Electronic Supplementary Material (ESI) for Chemical Communications

Supporting Information

Branched Peptide Boronic Acids (BPBAs): A Novel Mode of Binding Towards RNA

Wenyu Zhang, David I. Bryson, Jason B. Crumpton, Jessica Wynn, and Webster L. Santos*

Contents

1. Structure and Sequence of BPBAs in Fig. 3.

2. Experimental Section

Synthesis of Branched Peptide Boronic Acids Library. We used standard solid phase peptide synthesis techniques to generate the 3.3.4 library via the split and pool method using the previously described apparatus.[1] N-α-Fmoc protected L-amino acids (Novabiochem), PyOxim (Novabiochem) and N,N-Diisopropylethylamine (DIEA, Aldrich) were used in coupling reactions. The synthesis and full characterization of Fmoc-N-ε-(4-boronobenzoyl)-L-lysine (K_{BBA}), Fmoc-N-ε-benzoyl-L-lysine (K_{Bz}) and Fmoc-N-ε-(4-borono-3-fluorobenzoyl)-L-lysine (K_{FBA}) will be described elsewhere. Fmoc-L-4-boronophenylalanine (F_{BPA}) and Fmoc-ANP-OH were synthesized as previously reported.[2-4] Three copies of library were prepared simultaneously by using a three-fold excess of Tentagel Macrobead-NH_2 resin (2.13 g, 0.57 mmol/g) (Peptides International). The resin was swollen in DCM (20 mL, 2 x 15 min) followed by DMF (20 mL, 15 min). The photocleavable linker Fmoc-ANP-OH (739 mg, 1.71 mmol) was first coupled to the resin in DMF for 3 hr in the presence of PyOxim (901.53 mg, 1.71 mmol) and DIEA (493 µL, 2.85 mmol). After coupling, the resin was washed with DMF (20 mL, 1 min), DCM (20 mL, 1 min) and DMF (20 mL, 1 min). The same washing procedure was applied after every step. Then, 20% piperidine in DMF (20 mL, 2 x 10 min) was used for Fmoc deprotection. A Kaiser test was used after each coupling and deprotection step to confirm reaction completion. N-Fmoc amino acids (3 equiv), PyOxim (3 equiv), and DIEA (5 equiv) were added to each reaction vessel in DMF and coupled for 30 min. Fmoc-Lys(Fmoc)-OH was used as a branching unit, and molar equivalencies of reagents were doubled in coupling reactions after installation of the branching unit. After Fmoc deprotection of the N-terminal amino acids, the resin was bubbled in a phenylboronic acid solution (0.2 g/mL) overnight to remove the pinacol groups of boron-containing side chains. Finally, the resin was treated with 95:2.5:2.5 TFA (Trifluoroacetic acid, Acros)/H_2O/TIS (Triisopropylsilane, Acros) (v/v/v) for 3 hr. After deprotection, the resin was washed extensively with DMF, DCM, and MeOH before drying and storing at -20 °C.

On-bead Screening Assay. DY547 labeled HIV-1 RRE-IIB RNA (5′-DY547-GGC-UGG-UAU-GGG-CGC-AGC-GUC-AAU-GAC-GCU-GAC-GGU-ACA-GGC-CAG-CC-3′**)** (**Figure S1.**) was purchased from Dharmacon and prepared according to the manufacturer's protocol. To account for the

autofluorescence of Tentagel Macrobead-NH_2 resin, control peptide $(KYR)_2$*FDS was incubated in 100 nM DY547 labeled HIV-1 RRE-IIB RNA for 1 hr in phosphate buffer (10 mM potassium phosphate, 100 mM KCl, 1 mM $MgCl_2$, 20 mM NaCl, pH 7.0). These beads were washed extensively and placed into a sterile 96-well plate (Nunc) and imaged by a Zeiss Axiovert 200 fluorescent microscope under a rhodamine filter. The fluorescence intensity was adjusted to remove the background autofluorescence of untreated beads. Screening conditions were initially tested using about 2,400 beads from the library. The beads were placed into a 1.5 mL non-stick microfuge tube (Fisher) with a 500 µL final volume of phosphate buffer and mixed by a Barnstead/Thermolyne Labquake rotisserie shaker. The beads were first treated with 1 mg/mL bovine serum albumin (BSA) (New England BioLabs) and 0.8 mg/mL *E. coli* tRNA (Roche) (~3,144-fold molar excess to RRE IIB RNA) for 3 hr at rt to block nonspecific binding peptide sequences. Then, the beads were washed 5 times with phosphate buffer and incubated in 500 µL of 10 nM DY547 labeled RRE IIB RNA in phosphate buffer for 5 hr at rt. After the final incubation, the resin was extensively washed with buffer, and the beads were imaged under a fluorescence microscope in a 96-well plate using the previously optimized settings. The initial screening afforded 7 hits which were isolated, rinsed with DMF (5 x 500 µL) and MeOH (5 x 500 µL), and photocleaved in clear non-stick 0.5 mL microfuge tubes in 15 µL of 1:1 MeOH: H_2O (v/v) by irradiation at 365 nm with a 4W handheld UV lamp. The supernatant was retained and subjected to MALDI-TOF analysis. MALDI-MS/MS fragmentation analysis generated 6 sequences (one was deemed a false positive) providing a hit rate of 0.25%. The remaining beads of the 3.3.4 library were screened using more stringent conditions in an attempt to reduce the total number of hit beads. First the beads were blocked at rt for 3 hr in phosphate buffer with 1 mg/mL BSA, and 1.25 mg/mL tRNA (5000-fold molar excess to RRE stem IIB RNA). Next, the beads were washed in buffer and incubated with 10 nM DY547 labeled RRE-IIB RNA in phosphate buffer for 5 hr at RT. The resin was washed extensively after final incubation prior to screening. Another 5 beads (from approximately half of the total compound library) were found with elevated fluorescence using these more stringent conditions. These beads were photocleaved and sequenced using MALDI-TOF.

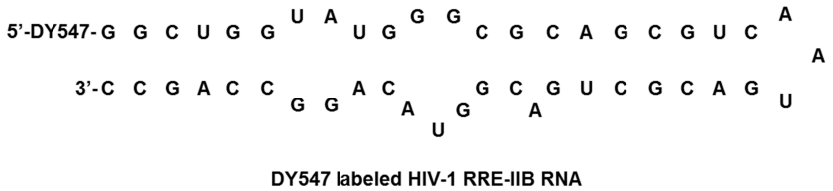


Figure S1. Secondary structure of DY547 labeled HIV-1 RRE-IIB RNA

Peptide Synthesis, Purification and Characterization. Unlabeled and fluorescein 5-isothiocyanate (FITC) (Sigma) labeled peptides were synthesized on Rink amide MBHA resin (100-200 mesh) (Novabiochem). In the preparation of the FITC-labeled peptides, Fmoc-Lys(ivDde)-OH (Novabiochem) was used as the branching unit.[5] Acetic anhydride in DMF (1:1 v/v) with 10 equivalents of DIEA was used to cap the first N-terminus. Then, ivDde was removed by treatment with 2% hydrazine in DMF for 1 hr, and the second N-terminus was synthesized through the ε-N of the Lys side chain. Fmoc-6-Ahx-OH (AnaSpec) was coupled to the N-terminal amino acid to provide a linker

for FITC, which prevents autocleavage of FITC under acidic conditions.[5, 6] All subsequent steps were protected from light. FITC (5 equiv) was reacted with the deprotected N-terminus of the peptides for 6 hr using DIEA (14 equiv). The boronic acid deprotection was performed as before. The supernatant was dried under reduced pressure, and the crude peptide was triturated from cold diethyl ether. The peptides were purified using a Jupiter 4 μm Proteo 90 Å semiprep column (Phenomenex) using a solvent gradient composed of 0.1% TFA in Milli-Q water and HPLC grade acetonitrile. Peptide purity was determined using a Jupiter 4 μm Proteo 90 Å analytical column (Phenomenex), and peptide identity was confirmed by MALDI-TOF analysis. Unlabeled peptide concentrations were measured in nuclease free water at 280 nm using their calculated extinction coefficients. FITC-labeled peptide concentrations were monitored at 495 nm using the extinction coefficient of FITC at 77,000 mol^{-1} cm^{-1} in 100 mM glycine, pH 9.0.

Preparation of ^{32}P-labeled RNA and DNA. Wild-type a RRE-IIB RNA was transcribed *in vitro* by T7 polymerase with the Ribomax T7 Express System (Promega) using previously reported techniques.[7, 8] The antisense template (5'-GGCTGGCCTGTACCGTCAGCGTCATTGACGCTGCG CCCATACCAGCCCTATAGTGAGTCGTATTACAT-3'), sense complementary strand (5'-ATGTAAT ACGACTCACT ATAGG-3') and RRE IIB reverse PCR primer (5'-GGCTGGCCTGTAC-3') were purchased from Integrated DNA Technologies. RRE IIB was PCR amplified using HotstarTaq DNA polymerase (Qiagen) followed by a clean-up procedure using a spin column kit (Qiagen). The antisense DNA template was annealed with the sense DNA complementary strand in reaction buffer at 95 °C for 2 min then cooled on ice for 4 min. T7 transcription proceeded at 42 °C for 1.5 hr. After transcription, DNA template was degraded with DNase at 37 °C for 45 min and the RNA was purified by a 12% polyacrylamide gel containing 7.5 M urea. The band corresponding to the RNA of interest was excised from the gel and eluted overnight in 1x TBE buffer at 4 °C. The sample was desalted using a Sep-Pak syringe cartridge (Waters Corporation) and lyophilized. The product was dephosphorylated with calf intestinal phosphatase (CIP) in NEBuffer 3 (New England Biolabs) according to manufacturer's protocol. The product was recovered by a standard phenol extraction followed by ethanol precipitation. Purified RNA was stored as a pellet at -80 °C. HIV-1 RRE IIB RNA was labeled at the 5'-end by treating 10 pmol of dephosphorylated RNA / DNA with 20 pmol of [γ-^{32}P] ATP (111 TBq mol^{-1}) and 20 units of T4 polynucleotide kinase in 70 mM Tris•HCl, 10 mM $MgCl_2$, and 5 mM dithiothreitol, pH 7.6. The mixture was incubated at 37 °C for 30 min, and then at rt for 20 min. The kinase was heat-inactivated at 65 °C for 10 min. The product was recovered by ethanol precipitation, and the purity was examined using 12% denaturing PAGE followed by autoradiography.

Dot Blot Assay. Dot blot assays were performed at rt using a Whatman Minifold I 96 well Dot Blot system and Whatman 0.45 μm pore size Protran nitrocellulose membranes. To determine the binding affinities, 0.4 nM radiolabeled RNA was titrated with peptide (0.001–100 μM). First, a solution of 0.8 nM ^{32}P-labeled RNA was refolded in 2x phosphate buffer (20 mM potassium phosphate, 200mM KCl, 1mM $MgCl_2$, 40mM NaCl, pH 7.0) by heating at 95 °C for 3 min and then slowly cooling at rt for 20 min. Next, 25 μL of the [^{32}P]-RNA solution was added to 25 μL of peptide in nuclease free water and incubated at rt for 4 hr. The 50 μL mixtures were filtered through the nitrocellulose membrane, which was immediately followed by two consecutive 50 μL washes with 1x phosphate buffer. Peptide binding was visualized by autoradiography using a storage phosphor screen (GE Healthcare) and a Typhoon Trio phosphorimager (GE Healthcare). Densitometry measurements were quantified using

ImageQuant TL (Amersham Biosciences). Binding curves were generated using a four parameter logistic equation with Kaleidagraph (Synergy Software): y=m1+(m2-m1)/(1+10^(log(m3)-x); m1=100; m2=1; m3=.000003, where y=percentage of RNA binding, x= log[peptide], m1=percentage of RNA binding affinity at infinite concentration (nonspecific binding), m2= percentage of RNA binding affinity at zero concentration, m3=peptide concentration at 50% binding (K_d). Each experiment was performed in triplicate and error bars represent the standard deviation calculated over three replicates.

3. Bead images for determining the optimal filter settings

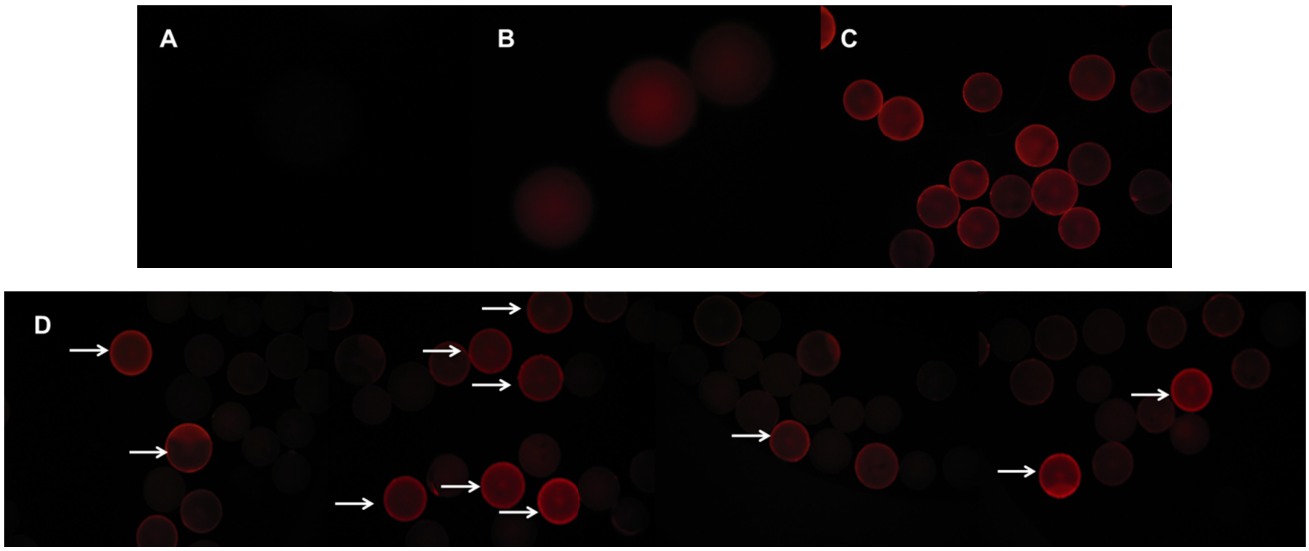


Figure S2. Bead images for determining the optimal filter settings. (A) Tentagel Macrobead-NH_2 resin prior to peptide coupling visualized with a fluorescence microscope using rhodamine filter, (B) Branched peptide boronic acid derivatized macrobeads, (C) Positive control peptide $(KYR)_2$*FDS with 100nM RRE incubated in phosphate buffer for 1 hour, washed and visualized under fluorescence microscope, and (D) eleven hits from optimized screening of the 3.3.4 branched boronic acid peptide library.

4. Isolated hit beads

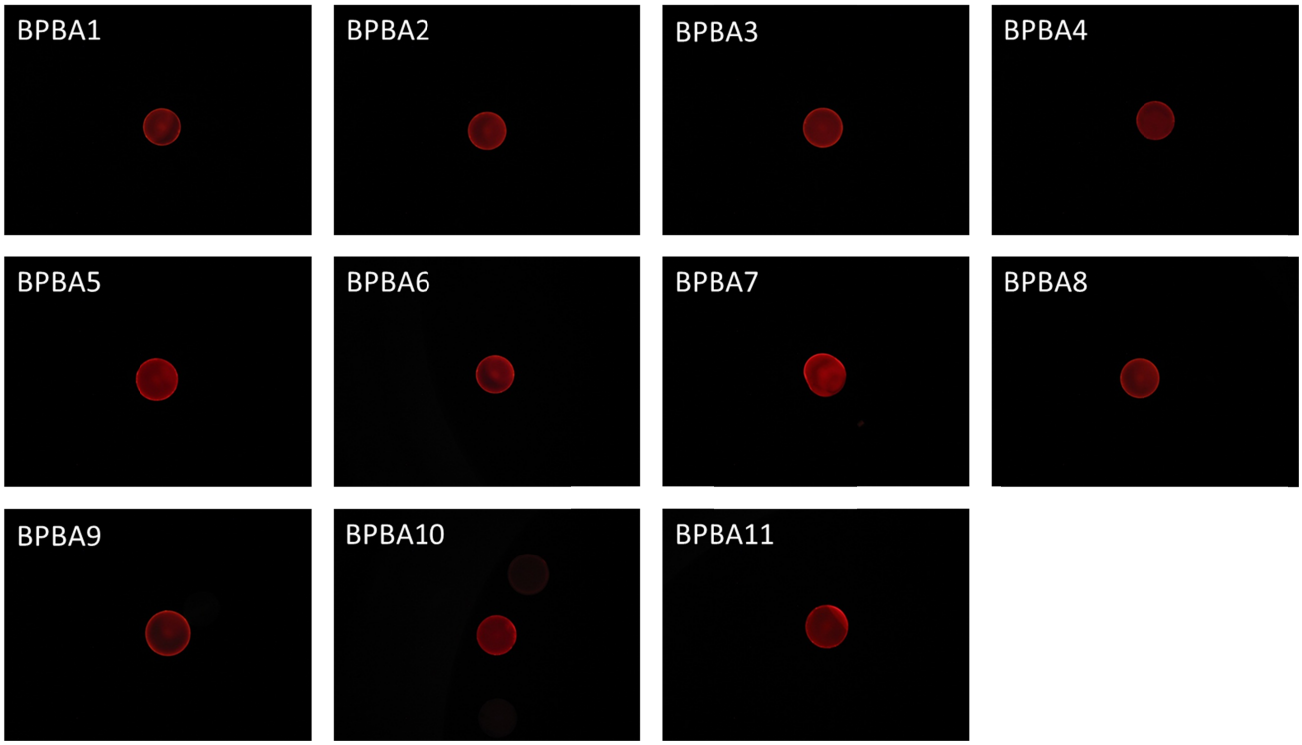


Figure S3. Images of isolated hits beads under fluorescence microscope using a rhodamine filter.

5. EMSA of BPBA1

EMSA were performed at room temperature. First, 4 nM ^{32}P labeled RNA in 2x phosphate buffer (20 mM potassium phosphate, 200mM KCl, 1mM $MgCl_2$, 40mM NaCl, pH 7.0) was refolded by heating at 95 °C for 3 min and cooling at room temperature for 20 min. A10 µL solution of the refolded RNA was added to 10 µL of peptide in nuclease free water and incubated at room temperature for 4 hours. The final concentration of peptide was varied from 0.001 to 100 µM. After incubation, 3 µL of 30% glycerol was added for loading. Peptide-RNA complexes were resolved on 10% non-denaturing PAGE, which had been pre-run for at least 1 hour. Gels were electrophoresed at 150 V for 35 min at room temperature. Gels were dried to filter paper and visualized by autoradiography. Each experiment was repeated 3 times. Data was measured as the percentage of bound RNA in each lane, and error bars represent the standard deviation calculated over three replicates.

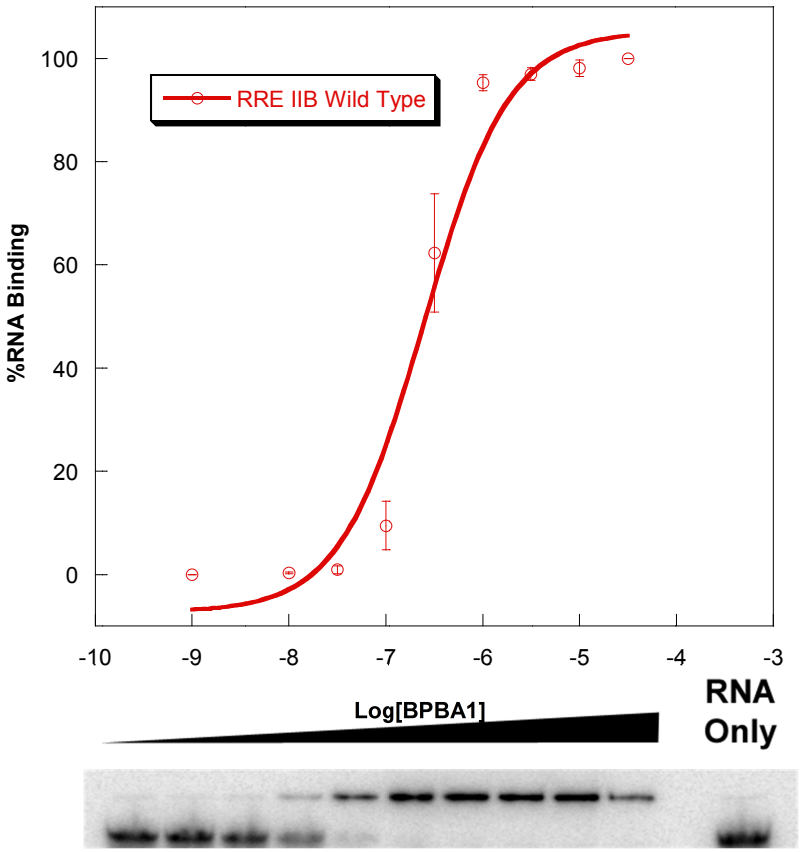
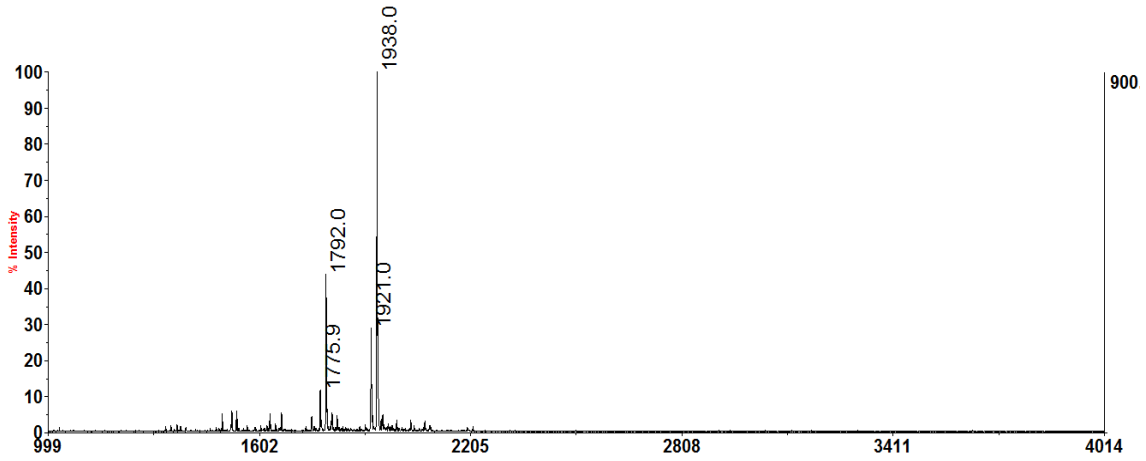


Figure S4. Binding curve and EMSA result (K_d=0.25 ± 0.08 µM) of **BPBA1** with RRE IIB RNA.

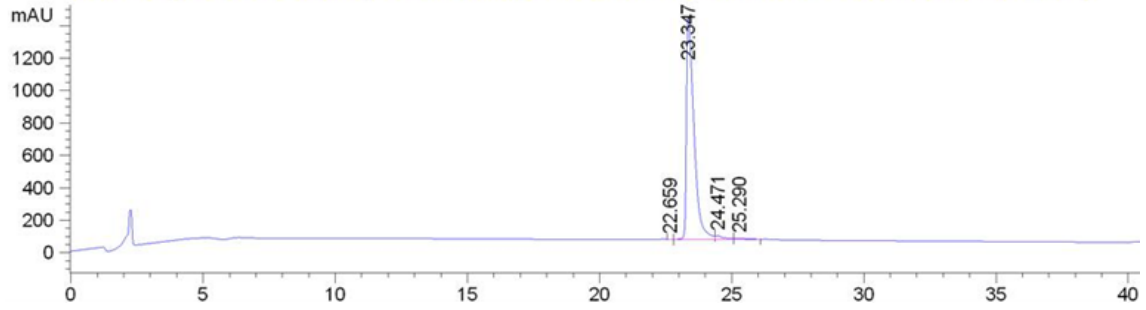
6. Branched Peptide Boronic Acid HPLC Condition, Purity, and Identification by MALDI/MS

HPLC Solvent A: 0.1% TFA in Milli-Q water, Solvent B: HPLC grade acetonitrile. Peptide purity was determined using a Jupiter 4 µm Proteo 90 Å analytical column (Phenomenex), and peptide identity was confirmed by MALDI-TOF analysis.

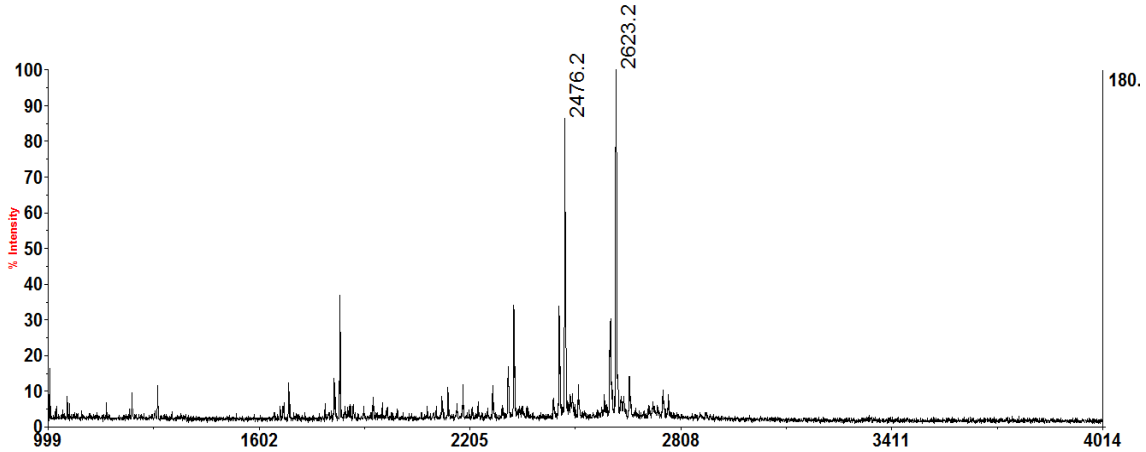
BPBA1 Predicted Mass: 1937.0, Mass Detected: 1938.0



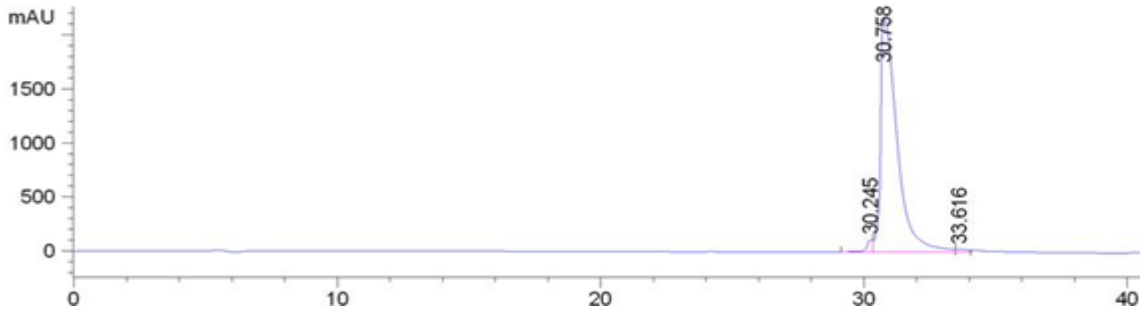
The Purity (96%) was determined by analytical HPLC employing a linear gradient of 12%-30% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



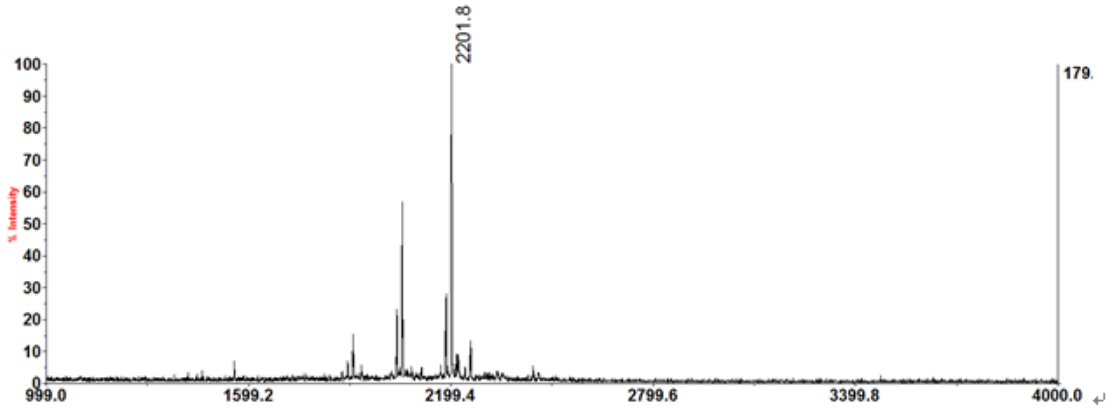
BPBA2 Predicted Mass: 2622.1, Mass Detected: 2623.2



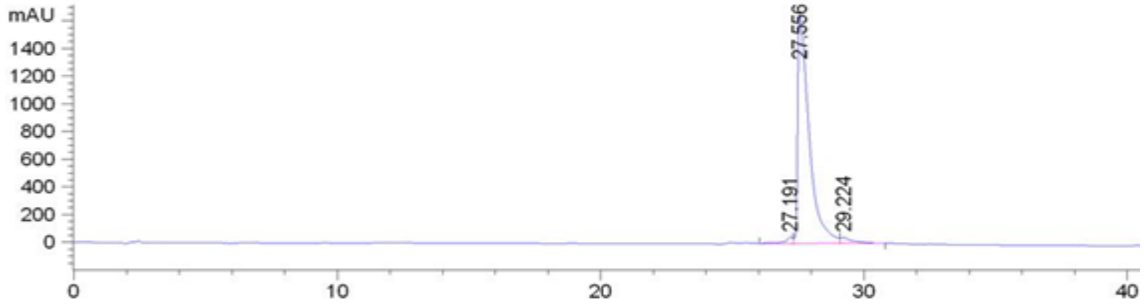
The Purity (97%) was determined by analytical HPLC employing a linear gradient of 7%-20% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



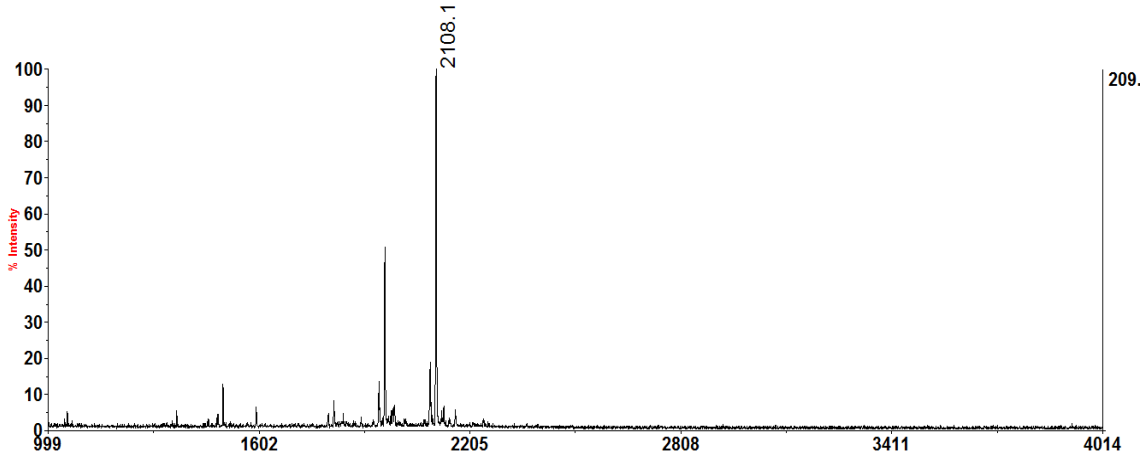
BPBA3 Predicted Mass: 2200.8, Mass Detected: 2201.8



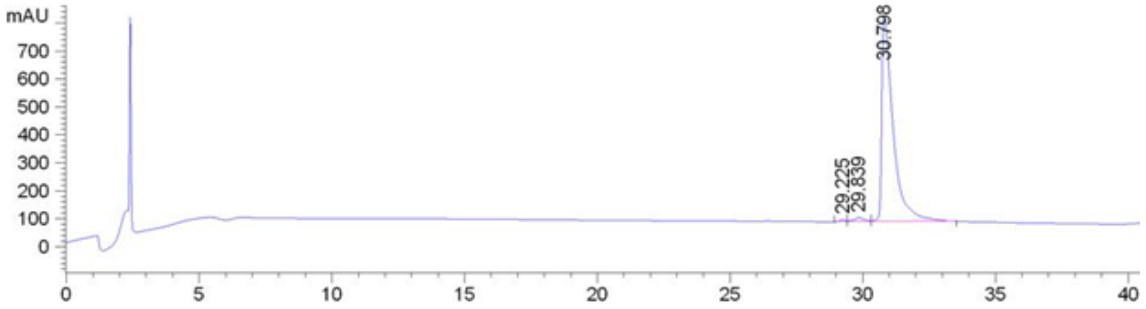
The Purity (95%) was determined by analytical HPLC employing a linear gradient of 10%-30% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



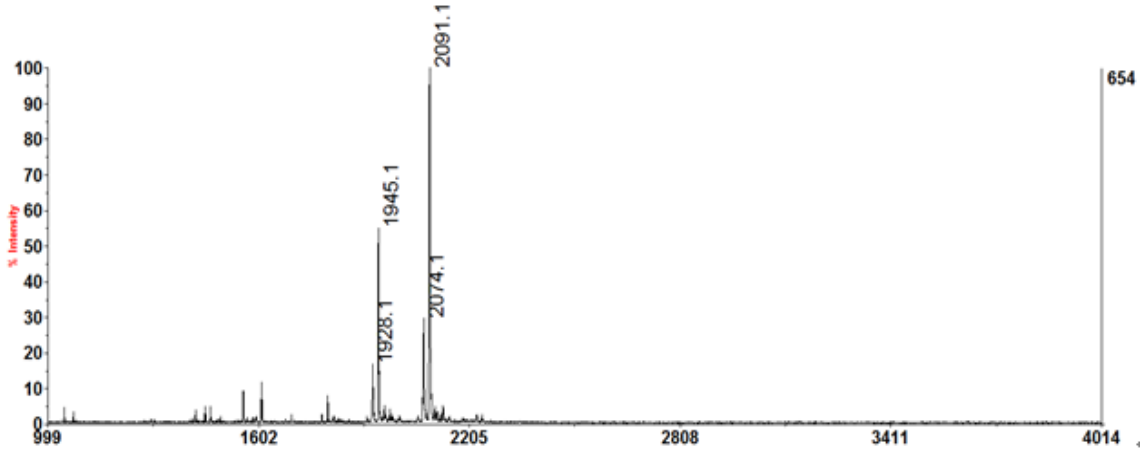
BPBA4 Predicted Mass: 2107.0, Mass Detected: 2108.1



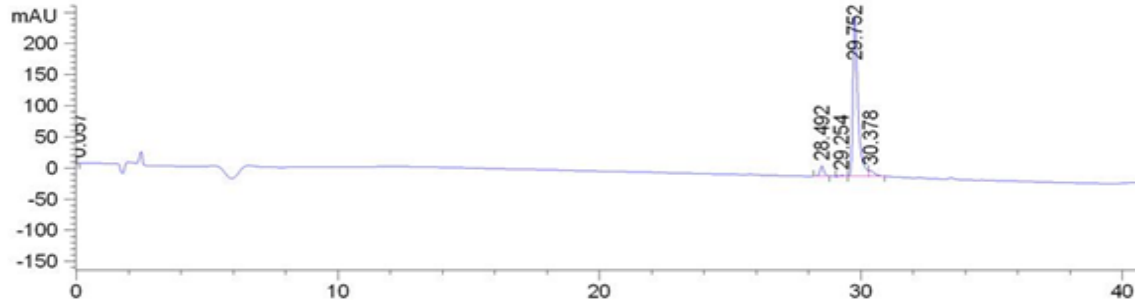
The Purity (97%) was determined by analytical HPLC employing a linear gradient of 5%-20% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



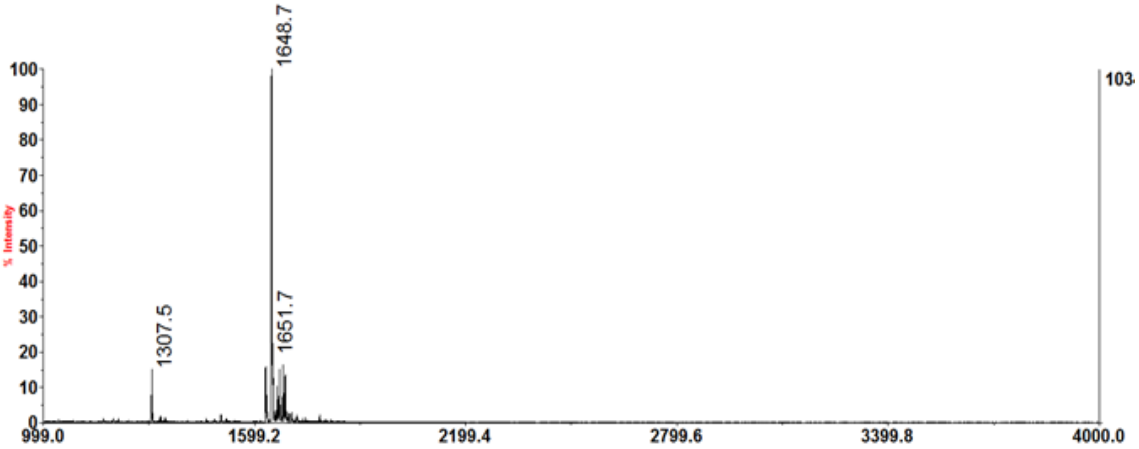
BPBA5 Predicted Mass: 2090.1, Mass Detected: 2090.1



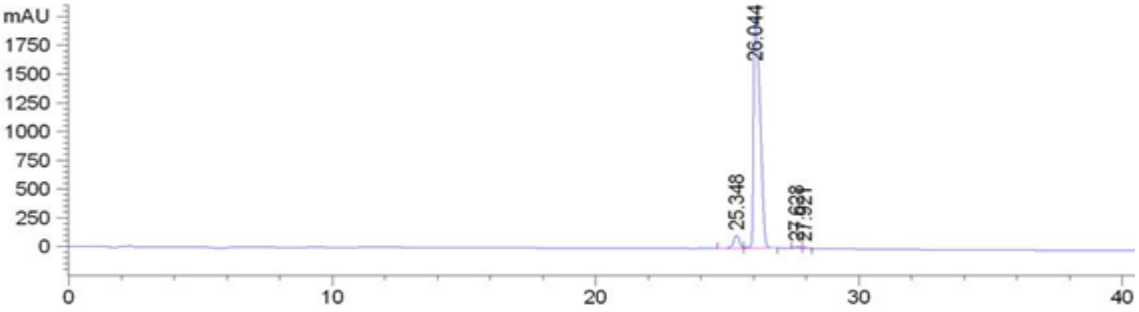
The Purity (90%) was determined by analytical HPLC employing a linear gradient of 8%-28% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



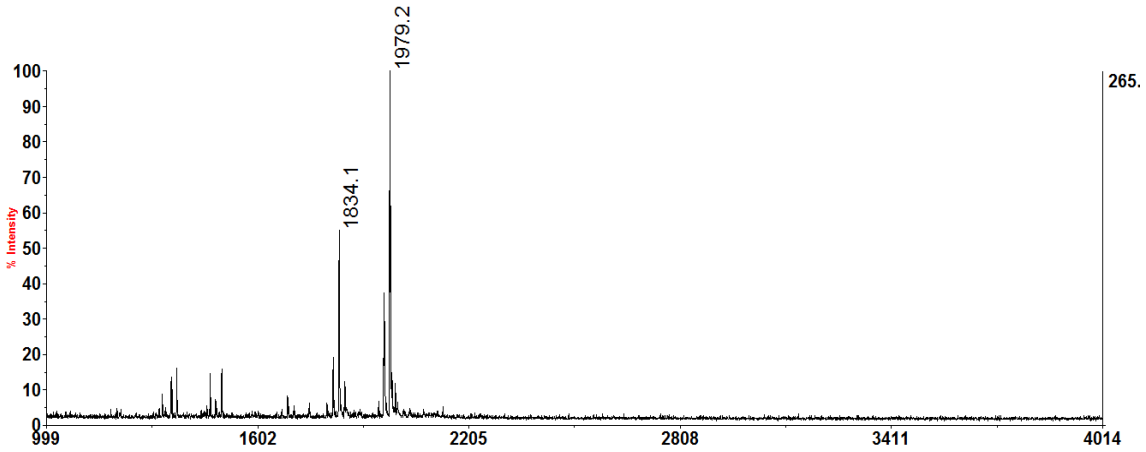
BPBA6 Predicted Mass: 1648.8, Mass Detected: 1648.7



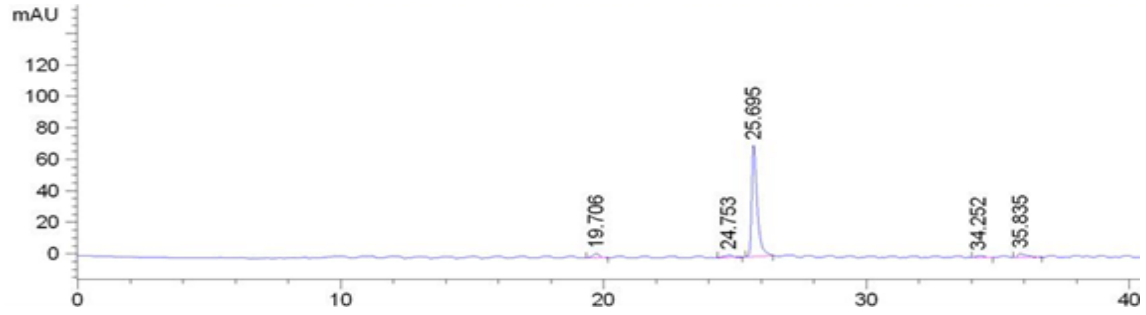
The Purity (94%) was determined by analytical HPLC employing a linear gradient of 14%-34% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



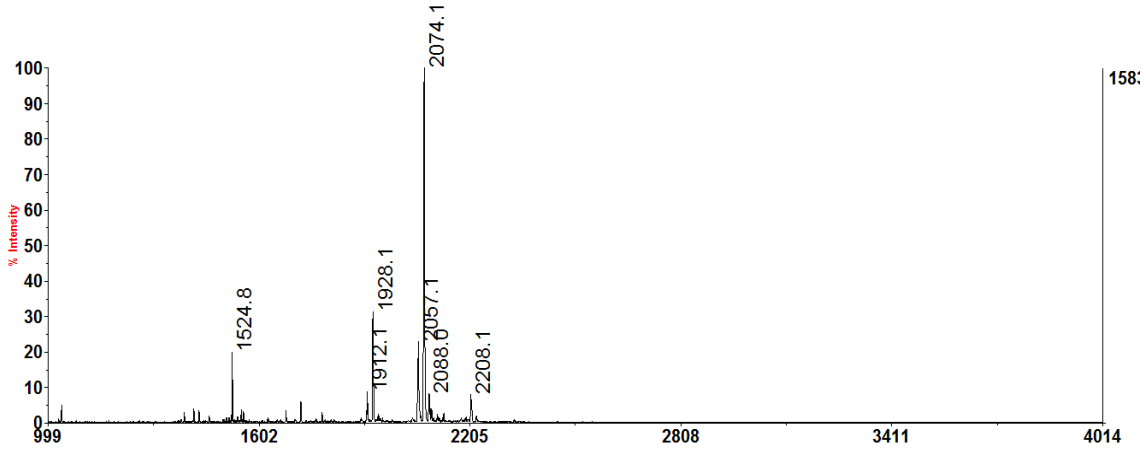
BPBA7 Predicted Mass: 1979.1, Mass Detected: 1979.2



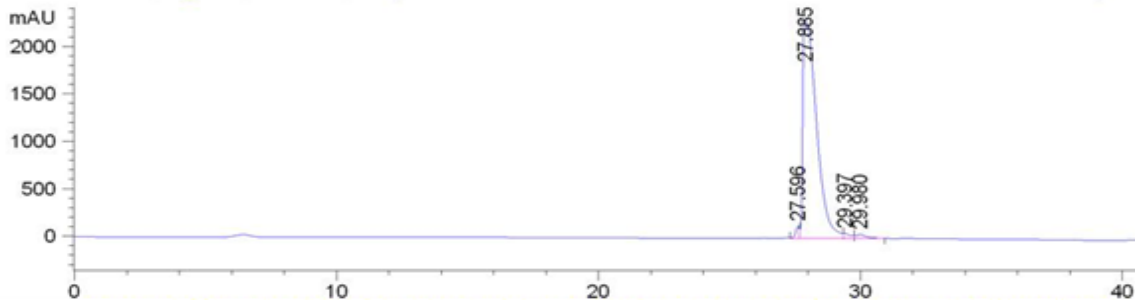
The Purity (92%) was determined by analytical HPLC employing a linear gradient of 7%-27% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



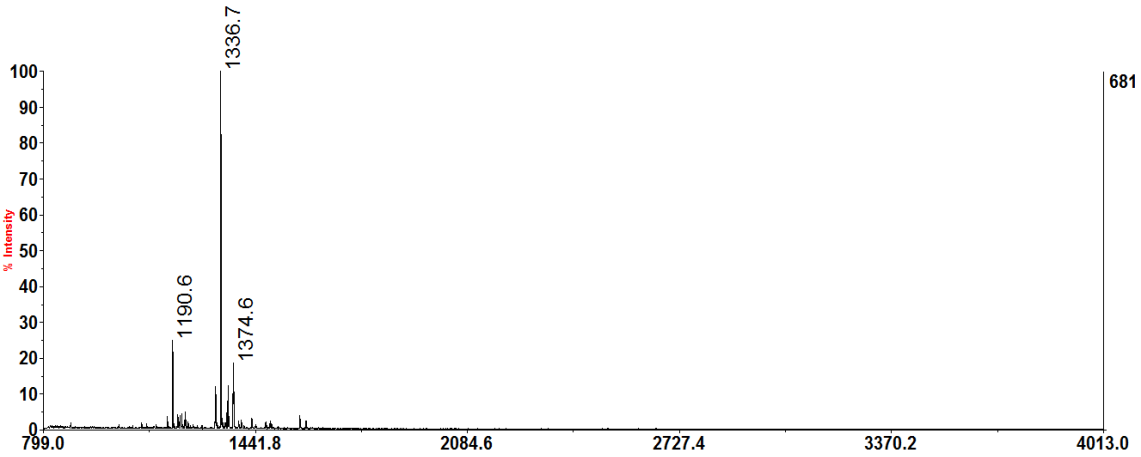
BPBA8 Predicted Mass: 2073.0, Mass Detected: 2074.1



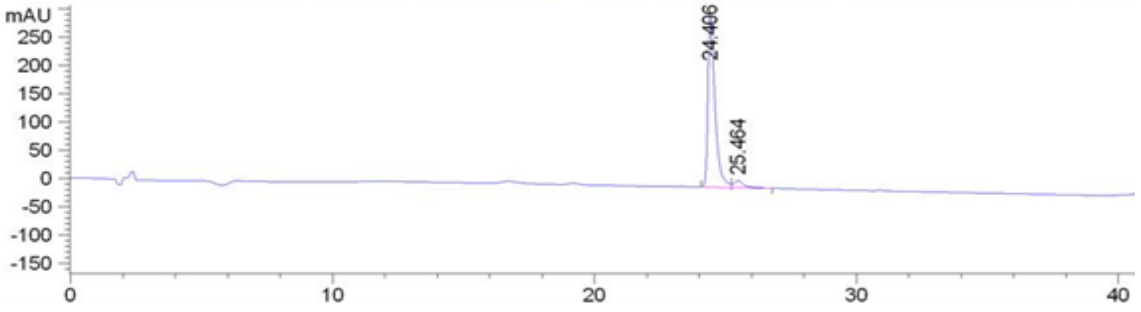
The Purity (95%) was determined by analytical HPLC employing a linear gradient of 5%-20% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



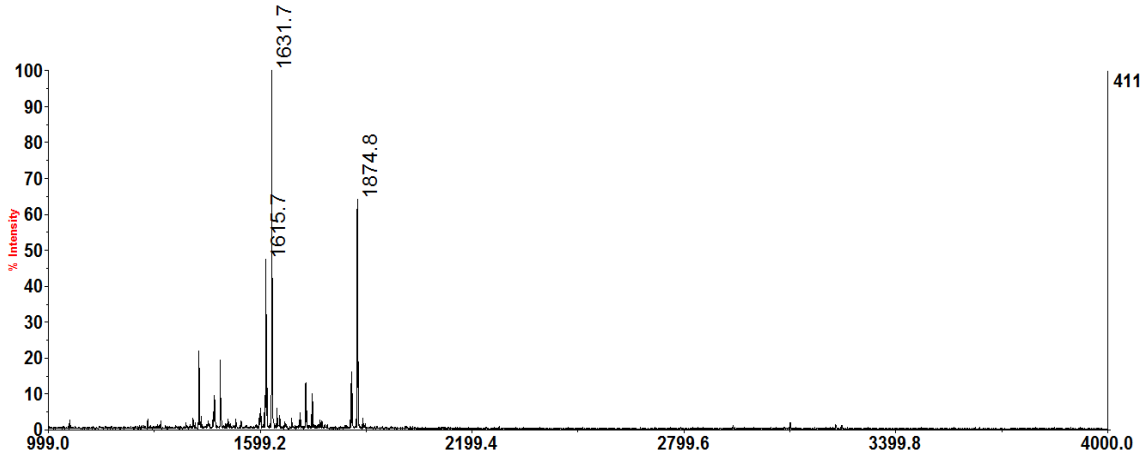
BPBA9 Predicted Mass: 1335.7, Mass Detected: 1336.7



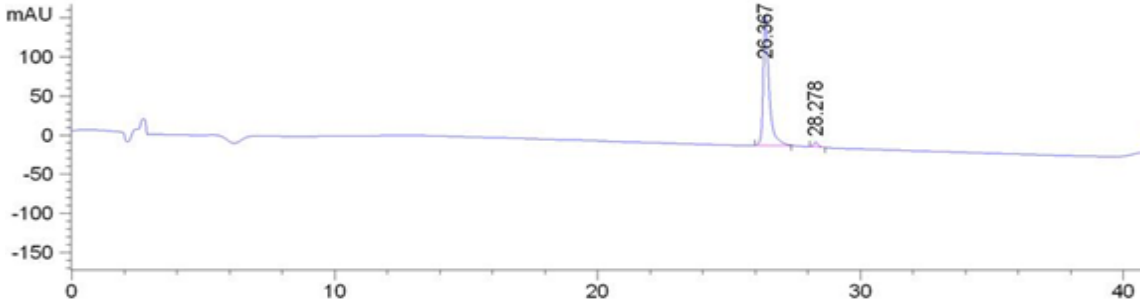
The Purity (91%) was determined by analytical HPLC employing a linear gradient of 12%-30% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



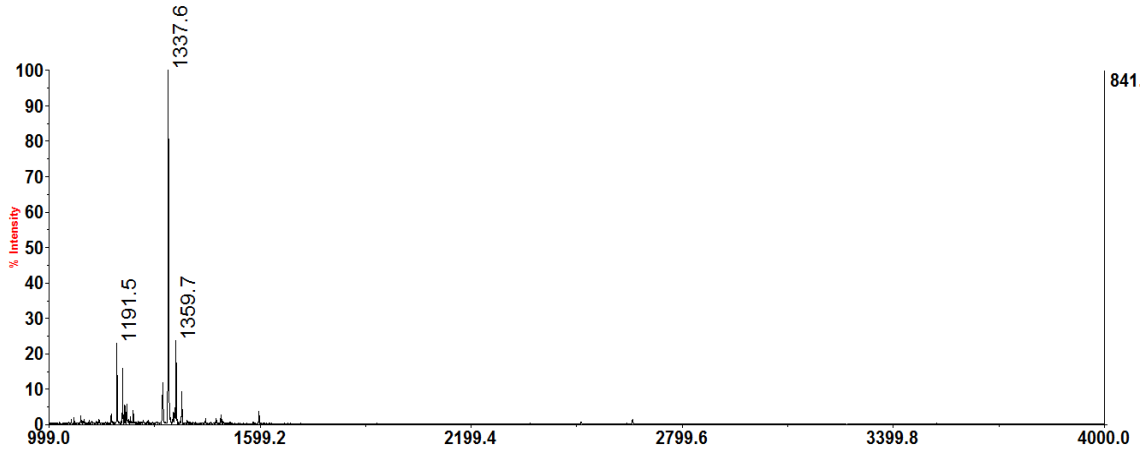
BPBA10 Predicted Mass: 1631.8, Mass Detected: 1631.7



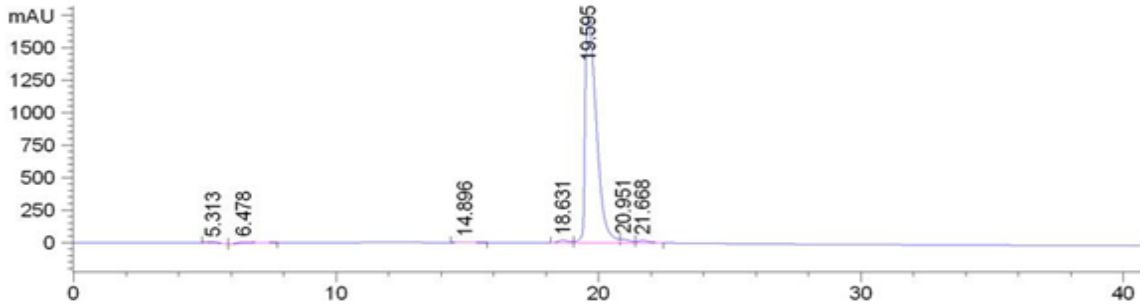
The Purity (97%) was determined by analytical HPLC employing a linear gradient of 2%-22% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



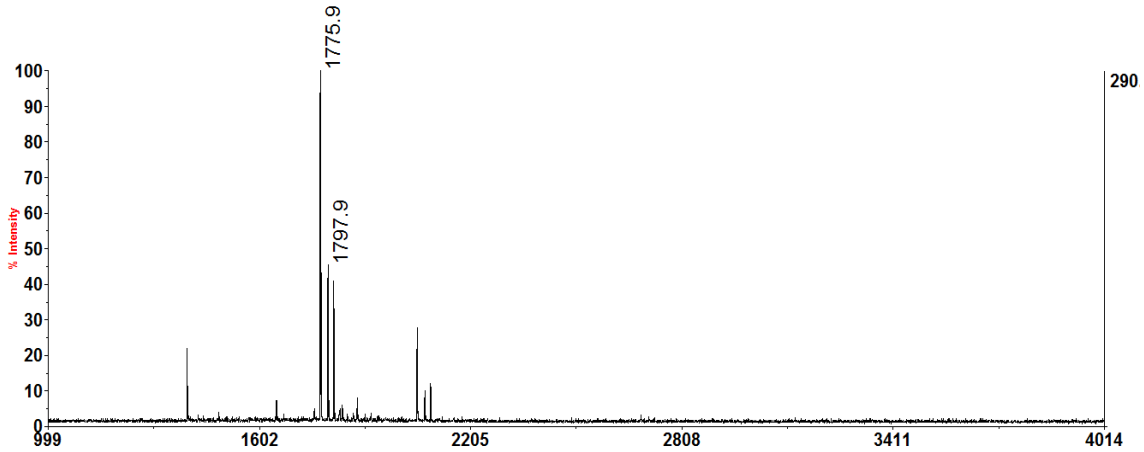
BPBA11 Predicted Mass: 1337.7, Mass Detected: 1337.6



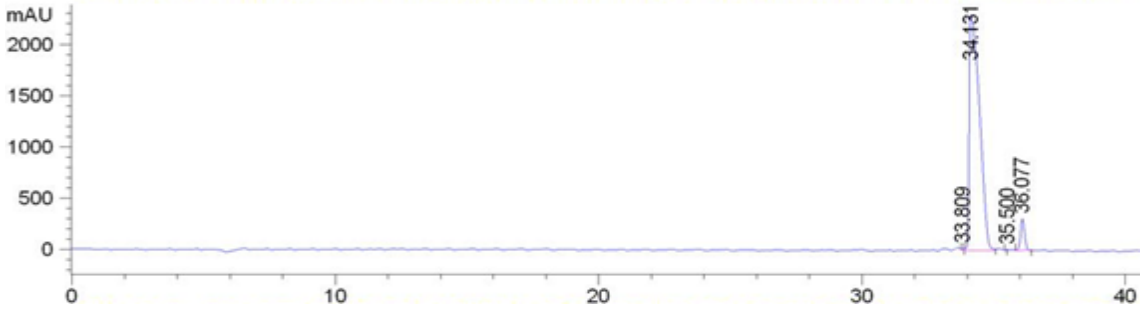
The Purity (94%) was determined by analytical HPLC employing a linear gradient of 10%-30% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



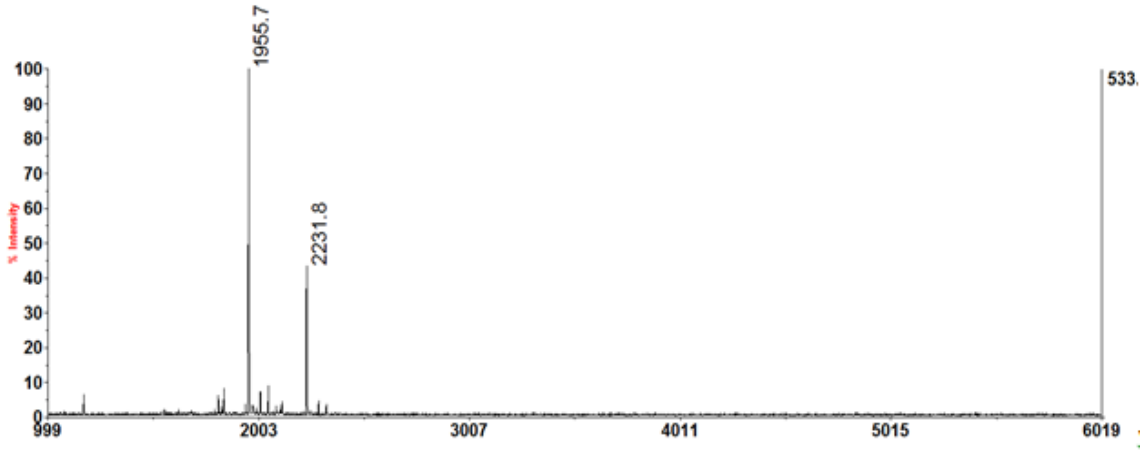
BPBA1.1 Predicted Mass: 1774.9, Mass Detected: 1775.9



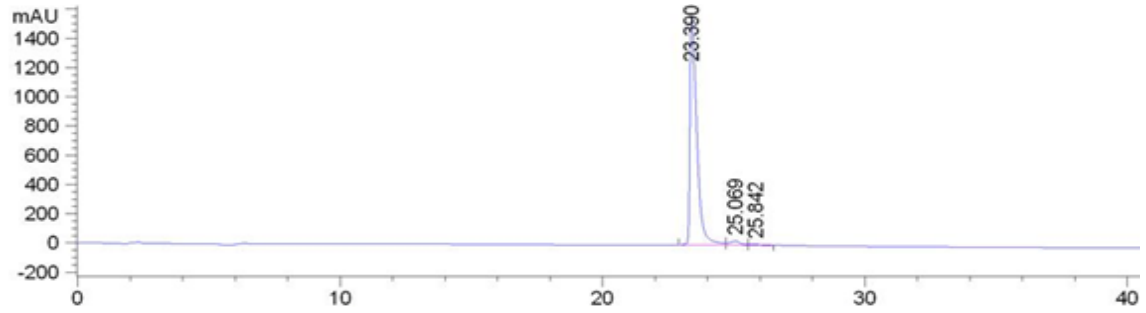
The Purity (94%) was determined by analytical HPLC employing a linear gradient of 12%-32% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



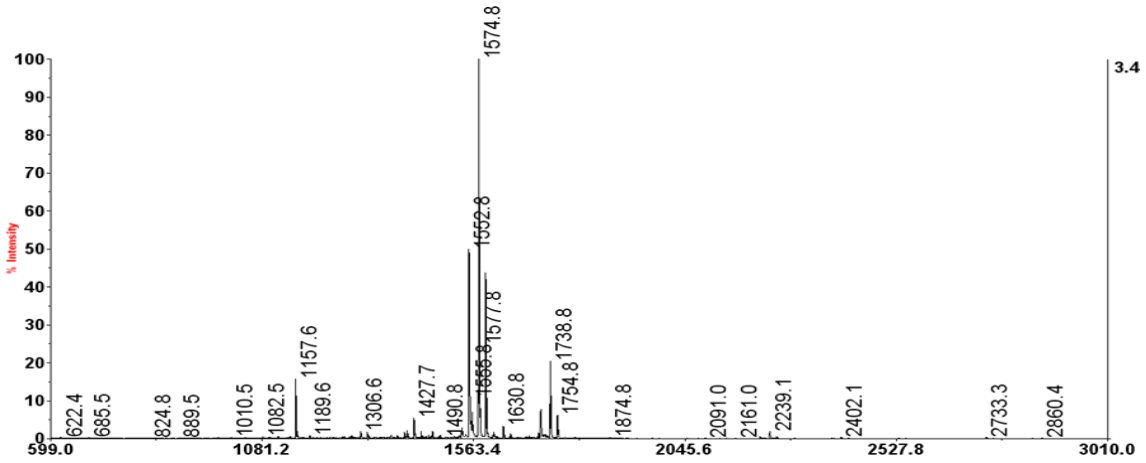
BPBA1.2 Predicted Mass: 1955.0, Mass Detected: 1955.7



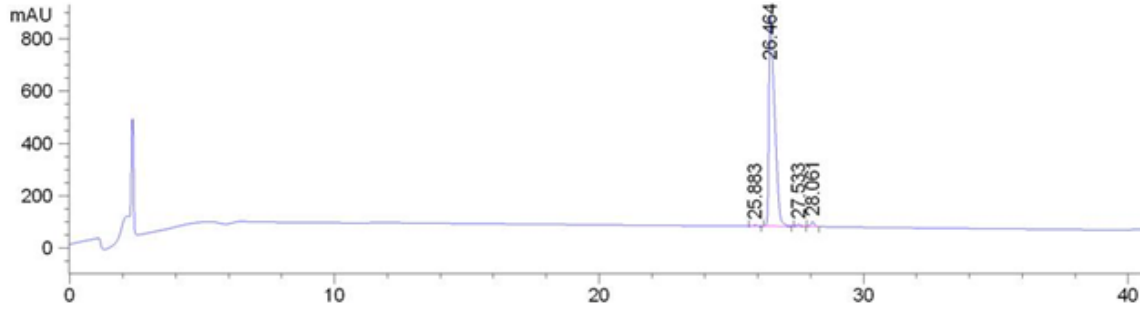
The Purity (96%) was determined by analytical HPLC employing a linear gradient of 14%-34% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



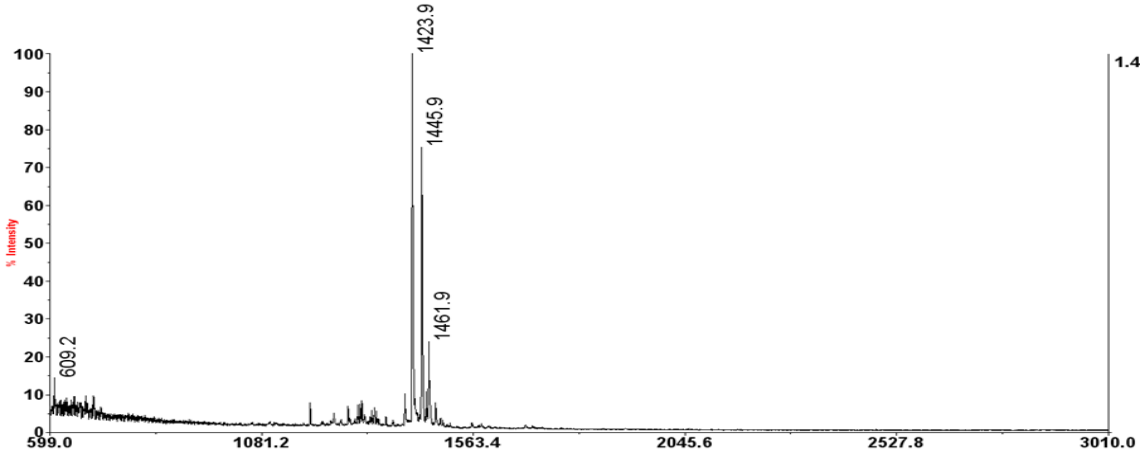
BPBA3.1 Predicted Mass: 1551.8, Mass Detected: 1552.8



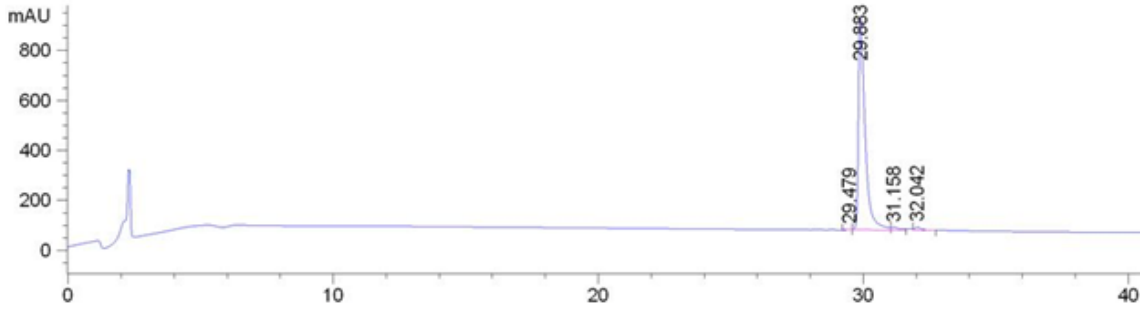
The Purity (97%) was determined by analytical HPLC employing a linear gradient of 7%-27% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



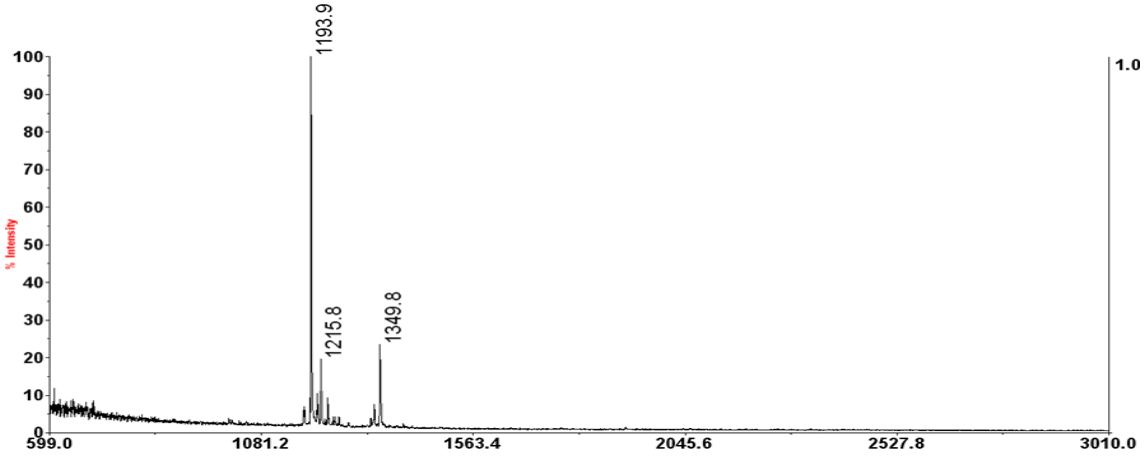
LPBA1 Predicted Mass: 1423.3, Mass Detected: 1423.8



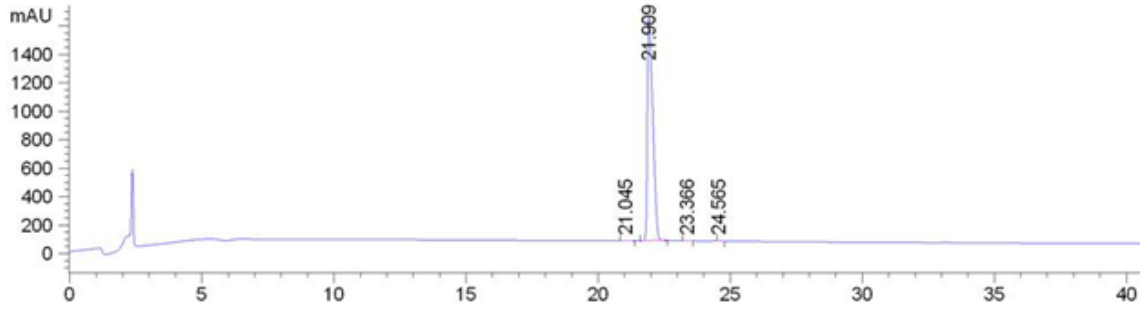
The Purity (97%) was determined by analytical HPLC employing a linear gradient of 12%-30% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



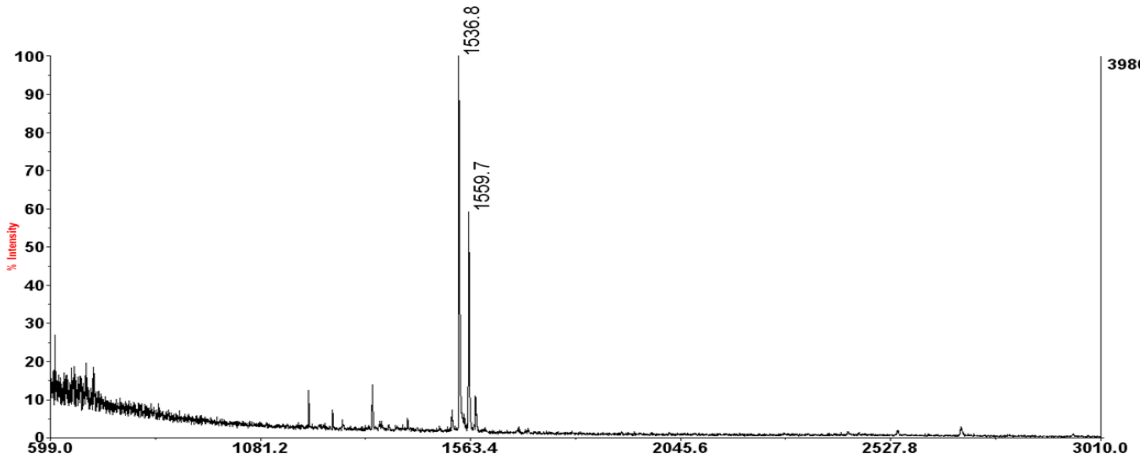
LPBA2 Predicted Mass: 1193.5, Mass Detected: 1193.8



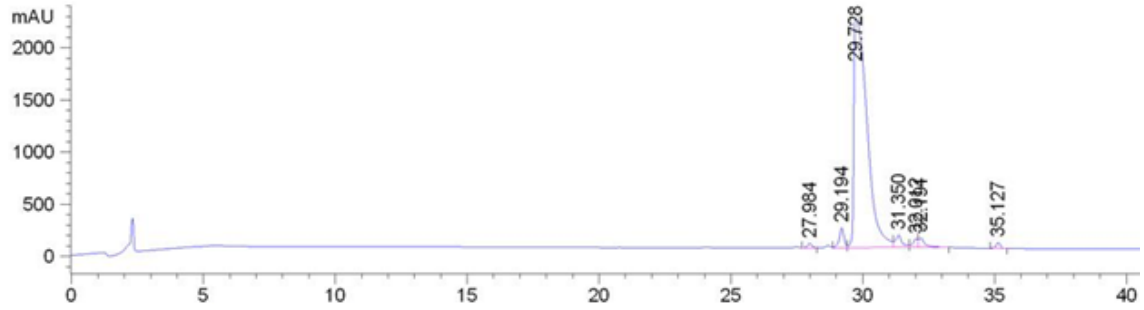
The Purity (99%) was determined by analytical HPLC employing a linear gradient of 8%-28% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



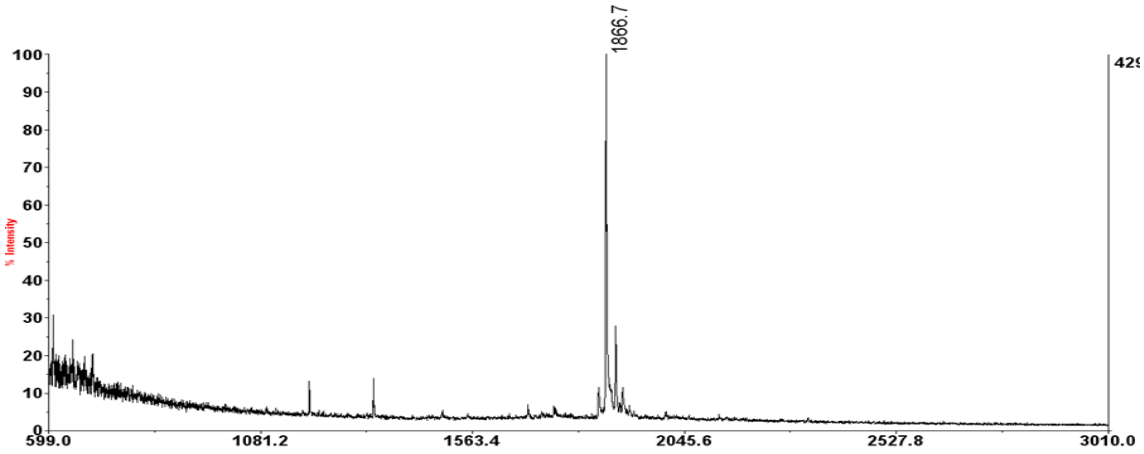
LPBA3 Predicted Mass: 1535.7, Mass Detected: 1536.8



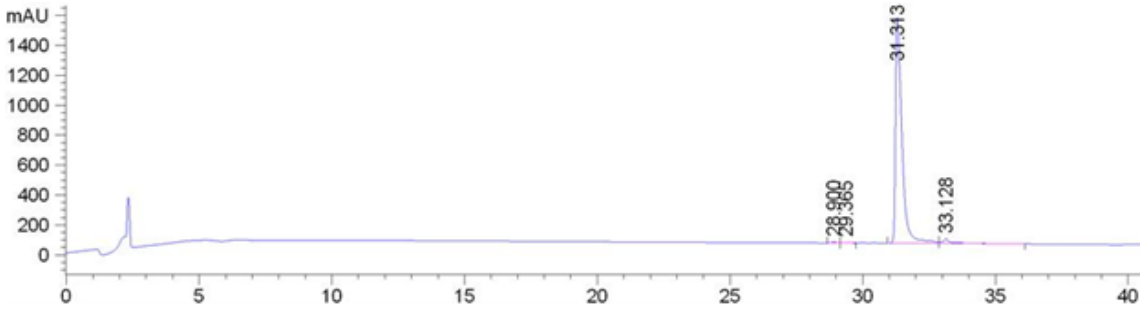
The Purity (90%) was determined by analytical HPLC employing a linear gradient of 10%-30% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



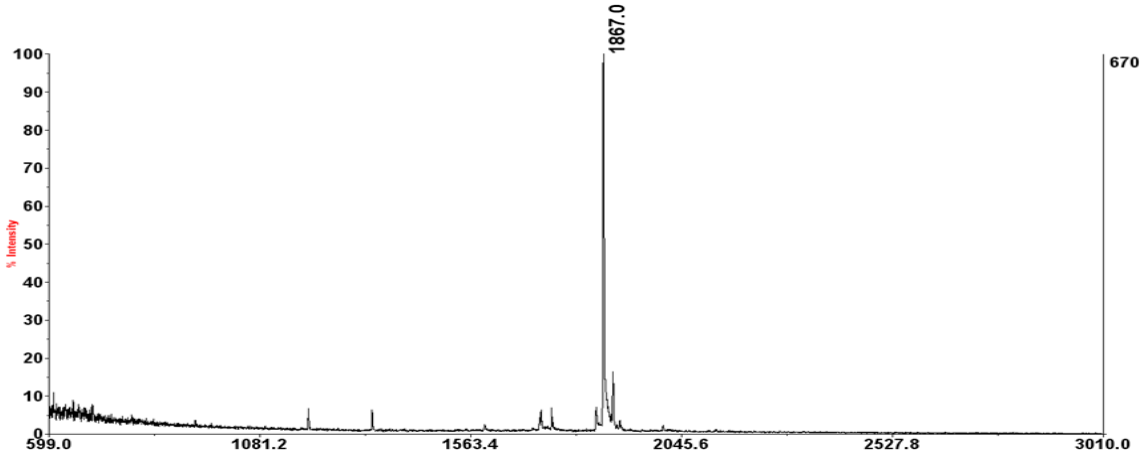
LPBA4 Predicted Mass: 1865.9, Mass Detected: 1866.6



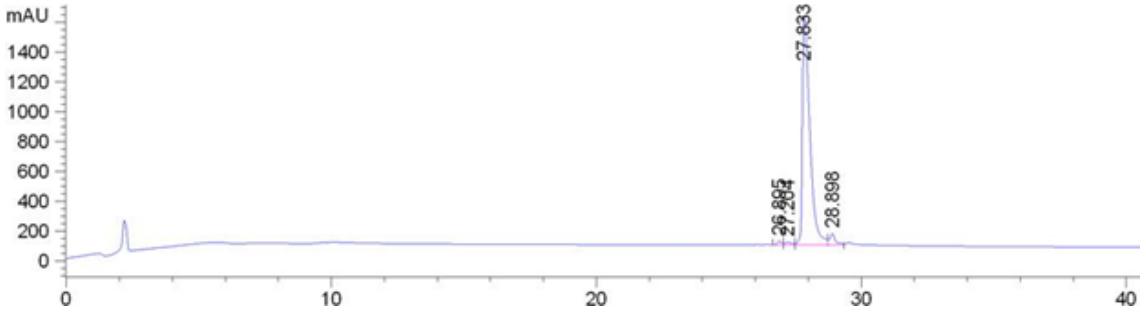
The Purity (96%) was determined by analytical HPLC employing a linear gradient of 10%-30% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



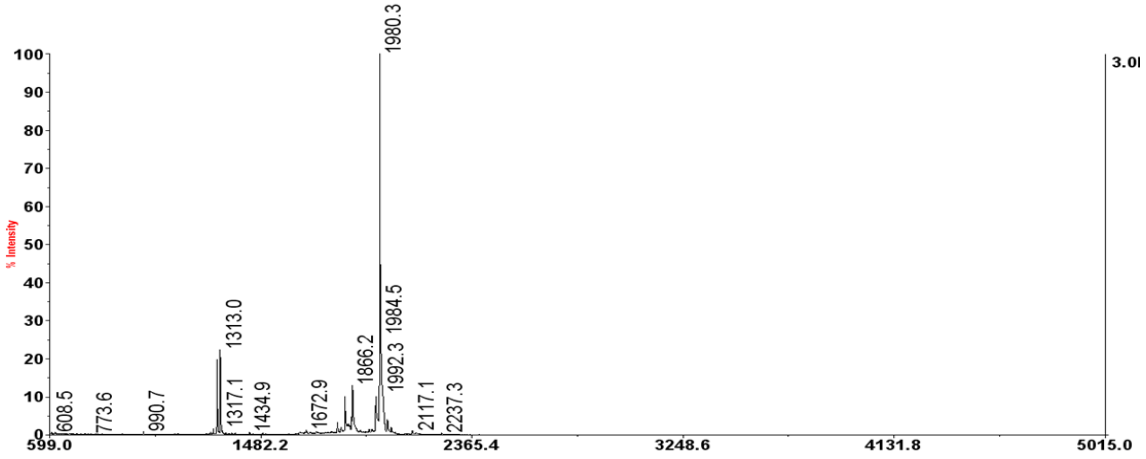
LPBA5 Predicted Mass: 1865.9, Mass Detected: 1867.0



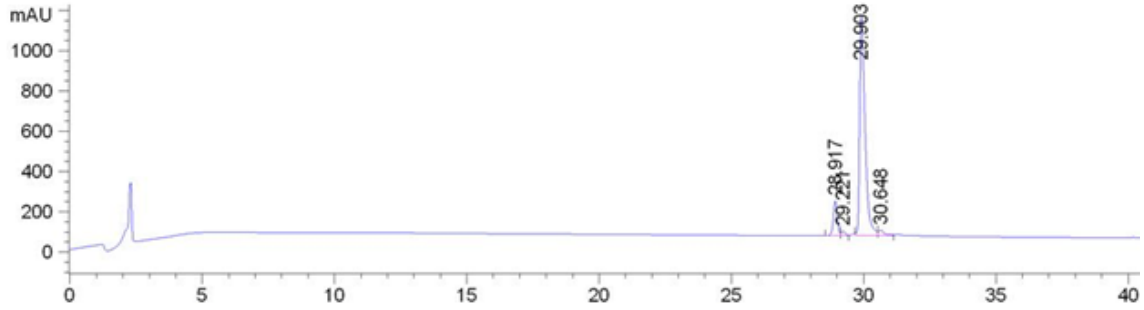
The Purity (94%) was determined by analytical HPLC employing a linear gradient of 12%-30% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



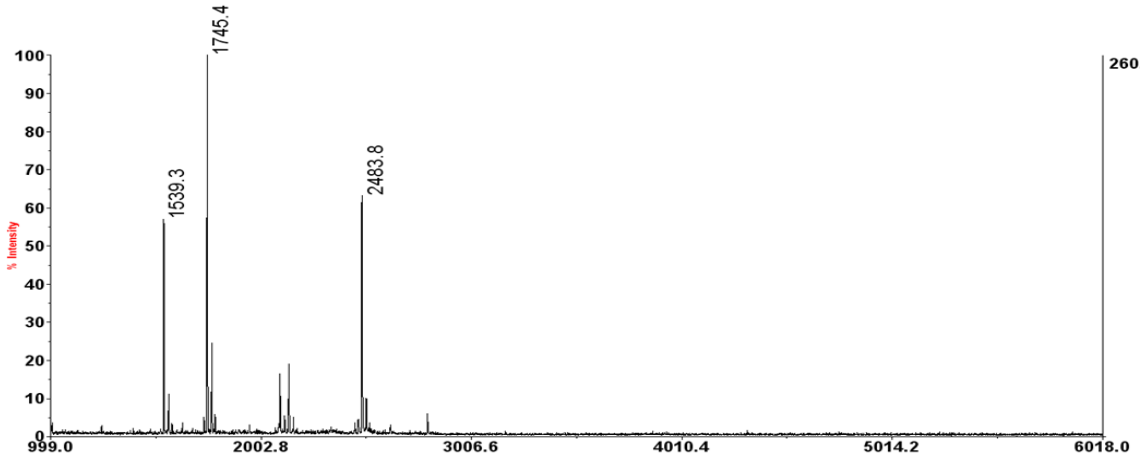
RPBA1 Predicted Mass: 1979.1, Mass Detected: 1980.2



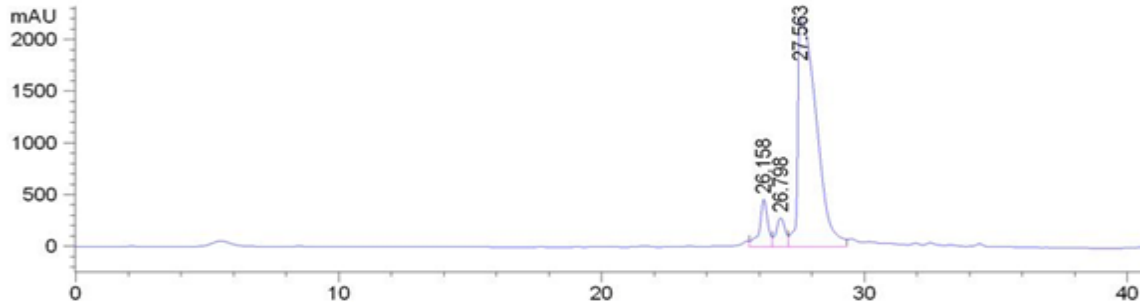
The Purity (90%) was determined by analytical HPLC employing a linear gradient of 12%-30% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



FPBA1 Predicted Mass: 2481.6, Mass Detected: 2483.8



The Purity (88%) was determined by analytical HPLC employing a linear gradient of 23%-35% B in A in 30min at a flow rate of 0.2mL/min following a wash procedure.



References

1 D. I. Bryson, W. Zhang, W. K. Ray and W. L. Santos, *Mol. BioSyst.*, 2009, **5**, 1070-1073.

2 C. Malan and C. Morin, *J. Org. Chem.*, 1998, **63**, 8019-8020.

3 P. J. Duggan and D. A. Offermann, *Aust. J. Chem.*, 2007, **60**, 829-834.

4 D. S. Tan, M. A. Foley, B. R. Stockwell, M. D. Shair and S. L. Schreiber, *J. Am. Chem. Soc.*, 1999, **121**, 9073-9087.

5 D. I. Bryson, W. Zhang, P. M. McLendon, T. M. Reineke and W. L. Santos, *ACS. Chem. Biol.*, 2012, **7**, 210-217.

6 M. Jullian, A. Hernandez, A. Maurras, K. Puget, M. Amblard, J. Martinez and G. Subra, *Tetrahedron Lett.*, 2009, **50**, 260-263.

7 D. I. Bryson, W. Zhang, P. M. McLendon, T. M. Reineke and W. L. Santos, *ACS Chem. Biol.*, 2012, **7**, 210-217.

8 J. F. Milligan, D. R. Groebe, G. W. Witherell and O. C. Uhlenbeck, *Nucleic Acids Res.*, 1987, **15**, 8783-8798.