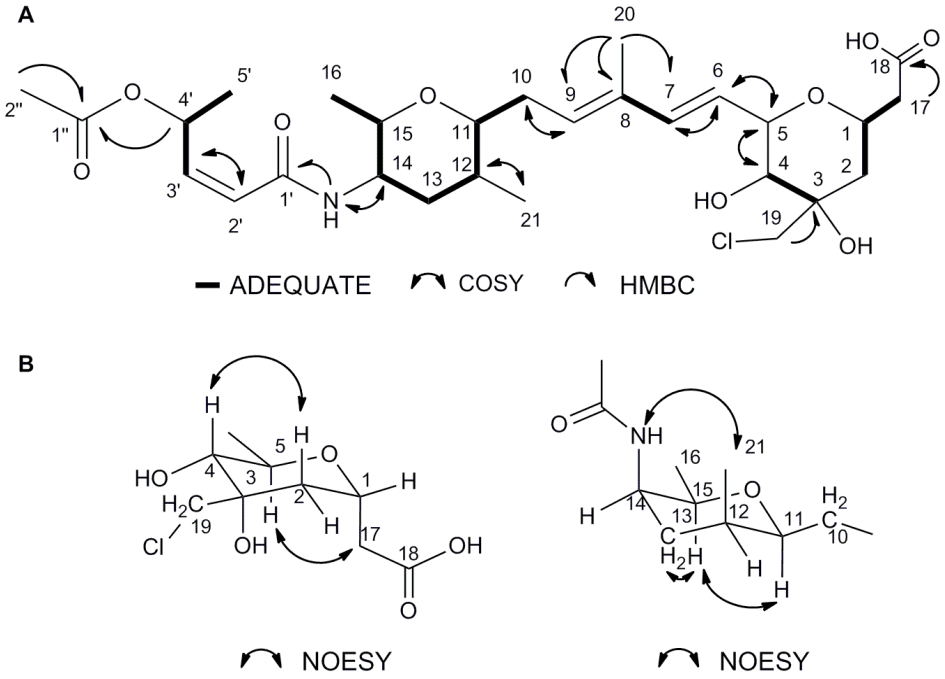



**Figure S35.** Diagnostic 2D NMR Correlations and Structure Fragments of Thailanstatin B.



**FR901464**

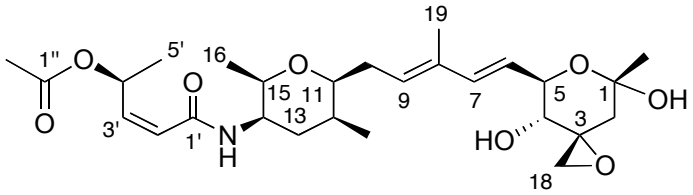



**Figure S36.** $^{13}C$ Spectrum of Freshly Purified Natural FR901464 ($CD_2Cl_2$, 298K, 125 MHz).

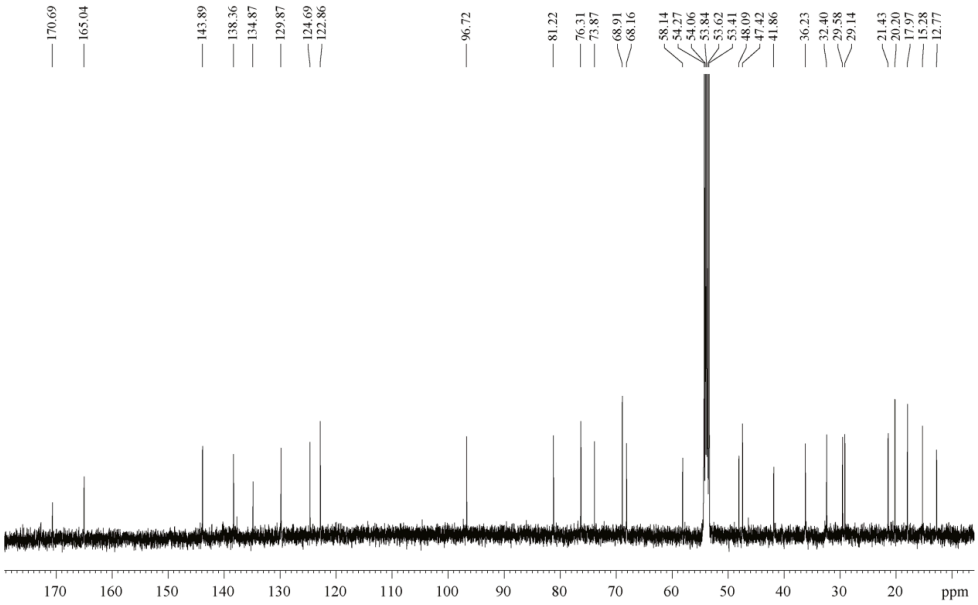



**Figure S37.** $^{1}H$ Spectrum of Freshly Purified Natural FR901464 ($CD_2Cl_2$, 298K, 500 MHz).

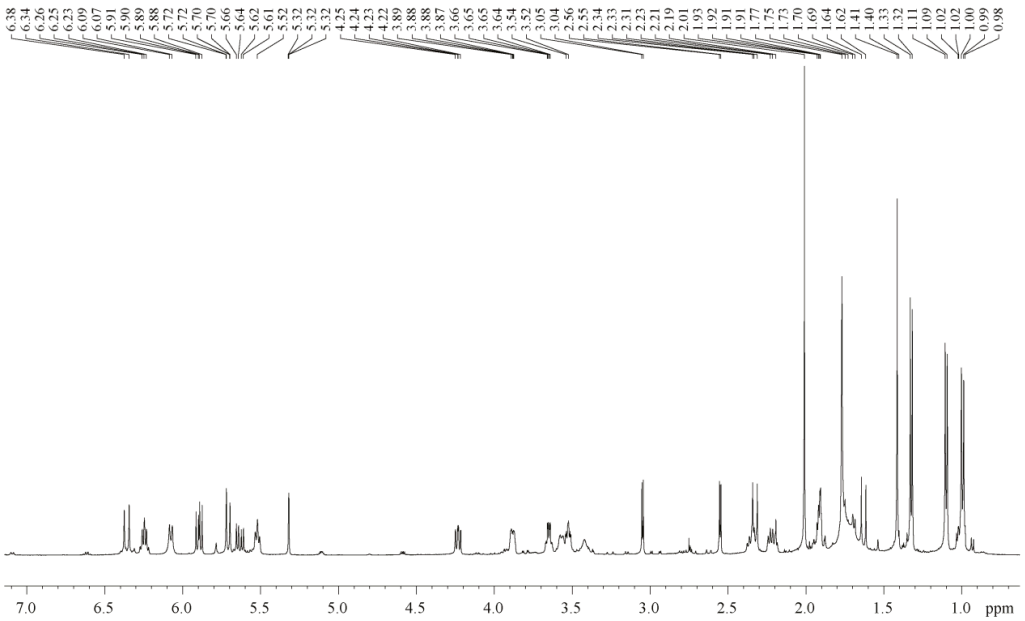


**Figure S38.** TLC Staining and Visualization of FR901464 (**a**) and Thailanstatins A (**b**), B (**c**) and C (**d**). The plate was sprayed with 0.1% Bromocresol Green (BCG) in ethanol solution. The visualization of a yellow spot on a blue background TLC plate supproted that Thailanstatins A, B and C contain a carboxy group.

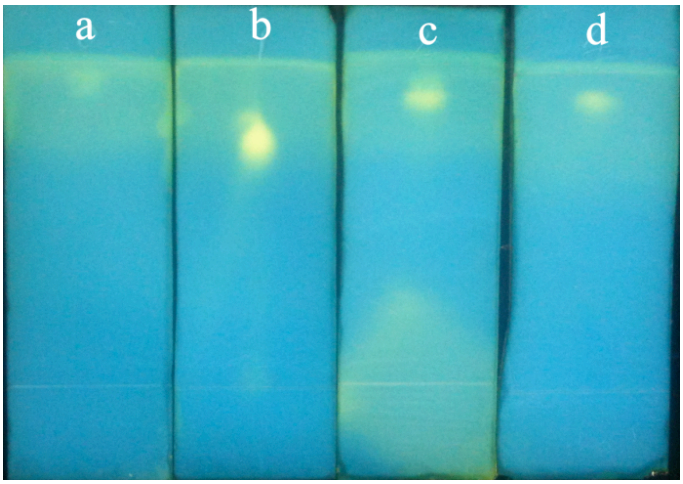


**Figure S39.** Amino Acid Alignments of Conserved Regions of PKS and NRPS Domains. (**A**) The PKS and NRPS domains of thailanstatin biosynthetic enzymes; (**B**) The conserved amino acid sites in the glyceryl transferase/phosphatase domain; (**C**) the conserved amino acid sites in the Baeyer-Villiger oxidase domain of TstGH. KS, ketosynthase; CP, carrier protein; AT, acyltransferase; DH, dehydratase; C, condensation; TE, thioesterase; KR, ketoreductase. The active sites or the key conserved amino acid residues are marked with *.

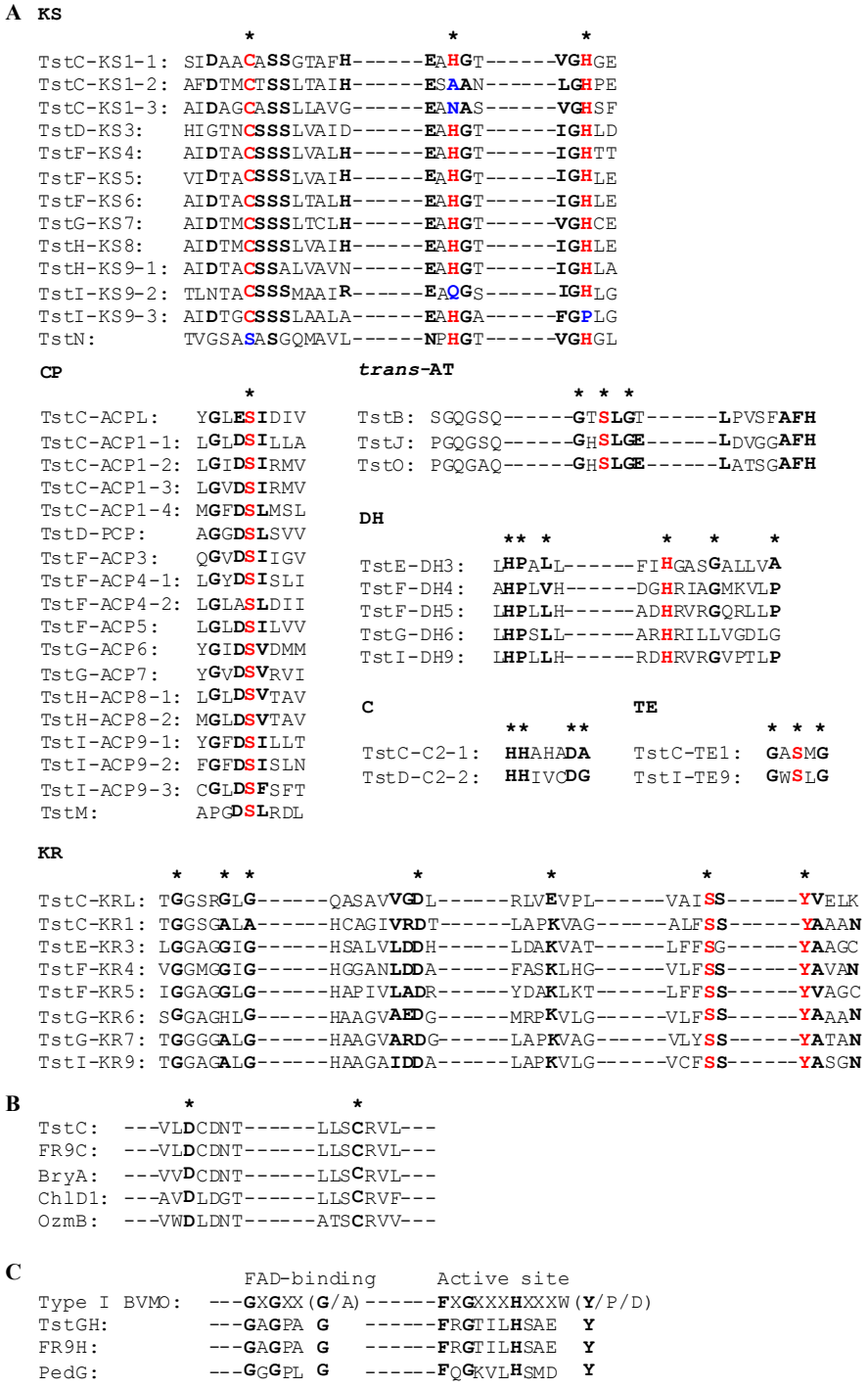


**Figure S40.** Decomposition Patterns of Thailanstatins and FR901464. HPLC profiling of thailanstatin A (TST-A) (**a**), thailanstatin B (TST-B) (**b**), thailanstatin C (TST-C) (**c**) and FR901464 (**d**) at 0 h (upper trace of each panel) and after 12 h incubation in phosphate buffer (pH 7.4) at 37ºC (lower trace of each panel). Benzoic acid ($PhCO_2H$) was included as an internal control. Note that an isocratic elution scheme was applied to TST-A whereas a gradient elution scheme was applied to other compounds. Each compound appeared to decompose into a specific product whose chemical identity requires further investigation.