Supporting information

Mechanism of Action of Thalassospiramides, A New Class of Calpain Inhibitors

Liang Lu,¹ Michael J. Meehan,² Shuo Gu,³ Zhilong Chen,³ Weipeng Zhang,¹ Gen Zhang,¹ Lingli Liu,¹

Xuhui Huang,³ Pieter C. Dorrestein,^{2, 4, 5} Ying Xu,^{*, 1, 6} Bradley S. Moore, ^{*, 2, 5} Pei-Yuan Qian^{*, 1}

¹ KAUST Global Collaborative Research, Division of Life Science, School of Science, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China

² Skaggs School of Pharmacy & Pharmaceutical Sciences, University of California at San Diego, La Jolla, California 92037, United States

³ Department of Chemistry, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China

⁴ Department of Pharmacology, University of California at San Diego, La Jolla, California 92037, United States

⁵ Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92037, United States

⁶ School of Life Science, Shenzhen University, Nanhai Ave 3688, Shenzhen, Guangdong Province, 518060, China

* Correspondence should be addressed to P.Y.Qian (boqianpy@ust.hk), B. Moore (bsmoore@ucsd.edu) or Y. Xu (boxuying@szu.edu.cn).

Table of contents

Supplementary Figures

Supplementary Fig. 1: MS/MS analysis of peak 3097.5 (peptide 105-133) by MALDI-TOF/TOF.

Supplementary Fig. 2: Sequence alignment of human calpain 1 and two expressed proteins in E.coli.

Supplementary Fig. 3: SDS-PAGE of µI-II and µI-IIC115A.

Supplementary Fig. 4: Identification of µI-II by MS/MS annotation with LC-HR-MS/MS (qTOF).

Supplementary Fig. 5: Identification of µI-IIC115A by MS/MS annotation with LC-HR-MS/MS (qTOF).

Supplementary Fig. 6: MS/MS annotation of β -ME modified peptides in μ I-II by FT-ICR-MS.

Supplementary Fig. 7: FT-ICR-MS data comparison of µI-II and µI-IIC115A samples.

Supplementary Fig. 8: MS/MS annotation of 1-modified peptide by HR-MS/MS (qTOF).

Supplementary Fig. 9: Representative pose of 4 (thalassospiramide C) in the calpain model.

Supplementary Fig. 10: Characterization of **1** by HRMS and ¹H NMR.

Supplementary Fig. 11: Characterization of 2 by HRMS and MSⁿ.

Supplementary Fig. 12: Characterization of 3 by HRMS and MSⁿ.

Supplementary Fig. 13: Characterization of 4 by HRMS and H-NMR.

Supplementary Table

Supplementary Table 1: Docking results for the representative thalassospiramides.

Supplementary Figures

TDICQGALGD <mark>C(115)</mark> WLLAAIASLTLNDTLLHR	Species	Cal. mass	Obs. mass	Error (Da)
	y1	175.12	175.11	0.01
MS1	у2	312.18	312.11	0.07
	у3	425.26	425.14	0.12
	y4	538.35	538.21	0.14
	у5	639.39	639.27	0.12
2800 3200 m/z	уб	754.42	754.30	0.12
y5	y7	868.46	868.35	0.11
MS2 A	y8	981.55	981.42	0.13
	у9	1082.6	1082.5	0.1
	y10	1195.7	1195.6	0.1
y1	y11	1282.7	1282.5	0.2
$\frac{y^{6}}{754,296}$	y12	1353.8	1353.6	0.2
$\frac{10000}{408.155}$ $\frac{408.155}{7}$ $\frac{1000}{7}$ $\frac{1000}{7}$ $\frac{1000}{7}$ $\frac{1000}{7}$ $\frac{1000}{7}$ $\frac{1000}{7}$ $\frac{1000}{7}$ $\frac{1000}{7}$ $\frac{1000}{7}$	y14	1537.9	1537.5	0.4
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	y15	1608.9	1608.6	0.3
	y18	2021.2	2021.1	0.1
500 1000 1500 2000 2500	y19	2124.2	2124.0	0.2

Supplementary Fig. 1 | MS/MS analysis of peak 3097.5 (peptide105-133) by MALDI-TOF/TOF (LIFT mode).

μI-II μI-IIC115A Human calpain 1	MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSGRHENAIKYLGQDYEQLRVRCLQSGT MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSGRHENAIKYLGQDYEQLRVRCLQSGT GRHENAIKYLGQDYEQLRVRCLQSGT
μI-II	LFRDEAFPPVPQSLGYKDLGPNSSKTYGIKWKRPTELLSNPQFIVDGATRTDICQGALGD
μI-IIC115A Human calpain 1	LFRDEAFPPVPQSLGYKDLGPNSSKTYGIKWKRPTELLSNPQFIVDGATRTDICQGALGD LFRDEAFPPVPQSLGYKDLGPNSSKTYGIKWKRPTELLSNPQFIVDGATRTDICQGALGD
μI-II	CWLLAAIASLTLNDTLLHRVVPHGQSFQNGYAGIFHFQLWQFGEWVDVVVDDLLPIKDGK
μI-IIC115A Human calpain 1	AWLLAAIASLTLNDTLLHRVVPHGQSFQNGYAGIFHFQLWQFGEWVDVVVDDLLPIKDGK CWLLAAIASLTLNDTLLHRVVPHGQSFQNGYAGIFHFQLWQFGEWVDVVVDDLLPIKDGK
μI-II	LVFVHSAEGNEFWSALLEKAYAKVNGSYEALSGGSTSEGFEDFTGGVTEWYELRKAPSDL
μI-IIC115A Human calpain 1	LVFVHSAEGNEFWSALLEKAYAKVNGSYEALSGGSTSEGFEDFTGGVTEWYELRKAPSDL LVFVHSAEGNEFWSALLEKAYAKVNGSYEALSGGSTSEGFEDFTGGVTEWYELRKAPSDL
μI-II	YQIILKALERGSLLGCSIDISSVLDMEAITFKKLVKGHAYSVTGAKQVNYRGQVVSLIRM
µI-IIC115A Human calpain 1	YQIILKALERGSLLGCSIDISSVLDMEAITFKKLVKGHAYSVTGAKQVNYRGQVVSLIRM YQIILKALERGSLLGCSIDISSVLDMEAITFKKLVKGHAYSVTGAKQVNYRGQVVSLIRM
μI-II	RNPWGEVEWTGAWSDSSSEWNNVDPYERDQLRVKMEDGEFWMSFRDFMREFTRLEICNLT
µI-IIC115A Human calpain 1	$RNPWGEVEWTGAWSDSSSEWNNVDPYERDQLRVKMEDGEFWMSFRDFMREFTRLEICNLT\\RNPWGEVEWTGAWSDSSSEWNNVDPYERDQLRVKMEDGEFWMSFRDFMREFTRLEICNLT\\$
μI-II	PDLDKLAAALEHHHHHH
µI-IIC115A	PDLDKLAAALEHHHHHH
Human calpain 1	PD

Supplementary Fig. 2 | Sequence alignment of human calpain 1 and two expressed proteins. The sequence of human calpain 1 was downloaded from *pubmed* (accession: 2ARY_A). For µI-IIC115A protein, the active site Cys115 was mutated to Ala (marked in red).



Supplementary Fig. 3 | SDS-PAGE of μ I-II and μ I-IIC115A protein (purified/ in the cell lysis).

AA sequence coverage of $\mu\text{I-II}\text{:}$

GSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSGRHENAIKYLGQDYEQLRVRCLQSGTLFRDEAFPPVPQSLGY KDLGPNSSKTYGIKWKRPTELLSNPQFIVDGATRTDICQGALGDC(115)WLLAAIASLTLNDTLLHRVVPHGQSFQNG YAGIFHFQLWQFGEWVDVVVDDLLPIKDGKLVFVHSAEGNEFWSALLEKAYAKVNGSYEALSGGSTSEGFEDFTG GVTEWYELRKAPSDLYQIILKALERGSLLGCSIDISSVLDMEAITFKKLVKGHAYSVTGAKQVNYRGQVVSLIRMRN PWGEVEWTGAWSDSSSEWNNVDPYERDQLRVKMEDGEFWMSFRDFMREFTRLEICNLTPDLDKLAAALEHHHHH H

Start	End	Amino acid sequence	Charge	Cal. Mass	Obs. Mass	Error(Da)
17	31	(R) GSHMASMTGGQQMGR(G)	3	512.56	512.56	0.00
17	31	(R) GSHMASMTGGQQMGR(G)	2	768.33	768.33	0.00
42	51	(K) YLGQDYEQLR(V)		642.82	642.82	0.00
54	62	(R) CLQSGTLFR(D)	2	512.77	512.77	0.00
63	76	(R) DEAFPPVPQSLGYK(D)	3	516.60	516.60	0.00
63	76	(R) DEAFPPVPQSLGYK(D)	2	774.39	774.40	-0.01
85	89	(K) TYGIK(W)	2	291.17	291.17	0.00
85	89	(K) TYGIK(W)	1	581.33	581.33	0.00
92	109	(K) RPTELLSNPQFIVDGATR(T)	3	672.03	672.03	0.00
110	138	(R) TDICQGALGDCWLLAAIASLTLNDTLLHR(V)	3	1033.2	1033.2	0.0
139	176	(R) VVPHGQSFQNGYAGIFHFQLWQFGEWVDVVVDDLLPIK(D)	4	1097.1	1097.1	0.0
180	198	(K) LVFVHSAEGNEFWSALLEK(A)	3	726.04	726.05	-0.01
203	233	(K) VNGSYEALSGGSTSEGFEDFTGGVTEWYELR(K)	4	836.88	836.88	0.00
203	233	(K) VNGSYEALSGGSTSEGFEDFTGGVTEWYELR(K)	3	1115.5	1115.5	0.0
235	245	(K) APSDLYQIILK(A)	2	630.86	630.87	-0.01
250	271	(R) GSLLGCSIDISSVLDMEAITFK(K)	3	767.06	767.06	0.00
291	298	(R) GQVVSLIR(M)	2	436.27	436.28	-0.01
301	327	(R) NPWGEVEWTGAWSDSSSEWNNVDPYER(D)	4	800.09	800.10	-0.01
301	327	(R) NPWGEVEWTGAWSDSSSEWNNVDPYER(D)	3	1066.5	1066.5	0.0
334	344	(K) MEDGEFWMSFR(D)	2	717.80	717.80	0.00
353	364	(R) LEICNLTPDLDK(L)	2	687.35	687.36	-0.01

Supplementary Fig. 4 | Identification of μ I-II by MS/MS annotation with LC-HR-MS/MS (qTOF). A) Amino acid sequence coverage of μ I-II and m/z's of all tryptic peptides verified by MS².

B.

Mei	Species	Cal. mass	Obs. mass	Error(Da)	Species	Cal. mass	Obs. mass	Error(Da)
1033.5	b2	217.08	217.08	0.00	y18+2	1011.1	1011.1	0.0
	b3	330.17	330.17	0.00	y18+2	1062.6	1062.6	0.0
1033.2	b4	433.18	433.18	0.00	y20+2	1120.1	1120.1	0.0
	b5	561.23	561.24	-0.01	y21+2	1148.6	1148.6	0.0
	b6	618.26	618.25	0.01	y22+2	1205.2	1205.2	0.0
	b7	689.29	689.29	0.00	y23+2	1240.7	1240.7	0.0
	b8	802.38	802.38	0.00	y24+2	1269.2	1269.2	0.0
	y2	312.18	312.18	0.00	y25+2	1333.2	1333.2	0.0
	y3	425.26	425.26	0.00	y25+2	889.14	889.13	0.01
1032 1034 1036	y4	538.35	538.36	-0.01	y26+2	1384.7	1384.7	0.0
m/z	y12+2	677.38	677.38	0.00	y26+2	923.48	923.47	0.01
	y16+2	861.50	861.50	0.00	y27+2	961.17	961.18	-0.01

T D I C Q G A L G D C W L L A A I A S L T L N D T L L H R

MS2



Supplementary Fig. 4 | Identification of µI-II by MS/MS annotation with LC-HR-MS/MS (qTOF). B) MS/MS annotation of peptide TDICQGALGDCWLLAAIASLTLNDTLLHR (m/z=1033.2, +3 charge).

AA sequence coverage of $\mu I\text{-}IIC115A$:

GSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSGRHENAIKYLGQDYEQLRVRCLQSGTLFRDEAFPPVPQSLGYKDLGP NSSKTYGIKWKRPTELLSNPQFIVDGATRTDICQGALGDA(115)WLLAAIASLTLNDTLLHRVVPHGQSFQNGYAGIFHFQLW QFGEWVDVVVDDLLPIKDGKLVFVHSAEGNEFWSALLEKAYAKVNGSYEALSGGSTSEGFEDFTGGVTEWYELRKAPSDL YQIILKALERGSLLGCSIDISSVLDMEAITFKKLVKGHAYSVTGAKQVNYRGQVVSLIRMRNPWGEVEWTGAWSDSSSEWN NVDPYERDQLRVKMEDGEFWMSFRDFMREFTRLEICNLTPDLDKLAAALEHHHHHH

Start	End	Amino acid sequence	Charge	Cal. Mass	Obs. Mass	Error(Da)
17	31	(R) GSHMASMTGGQQMGR(G)	3	512.56	512.56	0.00
17	31	(R) GSHMASMTGGQQMGR(G)	2	768.33	768.34	-0.01
42	51	(K) YLGQDYEQLR(V)	2	642.82	642.82	0.00
54	62	(R) CLQSGTLFR(D)	2	512.77	512.76	0.01
63	76	(R) DEAFPPVPQSLGYK(D)	3	516.60	516.60	0.00
63	76	(R) DEAFPPVPQSLGYK(D)	2	774.39	774.39	0.00
77	84	(K) DLGPNSSK(T)	2	409.21	409.21	0.00
85	89	(K) TYGIK(W)	2	291.17	291.17	0.00
85	89	(K) TYGIK(W)	1	581.33	581.33	0.00
92	109	(K) RPTELLSNPQFIVDGATR(T)	3	672.03	672.03	0.00
92	109	(K) RPTELLSNPQFIVDGATR(T)	2	1007.5	1007.5	0.0
110	138	(R) TDICQGALGDAWLLAAIASLTLNDTLLHR(V)	3	1022.5	1022.5	0.0
139	176	$(R) \ VVPHGQSFQNGYAGIFHFQLWQFGEWVDVVVDDLLPIK(D) \\$	4	1097.1	1097.1	0.0
180	198	(K)LVFVHSAEGNEFWSALLEK(A)	3	726.04	726.04	0.00
203	233	(K) VNGSYEALSGGSTSEGFEDFTGGVTEWYELR(K)	4	836.88	836.93	-0.05
203	233	(K) VNGSYEALSGGSTSEGFEDFTGGVTEWYELR(K)	3	1115.5	1115.5	0.0
235	245	(K) APSDLYQIILK(A)	2	630.86	630.87	-0.01
250	271	(R) GSLLGCSIDISSVLDMEAITFK(K)	3	767.06	767.06	0.00
291	298	(R) GQVVSLIR(M)	2	436.27	436.28	0.00
301	327	(R) NPWGEVEWTGAWSDSSSEWNNVDPYER(D)	4	800.09	800.10	-0.01
301	327	(R) NPWGEVEWTGAWSDSSSEWNNVDPYER(D)	3	1066.5	1066.5	0.0
334	344	(K) MEDGEFWMSFR(D)	2	717.80	717.80	0.00
345	348	(R) DFMR(E)	1	568.26	568.25	0.01
353	364	(R) LEICNLTPDLDK(L)	2	687.35	687.36	-0.01

Supplementary Fig. 5 | Identification of μ I-IIC115A by MS/MS annotation with LC-HR-MS/MS (qTOF). A) Amino acid sequence coverage of μ I-IIC115A and m/z's of all tryptic peptides verified by MS².

B.

T D I C Q G A L G D A W L L A A I A S L T L N D T L L H R

	Species	Cal. mass	Obs. mass	Error(Da)	Species	Cal. mass	Obs. mass	Error(Da)
	b2	217.08	217.08	0.00	y9+2	541.80	541.80	0.00
MS1	b3	330.17	330.17	0.00	y11+2	641.86	641.86	0.00
	b4	433.18	433.18	0.00	y12+2	677.38	677.38	0.00
1022.9	b5	561.23	561.23	0.00	y13+2	733.92	733.92	0.00
	b6	618.26	618.26	0.00	y10	1195.7	1195.7	0.0
1022.5	b7	689.29	689.29	0.00	y11	1282.7	1282.7	0.0
	b7+2	345.15	345.16	-0.01	y12	1353.8	1353.7	0.1
	b8	802.38	802.38	0.00	y14	1537.9	1537.9	0.0
	b9	859.40	859.40	0.00	y14+2	769.44	769.43	0.01
	b10	974.43	974.42	0.01	y15+2	804.96	804.96	0.00
	b11	1045.5	1045.5	0.0	y16+2	861.50	861.50	0.00
	b13+2	672.82	672.82	0.00	y17+2	918.04	918.04	0.00
	b14+2	729.36	729.36	0.00	y18+2	1011.1	1011.1	0.0
	b15+2	764.88	764.88	0.00	y19+2	1046.6	1046.6	0.0
	b16+2	800.40	800.40	0.00	y20+2	1104.1	1104.1	0.0
1022 1023 1024 m/z	b17+2	856.94	856.93	0.01	y21+2	1132.6	1132.6	0.0
	b18+2	892.46	892.45	0.01	y21+2	755.42	755.42	0.00
	y2	312.18	312.18	0.00	y22+2	1189.2	1189.2	0.0
	y3	425.26	425.26	0.00	y23+2	1224.7	1224.7	0.0
	y4	538.35	538.35	0.00	y24+2	1253.2	1253.2	0.0
	y5	639.39	639.40	-0.01	y25+2	1317.2	1317.2	0.0
	y6	754.42	754.42	0.00	y25+3	878.49	878.49	0.00
	y7	868.46	868.46	0.01	y26+2	1368.7	1368.7	0.0
	y8	981.55	981.54	0.01	y26+3	912.82	912.83	-0.01
	у9	1082.6	1082.6	0.0	y27+3	950.52	950.53	-0.01
					y28+3	988.86	988.87	-0.01



Supplementary Fig. 5 | Identification of μI-IIC115A by MS/MS annotation with LC-HR-MS/MS (qTOF). B) MS/MS annotation of peptide TDICQGALGDAWLLAAIASLTLNDTLLHR. (m/z = 1022.5, 3 charge).

CLQSGTLFRDEAFPPVPQSLGYK



Supplementary Fig. 6 | MS/MS annotation of β -ME modified peptides in μ I-II by FT-ICR-MS. A) A β -ME modification (+75.9983 Da) occurs at Cys49 of the peptide CLQSGTLFRDEAFPPVPQSLGYK.

B.

T D I C Q G A L G D C W L L A A I A S L T L N D T L L H R



Supplementary Fig. 6 | MS/MS annotation of β -ME modified peptides in μ I-II by FT-ICR-MS. B) β -ME modifications (+75.9983 Da) occur at Cys108 and Cys115 of the peptide TDICQGALGDCWLLAAIASLTLNDTLLHR.

μI-II

Calculated exact mass: 42467.9 Da, Observed exact mass: 42467.9 Da, Error: 0.0 Da



μI-II + 1 (shifted by 881.5 Da)

Calculated exact mass: 43349.4 Da, Observed exact mass: 43349.4 Da, Error: 0.0 Da



Supplementary Fig. 7 | FT-MS data comparison of μ I-II and μ I-IIC115A samples. A) μ I-II and μ I-II + 1 sample. A shift of 881.5 was observed in μ I-II + 1 sample. The calculated exact mass of μ I-II includes 3 β -ME adducts (75.9983 Da each). The calculated exact mass of μ I-II + 1 includes 2 β -ME adducts and the mass of 1 (957.5423 Da).

В.

µI-IIC115A

Calculated exact mass: 42359.9 Da, Observed exact mass: 42360.1 Da, Error: -0.2 Da



µI-IIC115A+1 (no shift)

Calculated exact mass: 42359.9 Da, Observed exact mass: 42360.1 Da, Error: -0.2 Da



Supplementary Fig. 7 | FT-MS data comparison of μ I-II and μ I-IIC115A samples. B) μ I-IIC115A and μ I-IIC115A + 1 sample. No shift was observed in μ I-IIC115A + 1 sample. The calculated exact masses of μ I-IIC115A and μ I-IIC115A + 1 both include 2 β -ME adducts (75.9983 Da each).

$T \ D \big| I \big| C \big| Q \big| G \big| A \big| L \big| G \big| D \ C \ W \big| L \big| L \big| A \big| A \big| I \big| A \big| S \big| L \big| T \big| L \big| N \big| D \big| T \big| L \big| L \big| H \ R$





Supplementary Fig. 8 | MS/MS annotation of **1**-modified peptide by Q-TOF (m/z=1014.5, +4). B ions of **1** were labeled by b' and red color.



Supplementary Fig. 9 | Representative pose of **4** (thalassospiramide C) in the calpain model. The lowest energy binding conformation of **4** shows the best pose to S1, S2 and S3 position of calpain in the model. The double bond in the ring has a short distance to Cys115 and the phenol group forms a hydrogen bond with Y202 (the enlarged view of S1 region shown on the right).

ESI-TOF-HRMS







Supplementary Fig. 10 | Characterization of **1** by HRMS and ¹H NMR (in methanol- d_4).

MS¹ (ESI-TOF-HRMS)



Supplementary Fig. 11 | Characterization of 2 by HRMS and MSⁿ.

ESI-TOF-MS





Supplementary Fig. 12 | Characterization of 3 by HRMS and MSⁿ.

ESI-TOF-HRMS







Supplementary Fig. 13 | Characterization of **4** by HRMS and ¹H NMR (in methanol- d_4).

Supplementary Tables

Compound	Minimum warhead distance (Å)	Number of docking poses with distance ^[a] <5Å	Minimum AutoDock score (kcal/mol)	IC ₅₀ (nM)
Thalassospiramide A	4.29	2	-6.2	57
Thalassospiramide A1	3.98	5	-7.3	42
Thalassospiramide A5	4.95	3	-7	22
Thalassospiramide D	4.64	3	-7.4	21
Thalassospiramide D1	3.7	8	-6.8	35
Thalassospiramide B	3.98	10	-6.8	29
Thalassospiramide C	3.39	20	-8.1	3
Thalassospiramide E1	3.60	6	-7.6	21

Supplementary Table 1: Docking results for the representative thalassospiramides.

[a]: the distance between warhead carbon and active site sulfur.