Structure-Activity Relationship Studies of Miniproteins Targeting the Androgen Receptor-Coactivator Interaction

Marta Dominguez Seoane,[b] Katja Petkau-Milroy,[a,b] Belen Vaz,[b] Sabine Möcklinghoff,[a,b] Simon Folkertsma,[c] Lech-Gustav Milroy,[a] and Luc Brunsveld,*[a,b]

[a] *Laboratory of Chemical Biology, Department of Biomedical Engineering, Technische Universiteit Eindhoven, Den Dolech , 5612 AZ Eindhoven (The Netherlands)*

[b] *Chemical Genomics Centre of the Max Planck Society, Otto-Hahn Straße 15, 44227 Dortmund (Germany)*

[c] *Computational Drug Discovery, Centre for Molecular and Biomolecular Informatics, Radboud University, Nijmegen (The Netherlands)*

Telephone: int+31(0)40 247 8367

FAX: int+31(0)40 247 3737

* e-mail: l.brunsveld@tue.nl

SUPPORTING INFORMATION

Table of contents

General

Rink Amide MBHA resin LL (100-200 mesh) with loading of 0.34 mmol/g was purchased from Novabiochem. Fmoc- and side chain-protected amino acids were purchased from Fluka, Biosolve and Novabiochem. All other reagents were purchased from Aldrich-Sigma, Fluka and Acros.

All automated peptide syntheses were performed on a Syro II automated peptide synthesizer from MultiSynTech GmbH, using standard *Fmoc*-chemistry. LC-ESI-MS was carried out by using an Agilent 1100 series binary pump together with a reversed phase HPLC column (Macherey-Nagel) and a Finnigan Thermoquest LCQ. The following gradient program was used for analytical LC-MS: flow: 1 mL/min, solvent A: 0.1% HCO_2H in H_2O, solvent B: 0.1% HCO_2H in CH_3CN, A/B: 90/10 (0-1 min) to 0/100 (over 10 min), 0/100 (12 min). Purification of products by RP-HPLC was performed in an Agilent 1100 Series Purification Platform using a NUCLEODUR® C18 Gravity preparative column from Macherey-Nagel (21 x 250 mm) and flow rate of 25 mL/min. The products were eluted using different solvent gradients of solvents A and B (solvent A = 0.1% TFA/H_2O; solvent B = 0.1% TFA/CH_3CN). UV signal at 210 nm was used for detection. All spectral data for competitive fluorescence polarization assays were acquired using a Tecan SAFIRE II system.

Miniprotein library design

Using Yasara, the mutated miniproteins (κ-hefutoxin[1], Om-toxin[2]) were overlaid with the AR-LBD surface in complex with a known cofactor motif[3] to identify a suitable fit, especially with respect to the positioning of the FXXLF motif and to observe the overall fit of the other parts of the miniprotein and the introduced mutations into the helix binding groove.

Table S1. Molecular weight and K_i values[4] of κ-hefutoxin, Om-toxin and derivatives **Het-2** and **Omt-1**. For **Het-2**, the FXXLF motif was introduced at positions 6, 9, and 10 (blue). For **Omt-1**, the FXXLF motif was introduced at positions 3, 6 and 7 (blue). n.a. = no activity.

Name	Sequence	M_w	K_i (μM)[4]
Om-toxin	N D P C E E V C I Q H T G D V K A C E E A C Q	2521.7	n.a.
Omt-1	N R **F** C E **L F** C I Q G T G D V K A C E E A C Q	2564.9	1.3 ± 0.2
κ-hefutoxin	G H A C Y R N C W R E G N D E E T C K E R C	2654.8	n.a.
Het-2	A C Y **F** N C **L F** E G N D E E T C K E R C	2369.6	1.6 ± 0.3

Table S2. Om-toxin miniprotein library. Mutations are red. The FXXLF motif is shown in bold. $F^{(oCl)}$ = 2-chloro-phenylalanine; $F^{(pCl)}$ = 4-chloro-phenylalanine; C_{ha}= β-cyclohexylalanine.

Name	Sequence	Yield / mg	Yield / %	Mass calculated $(M + 2H)^{2+}$	Found mass $(M + 2H)^{2+}$
Omt-1a	N R **F** C E **L F** C I Q G T G D V K A C E W A C Q	4.1	2.3	1309.57	1309.20
Omt-1b	N R **F** C E **L F** C I Q D T G D V K A C E W A C Q	4.3	1.6	1338.58	1338.20
Omt-1c	N R **F** C E **L F** C I E G T G D V K A C E W A C Q	11.5	3.2	1310.05	1309.73
Omt-1d	K R **F** C E **L F** C I Q G T G D V K A C E W A C Q	4.9	4.9	1316.58	1316.33
Omt-1e	N R **L** C E **L L** C I Q G T G D V K A C E W A C Q	8.5	4.9	1275.59	1275.27
Omt-1f	N **L F** C E **L F** C I Q G T G D V K A C E W A C Q	2.6	1.8	1288.05	1287.67
Omt-1g	S R **F** C E **L F** C I Q G T G D V K A C E W A C Q	3.1	1.8	1296.05	1296.13
Omt-1h	N R **F** C K **L F** C I Q G T G D V K A C E W A C Q	2	1.2	1309.08	1308.80
Omt-1i	N R **F** C E **M F** C I Q G T G D V K A C E W A C Q	1.9	2.0	1318.54	1318.27
Omt-1j	N R **F** C E **L F** C F Q G T G D V K A C E W A C Q	12.1	4.8	1326.55	1326.13
Omt-1k	N R **$F^{(pCl)}$** C E **L F** C I Q G T G D V K A C E W A C Q	0.9	1.5	1326.54	1326.80
Omt-1l	N R **F** C E **L $F^{(pCl)}$** C I Q G T G D V K A C E W A C Q	3.2	3.5	1326.54	1326.54
Omt-1m	N R **$F^{(oCl)}$** C E **L F** C I Q G T G D V K A C E W A C Q	1.7	0.9	1326.54	1326.73
Omt-1n	N R **F** C E **L $F^{(oCl)}$** C I Q G T G D V K A C E W A C Q	3.4	1.9	1326.54	1326.40
Omt-1o	N R **C^{ha}** C E **L F** C I Q G T G D V K A C E W A C Q	1.7	2.0	1312.58	1312.27
Omt-1p	N R **F** C E **L C^{ha}** C I Q G T G D V K A C E W A C Q	2.2	2.6	1312.58	1312.33
Omt-1q	N R **F** C E **L F** C I Q G T E D V K A C E W A C Q	8.1	3.0	1345.57	1345.33

Table S3. *κ*-hefutoxin miniprotein library. Mutations are red. The FxxLF motif is shown in bold. F_o^{Cl} = 2-chloro-phenylalanine; F_p^{Cl} = 4-chloro-phenylalanine; C^{ha}= β-cyclohexylalanine.

Name	Sequence	Yield / mg	Yield / %	Mass calculated $(M + 2H)^+$	Found mass $(M + 2H)^+$
Het-2a	A C Y **F** N C **L F** E G N D E E T C K E W C	27.4	16.8	1200.94	1200.60
Het-2b	A C L **F** N C **L F** E G N D E E T C K E W C	10	6.3	1175.95	1175.60
Het-2c	A C R **F** N C **L F** E G N D E E T C K E W C	24.1	14.8	1197.46	1197.13
Het-2d	A C Y **L** N C **L L** E G N D E E T C K E W C	6.3	4	1166.95	1166.60
Het-2e	A C Y **F** N C **L F** E G E D E E T C K E W C	12.5	7.6	1208.45	1208.13
Het-2f	A C Y **F** V C **L F** E G N D E E T C K E W C	5.2	3.2	1193.45	1193.07
Het-2g	A C Y **F** F C **L F** E G N D E E T C K E W C	8.7	5.3	1217.45	1217.07
Het-2h	A C Y **F** H C **L F** E G N D E E T C K E W C	28.6	16.4	1212.45	1212.13
Het-2i	A C Y **F** E C **L F** E G N D E E T C K E W C	3.1	1.9	1208.44	1208.07
Het-2j	A C Y **F**$^{(pCl)}$ N C **L F** E G N D E E T C K E W C	3.2	1.9	1217.92	1218.00
Het-2k	A C Y **F** N C **L F**$^{(pCl)}$ E G N D E E T C K E W C	0.5	0.3	1217.92	1217.53
Het-2l	A C Y **F** N C **L F**$^{(oCl)}$ E G N D E E T C K E W C	4.6	2.8	1217.92	1218.00
Het-2m	A C Y **C$_{ha}$** N C **L F** E G N D E E T C K E W C	5.7	3.5	1203.96	1203.67
Het-2n	A C Y **F** N C **L C$_{ha}$** E G N D E E T C K E W C	11.2	6.8	1203.96	1203.60

Peptide synthesis

All peptides were synthesized from C- to N-terminus on solid-phase using an automatic solid-phase peptide synthesizer (SyroXP, Multisyntech). Rink amide resin was used as the solid support. In each case, 200 mg of Fmoc-protected Rink amide MBHA resin (LL, 100-200 mesh, NovaBiochem, resin loading of 0.34 mmol/g) was used per reactor (10 mL plastic syringes equipped with a frit column plate). The coupling of amino acids was performed using standard 9-Fluorenylmethoxycarbonyl (Fmoc) chemistry[5] and HOBt/DIC amino acid activation.[6]

The following amino acid protective group strategy was used: Gln, Asn, Ser, Thr, Tyr(tBu); Trp, Lys(Boc); Arg(Pbf) and Asp, Cys, Glu, His(Trt).

Fmoc-protected amino acids (BioSolve, Fluka, NovaBiochem) were dissolved in a 0.3 M *N,N'*-dimethyl formamide (DMF) solution of HOBt to form a 0.3 M amino acid and HOBt solution. Due to reasons of poor solubility, amino acids Phe, Gly, Tyr and 4-Cl-Phe were instead dissolved in *N*-methylpyrolidine (NMP).

5 mL of DMF was added per reactor to swell the resin. After 20 min the solvent was removed via vacuum filtration. After swelling, resin deprotection was performed using a piperidine/DMF mixture

to cleave the Fmoc-protecting group, which was carried out as follows: addition of 2 mL of 40 % piperidine/DMF (v/v), gentle shaking for 3 min, solvent removal via vacuum filtration; addition of 2 mL of 20% piperidine/DMF (v/v), gentle shaking for 10 min, solvent removal via vacuum filtration; addition of 2.2 mL DMF (washing step), gentle shaking for 1 min and solvent removal via vacuum filtration; repetition of the last washing step five times.

To elongate the peptide chain, the following coupling conditions were used: addition of 900 µL of the 0.3 M Fmoc-protected amino acid/HOBt stock solution (see above); then addition of 1 mL of 0.3 M solution of DIC in DMF, followed by gentle shaking for 50 min and then solvent removal via vacuum filtration; addition of 2.7 mL DMF (washing step) and gentle shaking for 1 min before solvent removal; repetition of the washing step three times.

For the coupling of each individual amino acid, the Fmoc-deprotection and amino acid coupling steps were repeated as above except that for the coupling of non-natural amino acids, a longer reaction time was used at the amino acid coupling step (80 min instead of 50 min).

Before peptide cleavage, the resin was washed five times with dichloromethane (DCM) and five times with diethyl ether, and then dried in vacuo for at least 1 h. Cleavage from the resin and concomitant side chain deprotection were performed using 5 mL per reactor of a solution containing TFA, H_2O, EDT, and TIS (94:2.5:2.5:1, v/v/v/v) by shaking for 3 h at rt. The peptide solution was filtered from the resin and the resin washed three times with 2 mL of TFA. TFA was coevaporated with toluene to a volume of 1 mL. Next, the peptides were precipitated with 30 mL cold diethyl ether. After centrifugation, the supernatant was discarded and the pellet was resuspended in 30 mL diethyl ether and centrifuged again. This procedure was repeated three times. The supernatant was then dissolved in a mixture of water and acetonitrile (MeCN) and lyophilized in readiness for oxidization and desalting prior to purification by reverse-phase HPLC.

Miniprotein folding and purification

The crude lyophilates were dissolved in a 2:1-mixture of phosphate buffer (pH 8) and trifluoroethanol (TFE). TFE was used as it is known to support formation of helices in peptides. Then crude products were subjected to reduction of randomly formed disulfide bridges using Tris-[2-carboxyethyl]phosphine hydrochloride (TCEP) and subsequently, after addition of 10 vol% of dimethylsulfoxide, oxidized by exposure to air for 24 - 48 h. The oxidation was followed by LC-MS. When the oxidation was complete, TFE was removed by evaporation and the aqueous residue was lyophilized. The crude product was dissolved in water and desalted using small SepPak® Vac C18 columns. After subsequent lyophilization the peptides were purified by reversed phase - high performance liquid chromatography (RP-HPLC). Purities were determined by inspection of UV spectrum of the LC-MS measured at 210 nm. The peptides were obtained in >90% purity and the yields varied between 1-20 % (for selected LC-MS spectra see Figure S 18 - Figure S 29 in the appendix).

Protein expression and purification

The cloning, expression and purification of the human androgen receptor - ligand binding domain (hAR-LBD) is described in reference 4.

Competitive fluorescence polarization assay

For the inhibition studies of the androgen receptor a high affinity peptide obtained by phage display by Fletterick *et. al*[3] was used as reference. For the synthesis of this fluorescein labeled reference peptide (Fluorescein-CSSRFESLFAGEKESR, **1**) see reference 4. The ability of the synthesized miniproteins to displace the fluorescein labeled reference peptide **1** from the dihydrotestosterone - AR-LBD complex was assessed using fluorescence polarization (FP) equilibrium competition assays. A protein activity control using the unlabeled reference peptide as inhibitor was performed prior fluorescence polarization (FP) equilibrium competition assays (see reference 4).

The competition assays were performed in 384-well plates (Perkin Elmer, Optiplate-384 F). 2 µM GST-hAR-LBD and 0.1 µM **1** were premixed in the assay-buffer (HEPES buffer containing 10 µM DHT) and pre-equilibrated on ice for at least 30 min. A concentration range (0.1 mM to 1 nM) of miniproteins was prepared by diluting the 1 mM stock solutions with the assay buffer. The concentration of the miniprotein stocks was determined measuring UV absorption at 280 nm and using the calculated extinction coefficient (Expasy ProtParam tool). 40 µL of the GST-hAR-LBD and **1** mixture were added to 10 µL of increasing amounts of miniproteins. The reaction mixture of GST-hAR-LBD and **1** without miniproteins and mixtures containing only **1** were used as controls. After

centrifugation for 1 min at 4 °C and 2200 rcf the fluorescence polarization was measured. No significant difference in IC_{50} values were observed when the polarization was measured at a later time point (1h, 24 hs) therefore all measurements were performed after 1 hour. The fluorescence polarization was measured at room temperature (λ_{ex} = 470 nm and λ_{em} = 519 nm) and plotted (in millipolarization, mP) against increasing concentrations of the miniprotein and fitted with a Klotz binding model to a sigmoid curve using ORIGIN 8.5 (Scientific Graphing and Analysis Software, OriginLab Corp.) to determine the IC_{50} value K_i of the miniproteins [For an overview of all curves, see Figure S 1 and Figure S 2 below; for individual titration curves (mean value plus standard deviation) see Figure S 3 - Figure S 17 in the appendix]. For the used equations and the K_d of the reference peptide **1** see reference 4. The competitive binding of the miniproteins were measured in triplicate in at least two independent experiments.

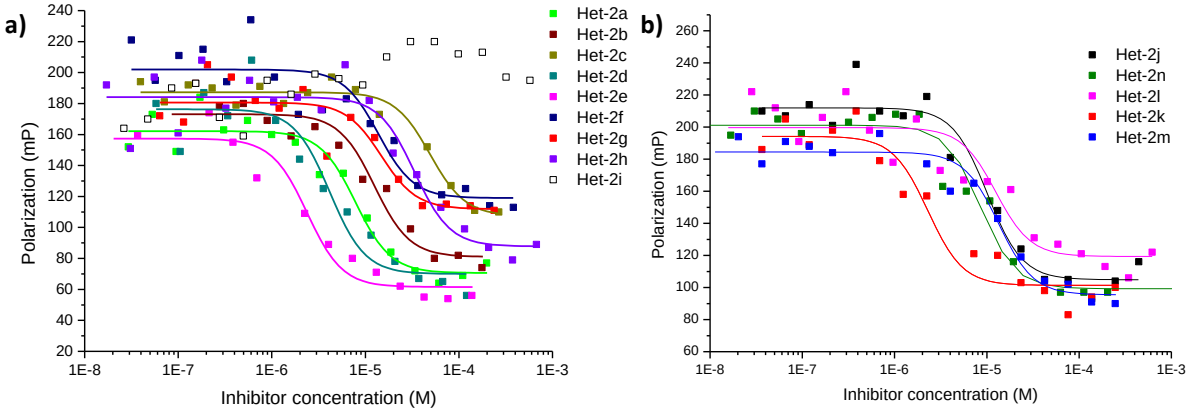

Figure S 1. Serial dilution curves for the **Het-2**-derived miniprotein library: a) Analogues **Het-2a** to **Het-2i**. b) Analogues **Het-2j** to **Het-2n**.

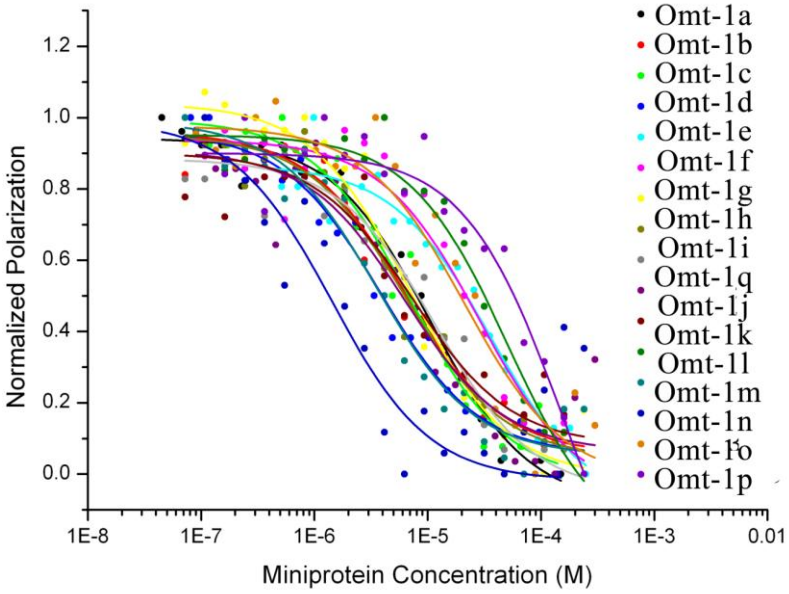

Figure S 2. Normalized serial dilution curves for the **Omt-1**-derived miniprotein library.

Appendix

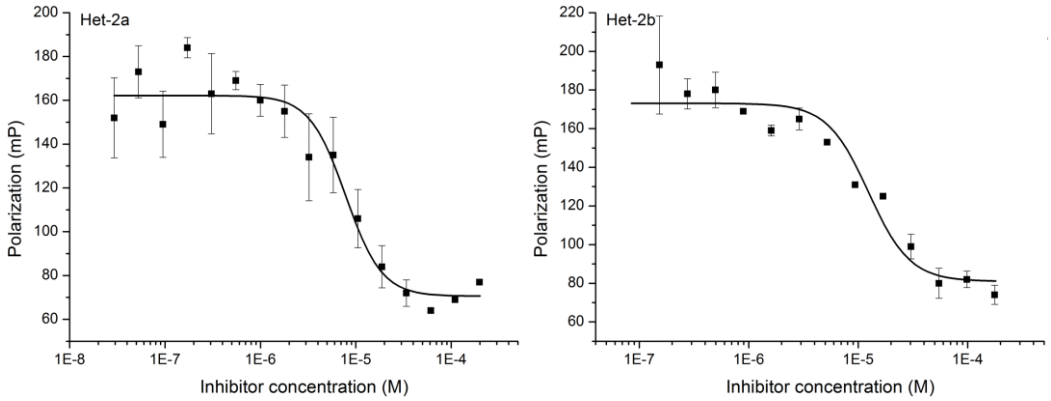

Figure S 3. Serial dilution curves for **Het-2a** (left) and **Het-2b** (right).

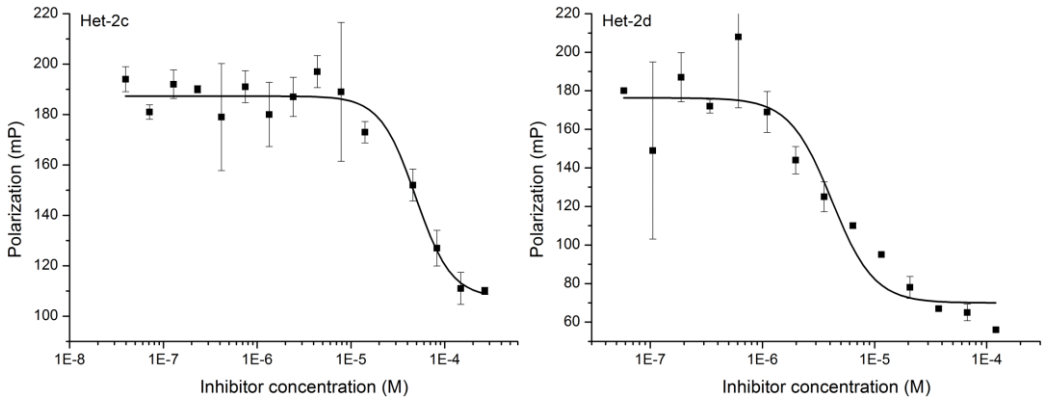

Figure S 4. Serial dilution curves for **Het-2c** (left) and **Het-2d** (right).

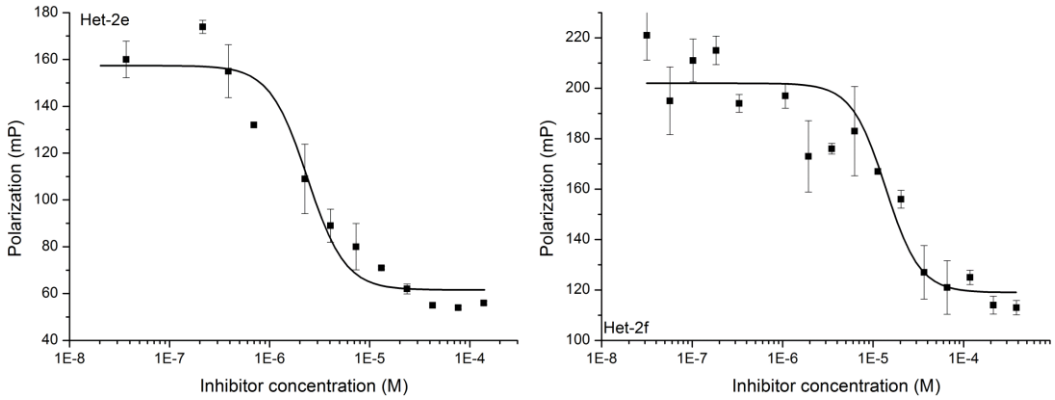

Figure S 5. Serial dilution curves for **Het-2e** (left) and **Het-2f** (right).

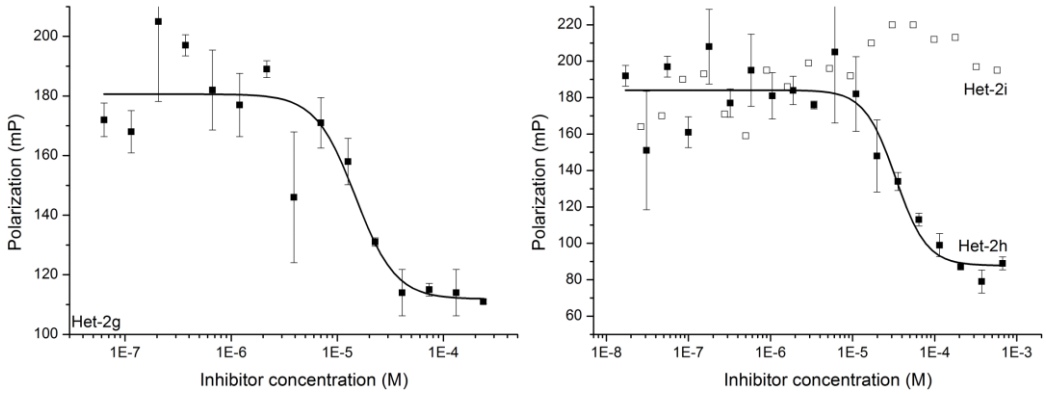

Figure S 6. Serial dilution curves for **Het-2g** (left) and **Het-2h** and **Het-2i** (right).

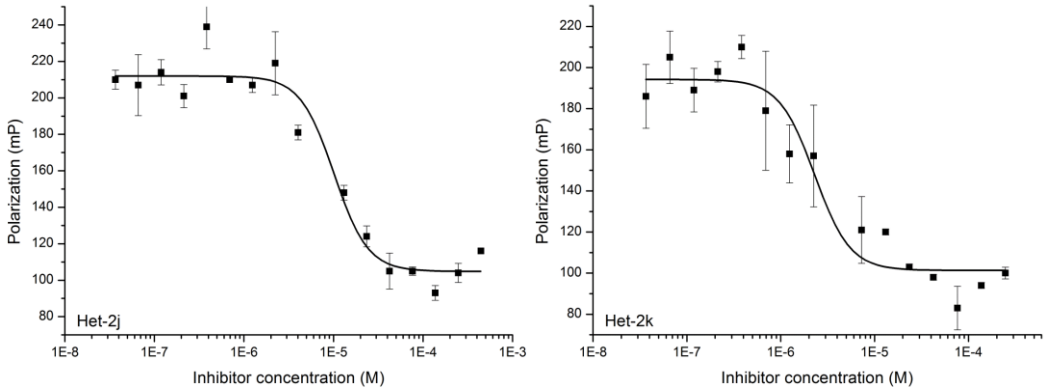

Figure S 7. Serial dilution curves for **Het-2j** (left) and **Het-2k** (right).

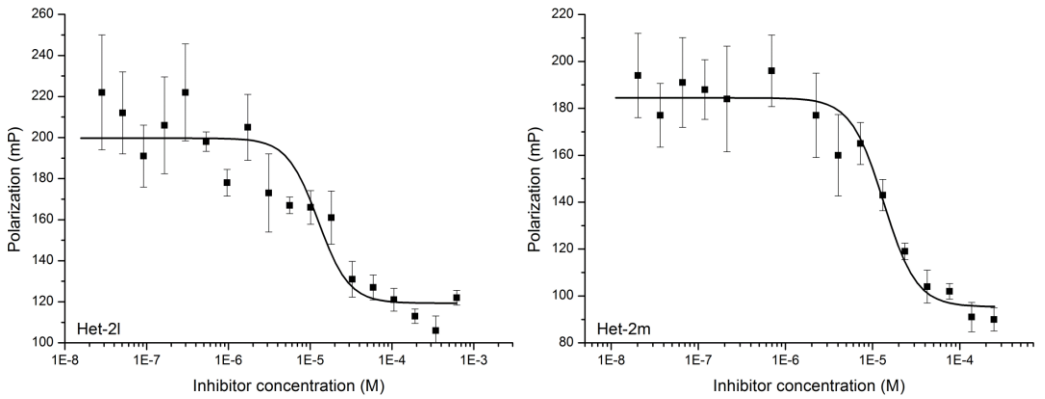

Figure S 8. Serial dilution curves for **Het-2l** (left) and **Het-2m** (right).

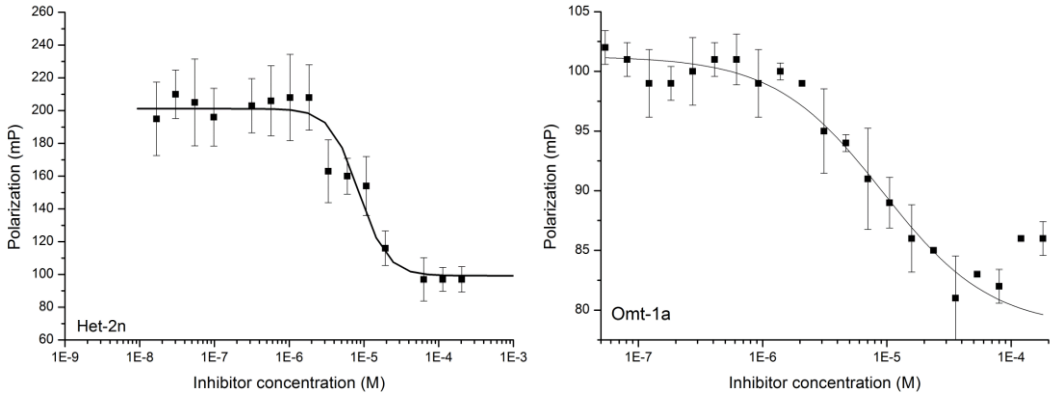

Figure S 9. Serial dilution curves for **Het-2n** (left) and **Omt-1a** (right).

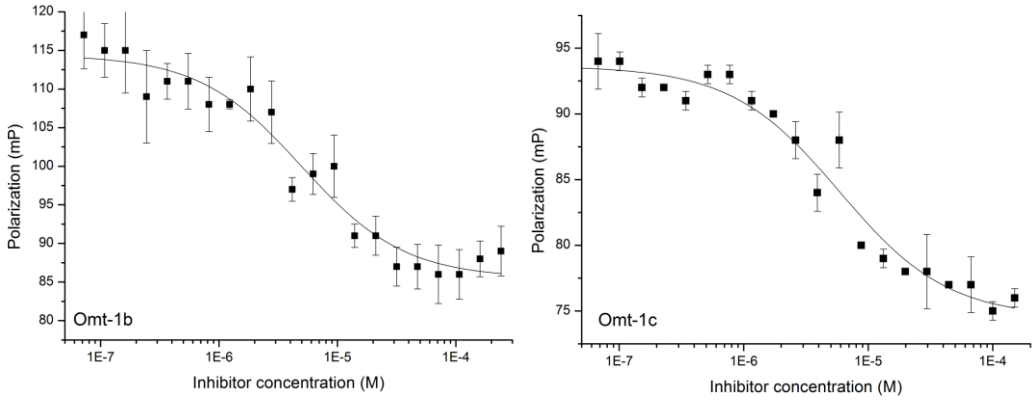

Figure S 10. Serial dilution curves for **Omt-1b** (left) and **Omt-1c** (right).

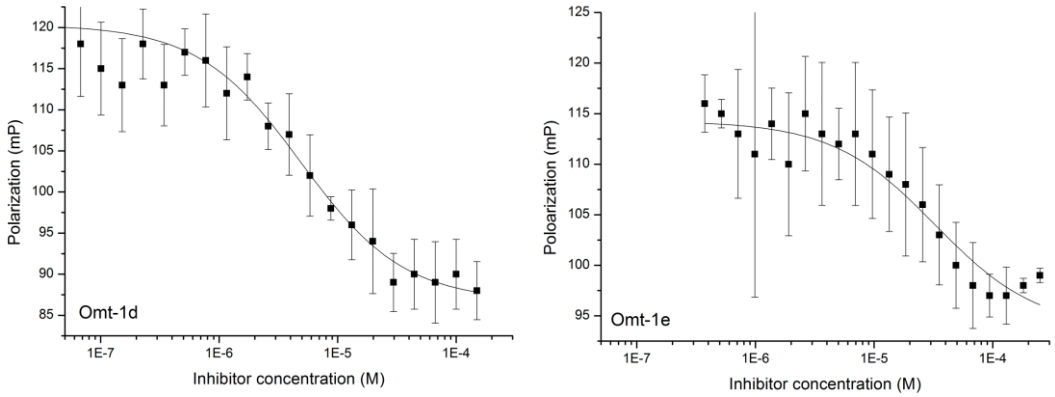

Figure S 11. Serial dilution curves for **Omt-1d** (left) and **Omt-1e** (right).

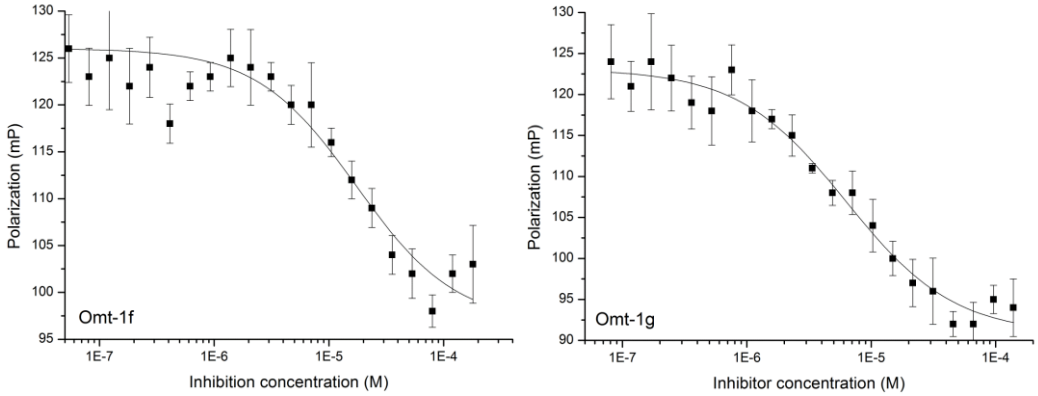

Figure S 12. Serial dilution curves for **Omt-1f** (left) and **Omt-1g** (right).

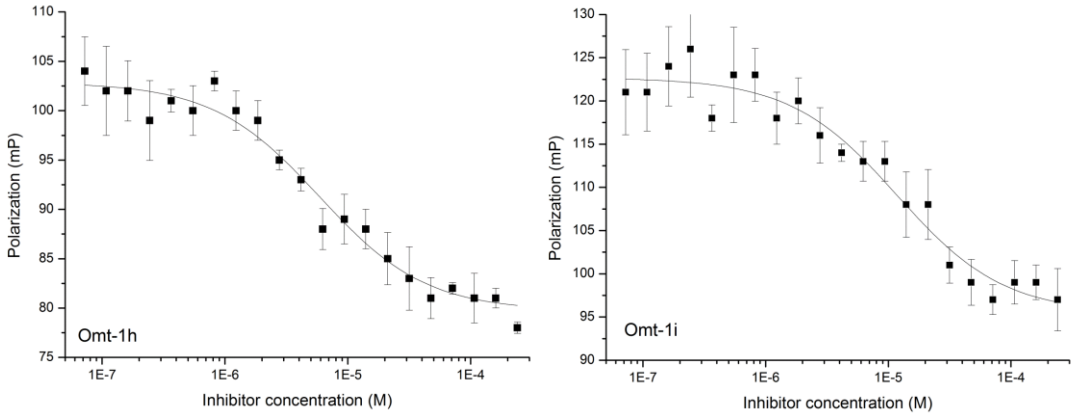

Figure S 13. Serial dilution curves for **Omt-1h** (left) and **Omt-1i** (right).

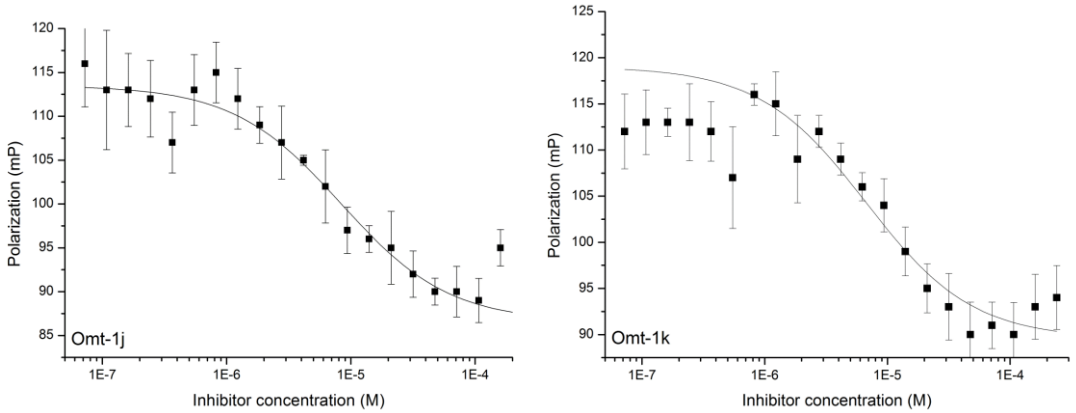

Figure S 14. Serial dilution curves for **Omt-1j** (left) and **Omt-1k** (right).

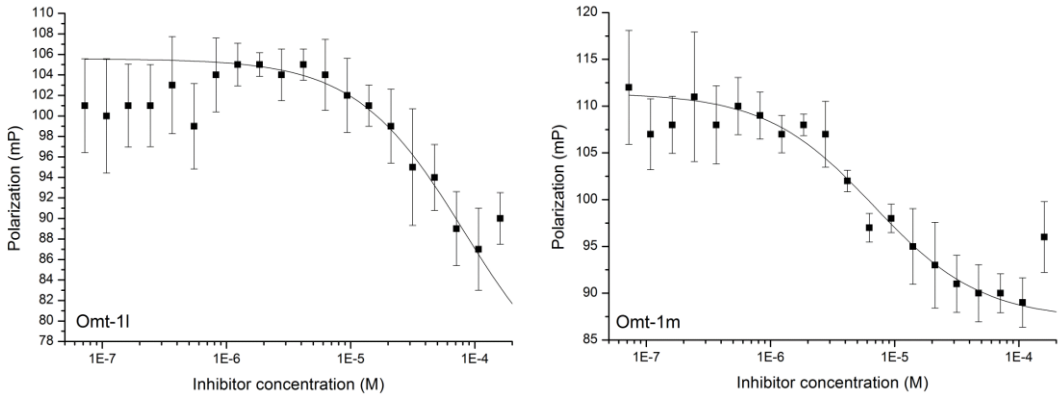

Figure S 15. Serial dilution curves for **Omt-1l** (left) and **Omt-1m** (right).

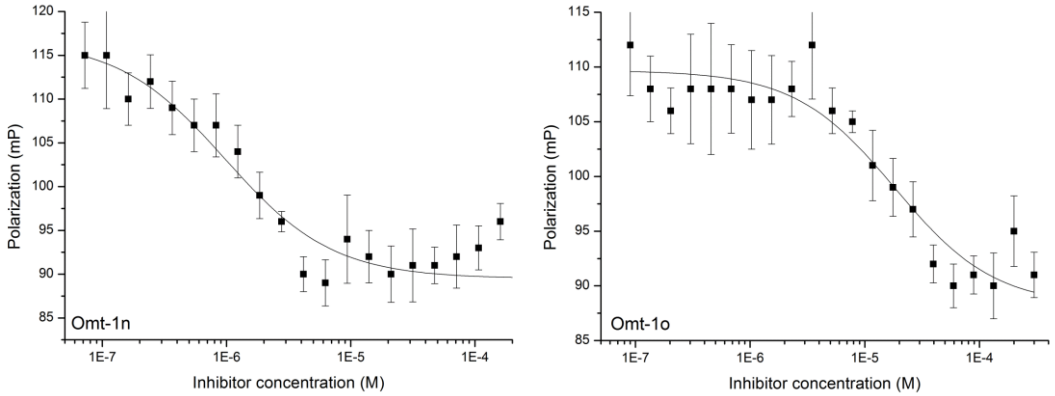

Figure S 16. Serial dilution curves for **Omt-1n** (left) and **Omt-1o** (right).

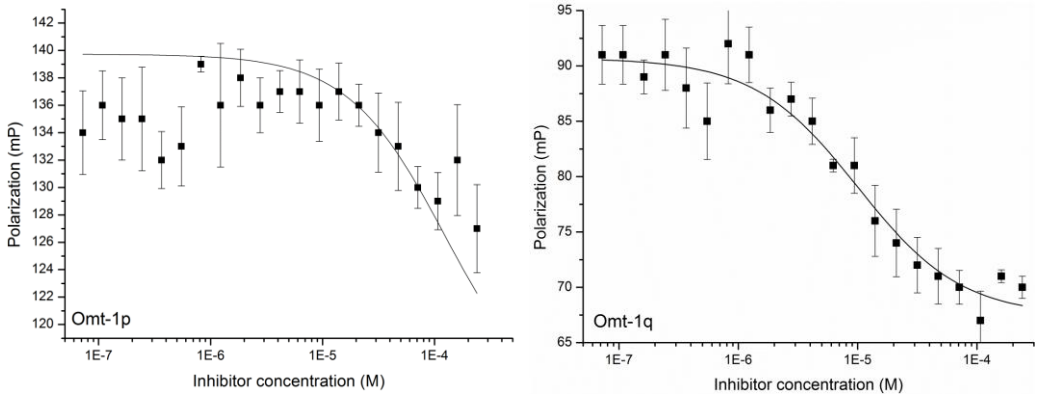

Figure S 17. Serial dilution curves for **Omt-1p** (left) and **Omt-1q** (right).

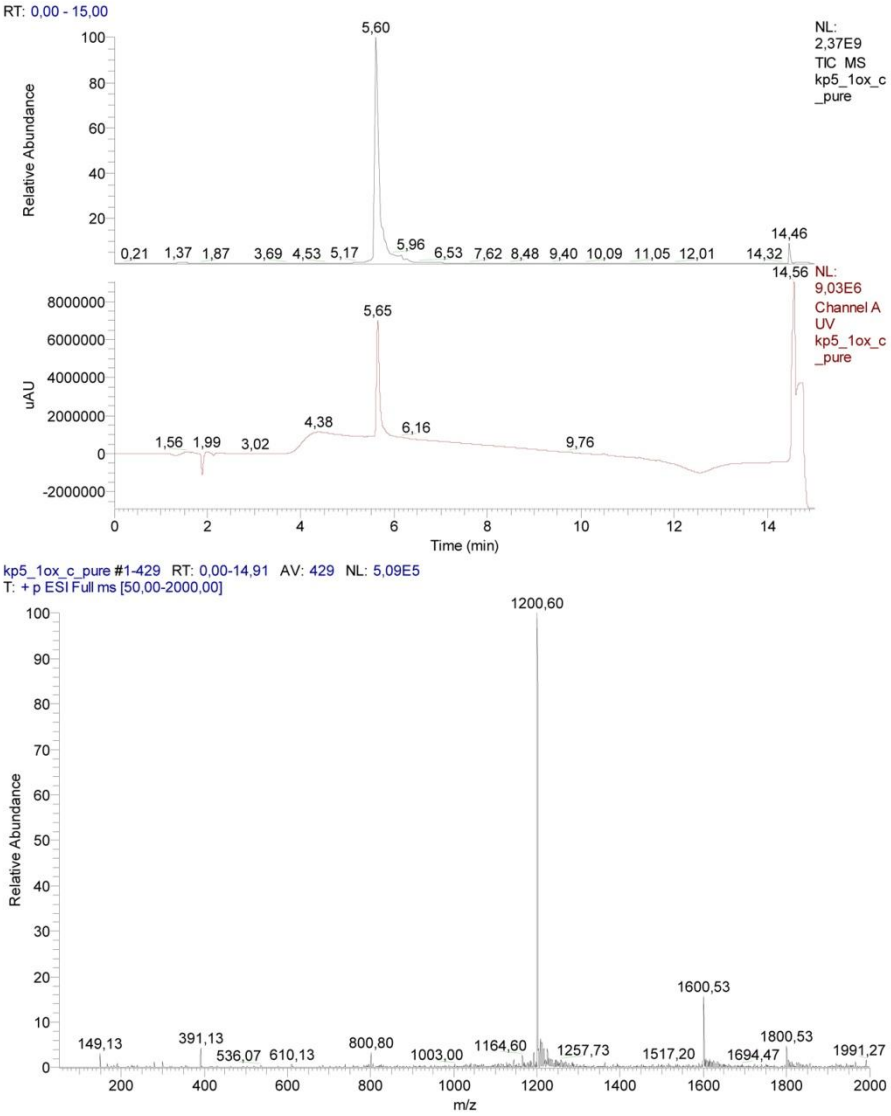

Figure S 18. LC-MS data of **Het-2a** after purification by reverse-phase preparative HPLC.

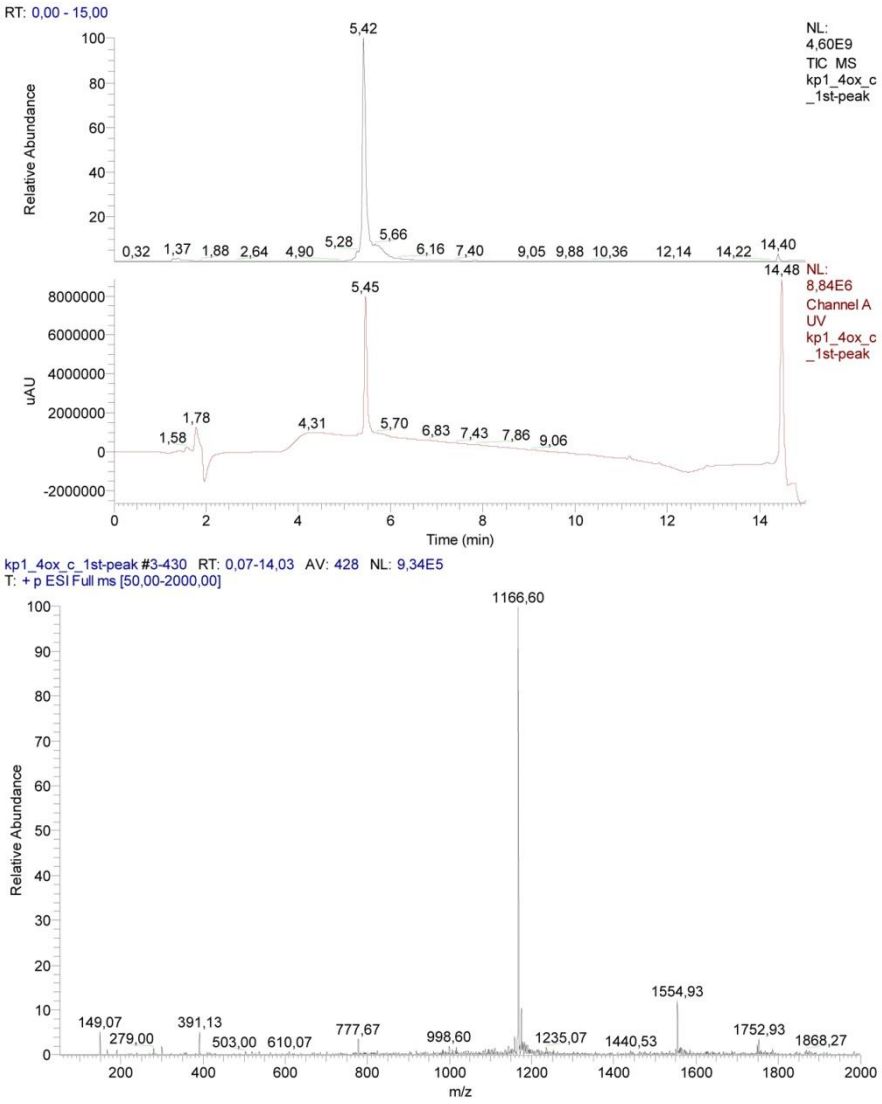

Figure S 19. LC-MS data of **Het-2d** after purification by reverse-phase preparative HPLC.

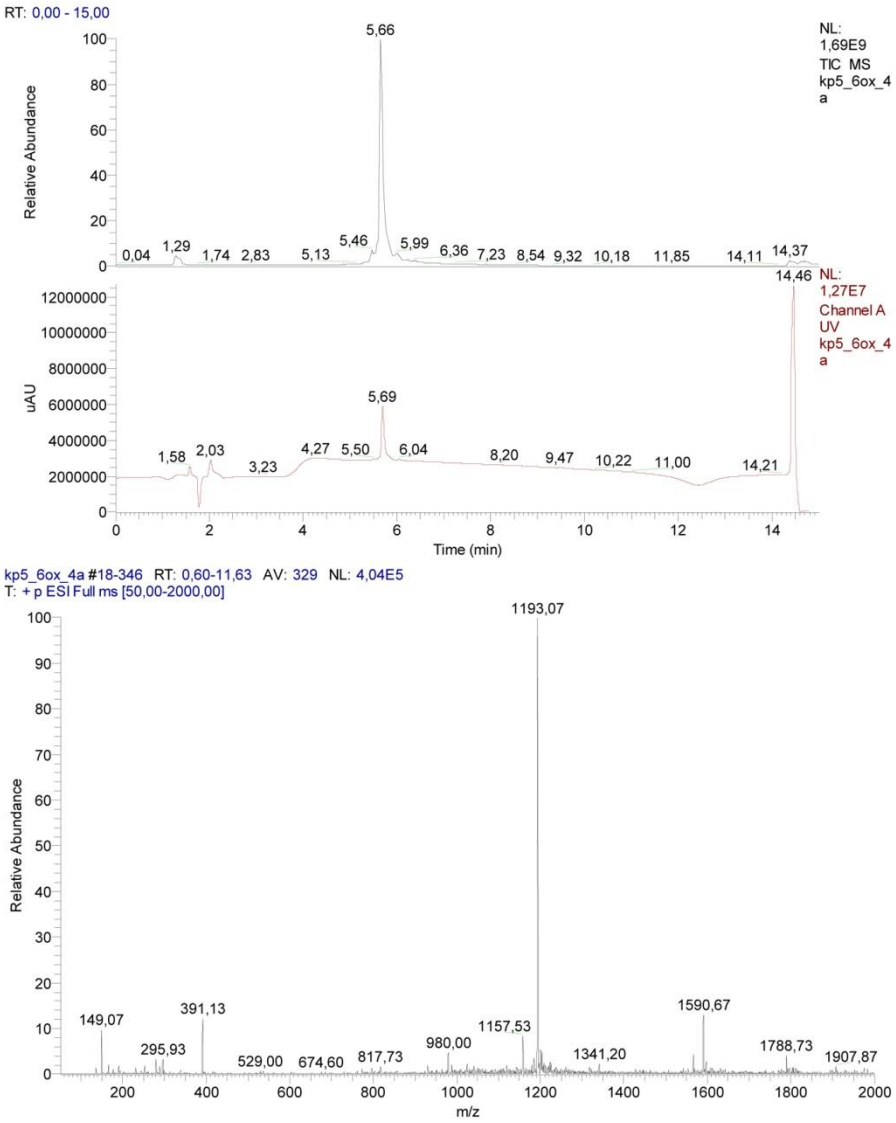

Figure S 20. LC-MS data of **Het-2e** after purification by reverse-phase preparative HPLC.

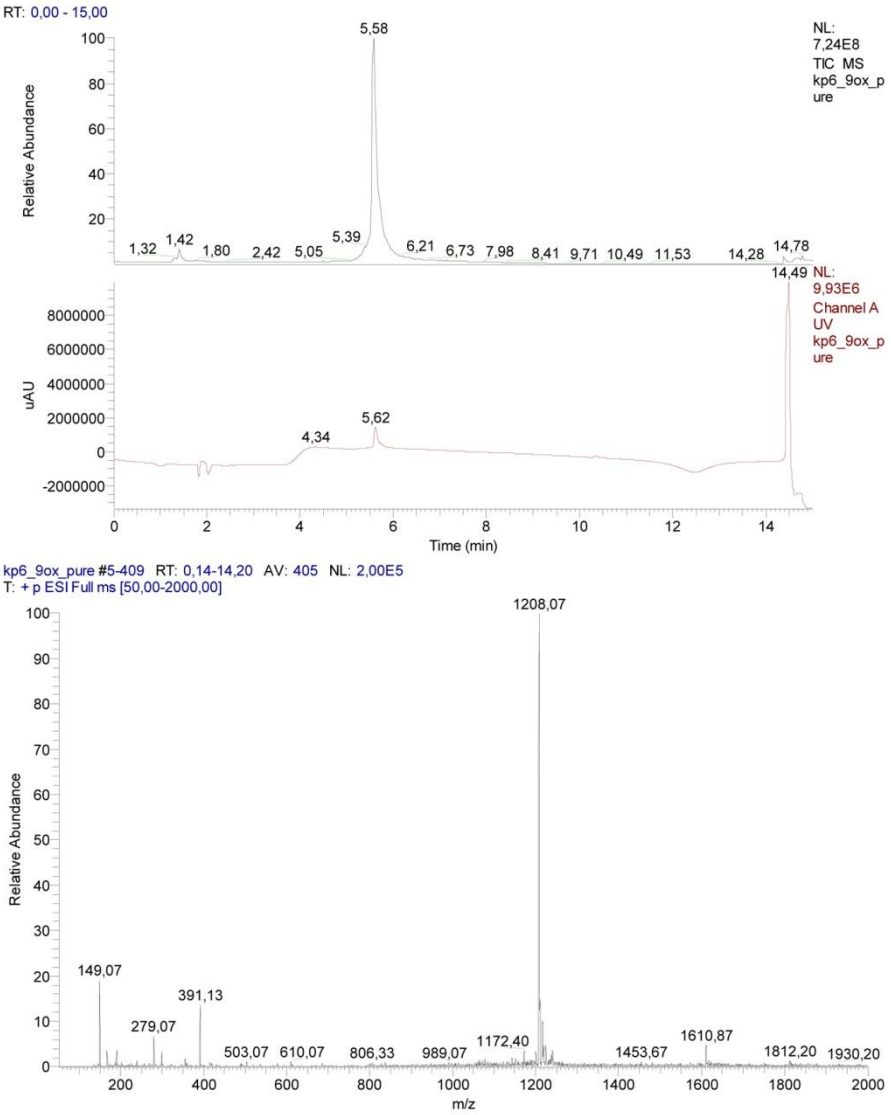

Figure S 21. LC-MS data of **Het-2i** after purification by reverse-phase preparative HPLC.

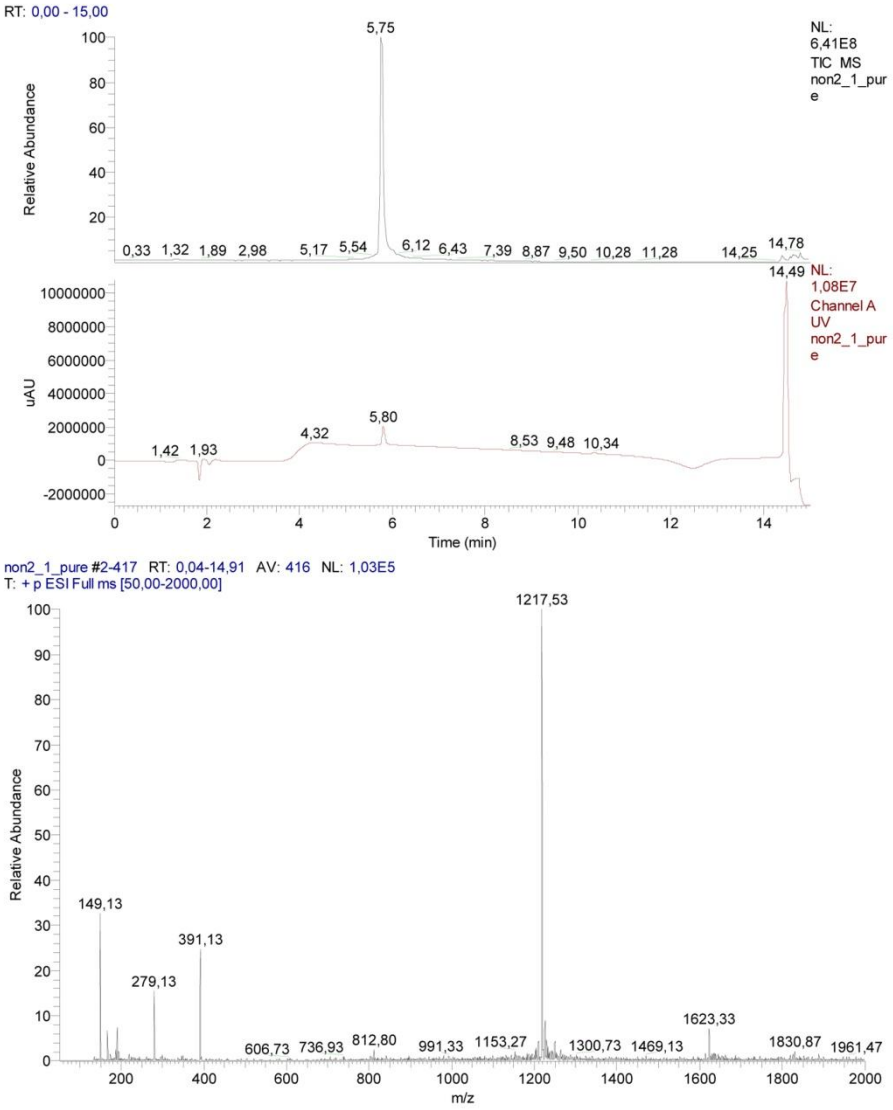

Figure S 22. LC-MS data of **Het-2k** after purification by reverse-phase preparative HPLC.

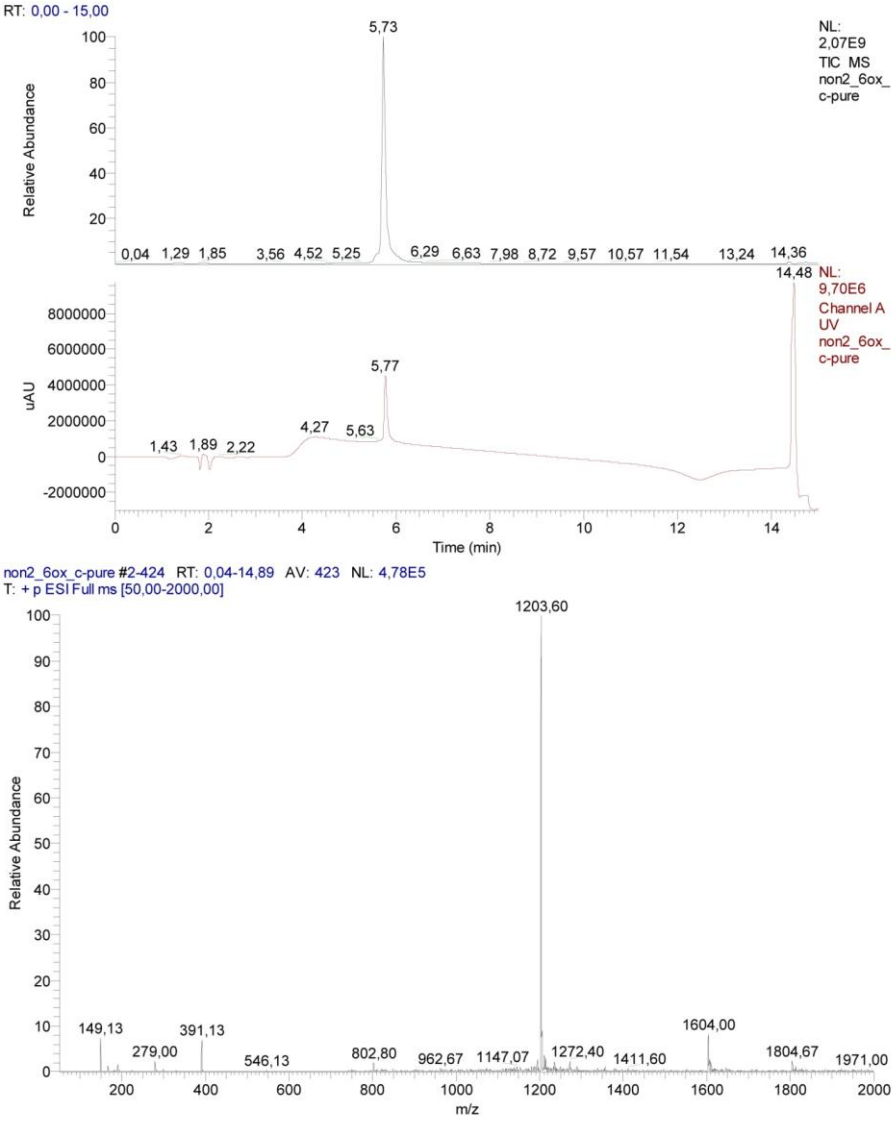

Figure S 23. LC-MS data of **Het-2n** after purification by reverse-phase preparative HPLC.

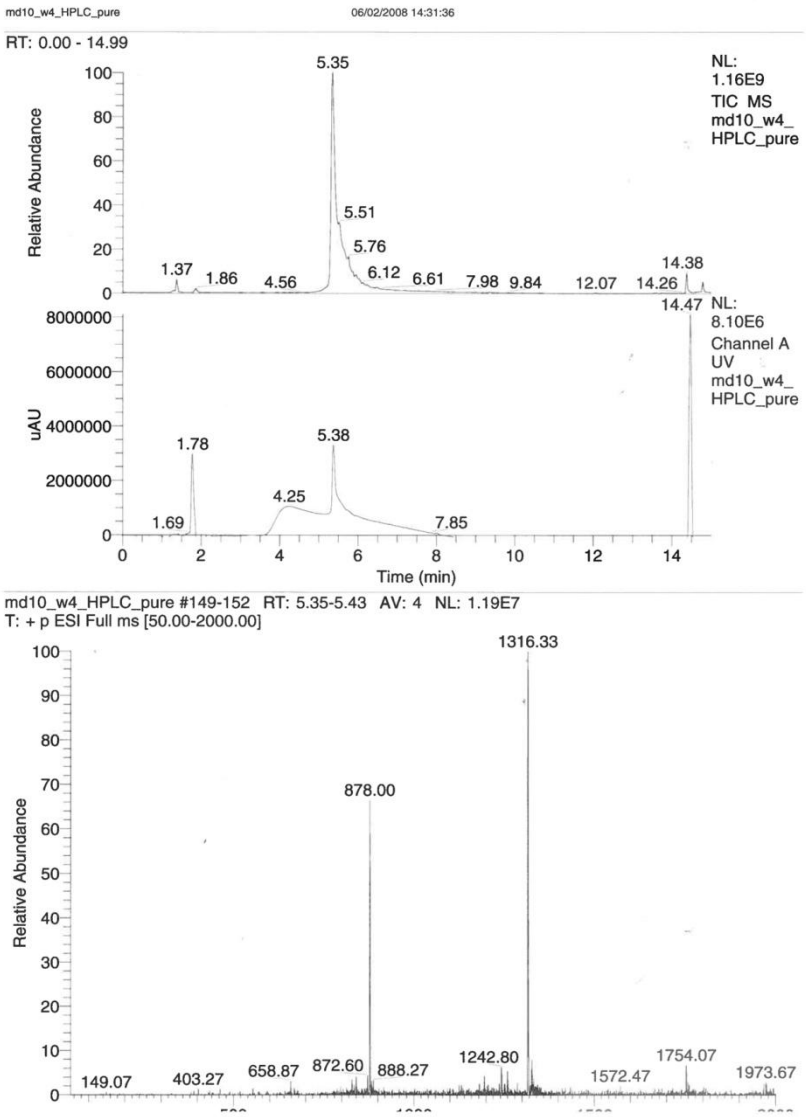

Figure S 24. LC-MS data of **Omt-1d** after purification by reverse-phase preparative HPLC.

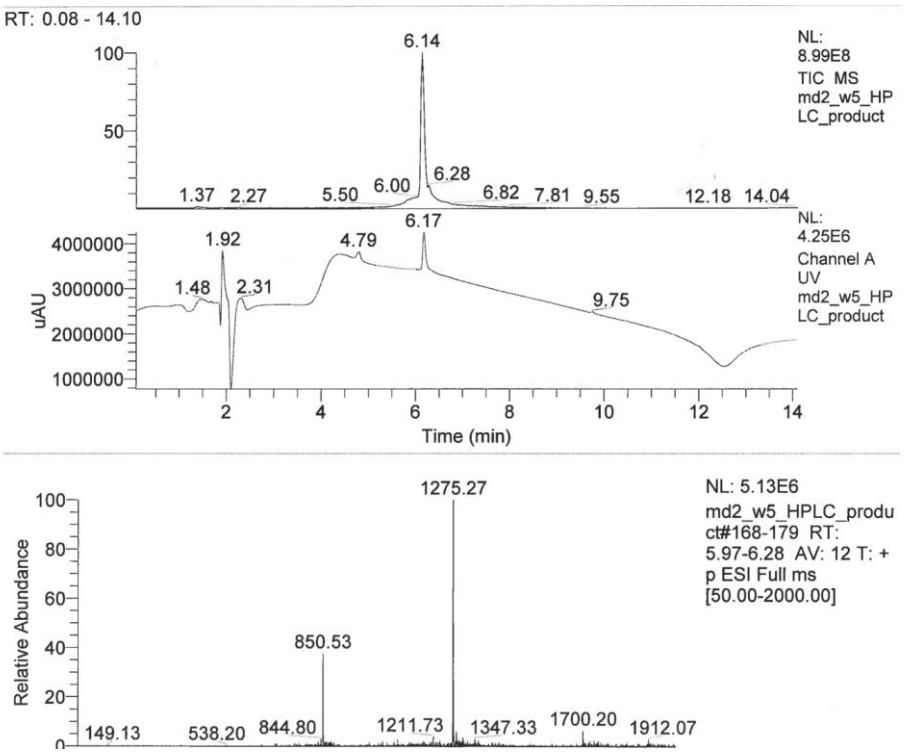

Figure S 25. LC-MS data of **Omt-1e** after purification by reverse-phase preparative HPLC.

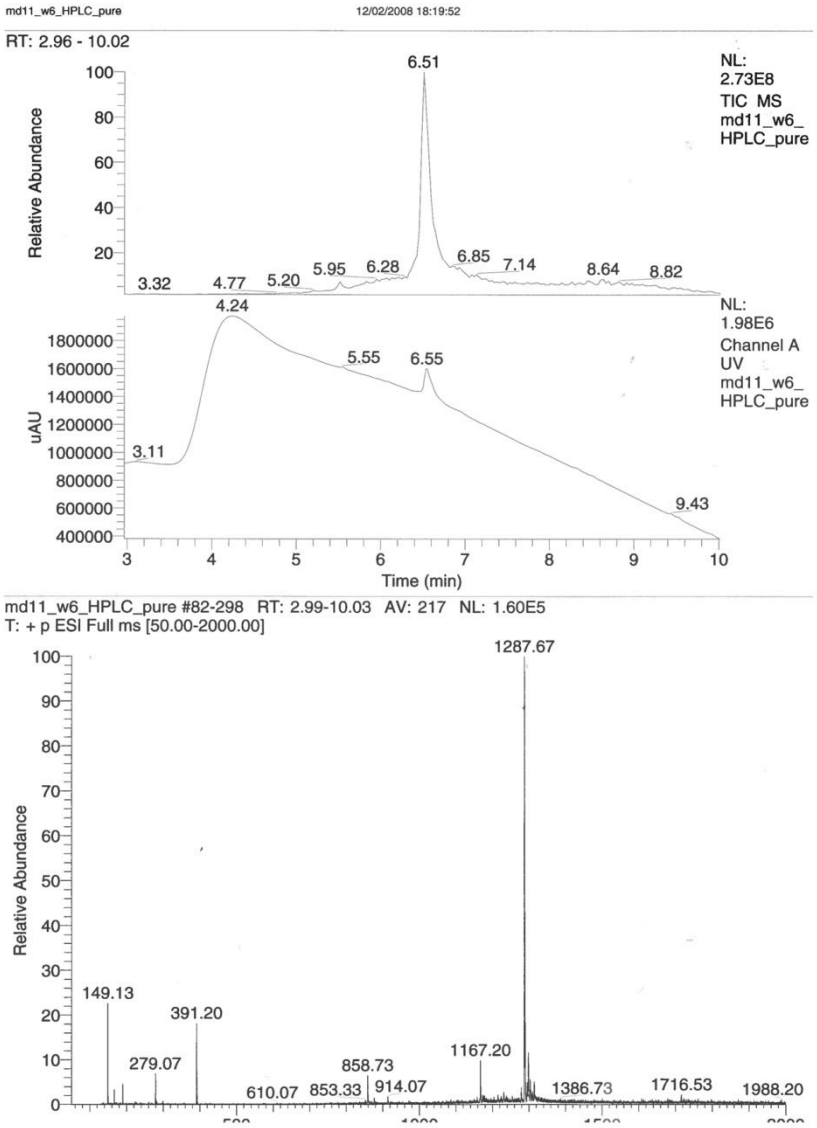

Figure S 26. LC-MS data of **Omt-1f** after purification by reverse-phase preparative HPLC.

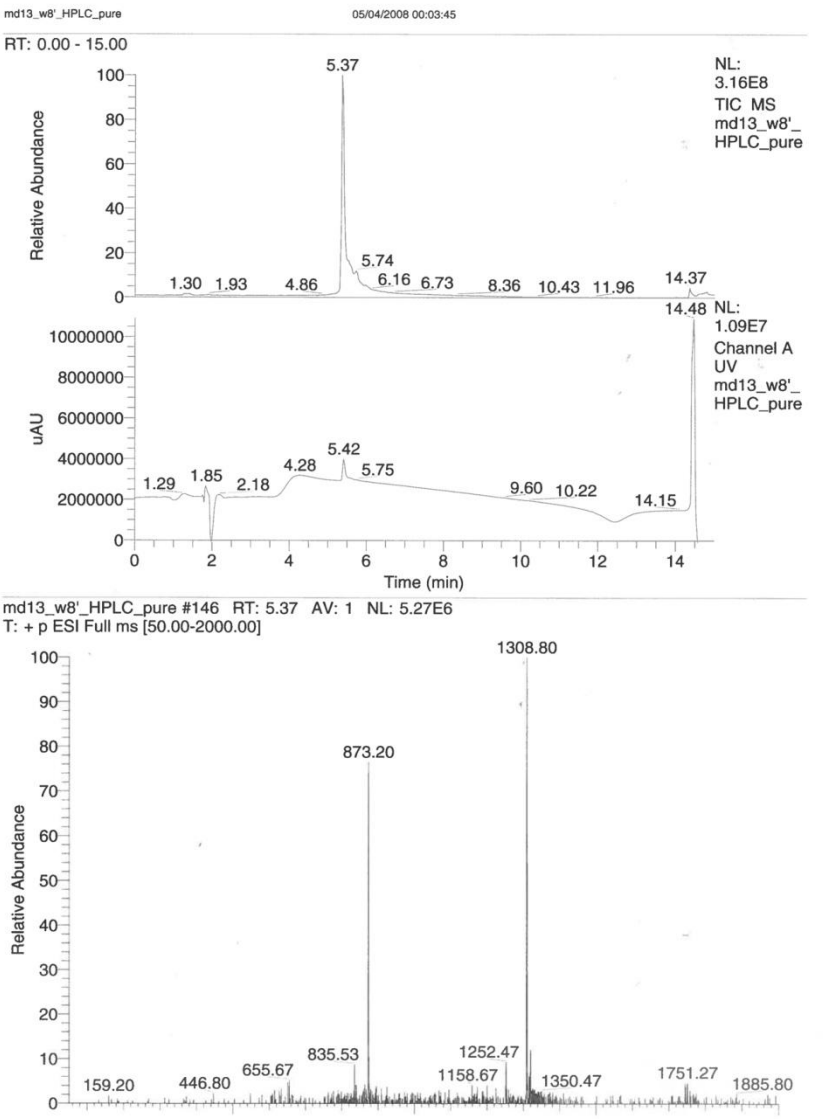

Figure S 27. LC-MS data of **Omt-1h** after purification by reverse-phase preparative HPLC.

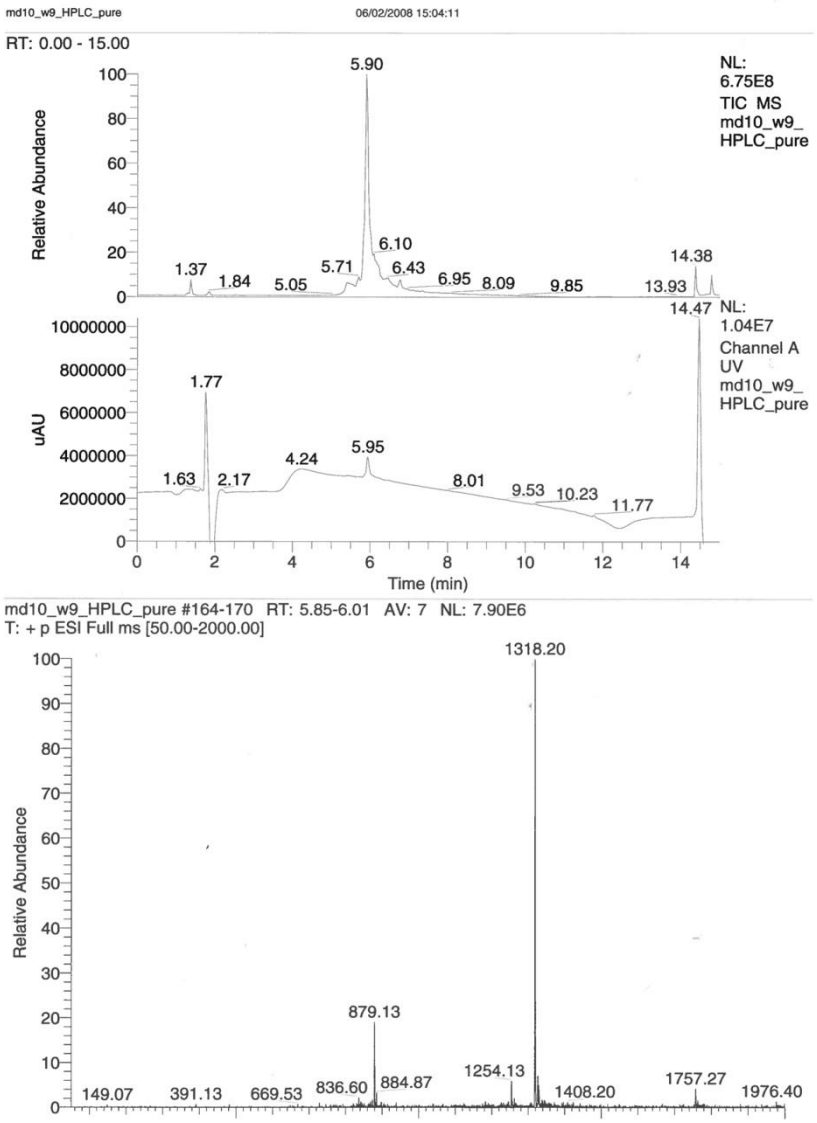

Figure S 28. LC-MS data of **Omt-1i** after purification by reverse-phase preparative HPLC.

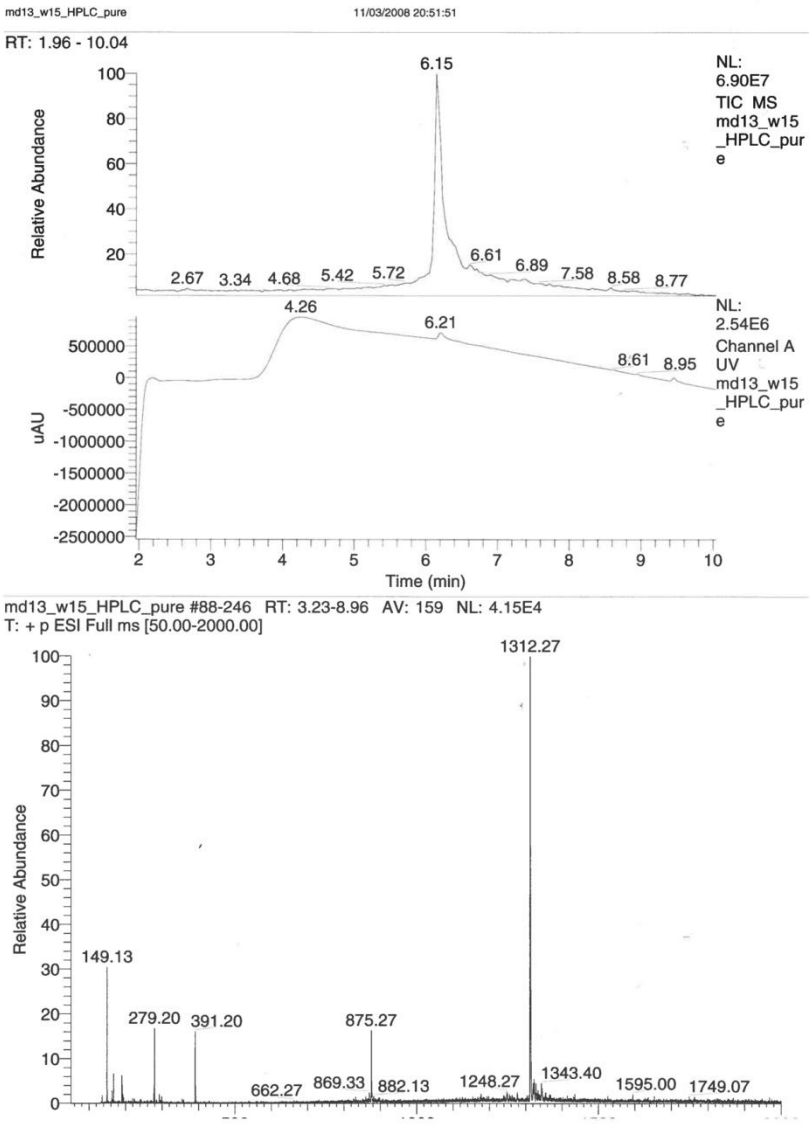

Figure S 29. LC-MS data of **Omt-1o** after purification by reverse-phase preparative HPLC.

References

1. K. N. Srinivasan, V. Sivaraja, I. Huys, T. Sasaki, B. Cheng, T. K. S. Kumar, K. Sato, J. Tytgat, C. Yu, B. C. C. San, S. Ranganathan, H. J. Bowie, R. M. Kini, and P. Gopalakrishnakone, *J. Biol. Chem.*, 2002, **277**, 30040–30047.
2. B. Chagot, C. Pimentel, L. Dai, J. Pil, J. Tytgat, T. Nakajima, G. Corzo, H. Darbon, and G. Ferrat, *Biochemical Journal*, 2005, **388**, 263.
3. E. Hur, S. J. Pfaff, E. S. Payne, H. Grøn, B. M. Buehrer, and R. J. Fletterick, *PLoS Biol*, 2004, **2**, e274.
4. B. Vaz, S. Möcklinghoff, S. Folkertsma, S. Lusher, J. de Vlieg, and L. Brunsveld, *Chem. Commun.*, 2009, 5377.
5. L. A. Carpino and G. Y. Han, *J. Am. Chem. Soc.*, 1970, **92**, 5748–5749.
6. J. C. Sheehan and G. P. Hess, *J. Am. Chem. Soc.*, 1955, **77**, 1067–1068.