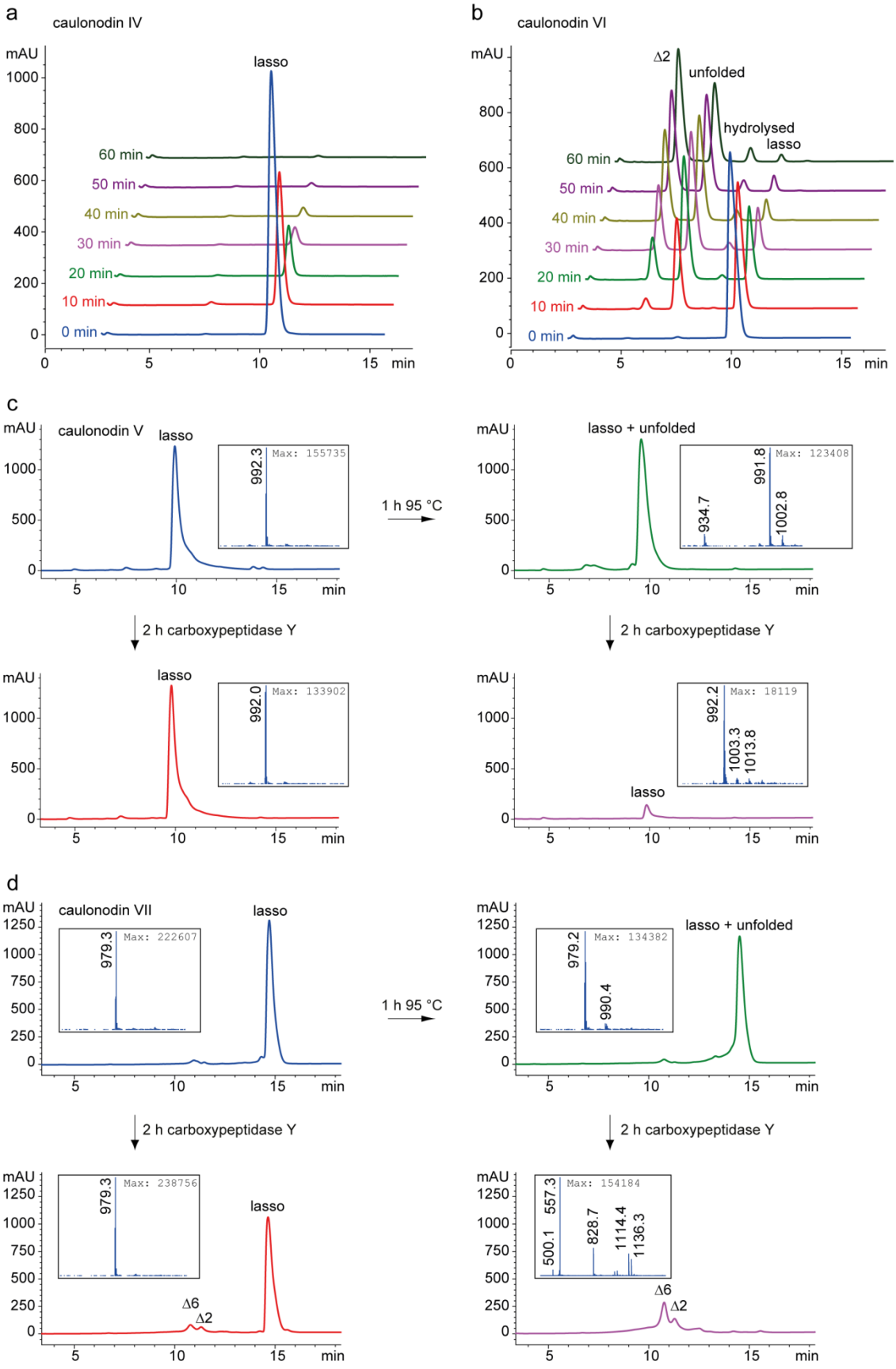
Characterization of Caulonodin Lasso Peptides Revealed Unprecedented N-Terminal Residues and a Precursor Motif Essential for Peptide Maturation

*Marcel Zimmermann, Julian D. Hegemann, Xiulan Xie and Mohamed A. Marahiel**

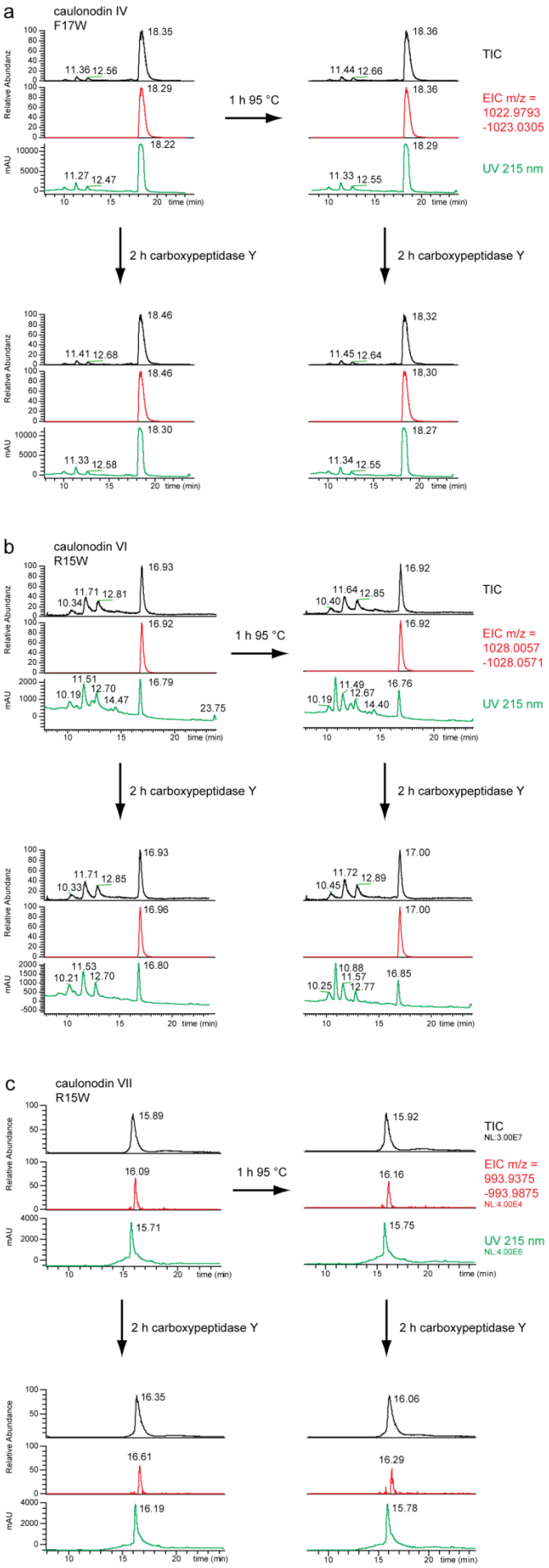
Department of Chemistry, Biochemistry, Philipps-University Marburg, Hans-Meerwein-Strasse 4 and LOEWE-Center for Synthetic Microbiology, D-35032, Marburg, Germany

*marahiel@staff.uni-marburg.de

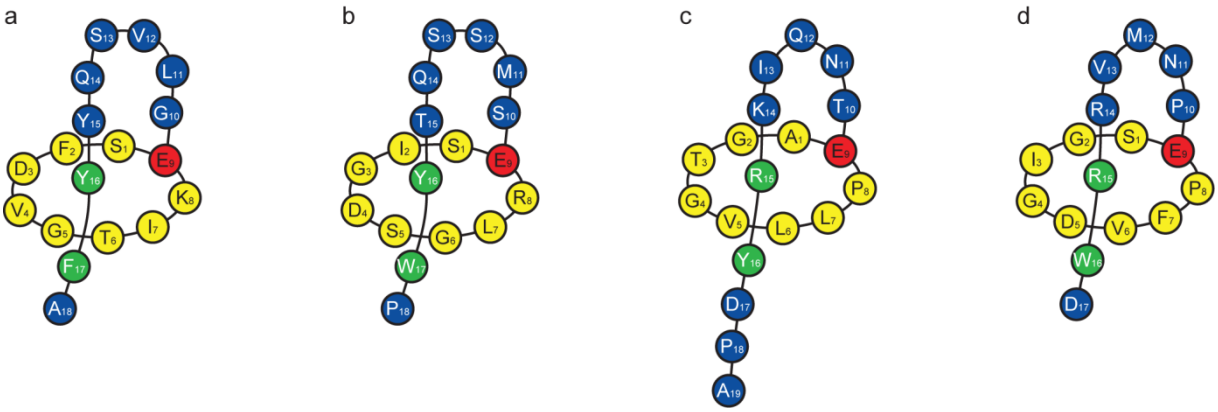
Supporting Information



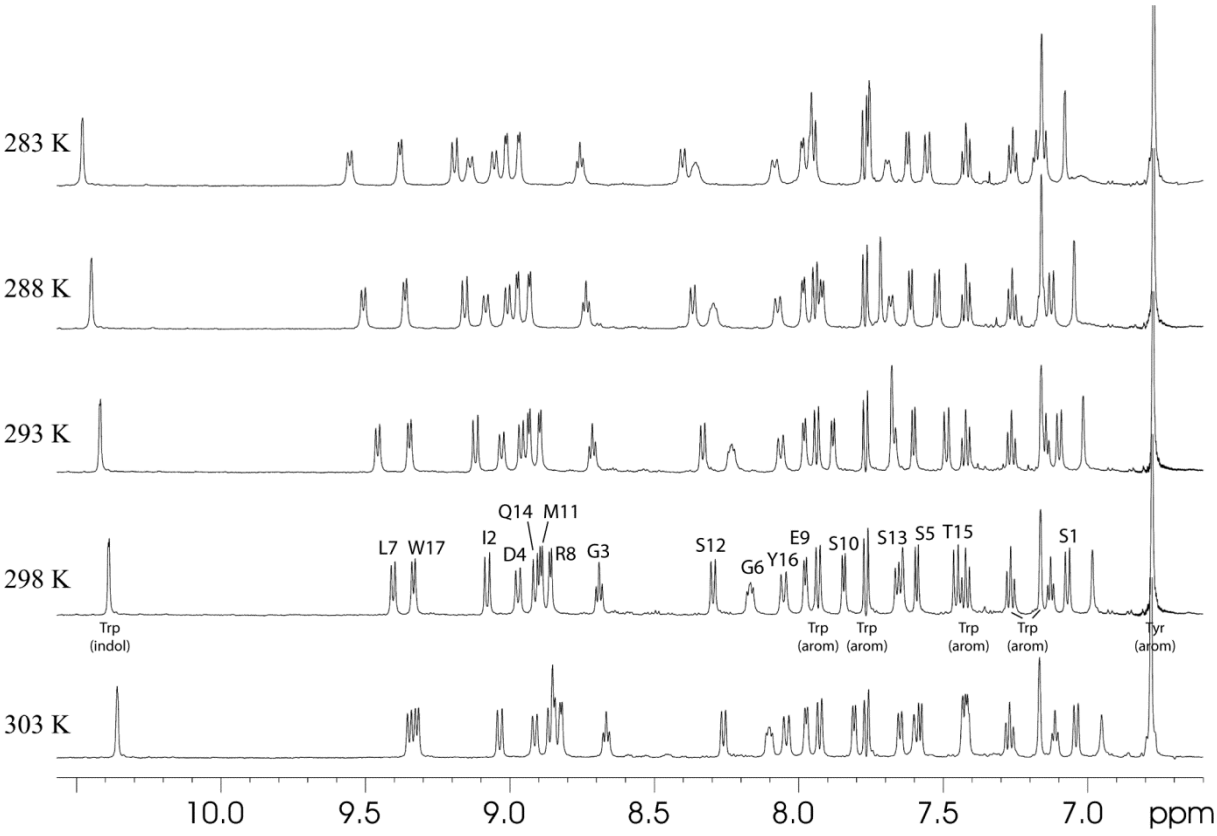
Supporting figure 1 – heat stability of **(a)** caulonodin IV and **(b)** caulonodin VI, heat and carboxypeptidase stability of **(c)** caulonodin V and **(d)** caulonodin VII shown as UV traces and mass signals measured on the low-resolution HPLC-MS system



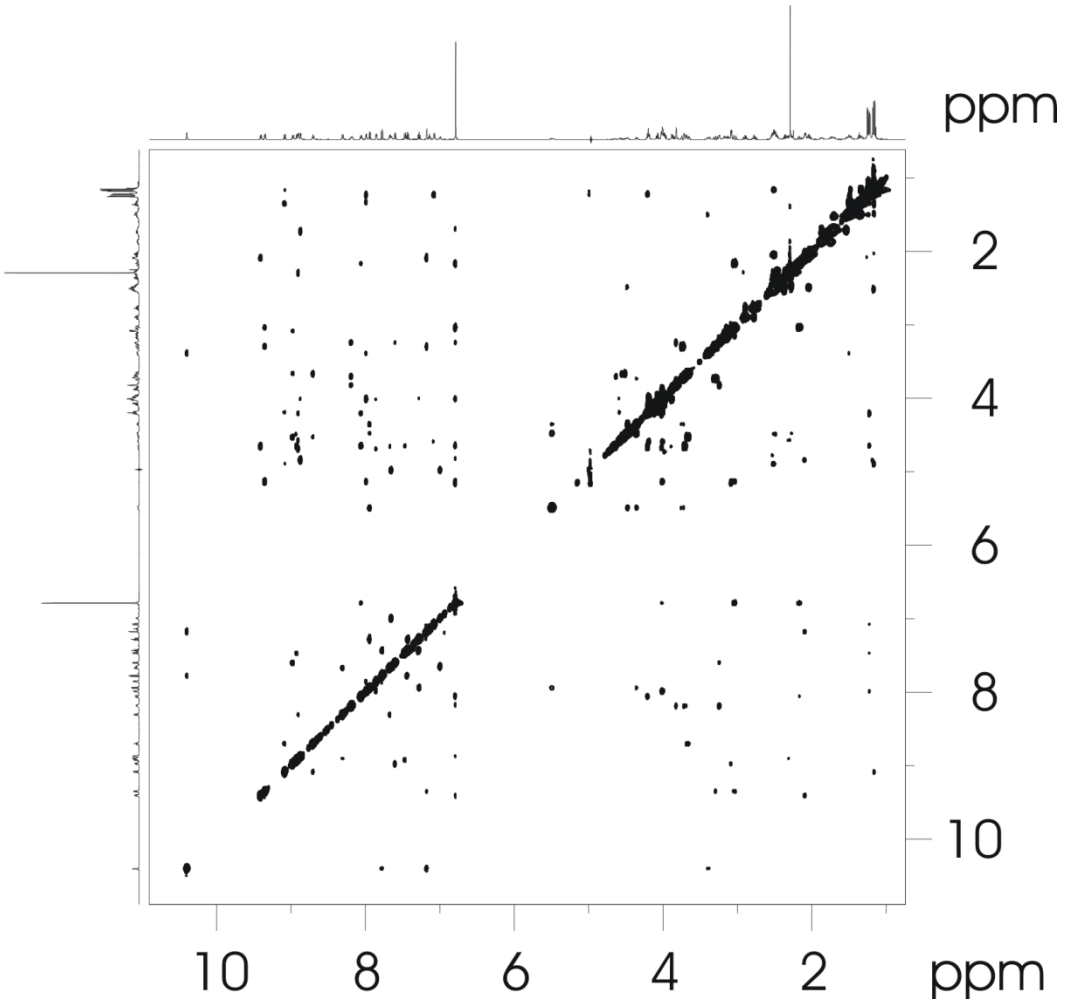
Supporting figure 2 – heat stable mutants of **(a)** caulonodin IV, **(b)** caulonodin VI and **(c)** caulonodin VII shown as total ion current (TIC), extracted ion current (EIC) and UV traces measured on the high-resolution HPLC-MS system



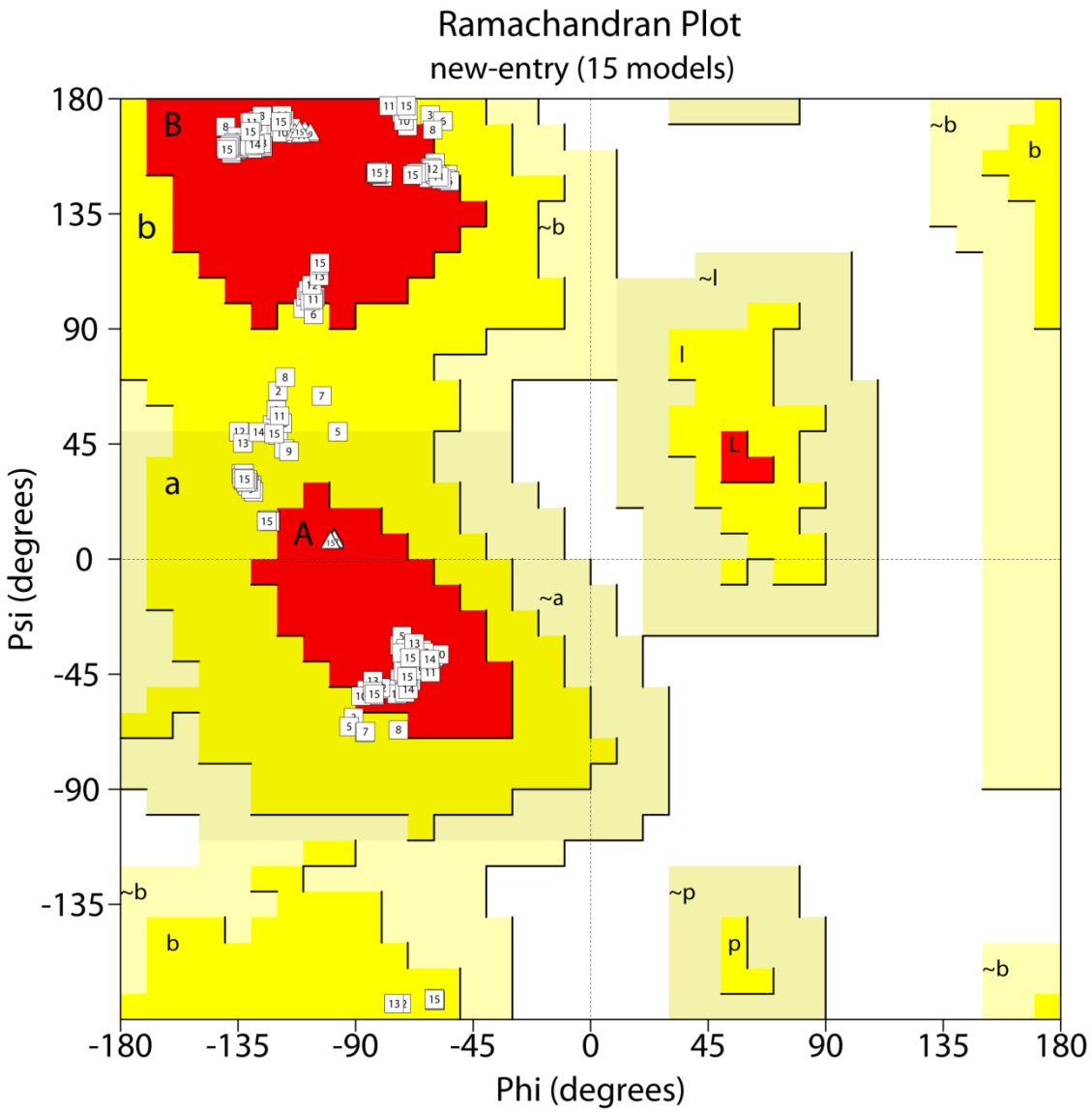
Supporting figure 3 – 3D fold predictions of the caulonodins **(a)** IV, **(b)** V, **(c)** VI and **(d)** VII in a schematic representation based on the plug mutants; aa involved in the formation of the macrolactam ring shown in red, other aa of the ring shown in yellow, aa in the tail shown in blue, plug aa shown in green.



Supporting figure 4 – ^{1}H NMR spectra of caulonodin V in H_2O/D_2O (9:1) at variable temperatures in the range from 10.6 to 6.6 ppm



Supporting figure 5 – 2D ^{1}H NOESY spectrum of caulonodin V at 300 ms mixing time

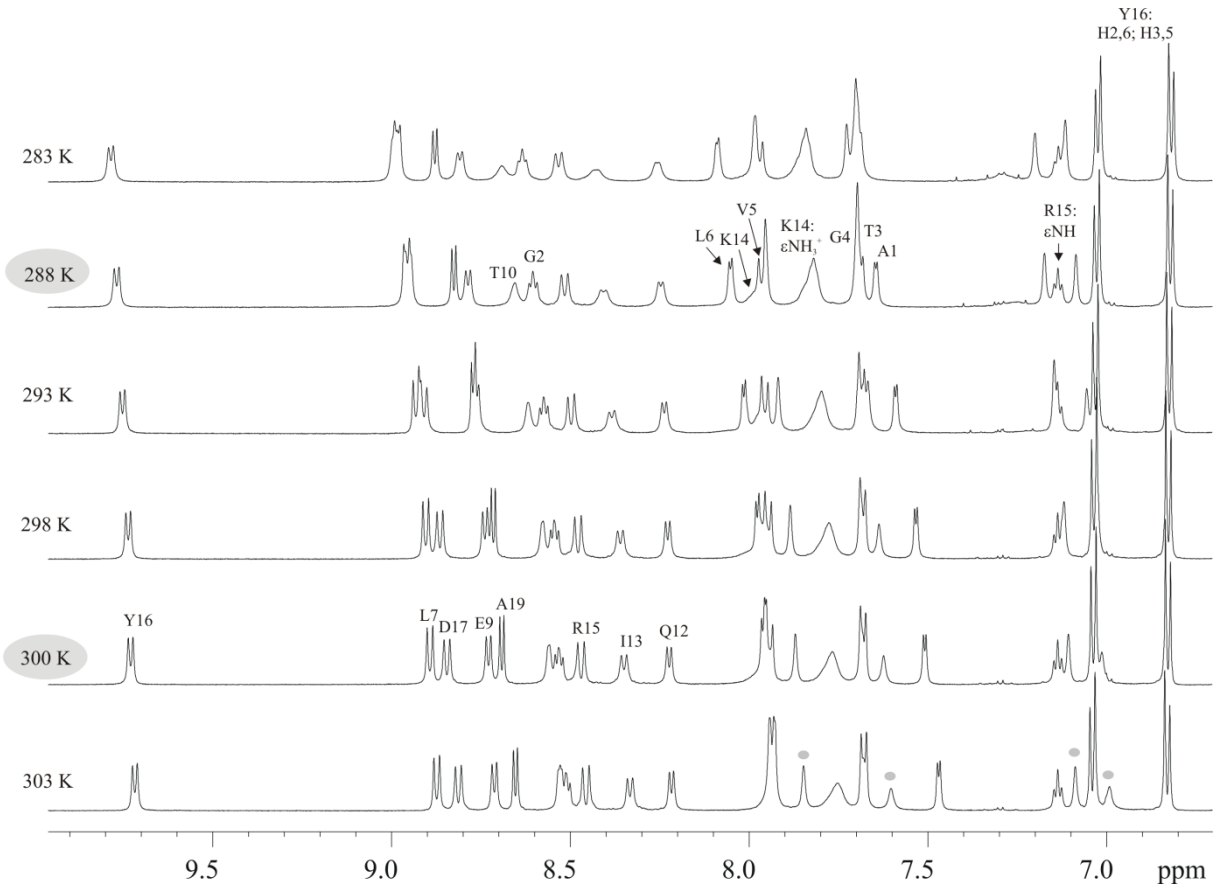


Plot statistics

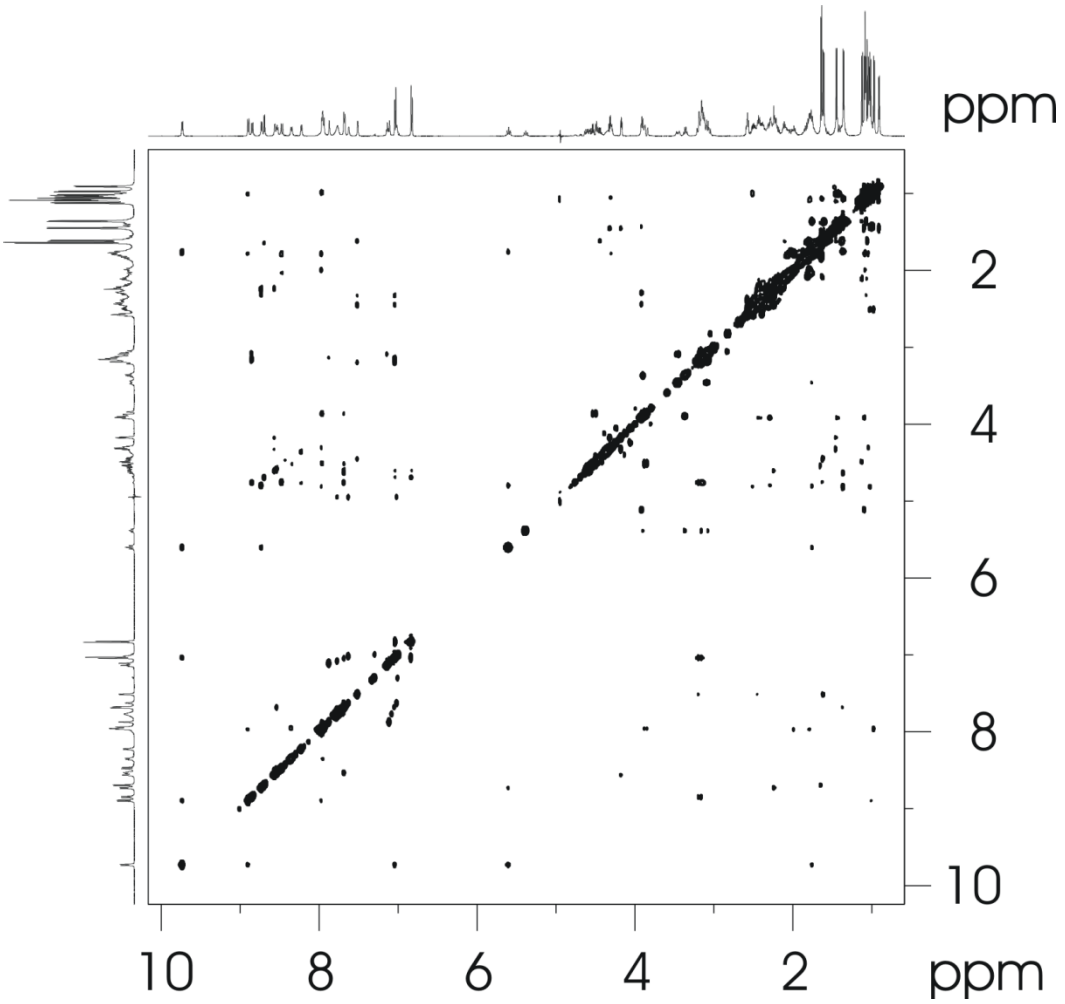
Residues in most favored regions [A,B,L]	132	62.9%
Residues in additional allowed regions [a,b,l,p]	78	37.1%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	210	100.0%
Number of end-residues (excl. Gly and Pro)	15	
Number of glycine residues (shown as triangles)	30	
Number of proline residues	15	
Total number of residues	270	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.
Model numbers shown inside each data point.

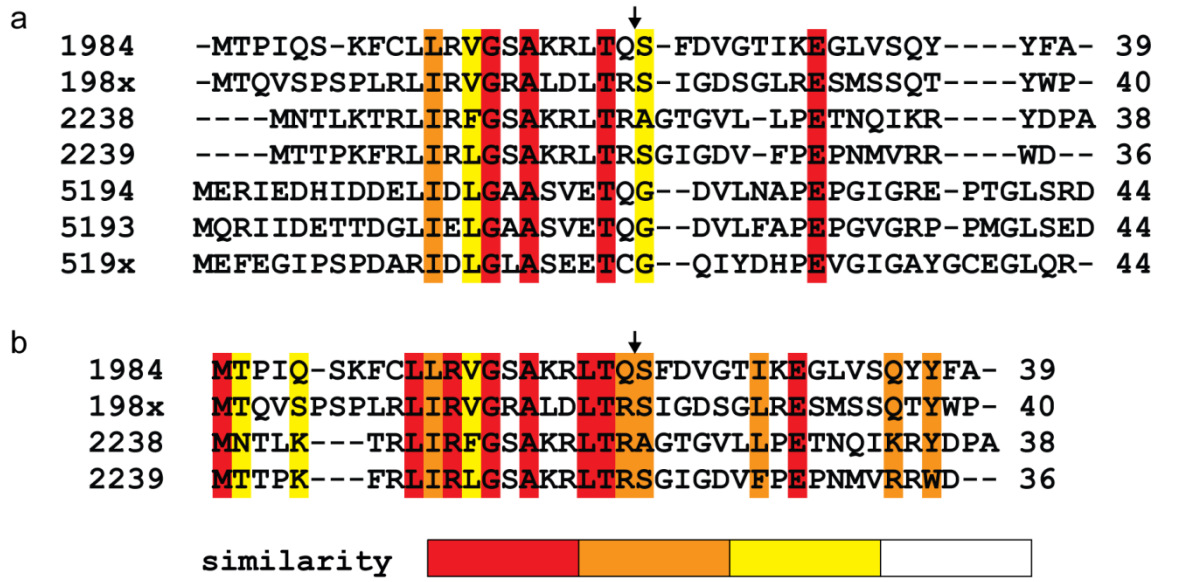
Supporting figure 6 – Quality of the structure determination. Ramachandran Plot of the 15 minimum energy structures showing all residues in most favored and additionally allowed regions.



Supporting figure 7 – ^{1}H NMR spectra of caulonodin VI in H_2O/D_2O (9:1) at variable temperatures in the range from 10.0 to 6.7 ppm; labeling according to signal assignments at 288 and 300 K; circles denote side chain NH2 of N11 and Q12.



Supporting figure 8 – 2D ^{1}H NOESY spectrum of caulonodin VI at 300 ms mixing time



Supporting figure 9 – Multiple alignments using the gonnet protein weight matrix of **(a)** all seven caulonodin precursors and **(b)** the four chromosomal precursors; coloring according to similarity from red for identical and orange to yellow and white for different; arrow marks proteolytic cleavage site between leader and core peptide

Name	Start	*p*-value	Sites
Caulonodin_I	10	8.15e-15	MERIEDHID DELIDLGAASVETQ
Caulonodin_II	10	7.81e-14	MQRIIDETT DGLIELGAASVETQ
Caulonodin_VII	6	2.77e-12	MTTPK FRLIRLGSAKRLTR
Sphingopyxin_II	8	5.99e-12	MERTEVI EEVIDLGKASVETK
Caulosegnin_II	5	1.24e-11	MTKT HRLIRLGDAQRLTQ
Sphingopyxin_I	5	1.43e-11	MKDF NELIDLGAISVETR
Astexin-1	15	1.43e-11	IISETVQPKT AGLIVLGKASAETR
Sphingonodin_I	5	5.42e-11	MERD NDVIELGAVSVETK
Xanthomonin_III	16	8.89e-11	DSAPSQEDHG KDIIVLGIASVETQ
Caulosegnin_I	10	1.61e-10	MTKKNATQA PRLVRVGDAHRLTQ
Astexin-3	12	2.28e-10	TKRTTIAARR VGLIDLGKATRQTK
Caulonodin_IV	8	3.19e-10	MTPIQSK FCLLRVGSAKRLTQ
Syanodin	10	4.43e-10	MERNSEDRR DDVVELGAVSVETK
Sphingonodin_II	10	4.94e-10	MDRHDNSEV DEIIDLGTASAVTQ
Xanthomonin_II	15	5.49e-10	NNDARTTALD QDLIVLGVASLDTQ
Caulonodin_VI	6	5.49e-10	MNTLK TRLIRFGSAKRLTR
Rhodanodin	21	1.03e-09	TNENIRSNAQ DDVIELSVASVETK
Rubrivinodin	12	1.26e-09	KEFAMDEELE LEIVDLGDAKELTQ
Caulosegnin_III	5	1.54e-09	MTSR FQLLRLGKADRLTR
SSV-2083	23	1.87e-09	MTSTDELYEA PELIEIGDYAELTR
Caulonodin_III	10	2.51e-09	MEFEGIPSP DARIDLGLASEETC
Xanthomonin_I	13	2.16e-08	SNDTTHSDAS NEITVLGVASTDTK
Caulonodin_V	9	6.72e-08	MTQVSPSP LRLIRVGRALDLTR
Astexin-2	12	1.81e-07	RTYNRSLPAR AGLTDLGKVTTHTK
Zucinodin	10	4.61e-07	MTRLLNLMS VRLLGFGSAKAATN
Lariatin	13	8.55e-07	SQPSKKTYNA PSLVQRGKFARTTA
SRO15-2005	15	4.01e-06	IKQQKKAYVK PSMFQQGDFSKKTA

Supporting figure 10 – Alignment of the 27 leader peptides from the 30 known, functional lasso peptide biosynthetic gene clusters, wherein the motif was identified using the MEME algorithm; leader peptide of caulonodin V marked with a blue arrow, of rhodanodin with a red arrow

Supporting table 1 – Oligonucleotide primers

name	sequence
CK31_198x_A1_NdeI_FP	GGATTAC CAT ATG ACC CCG ATC CAA TCC AAG TTC TGC CTT CTG
CK31_198x_C_XhoI_RP	ATAT CTC GAG TCA TGA GAC CCT GAG CTT GGC TTC CCA GG
CK31_2238_A1_NdeI_FP	GGATTAC CAT ATG AAT ACG CTC AAG ACC CGG CTG ATC CG
CK31_2241_C_XhoI_RP	ATAT CTC GAG TTA GGC GCC GCG CGC TTC CCA G
CK31_198xB_SLIM_FPS	CTT TGG CTC GCC CAG GAC GTC CAT GC
CK31_198xB_rbs_SLIM_FPL	AG*AGGAGA*AATTAACC ATG CCC CTT TGG CTC GCC CAG GAC GTC CAT GC
CK31_198xA2_SLIM_RPS	TTA GGG CCA GTA GGT CTG CGA GCT CAT G
CK31_198xA2_rbs_SLIM_RPL	GGG CAT GGTTAATT*TCTCCT*CT TTA GGG CCA GTA GGT CTG CGA GCT CAT G
CK31_198xA2Del_SLIM_RPS	TCA GGC GAA GTA GTA TTG GCT GAC GAG ACC
CK31_198xA2Del_SLIM_RPL	GGG CAT GGTTAATT*TCTCCT*CT TCA GGC GAA GTA GTA TTG GCT GAC GAG ACC
CK31_198xA1Del_SLIM_FPS	ATG ACT CAA GTC TCT CCC TCG CCC CTG C
CK31_198xA1Del_SLIM_FPL	CTT TAA GAA GGA GAT ATA CAT ATG ACT CAA GTC TCT CCC TCG CCC CTG C
CK31_bothA1Del_SLIM_RPS	TTA AAC AAA ATT ATT TCT AGA GGG GAA TTG TTA TCC GCT CAC
CK31_bothA1Del_SLIM_RPL	ATG TAT ATC TCC TTC TTA AAG TTA AAC AAA ATT ATT TCT AGA GGG GAA TTG TTA TCC GCT CAC
CK31_22xxB_SLIM_FPS	CTA ACC TGG CGG CCC GGC GTT CAC
CK31_22xxB_SLIM_FPL	AG*AGGAGA*AATTAACC ATG ACG CTA ACC TGG CGG CCC GGC GTT CAC
CK31_22xxA2_SLIM_RPS	TCA GTC CCA ACG GCG AAC CAT GTT GGG C
CK31_22xxA2_SLIM_RPL	CGT CAT GGTTAATT*TCTCCT*CT TCA GTC CCA ACG GCG AAC CAT GTT GGG C
CK31_22xxA2Del_SLIM_RPS	TCA GGC CGG GTC GTA CCG CTT GAT CTG
CK31_22xxA2Del_SLIM_RPL	GGG CAT GGTTAATT*TCTCCT*CT TCA GGC CGG GTC GTA CCG CTT GAT CTG
CK31_22xxA1Del_SLIM_FPS	ATG ACG ACT CCC AAG TTT CGA CTG ATC CGT CTG
CK31_22xxA1Del_SLIM_FPL	CTT TAA GAA GGA GAT ATA CAT ATG ACG ACT CCC AAG TTT CGA CTG ATC CGT CTG
198xA12_SLIM_FPS	TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
1984A1_SLIM_RPS	GAG ACC TTC CTT GAT GGT GCC GAC ATC GAA G
1984A1_Y15A_FPL	GTC AGC CAA **GC**C TAC TTC GCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
1984A1_Y15A_RPL	GGC GAA GTA G**GC** TTG GCT GAC GAG ACC TTC CTT GAT GGT GCC GAC ATC GAA G
1984A1_Y16A_FPL	GTC AGC CAA TAC **GC**C TTC GCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
1984A1_Y16A_RPL	GGC GAA G**GC** GTA TTG GCT GAC GAG ACC TTC CTT GAT GGT GCC GAC ATC GAA G
1984A1_F17A_FPL	GTC AGC CAA TAC TAC **GC**C GCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
1984A1_F17A_RPL	GGC G**GC** GTA GTA TTG GCT GAC GAG ACC TTC CTT GAT GGT GCC GAC ATC GAA G
1984A1_Y15W_FPL	GTC AGC CAA T**GG** TAC TTC GCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
1984A1_Y15W_RPL	GGC GAA GTA **CC**A TTG GCT GAC GAG ACC TTC CTT GAT GGT GCC GAC ATC GAA G
1984A1_Y16W_FPL	GTC AGC CAA TAC T**GG** TTC GCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
1984A1_Y16W_RPL	GGC GAA **CC**A GTA TTG GCT GAC GAG ACC TTC CTT GAT GGT GCC GAC ATC GAA G
1984A1_F17W_FPL	GTC AGC CAA TAC TAC T**GG** GCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C

name	sequence
1984A1_F17W_RPL	GGC **CC**A GTA GTA TTG GCT GAC GAG ACC TTC CTT GAT GGT GCC GAC ATC GAA G
198xA2_SLIM_RPS	CAT GGA CTC CCG CAG ACC GGA GTC G
198xA2_T15A_FPL	AGC TCG CAG **G**CC TAC TGG CCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
198xA2_T15A_RPL	GGG CCA GTA GG**C** CTG CGA GCT CAT GGA CTC CCG CAG ACC GGA GTC G
198xA2_Y16A_FPL	AGC TCG CAG ACC **GC**C TGG CCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
198xA2_Y16A_RPL	GGG CCA G**GC** GGT CTG CGA GCT CAT GGA CTC CCG CAG ACC GGA GTC G
198xA2_Y16F_FPL	AGC TCG CAG ACC T**T**C TGG CCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
198xA2_Y16F_RPL	GGG CCA G**A**A GGT CTG CGA GCT CAT GGA CTC CCG CAG ACC GGA GTC G
198xA2_W17A_FPL	AGC TCG CAG ACC TAC **GC**G CCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
198xA2_W17A_RPL	GGG C**GC** GTA GGT CTG CGA GCT CAT GGA CTC CCG CAG ACC GGA GTC G
198xA2_W17F_FPL	AGC TCG CAG ACC TAC T**TC** CCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
198xA2_W17F_RPL	GGG **GA**A GTA GGT CTG CGA GCT CAT GGA CTC CCG CAG ACC GGA GTC G
198xA2_W17Y_FPL	AGC TCG CAG ACC TAC T**AC** CCC TGA AGA GGA GAA ATT AAC CAT GCC CCT TTG GCT C
198xA2_W17Y_RPL	GGG **GT**A GTA GGT CTG CGA GCT CAT GGA CTC CCG CAG ACC GGA GTC G
CK31_223xA12_SLIM2_FPS	TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2238_SLIM2_RPS	CTG GTT GGT TTC CGG GAG AAG CAC CCC G
CK31_2238_K14A_FPL	ATC **GC**G CGG TAC GAC CCG GCC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2238_K14A_RPL	GGC CGG GTC GTA CCG C**GC** GAT CTG GTT GGT TTC CGG GAG AAG CAC CCC G
CK31_2238_R15A_FPL	ATC AAG **GC**G TAC GAC CCG GCC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2238_R15A_RPL	GGC CGG GTC GTA C**GC** CTT GAT CTG GTT GGT TTC CGG GAG AAG CAC CCC G
CK31_2238_Y16A_FPL	ATC AAG CGG **GC**C GAC CCG GCC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2238_Y16A_RPL	GGC CGG GTC G**GC** CCG CTT GAT CTG GTT GGT TTC CGG GAG AAG CAC CCC G
CK31_2238_K14W_FPL	ATC **TG**G CGG TAC GAC CCG GCC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2238_K14W_RPL	GGC CGG GTC GTA CCG C**CA** GAT CTG GTT GGT TTC CGG GAG AAG CAC CCC G
CK31_2238_R15W_FPL	ATC AAG **T**GG TAC GAC CCG GCC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2238_R15W_RPL	GGC CGG GTC GTA CC**A** CTT GAT CTG GTT GGT TTC CGG GAG AAG CAC CCC G
CK31_2238_Y16W_FPL	ATC AAG CGG T**GG** GAC CCG GCC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2238_Y16W_RPL	GGC CGG GTC **CC**A CCG CTT GAT CTG GTT GGT TTC CGG GAG AAG CAC CCC G
CK31_2239_SLIM_RPS	GGG CTC GGG GAA CAC GTC CCC GAT CC
CK31_2239_R14A_FPL	AAC ATG GTT **GC**C CGT TGG GAC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2239_R14A_RPL	GTC CCA ACG G**GC** AAC CAT GTT GGG CTC GGG GAA CAC GTC CCC GAT CC
CK31_2239_R14W_FPL	AAC ATG GTT **TG**G CGT TGG GAC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC

name	sequence
CK31_2239_R14W_RPL	GTC CCA ACG **C**C**A** AAC CAT GTT GGG CTC GGG GAA CAC GTC CCC GAT CC
CK31_2239_R15A_FPL	AAC ATG GTT CGC **GC**T TGG GAC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2239_R15A_RPL	GTC CCA A**GC** GCG AAC CAT GTT GGG CTC GGG GAA CAC GTC CCC GAT CC
CK31_2239_R15K_FPL	AAC ATG GTT CGC **AAA** TGG GAC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2239_R15K_RPL	GTC CCA **TTT** GCG AAC CAT GTT GGG CTC GGG GAA CAC GTC CCC GAT CC
CK31_2239_R15W_FPL	AAC ATG GTT CGC **T**G**G** TGG GAC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2239_R15W_RPL	GTC CCA **C**C**A** GCG AAC CAT GTT GGG CTC GGG GAA CAC GTC CCC GAT CC
CK31_2239_W16A_FPL	AAC ATG GTT CGC CGT **GC**G GAC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2239_W16A_RPL	GTC C**GC** ACG GCG AAC CAT GTT GGG CTC GGG GAA CAC GTC CCC GAT CC
CK31_2239_W16Y_FPL	AAC ATG GTT CGC CGT T**AC** GAC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2239_W16Y_RPL	GTC **GT**A ACG GCG AAC CAT GTT GGG CTC GGG GAA CAC GTC CCC GAT CC
CK31_2239_W16F_FPL	AAC ATG GTT CGC CGT T**TC** GAC TGA AGA GGA GAA ATT AAC CAT GAC GCT AAC CTG GC
CK31_2239_W16F_RPL	GTC **GA**A ACG GCG AAC CAT GTT GGG CTC GGG GAA CAC GTC CCC GAT CC
198xA2_SLIM_RING_FPS	TCC GGT CTG CGG GAG TCC ATG AGC TC
198xA2_SLIM_RING_RPS	ATC GAG GGC GCG CCC GAC ACG
198xA2_S1G_FP	CTG ACC CGT **GG**G ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xA2_S1G_RP	GTC GCC GAT C**CC** ACG GGT CAG ATC GAG GGC GCG CCC GAC ACG
198xA2 _S1A_FP	CTG ACC CGT **G**CG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xA2 _S1A_RP	GTC GCC GAT CG**C** ACG GGT CAG ATC GAG GGC GCG CCC GAC ACG
198xA2 _S1C_FP	CTG ACC CGT T**GC** ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xA2 _S1C_RP	GTC GCC GAT **GC**A ACG GGT CAG ATC GAG GGC GCG CCC GAC ACG
198xA2 _S1F_FP	CTG ACC CGT T**TC** ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xA2 _S1F_RP	GTC GCC GAT **GA**A ACG GGT CAG ATC GAG GGC GCG CCC GAC ACG
198xA2_S1V_FP	CTG ACC CGT **GT**G ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xA2_S1V_RP	GTC GCC GAT C**AC** ACG GGT CAG ATC GAG GGC GCG CCC GAC ACG
198xA2_S1T_FP	CTG ACC CGT **A**CG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xA2_S1T_RP	GTC GCC GAT CG**T** ACG GGT CAG ATC GAG GGC GCG CCC GAC ACG
CK31_2238_SLIM1_FPS	GTG CTT CTC CCG GAA ACC AAC CAG ATC AAG C
CK31_2238_SLIM1_RPS	GCG CTT GGC CGA GCC GAA ACG G
CK31_2238_A1G_FPL	CTG ACC CGC G**G**C GGC ACC GGG GTG CTT CTC CCG GAA ACC AAC CAG ATC AAG C
CK31_2238_A1G_RPL	CCC GGT GCC G**C**C GCG GGT CAG GCG CTT GGC CGA GCC GAA ACG G
CK31_2238_A1S_FPL	CTG ACC CGC **T**CC GGC ACC GGG GTG CTT CTC CCG GAA ACC AAC CAG ATC AAG C

name	sequence
CK31_2238_A1S_RPL	CCC GGT GCC GG**A** GCG GGT CAG GCG CTT GGC CGA GCC GAA ACG G
CK31_2238_A1C_FPL	CTG ACC CGC **TG**C GGC ACC GGG GTG CTT CTC CCG GAA ACC AAC CAG ATC AAG C
CK31_2238_A1C_RPL	CCC GGT GCC G**CA** GCG GGT CAG GCG CTT GGC CGA GCC GAA ACG G
CK31_2238_A1F_FPL	CTG ACC CGC **TT**C GGC ACC GGG GTG CTT CTC CCG GAA ACC AAC CAG ATC AAG C
CK31_2238_A1F_RPL	CCC GGT GCC G**AA** GCG GGT CAG GCG CTT GGC CGA GCC GAA ACG G
CK31_2238_A1V_FPL	CTG ACC CGC G**T**C GGC ACC GGG GTG CTT CTC CCG GAA ACC AAC CAG ATC AAG C
CK31_2238_A1V_RPL	CCC GGT GCC G**A**C GCG GGT CAG GCG CTT GGC CGA GCC GAA ACG G
CK31_2238_A1T_FPL	CTG ACC CGC **A**CC GGC ACC GGG GTG CTT CTC CCG GAA ACC AAC CAG ATC AAG C
CK31_2238_A1T_RPL	CCC GGT GCC GG**T** GCG GGT CAG GCG CTT GGC CGA GCC GAA ACG G
198xLeaderDel1_RtH_RP	CAT ATG TAT ATC TCC TTC TTA AAG TTA AAC AAA ATT ATT TCT AGA GGG GAA TTG
198xLeaderDel1_RtH_FP	CTG CGG CTC ATC CGT GTC GGG CG
198xLeaderDel12_RtH_FP	GGG CGC GCC CTC GAT CTG ACC CG
198xLeaderDel123_RtH_FP	ACC CGT TCG ATC GGC GAC TCC GGT CTG
198xLeaderDel2FPS	CTC GAT CTG ACC CGT TCG ATC GGC GAC
198xLeaderDel2RPS	AGA GAC TTG AGT CAT ATG TAT ATC TCC TTC TTA AAG TTA AAC AAA ATT ATT TC
198xLeaderDel2FPL	CCC TCG CCC GGG CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC
198xLeaderDel2RPL	GGC GCG CCC GGG CGA GGG AGA GAC TTG AGT CAT ATG TAT ATC TCC TTC TTA AAG TTA AAC AAA ATT ATT TC
198xLeaderDel3FPS	ATC GGC GAC TCC GGT CTG CGG GAG TCC
198xLeaderDel3RPS	GAG CCG CAG GGG CGA GGG AGA GAC TTG
198xLeaderDel3FPL	ATC CGT GTC ACC CGT TCG ATC GGC GAC TCC GGT CTG CGG GAG TCC
198xLeaderDel3RPL	CGA ACG GGT GAC ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198xLeaderBlock1 SLIM FPS	GAT CTG ACC CGT TCG ATC GGC GAC TCC
198xLeaderBlock1 SLIM RPS	CCG CAG GGG CGA GGG AGA GAC TTG
198xLeaderLIVAtoRRRR FPL	**CGC CGT** CGT **CGC** GGG CGC **CGT** CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198xLeaderLIVAtoRRRR RPL	GAG **ACG** GCG CCC **GCG** ACG **ACG GCG** CCG CAG GGG CGA GGG AGA GAC TTG
198xLeaderLIVAtoEEEE FPL	**GAA GAG** CGT **GAA** GGG CGC **GAA** CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198xLeaderLIVAtoEEEE RPL	GAG **TTC** GCG CCC **TTC** ACG **CTC TTC** CCG CAG GGG CGA GGG AGA GAC TTG
198x_LI-11-12toRR FPL	**CGC CGT** CGT GTC GGG CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_LI-11-12toRR RPL	GAG GGC GCG CCC GAC ACG **ACG GCG** CCG CAG GGG CGA GGG AGA GAC TTG
198x_LI-11-12toEE FPL	**GAA GAG** CGT GTC GGG CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_LI-11-12toEE RPL	GAG GGC GCG CCC GAC ACG **CTC TTC** CCG CAG GGG CGA GGG AGA GAC TTG
198x_LI-11-12toQQ FPL	**CAA CAG** CGT GTC GGG CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_LI-11-12toQQ RPL	GAG GGC GCG CCC GAC ACG **CTG TTG** CCG CAG GGG CGA GGG AGA GAC TTG
198x_VA-9-6toRR FPL	CTC ATC CGT **CGC** GGG CGC **CGT** CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC

name	sequence
198x_VA-9-6toRR RPL	GAG **ACG** GCG CCC **GCG** ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_VA-9-6toEE FPL	CTC ATC CGT **GAA** GGG CGC **GAA** CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_VA-9-6toEE RPL	GAG **TTC** GCG CCC **TTC** ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_VA-9-6toQQ FPL	CTC ATC CGT **CAG** GGG CGC **CAG** CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_VA-9-6toQQ RPL	GAG **CTG** GCG CCC **CTG** ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_V-9R FPL	CTC ATC CGT **CGC** GGG CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_V-9R RPL	GAG GGC GCG CCC **GCG** ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_V-9E FPL	CTC ATC CGT **GAA** GGG CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_V-9E RPL	GAG GGC GCG CCC **TTC** ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_A-6R FPL	CTC ATC CGT GTC GGG CGC **CGT** CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_A-6R RPL	GAG **ACG** GCG CCC GAC ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_A-6E FPL	CTC ATC CGT GTC GGG CGC **GAA** CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_A-6E RPL	GAG **TTC** GCG CCC GAC ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_G-8F FPL	CTC ATC CGT GTC **TTC** CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_G-8F RPL	GAG GGC GCG **GAA** GAC ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_G-8R FPL	CTC ATC CGT GTC **CGT** CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_G-8R RPL	GAG GGC GCG **ACG** GAC ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_G-8E FPL	CTC ATC CGT GTC **GAA** CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_G-8E RPL	GAG GGC GCG **TTC** GAC ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_G-8A FPL	CTC ATC CGT GTC **GCG** CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_G-8A RPL	GAG GGC GCG **CGC** GAC ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_G-8V FPL	CTC ATC CGT GTC **GTG** CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TC
198x_G-8V RPL	GAG GGC GCG **CAC** GAC ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_G-8P FPL	CTC ATC CGT GTC **CCG** CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_G-8P RPL	GAG GGC GCG **CGG** GAC ACG GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_R-10A FPL	CTC ATC **GCT** GTC GGG CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_R-10A RPL	GAG GGC GCG CCC GAC **AGC** GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_R-10F FPL	CTC ATC **TTT** GTC GGG CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC
198x_R-10F RPL	GAG GGC GCG CCC GAC **AAA** GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198x_R-10E FPL	CTC ATC **GAA** GTC GGG CGC GCC CTC GAT CTG ACC CGT TCG ATC GGC GAC TCC

name	sequence
198x_R-10E RPL	GAG GGC GCG CCC GAC **TTC** GAT GAG CCG CAG GGG CGA GGG AGA GAC TTG
198xLeaderBlock 2SLIM FPS	TCC GGT CTG CGG GAG TCC ATG AGC TC
198xLeaderBlock 2SLIM RPS	GGC GCG CCC GAC ACG GAT GAG C
198xR-1A_ FPL	CTC GAT CTG ACC **GCC** TCG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xR-1A_ RPL	GTC GCC GAT CGA **GGC** GGT CAG ATC GAG GGC GCG CCC GAC ACG GAT GAG C
198xR-1Q_ FPL	CTC GAT CTG ACC **CAG** TCG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xR-1Q_ RPL	GTC GCC GAT CGA **CTG** GGT CAG ATC GAG GGC GCG CCC GAC ACG GAT GAG C
198xR-1N_ FPL	CTC GAT CTG ACC **AAC** TCG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xR-1N_ RPL	GTC GCC GAT CGA **GTT** GGT CAG ATC GAG GGC GCG CCC GAC ACG GAT GAG C
198xR-1K_ FPL	CTC GAT CTG ACC **AAG** TCG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xR-1K_ RPL	GTC GCC GAT CGA **CTT** GGT CAG ATC GAG GGC GCG CCC GAC ACG GAT GAG C
198xR-1H_ FPL	CTC GAT CTG ACC **CAC** TCG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xR-1H_ RPL	GTC GCC GAT CGA **GTG** GGT CAG ATC GAG GGC GCG CCC GAC ACG GAT GAG C
198xR-1F_ FPL	CTC GAT CTG ACC **TTC** TCG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xR-1F_ RPL	GTC GCC GAT CGA **GAA** GGT CAG ATC GAG GGC GCG CCC GAC ACG GAT GAG C
198xR-1E_ FPL	CTC GAT CTG ACC **GAA** TCG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xR-1E_ RPL	GTC GCC GAT CGA **TTC** GGT CAG ATC GAG GGC GCG CCC GAC ACG GAT GAG C
198xT-2A_ FPL	CTC GAT CTG **GCC** CGT TCG ATC GGC GAC TCC GGT CTG CGG GAG TCC ATG AGC TC
198xT-2A_ RPL	GTC GCC GAT CGA ACG **GGC** CAG ATC GAG GGC GCG CCC GAC ACG GAT GAG C
FP_Rhothi_Leader_SLIM	GGA GTA TTG CCC ATC GGC AAT GAG TTC ATG GGC CAC G
RP_Rhothi_Leader_SLIM	GAG CTC GAT GAC ATC GTC CTG CGC GTT GCT G
FPTail_Rhothi_S-8G_SLIM	**GGC** GTC GCC AGT GTC GAG ACC AAA GGA GTA TTG CCC ATC GGC AAT GAG TTC ATG GGC CAC G
RPTail_Rhothi_S-8G_SLIM	TTT GGT CTC GAC ACT GGC GAC **GCC** GAG CTC GAT GAC ATC GTC CTG CGC GTT GCT G

Introduced restriction enzyme sites are underlined; introduced ribosomal binding sites are shown in italics; introduced mutations are shown in bold

Supporting table 2 – M9 Vitamin Mix

component	amount
biotin	0.2 g
choline chloride	1.0 g
disodium adenosine 5’-triphosphate	0.3 g
folic acid	1.0 g
myo-inositol	2.0 g
nicotinamide	1.0 g
panthothenic acid	1.0 g
pyridoxal hydrochloride	1.0 g
riboflavin	0.1 g
thiamine	1.0 g
H_2O	ad 300 mL

10 M NaOH were added to the resulting suspension until everything was dissolved. The solution was then sterile filtered and stored at 4°C

Supporting Table 3 – Tail sequences of the caulonodins with possible plug residues highlighted

peptide	14	15	16	17	18	19
caulonodin IV	Q	**Y**	**Y**	**F**	A	-
caulonodin V	Q	**T**	**Y**	**W**	P	-
caulonodin VI	**K**	**R**	**Y**	D	P	A
caulonodin VII	**R**	**R**	**W**	D	-	-

Positions picked for mutation in bold black, most likely plugs in red

Supporting Table 4 – Production amounts of all constructed caulonodin and rhodanodin mutants; mutant numbering according to figure 1a, positive numbers for mutations in the core peptide, negative numbers for mutations in the leader peptide.

lasso peptide mutant	relative production to corresponding wild type in %	heat stable
caulonodin IV WT	100 ± 15	no
Y15A	49 ± 2	no
Y15W	167 ± 12	no
Y16A	14.3 ± 0.7	no
Y16W	84 ± 6	no
F17A	not detected	-
F17W	132 ± 9	yes
caulonodin V WT	100 ± 23	no
S1G	4.0 ± 0.3	-
S1A	132.2 ± 1.6	-
S1C	detected	-
S1V	detected	-
S1T	detected	-
S1F	detected	-
T15A	2.7 ± 0.1	-
Y16A	detected	-
Y16F	detected	-
W17A	detected	-
W17F	detected	-
W17Y	8.5 ± 0.6	-
Del1	7.8 ± 0.5	-
Del2	detected	-
Del3	detected	-
Del12	not detected	-
Del123	not detected	-
L-12R_I-11R_V-9R_A-6R	not detected	-
L-12E_I-11E_V-9E_A-6E	not detected	-
L-12R_I-11R	0.3 ± 0.1	-
L-12E_I-11E	0.4 ± 0.1	-
L-12Q_I-11Q	7.6 ± 0.1	-
V-9R_A-6R	not detected	-
V-9E_A-6E	not detected	-
V-9Q_A-6Q	not detected	-
R-10A	97.8 ± 1.8	-
R-10F	71.2 ± 1.5	-
R-10E	63.0 ± 1.4	-
V-9E	2.0 ± 0.1	-
V-9R	3.6 ± 0.2	-
G-8A	detected	-

lasso peptide mutant	relative production to corresponding wild type in %	heat stable
G-8F	6.0 ± 0.5	-
G-8E	4.7 ± 0.1	-
G-8V	0.5 ± 0.1	-
G-8R	not detected	-
G-8P	0.3 ± 0.1	-
A-6E	detected	-
A-6R	detected	-
T-2A	detected	-
R-1A	80 ± 15	-
R-1F	1.0 ± 0.1	-
R-1E	33 ± 14	-
R-1Q	84 ± 3	-
R-1N	62 ± 12	-
R-1K	90 ± 8	-
R-1H	15.7 ± 1.0	-
caulonodin VI WT	100 ± 7	no
A1G	7.1 ± 0.5	-
A1S	122 ± 30	-
A1C	detected	-
A1V	detected	-
A1T	1.3 ± 0.1	-
A1F	detected	-
K14A	38 ± 6	no
K14W	22.9 ± 1.4	no
R15A	6.1 ± 0.5	no
R15W	2.0 ± 0.1	yes
Y16A	detected	-
Y16W	99 ± 14	no
caulonodin VII WT	100 ± 29	no
R14A	29.7 ± 0.6	no
R15A	0.8 ± 0.1	no
R15K	123 ± 5	no
R15W	detected	yes
W16A	not detected	-
W16F	detected	-
W16Y	73 ± 12	no
rhodanodin WT	100	-
S-8G	1460	-

Supporting table 5 – Chemical shift Assignment of ^{1}H signals (ppm) of caulonodin V in H_2O/D_2O (9:1) at 298 K

aa	NH	αH	βH	others
Ser1	7.075	4.560	4.190; 3.995	/
Ile2	9.084	4.891	2.509	γCH_3: 1.159
				γCH_2: 1.478; 1.342
				δCH_3: 1.151
Gly3	8.702	4.528; 3.664	/	/
Asp4	8.973	5.155	3.082	/
Ser5	7.600	3.240	3.821	/
Gly6	8.185	4.651; 3.701	/	/
Leu7	9.406	4.841	2.089	γCH: 2.044
				δCH_3: 1.244; 1.172
Arg8	8.872	4.007	1.863; 1.737	γCH_2: 1.700; 1.513
				δCH_2: 3.170; 3.132
				εNH: 7.140
Glu9	7.989	3.386	1.485; 1.327	γCH_2: 1.225
Ser10	7.854	4.691	4.200; 4.008	/
Met11	8.900	4.574	2.304	γCH_2: 2.894; 2.763
				εCH_3: 2.290
Ser12	8.305	4.730	4.078; 3.969	/
Ser13	7.671	4.653	4.013; 3.871	/
Gln14	8.921	4.488	2.039	γCH_2: 2.494
				εNH_2: 7.649; 6.992
Thr15	7.468	4.643	4.204	γCH_3: 1.221
Tyr16	8.058	5.136	3.033; 2.165	2. 6H: 6.787
				3. 5H: 6.787
Trp17	9.350	5.488	3.731; 3.293	2H: 7.176; 4H: 7.941
				5H: 7.278; 6H: 7.434
				7H: 7.779; NH: 10.401
Pro18	/	4.774	2.516; 2.355	γCH2: 2.267; 2.468
				δCH_2: 4.483; 4.351

Supporting table 6 – Structural statistics for the family of 15 structures selected to represent the structure of caulonodin V in H_2O/D_2O (9:1)

Restraining Constraints	Constraints Violations
Total: 166	Distance violations, >0.5 Å: 0
distance, i=j: 79	RMS deviations: 0.02 Å
distance, |i-j|=1: 35	Dihedral violations, > 5°: 0
distance, |i-j|>1: 22	RMS deviation: 1.8°
dihedral: 22	Average pairwise RMS deviation (Ile^2-Trp^{17})
hydrogen bond: 0	Backbone atoms: 0.06 Å
Constraints/residue: 9.2	All heavy atoms: 0.45 Å

Supporting table 7 – Chemical shift Assignment of ^{1}H signals (ppm) of caulonodin VI in H_2O/D_2O (9:1) at 288 K

aa	NH	αH	βH	others
Ala1	7.654	4.426	1.617	/
Gly2	8.608	4.576; 4.602		
Thr3	7.695	4.623	4.437	γCH3: 1.358
Gly4	7.700	3.844; 4.524		
Val5	7.960	4.829	2.516	γCH3: 0.964; 1.008
Leu6	8.027	4.289	amb.*	γCH2: amb.
				δCH3: 1.038; 1.078
Leu7	8.962	5.131	amb.	γCH2 amb.
				δCH3: 0.911; 1.096
Pro8	/	4.776	amb.	Amb.
Glu9	8.790	4.586	2.242	γCH2: 2.275; 2.424
Thr10	8.659	4.293	4.146	γCH3: 1.444
Asn11	amb.	amb.	amb.	amb.
Gln12	8.252	4.384	2.340; 2.547	γCH2: 2.505; 2.575
				εNH2: amb.
Ile13	8.412	4.494	amb.	γCH2: amb.
				γCH3: 1.120; δCH3: 1.056
Lys14	7.989	4.756	amb.	γCH2: amb.; δCH2: amb.
				εCH2: 3.162; $\varepsilon NH3^{+}$: 7.825
Arg15	8.521	5.593	amb.	γCH2: amb.
				δCH2: 3.077; 3.479
				NH: 7.143
Tyr16	9.772	4.727	3.135; 3.183	2, 6H: 7.027
				3, 5H: 6.825
Asp17	8.959	5.363	amb.	
Pro18	/	4.693	amb.	amb.
Ala19	8.825	4.532	1.644	

*amb. = ambiguous