Biosynthesis of the sactipeptide Ruminococcin C by the human microbiome: Mechanistic

insights into thioether bond formation by radical SAM enzymes

Clémence Balty¹, Alain Guillot¹, Laura Fradale¹, Clémence Brewee¹, Benjamin Lefranc²,

Christian Herrero³, Corine Sandström⁴, Jérôme Leprince², Olivier Berteau^{1*} & Alhosna

Benjdia¹*

¹ Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, ChemSyBio, 78350 Jouy-en-Josas, France

² Université de Normandie-Rouen, INSERM U1239, PRIMACEN, 76000 Rouen, France

³ Université Paris-Saclay, CNRS, ICMMO, 91405, Orsay, France

⁴ Department of Molecular Sciences, Uppsala BioCenter, Swedish University of Agricultural Sciences, P.O. Box 7015, SE-750 07 Uppsala, Sweden

* : Corresponding authors

| Figure S1 - Sequence alignment of RumMC1 & RumMC2 with other SPASM-domain radical SAM |
|--|
| enzymes2 |
| Figure S2 – SDS-PAGE analysis and UV-visible spectrum of RumMC1 before (pink trace) and after |
| (red trace) iron-sulfur cluster reconstitution |
| Figure S3 – LC-MS analysis of the reaction performed with the C2 ₂₈₋₆₃ peptide |
| Figure S4 - Sequence of the C1 to C5 peptides encoded by the RumC operon |
| Figure S5 - MS/MS spectrum of the P2d peptide obtained after incubation with RumMC2 |
| Figure S6 - MS/MS spectrum of the P3d peptide obtained after incubation with RumMC27 |
| Figure S7 – MS/MS spectrum of the P4d peptide obtained after incubation with RumMC27 |
| Figure S8 – MS/MS spectrum of the P6s peptide |
| Figure S9 – MS/MS spectrum of the P6d peptide |
| Figure S10 – MS/MS analysis of the C2 peptide after incubation in the presence of RumMC2 Δ RRE |
| and RRE-MC2 |
| Table S1 – Peptides used in this study and their theoretical and observed (obs) masses 10 |
| Table S2 - Theoretical mass fragments of peptide C2 ₂₈₋₆₃ [M+H] ⁺ 11 |
| Table S3 - Theoretical mass fragments of peptide $C2_{28-63}K^{53}K^{61}$ [M+H] ⁺ 12 |
| Table S4 - Theoretical mass fragments of peptide C2 ₂₈₋₆₃ A ⁶¹ [M+H] ⁺ |
| Table S5 - Theoretical mass fragments of peptide $C2_{28-63}T^{61}$ [M+H] ⁺ |
| Table S6 - Theoretical mass fragments of peptide $C2_{28-63}G^{61}$ [M+H]+15 |
| Table S7 - Theoretical mass fragments of peptide $C2_{28-63}A^{61}d4$ |
| Table S8 - Theoretical Mass fragments of peptide C228-63A61d317 |

| RumMC1/1-513 | 1 · · MKEHLKKIKETIDVERGYSLGTVFTTRNKQYMYDTGTGKVFECGVNEYRILKSL · · FEQTSLPDK · V · DG · · · · · VSEE | EL 72 |
|----------------|--|---------|
| RumMC2/1-515 | 1 MKAKEHLKKIKGTIDVERGYSLGTVFTTRNKQYMYDTGTGKVFECGANEYQILKSL - · FEQTSLPEK · V · EG · · · · · SSEE | EL 74 |
| An SME/1-370 | | |
| AIL A/1-448 | 1 MELEOMERE IN SVRVHOL REGGVI ELDVI RDNVSI SDEEVI DI NKTAVEL CMRMDGOKTAEO LI AEO CAVVDESPE | DH 70 |
| ABC/4 303 | | 011 70 |
| NIIC 1-392 | | • • |
| PqqE/1-384 | | · · . |
| SuiB/1-439 | 1 · · · · · · MRT ISED ILFRLEKFGG ILINKTNFER IELDETEAFFLYLVQNHG IEIATSFFKKEIEMGK · LERALSLNIYSDN | NI 76 |
| CteB/1-450 | 1 · · · · · · · · · · · · · · · · · · · | EI 58 |
| | | |
| RumMC1/1-513 | 73 EAAYRNIWEMVEAEHIL-QVSPNLKFVRETDETLRDLLRYDLQQVILELTEQCNMRCRYCIYNEHNEGYRNFSPKAMTWDV | AK 154 |
| RumMC2/1-515 | 75 EAAYRNIWEMIETEHIL-QISPDLKFVKETDETLRDLLRYDLQQVILELTEQ <mark>CNMRCRYC</mark> IYNEHNEGYRNFSPKAMTWEV | AK 156 |
| An SME/1-370 | 1MPPLSILLKPASSGONLKCTYCFYHSISDNRNVKSYGIMRDEV | LE 45 |
| Alb A/1-448 | 80 KOWYYDMI NMI ONKOVI OL GNRASRHTITTSG. SNEEPMPL, HATFEL THRONI KOAHOYLESSPEALGTVS | EK 154 |
| MAC/1 292 | | CV 57 |
| Nin Cor 1=352 | | UK 57 |
| PqqE/1-384 | 1DVIFAFV.GLLAELIHRUFLKUFTUSN.FLEEDRKSAEEDIQI | WL 54 |
| SuiB/1-439 | 77 EDSLNNPYETLQNARKHVAKLKKHNTLSFPLELVTYPSMY.CDLKCGFCFL.ANREDRNAKPAKD | WE 141 |
| CteB/1-450 | 59 DEALREI · ESLEAEGLLF SEDPYKEYVSSMDR · · · · · KSVVKALCLH I SHD <mark>CNLRCKYC</mark> FASTGNFGGQRNM · · · MSLEV | GK 131 |
| | | - |
| RumMC1/1-513 | I 155 RAVEYARDNSGDK • • VAISFY <mark>GGE</mark> PLVQ • FELMKK <mark>TI</mark> DYSRQIIK • • GKELTFSFS <mark>TN</mark> LTL <mark>V</mark> TPEIAAYVAGVE • GMS • VL | A S 230 |
| RumMC2/1-515 | i 157 RAVEYARDNSGDK - • VAVSFY <mark>GGE</mark> PLVQ - FELMKKTIDYSRQIIK - • GKELTFSFS <mark>TN</mark> LTL <mark>V</mark> TPEIAAYVAGVE - GMS • VL | A S 232 |
| An SME/1-370 | 48 SMVKRVLNEANGH ·· CSFAFQGGEPTLAGLEFFEKLMELQRKHNYK · NLKIYNSLQTNGTLIDESWAKFLS ·· E· NKFLVG | L S 122 |
| AILA/1-448 | 155 KTADMIEDNGVITCEITGGELEVH. PNANE. ILDVVCKKEKKVAVI TNGTIMRKESLELIKTYK. OKIIVG | 1 5 224 |
| M#C/1-392 | 58 DILLDELERMOV | 1 5 126 |
| Dan E /1 201 | | 1 6 120 |
| PqqE/1-384 | 55 RVLTEARGEGVENVELSGEFTAR.FDIVETARGE | 123 |
| SuiB/1-439 | 142 RTERGARDINGVESVSTEGGEPTRY.FDTDNELTAGEEKTRITTTINAGETKKSTVETEAKSKYTTPV | LS 210 |
| CteB/1-450 | 132 KAIDFLISESGNRKNLEIDFFGGEPMMN·FDVVKGILEYARQKEKEHNKNFRFTLTINGLLLNDENIKYIN··E·NMQNIV | LS 210 |
| | | |
| RumMC1/1-513 | 231 IDGPE-GIH-DAYRVMSGGKGSFEKAIQGLKYLVEAFGERAKESIVINTVVCPPFSAKKLDAIKEFFEGLSWLPKEMV | KK 308 |
| RumMC2/1-515 | I 233 IDGPE·KIH·DAYRVMSGGKGSFGKAIQGLKYLVEAFGERAKESIVI···NTVVCPPFSAKKLDAIKEFFEGLNWLPKEMV | KK 310 |
| An SME/1-370 | 123 MDGPK · EIH · NLNRKDCCGLDTFSKVERAAELFKKYKVE · FNILCVVTSNTARHVNKVYKYFKEKDFKFLQFINCLDPLYE | EK 202 |
| AIbA/1-448 | 225 LDSVNSEVH-DSFRGRKGSFAQTCKTIKLLSDHGIF-VRVAMSVFEKNMWEIHDMAQKVRDLGAKAFS-YNWVDDFGR | GR 301 |
| MAC/1-392 | 127 LDGANAEVN - DAVR GKGSF DMAVRALENLSNAGF TDAKI SVVVTRQNVDQLDEF AALAARYGATL RITRLRPSGR | GA 202 |
| PggE/1-384 | 124 VQGVDAANA-EKIGGLKNAQPQKMQFAARVTELGLP-LTLNSVIHRGNIHEVPGFIDLAVKLGAKRLEVAHTQYYGWA | YV 201 |
| SuiB/1-439 | 211 LOTLDSKLNFELMGVRPDRQIKLAKYFNEVGKK.CRINAVYTKQSYEQIJELVDFCJENKIDRFSVANYS | EV 281 |
| CteB/1-450 | 211 IDGRK-EVN-DRMRIRIDGSGCYDDILPKFKYVAESRNQDNYYVRGTFTRENMDFSNDVLHLADEGFRQISVEPVVAA | KD 288 |
| | | |
| RumMC1/1-513 | 309 CDYVEYGSVREEDISMEYAGDGEFIGEELDGFTLDAIEGWALARDLEEED - PKSYVAGIVADKLVRIHNRROTOEPCKDL | R R 389 |
| RumMC2/1-515 | 311 CDYVEYGSVREED I SMEYAGDGEFVGEEL DGFTL DA IEGWAL TRDLEEQD - PKSYVAG I VTDKLVRI HNRRQTQEPCKDL | RR 391 |
| An SME/1-370 | 203 G KYNY, SI KPKDYTKEL | KS 252 |
| AIL A/1-449 | | PA 241 |
| AIDA 1-440 | | AL 050 |
| 10/10/1-392 | 203 DVWDDERFTAEQQRQL | GE 200 |
| PqqE/1-384 | 202 NRAALMPDKSQVDESTRI | YP 245 |
| SuiB/1-439 | 282 TGYTKIKKKYDLADLRRLNEYVTDYITQREANLNFATEGCHLFTAYPELINNSIEFSEFDE | MY 344 |
| CteB/1-450 | 289 SGY DLREEDLPRLF EEYEKLAYEY VKRRKEGN WFNFFHFMID LTQGPCIVK | RL 341 |
| | | |
| RumMC1/1-513 | I 390 NG <mark>C</mark> CIPGNRRVY <mark>V</mark> KTD <mark>G</mark> KFLL <mark>C</mark> EK-TG-DAPDIGNVFEGADLEK <mark>I</mark> KKYYIEEYDEKSITRCNE <mark>C</mark> WARN | -L 456 |
| RumMC2/1-515 | i 392 NG <mark>C</mark> CIPGNRR <mark>VYV</mark> KAD <mark>G</mark> KFLL <mark>C</mark> EK-TG-DAPDIGNVFEGADLEK <mark>I</mark> KK <u>YY</u> IEEYDEKSLTR <mark>C</mark> NE <mark>C</mark> WARN | -L 458 |
| An SME/1-370 | 253 SSCGMNGTCTCQFVVESDGSVYPCDFYVL-DKWRLGNIQDMTMKELFETNKNHEFIKLSFKVHEECKKCKWFR | - L 325 |
| AIbA/1-448 | 342 ANCG - AGW - KSIVISPF GEVRPCALFP - KEFSLGNIF HDSYESIF NSPLVHKLWQAQ - APRFSEHCMKDKCPFS | GY 413 |
| MAC/1-392 | 251 NLCG AGR ·· VVCLIDPVGDVYACPFAIH· DKFLAGNIL· ·· SDGGFONVWQHSELFRELREPQ· SAGACASCGHFD· ·· · | - A 320 |
| PagE/1-384 | 246 KAC, AGGWGRKIMNVTPOGKVIPCHA AFTIPGIFFWYVTDHAIGFIWTKSPAFAAYRGTSWMKEPCRSCDRF | . K 318 |
| Sui B/1_4 29 | 245 YEER, AKY, TEMETING DIL DILAFLE, VNOTKONAF, EKNI DVWYDDDI YGEIDSEDTKNSKOLSEELLK | . 1 414 |
| Cha B/4 450 | | V 400 |
| Geb/ 1-450 | 34 TO SOL TELA TEODITIC TO TELA TEODITIC TO TELA TELA TELA TELA TELA TELA TELA TELA | · 1 408 |
| Dum MC 4/4 E42 | | 510 |
| Rum/wC1/1-513 | 46/ C. OLUTAACTEAEGIDMERKERVU | 513 |
| RumMC2/1-515 | 499 C. GLUYAAGYEAEGIDMERKEKVOGAHRYATKGELISYYSILEEKPEVIEEIDAVPYY | 515 |
| An SME/1-370 | 326 CKGGCR.RCRDSKEDSALELNYYCQSYKEFFEYAFPRLINVANNIK | 370 |
| AIbA/1-448 | 414 C • GGC YLKGL • • • NSNKYHRKNIC • • • • • • • • • • • • • • • • • • • | 448 |
| MftC/1-392 | 321 CROGOMAAKF · FTGLPLDGPDPEC · · · · · VEGWGAPALEKERVKPKPSGDHSRGTKQGPVALKLLTKPPARFCNESPV · · · | 392 |
| PqqE/1-384 | 319 DW <mark>GGG</mark> CCALALTGDAANT · DPA <mark>C</mark> SLSPLHAKMRDLAKEEAAETPPD · · · · · · · · · · · · · · · · · · | 384 |
| SuiB/1-439 | 415 CE00CYVNLI·KEKSPEYFRDSVD | 439 |
| CteB/1-450 | 409 C S C G C A A N S Y N F H K D I N T V Y K V G C E L E K K R V E C A L W I K A Q E M W I K A Q E M | 450 |
| | | |

Figure S1 - Sequence alignment of RumMC1 & RumMC2 with other SPASM-domain radical SAM enzymes.

Sequences were aligned using Jalview (1). Blue highlighting is proportional to the conservation of amino acids between protein sequences. The canonical CX_3CX_2C motif is indicated by red arrows and the seven cysteine residues implicated in the SPASM motif are indicated by green arrows. For AlbA, a cysteine residue tentatively proposed to be involved in the coordination of the AuxII [4Fe-4S] cluster is squared in blue. The eighth cysteine residue, involved in the complete coordination of the AuxI [4Fe-4S] cluster in AnSME, MftC, PqqE and SuiB is squared in orange.



Figure S2 – SDS-PAGE analysis and UV-visible spectrum of RumMC1 before (pink trace) and after (red trace) iron-sulfur cluster reconstitution.



Α

Β

Figure S3 – LC-MS analysis of the reaction performed with the C2₂₈₋₆₃ **peptide** (A) **MS/MS analysis of P1s.** Relevant peptide fragments are indicated showing that **P1s** contains a Cys⁴⁵-to- Arg⁵³ thioether bond. (B) **MS/MS analysis of P1d.** Relevant peptide fragments are indicated showing that **P1d** contains a Cys⁴¹-to- Arg⁶¹ thioether bond in addition to the Cys⁴⁵-to- Arg⁵³ thioether bond (*see table S2* for complete assignment). * indicates loss of ammonia (-17.02 Da).

| C1 | MRKIVAGKLQTGADFEGSKWG | VC | SGSTAV | ANSHNA | AGPAY | CVGY | <mark>C</mark> GNNGVVT <mark>I</mark> | NANANVA | <mark>к</mark> та |
|----|-----------------------|----|---------|--------|-------|------|---------------------------------------|---------|-------------------|
| C2 | MRKIVAGKLQTGADFEGSKGG | KC | SGGAVVI | ENSHNA | AGPAY | CVGY | <mark>C</mark> GNNGVVT <mark>I</mark> | NANANLA | RTK |
| C3 | MKLVETKTTKTGTNFEGNRAG | | NGTVAV | ANSHNA | AGPAY | CVGY | <mark>C</mark> GNSGVVT <mark>F</mark> | NANANVA | <mark>к</mark> та |
| C4 | MRLVQSKRIATGFNFEGSKAG | VC | SGTVAV | ANSHNA | AGPAY | CVGY | <mark>C</mark> GNNGEVT <mark>F</mark> | NANYNIA | RRS |
| C5 | MKLVTSKTMKTGTNFEGNKAG | IC | SGSVAV | ANSHNA | AGPAY | CVGY | <mark>C</mark> GNNGAVT <mark>F</mark> | NANANLA | R <mark>TA</mark> |

Figure S4 - Sequence of the C1 to C5 peptides encoded by the RumC operon

Arginine and lysine residues are highlighted in red and green respectively. Conserved cysteine residues are highlighted in blue.



Figure S5 - MS/MS spectrum of the P2d peptide obtained after incubation with RumMC2

Relevant peptide fragments are indicated confirming that this peptide contains two thioether bridges (*i.e.* Lys⁵³-Cys⁴⁵ and Lys⁶¹- Cys⁴¹ bridges) (*see table S3* for complete assignment). * indicates loss of ammonia (-17.02 Da).



Figure S6 - MS/MS spectrum of the P3d peptide obtained after incubation with RumMC2

Relevant peptide fragments are indicated showing that this peptide contains two thioether bridges (*i.e.* Arg^{53} -Cys⁴⁵ and Ala^{61} -Cys⁴¹ bridges) (*see table S4* for complete assignment). * indicates loss of ammonia (-17.02 Da).



Figure S7 – MS/MS spectrum of the P4d peptide obtained after incubation with RumMC2

Relevant peptide fragments are indicated showing that this peptide contains two thioether bridges (*i.e.* Arg⁵³-Cys⁴⁵ and Thr⁶¹-Cys⁴¹ bridges) (*see table S5* for complete assignment). * indicates loss of ammonia (-17.02 Da).



Figure S8 – MS/MS spectrum of the P6s peptide









Figure S10 – MS/MS analysis of the C2 peptide after incubation in the presence of RumMC2ΔRRE and RRE-MC2.

C2 (250 μ M) was incubated for 4 hours with RumMC2 Δ RRE (250 μ M), RRE-MC2 (250 μ M), SAM (2 mM) and sodium dithionite (3 mM). * indicates loss of ammonia (-17.02 Da).

| Peptide | Formula | [M+4H] ⁴⁺ | [M+4H] ⁴⁺ obs |
|---|----------------------|----------------------|--------------------------|
| C2 ₂₈₋₆₃ | C154 H246 N52 O51 S2 | 926.9497 | 926.9533 |
| C2 ₂₈₋₆₃ -2H | C154 H244 N52 O51 S2 | 926.4457 | 926.4463 |
| C2 ₂₈₋₆₃ -4H | C154 H242 N52 O51 S2 | 925.9419 | 925.9443 |
| C2 ₂₈₋₆₃ K ⁵³ K ⁶¹ | C154 H246 N48 O51 S2 | 912.9466 | 912.9545 |
| C2 ₂₈₋₆₃ K ⁵³ K ⁶¹ -2H | C154 H244 N48 O51 S2 | 912.4427 | 912.4449 |
| C2 ₂₈₋₆₃ K ⁵³ K ⁶¹ -4H | C154 H242 N48 O51 S2 | 911.9388 | 91.9415 |
| C2 ₂₈₋₆₃ A61 | C151 H239 N49 O51 S2 | 905.6837 | 905.6848 |
| C2 ₂₈₋₆₃ A61 -2H | C151 H237 N49 O51 S2 | 905.1797 | 905.1831 |
| C2 ₂₈₋₆₃ A61 -4H | C151 H235 N49 O51 S2 | 904.6759 | 904.6893 |
| C2 ₂₈₋₆₃ T ⁶¹ | C152 H241 N49 O52 S2 | 913.1863 | 913.1920 |
| C2 ₂₈₋₆₃ T ⁶¹ - 2H | C152 H239 N49 O52 S2 | 912.6824 | 912.6920 |
| C2 ₂₈₋₆₃ T ⁶¹ - 4H | C152 H237 N49 O52 S2 | 912.1785 | 912.1822 |
| C2 ₂₈₋₆₃ G61 | C150 H237 N49 O51 S2 | 902.1798 | 902.1845 |
| C2 ₂₈₋₆₃ G61 -2H | C150 H237 N49 O51 S2 | 901.6758 | 901.6862 |
| C2 ₂₈₋₆₃ G61 -4H | C150 H237 N49 O51 S2 | 901.1719 | 901.1770 |

$Table \ S1-Peptides \ used \ in \ this \ study \ and \ their \ theoretical \ and \ observed \ (obs) \ masses$

| Peptide | Formula | [M+4H] ⁴⁺ | [M+4H] ^{4+obs} |
|----------------------------|-------------------------|----------------------|-------------------------|
| C2 ₂₈₋₆₃ d3 | C151 H236 N49 O51 S2 D3 | 906.4384 | 906.4409 |
| C2 ₂₈₋₆₃ d3 -2H | C151 H234 N49 O51 S2 D3 | 905.9344 | 905.9334 |
| C2 ₂₈₋₆₃ d3 -4H | C151 H232 N49 O51 S2 D3 | 905.4306 | 905.4329 |

| Peptide | Formula | [M+4H] ⁴⁺ | [M+4H] ^{4+obs} |
|----------------------------|-------------------------|----------------------|-------------------------|
| C2 ₂₈₋₆₃ d4 | C151 H235 N49 O51 S2 D4 | 906.6900 | 906.6928 |
| C2 ₂₈₋₆₃ d4 -2H | C151 H233 N49 O51 S2 D4 | 906.1860 | 906.1890 |
| C2 ₂₈₋₆₃ d4 -4H | C151 H231 N49 O51 S2 D4 | 905.4306 | 905.4344 |

Table S2 - Theoretical mass fragments of peptide C2₂₈₋₆₃ $[M{+}H]^{+}$

| | | | b | у | |
|---|----|----|------------|------------|----|
| А | 28 | 1 | 72.04444 | 3704.77691 | 36 |
| v | 29 | 2 | 171.11285 | 3633.73979 | 35 |
| v | 30 | 3 | 270.18126 | 3534.67138 | 34 |
| Е | 31 | 4 | 399.22386 | 3435.60297 | 33 |
| Ν | 32 | 5 | 513.26678 | 3306.56037 | 32 |
| S | 33 | 6 | 600.29881 | 3192.51745 | 31 |
| Н | 34 | 7 | 737.35772 | 3105.48542 | 30 |
| Ν | 35 | 8 | 851.40065 | 2968.42651 | 29 |
| А | 36 | 9 | 922.43776 | 2854.38358 | 28 |
| G | 37 | 10 | 979.45923 | 2783.34647 | 27 |
| Р | 38 | 11 | 1076.51199 | 2726.325 | 26 |
| А | 39 | 12 | 1147.5491 | 2629.27224 | 25 |
| Y | 40 | 13 | 1310.61243 | 2558.23513 | 24 |
| С | 41 | 14 | 1413.62162 | 2395.1718 | 23 |
| v | 42 | 15 | 1512.69003 | 2292.16261 | 22 |
| G | 43 | 16 | 1569.71149 | 2193.0942 | 21 |
| Y | 44 | 17 | 1732.77482 | 2136.07273 | 20 |
| С | 45 | 18 | 1835.78401 | 1973.00941 | 19 |
| G | 46 | 19 | 1892.80547 | 1870.00022 | 18 |
| Ν | 47 | 20 | 2006.8484 | 1812.97876 | 17 |
| Ν | 48 | 21 | 2120.89133 | 1698.93583 | 16 |
| G | 49 | 22 | 2177.91279 | 1584.8929 | 15 |
| v | 50 | 23 | 2276.9812 | 1527.87144 | 14 |
| v | 51 | 24 | 2376.04962 | 1428.80303 | 13 |
| Т | 52 | 25 | 2477.0973 | 1329.73461 | 12 |
| R | 53 | 26 | 2633.19841 | 1228.68693 | 11 |
| Ν | 54 | 27 | 2747.24133 | 1072.58582 | 10 |
| А | 55 | 28 | 2818.27845 | 958.5429 | 9 |
| Ν | 56 | 29 | 2932.32137 | 887.50578 | 8 |
| А | 57 | 30 | 3003.35849 | 773.46286 | 7 |
| Ν | 58 | 31 | 3117.40141 | 702.42574 | 6 |
| L | 59 | 32 | 3230.48548 | 588.38281 | 5 |
| А | 60 | 33 | 3301.52259 | 475.29875 | 4 |
| R | 61 | 34 | 3457.6237 | 404.26164 | 3 |
| Т | 62 | 35 | 3558.67138 | 248.16053 | 2 |
| Κ | 63 | 36 | 3686.76634 | 147.11285 | 1 |

Table S3 - Theoretical mass fragments of peptide $C2_{28\text{-}63}K^{53}K^{61}\,[M\text{+}H]^+$

| | | | b | У | |
|---|----|----|------------|------------|----|
| А | 28 | 1 | 72.04444 | 3648.76461 | 36 |
| v | 29 | 2 | 171.11285 | 3577.7275 | 35 |
| v | 30 | 3 | 270.18126 | 3478.65908 | 34 |
| Е | 31 | 4 | 399.22386 | 3379.59067 | 33 |
| Ν | 32 | 5 | 513.26678 | 3250.54808 | 32 |
| S | 33 | 6 | 600.29881 | 3136.50515 | 31 |
| Н | 34 | 7 | 737.35772 | 3049.47312 | 30 |
| Ν | 35 | 8 | 851.40065 | 2912.41421 | 29 |
| Α | 36 | 9 | 922.43776 | 2798.37128 | 28 |
| G | 37 | 10 | 979.45923 | 2727.33417 | 27 |
| Р | 38 | 11 | 1076.51199 | 2670.31271 | 26 |
| Α | 39 | 12 | 1147.5491 | 2573.25994 | 25 |
| Y | 40 | 13 | 1310.61243 | 2502.22283 | 24 |
| С | 41 | 14 | 1413.62162 | 2339.1595 | 23 |
| v | 42 | 15 | 1512.69003 | 2236.15032 | 22 |
| G | 43 | 16 | 1569.71149 | 2137.0819 | 21 |
| Y | 44 | 17 | 1732.77482 | 2080.06044 | 20 |
| С | 45 | 18 | 1835.78401 | 1916.99711 | 19 |
| G | 46 | 19 | 1892.80547 | 1813.98793 | 18 |
| Ν | 47 | 20 | 2006.8484 | 1756.96646 | 17 |
| Ν | 48 | 21 | 2120.89133 | 1642.92353 | 16 |
| G | 49 | 22 | 2177.91279 | 1528.88061 | 15 |
| v | 50 | 23 | 2276.9812 | 1471.85914 | 14 |
| v | 51 | 24 | 2376.04962 | 1372.79073 | 13 |
| Т | 52 | 25 | 2477.0973 | 1273.72232 | 12 |
| К | 53 | 26 | 2605.19226 | 1172.67464 | 11 |
| Ν | 54 | 27 | 2719.23518 | 1044.57968 | 10 |
| Α | 55 | 28 | 2790.2723 | 930.53675 | 9 |
| Ν | 56 | 29 | 2904.31523 | 859.49963 | 8 |
| Α | 57 | 30 | 2975.35234 | 745.45671 | 7 |
| Ν | 58 | 31 | 3089.39527 | 674.41959 | 6 |
| L | 59 | 32 | 3202.47933 | 560.37667 | 5 |
| Α | 60 | 33 | 3273.51644 | 447.2926 | 4 |
| Κ | 61 | 34 | 3401.61141 | 376.25549 | 3 |
| Т | 62 | 35 | 3502.65908 | 248.16053 | 2 |
| к | 63 | 36 | 3630 75405 | 147 11285 | 1 |

Table S4 - Theoretical mass fragments of peptide $C2_{28\text{-}63}A^{61}\ [M\text{+}H]^+$

| | | | b | у | |
|---|----|----|------------|------------|----|
| А | 28 | 1 | 72.04444 | 3619.71291 | 36 |
| v | 29 | 2 | 171.11285 | 3548.6758 | 35 |
| v | 30 | 3 | 270.18126 | 3449.60738 | 34 |
| Е | 31 | 4 | 399.22386 | 3350.53897 | 33 |
| Ν | 32 | 5 | 513.26678 | 3221.49638 | 32 |
| S | 33 | 6 | 600.29881 | 3107.45345 | 31 |
| Н | 34 | 7 | 737.35772 | 3020.42142 | 30 |
| Ν | 35 | 8 | 851.40065 | 2883.36251 | 29 |
| А | 36 | 9 | 922.43776 | 2769.31958 | 28 |
| G | 37 | 10 | 979.45923 | 2698.28247 | 27 |
| Р | 38 | 11 | 1076.51199 | 2641.26101 | 26 |
| А | 39 | 12 | 1147.5491 | 2544.20824 | 25 |
| Y | 40 | 13 | 1310.61243 | 2473.17113 | 24 |
| С | 41 | 14 | 1413.62162 | 2310.1078 | 23 |
| v | 42 | 15 | 1512.69003 | 2207.09861 | 22 |
| G | 43 | 16 | 1569.71149 | 2108.0302 | 21 |
| Y | 44 | 17 | 1732.77482 | 2051.00874 | 20 |
| С | 45 | 18 | 1835.78401 | 1887.94541 | 19 |
| G | 46 | 19 | 1892.80547 | 1784.93622 | 18 |
| Ν | 47 | 20 | 2006.8484 | 1727.91476 | 17 |
| Ν | 48 | 21 | 2120.89133 | 1613.87183 | 16 |
| G | 49 | 22 | 2177.91279 | 1499.82891 | 15 |
| v | 50 | 23 | 2276.9812 | 1442.80744 | 14 |
| v | 51 | 24 | 2376.04962 | 1343.73903 | 13 |
| Т | 52 | 25 | 2477.0973 | 1244.67062 | 12 |
| R | 53 | 26 | 2633.19841 | 1143.62294 | 11 |
| Ν | 54 | 27 | 2747.24133 | 987.52183 | 10 |
| А | 55 | 28 | 2818.27845 | 873.4789 | 9 |
| Ν | 56 | 29 | 2932.32137 | 802.44179 | 8 |
| А | 57 | 30 | 3003.35849 | 688.39886 | 7 |
| Ν | 58 | 31 | 3117.40141 | 617.36174 | 6 |
| L | 59 | 32 | 3230.48548 | 503.31882 | 5 |
| А | 60 | 33 | 3301.52259 | 390.23475 | 4 |
| А | 61 | 34 | 3372.55971 | 319.19764 | 3 |
| Т | 62 | 35 | 3473.60738 | 248.16053 | 2 |
| Κ | 63 | 36 | 3601.70235 | 147.11285 | 1 |

| Table S5 - | Theoretical | mass fragments o | f peptide | $C2_{28-63}T^{61}$ | [M+H] ⁺ |
|------------|-------------|------------------|-----------|--------------------|--------------------|
|------------|-------------|------------------|-----------|--------------------|--------------------|

| | | | b | У | |
|---|----|----|------------|------------|----|
| А | 28 | 1 | 72.04444 | 3649.72348 | 36 |
| v | 29 | 2 | 171.11285 | 3578.68636 | 35 |
| v | 30 | 3 | 270.18126 | 3479.61795 | 34 |
| Е | 31 | 4 | 399.22386 | 3380.54953 | 33 |
| Ν | 32 | 5 | 513.26678 | 3251.50694 | 32 |
| S | 33 | 6 | 600.29881 | 3137.46401 | 31 |
| Н | 34 | 7 | 737.35772 | 3050.43199 | 30 |
| Ν | 35 | 8 | 851.40065 | 2913.37307 | 29 |
| А | 36 | 9 | 922.43776 | 2799.33015 | 28 |
| G | 37 | 10 | 979.45923 | 2728.29303 | 27 |
| Р | 38 | 11 | 1076.51199 | 2671.27157 | 26 |
| А | 39 | 12 | 1147.5491 | 2574.21881 | 25 |
| Y | 40 | 13 | 1310.61243 | 2503.18169 | 24 |
| С | 41 | 14 | 1413.62162 | 2340.11836 | 23 |
| v | 42 | 15 | 1512.69003 | 2237.10918 | 22 |
| G | 43 | 16 | 1569.71149 | 2138.04077 | 21 |
| Y | 44 | 17 | 1732.77482 | 2081.0193 | 20 |
| С | 45 | 18 | 1835.78401 | 1917.95597 | 19 |
| G | 46 | 19 | 1892.80547 | 1814.94679 | 18 |
| Ν | 47 | 20 | 2006.8484 | 1757.92533 | 17 |
| Ν | 48 | 21 | 2120.89133 | 1643.8824 | 16 |
| G | 49 | 22 | 2177.91279 | 1529.83947 | 15 |
| v | 50 | 23 | 2276.9812 | 1472.81801 | 14 |
| v | 51 | 24 | 2376.04962 | 1373.74959 | 13 |
| Т | 52 | 25 | 2477.0973 | 1274.68118 | 12 |
| R | 53 | 26 | 2633.19841 | 1173.6335 | 11 |
| Ν | 54 | 27 | 2747.24133 | 1017.53239 | 10 |
| А | 55 | 28 | 2818.27845 | 903.48946 | 9 |
| Ν | 56 | 29 | 2932.32137 | 832.45235 | 8 |
| А | 57 | 30 | 3003.35849 | 718.40942 | 7 |
| Ν | 58 | 31 | 3117.40141 | 647.37231 | 6 |
| L | 59 | 32 | 3230.48548 | 533.32938 | 5 |
| А | 60 | 33 | 3301.52259 | 420.24532 | 4 |
| Т | 61 | 34 | 3402.57027 | 349.20821 | 3 |
| Т | 62 | 35 | 3503.61795 | 248.16053 | 2 |
| Κ | 63 | 36 | 3631.71291 | 147.11285 | 1 |

Table S6 - Theoretical mass fragments of peptide $C2_{28\mathchar`eq}G^{61}~[M\mathchar`eq}H]^+$

| | | | b | у | |
|---|----|----|------------|------------|----|
| Α | 28 | 1 | 72.04444 | 3605.69726 | 36 |
| v | 29 | 2 | 171.11285 | 3534.66015 | 35 |
| v | 30 | 3 | 270.18126 | 3435.59173 | 34 |
| Е | 31 | 4 | 399.22386 | 3336.52332 | 33 |
| Ν | 32 | 5 | 513.26678 | 3207.48073 | 32 |
| s | 33 | 6 | 600.29881 | 3093.4378 | 31 |
| н | 34 | 7 | 737.35772 | 3006.40577 | 30 |
| Ν | 35 | 8 | 851.40065 | 2869.34686 | 29 |
| Α | 36 | 9 | 922.43776 | 2755.30393 | 28 |
| G | 37 | 10 | 979.45923 | 2684.26682 | 27 |
| Р | 38 | 11 | 1076.51199 | 2627.24536 | 26 |
| Α | 39 | 12 | 1147.5491 | 2530.19259 | 25 |
| Y | 40 | 13 | 1310.61243 | 2459.15548 | 24 |
| С | 41 | 14 | 1413.62162 | 2296.09215 | 23 |
| v | 42 | 15 | 1512.69003 | 2193.08296 | 22 |
| G | 43 | 16 | 1569.71149 | 2094.01455 | 21 |
| Y | 44 | 17 | 1732.77482 | 2036.99309 | 20 |
| С | 45 | 18 | 1835.78401 | 1873.92976 | 19 |
| G | 46 | 19 | 1892.80547 | 1770.92057 | 18 |
| Ν | 47 | 20 | 2006.8484 | 1713.89911 | 17 |
| Ν | 48 | 21 | 2120.89133 | 1599.85618 | 16 |
| G | 49 | 22 | 2177.91279 | 1485.81326 | 15 |
| v | 50 | 23 | 2276.9812 | 1428.79179 | 14 |
| v | 51 | 24 | 2376.04962 | 1329.72338 | 13 |
| Т | 52 | 25 | 2477.0973 | 1230.65497 | 12 |
| R | 53 | 26 | 2633.19841 | 1129.60729 | 11 |
| Ν | 54 | 27 | 2747.24133 | 973.50618 | 10 |
| Α | 55 | 28 | 2818.27845 | 859.46325 | 9 |
| Ν | 56 | 29 | 2932.32137 | 788.42614 | 8 |
| Α | 57 | 30 | 3003.35849 | 674.38321 | 7 |
| Ν | 58 | 31 | 3117.40141 | 603.34609 | 6 |
| L | 59 | 32 | 3230.48548 | 489.30317 | 5 |
| Α | 60 | 33 | 3301.52259 | 376.2191 | 4 |
| G | 61 | 34 | 3358.54406 | 305.18199 | 3 |
| Т | 62 | 35 | 3459.59173 | 248.16053 | 2 |
| К | 63 | 36 | 3587.6867 | 147.11285 | 1 |

| | b | У | |
|----|--|---|--|
| 1 | 72.04444 | 3623.73802 | 36 |
| 2 | 171.11285 | 3552.70091 | 35 |
| 3 | 270.18126 | 3453.63249 | 34 |
| 4 | 399.22386 | 3354.56408 | 33 |
| 5 | 513.26678 | 3225.52149 | 32 |
| 6 | 600.29881 | 3111.47856 | 31 |
| 7 | 737.35772 | 3024.44653 | 30 |
| 8 | 851.40065 | 2887.38762 | 29 |
| 9 | 922.43776 | 2773.34469 | 28 |
| 10 | 979.45923 | 2702.30758 | 27 |
| 11 | 1076.51199 | 2645.28612 | 26 |
| 12 | 1147.5491 | 2548.23335 | 25 |
| 13 | 1310.61243 | 2477.19624 | 24 |
| 14 | 1413.62162 | 2314.13291 | 23 |
| 15 | 1512.69003 | 2211.12372 | 22 |
| 16 | 1569.71149 | 2112.05531 | 21 |
| 17 | 1732.77482 | 2055.03385 | 20 |
| 18 | 1835.78401 | 1891.97052 | 19 |
| 19 | 1892.80547 | 1788.96133 | 18 |
| 20 | 2006.8484 | 1731.93987 | 17 |
| 21 | 2120.89133 | 1617.89694 | 16 |
| 22 | 2177.91279 | 1503.85402 | 15 |
| 23 | 2276.9812 | 1446.83255 | 14 |
| 24 | 2376.04962 | 1347.76414 | 13 |
| 25 | 2477.0973 | 1248.69573 | 12 |
| 26 | 2633.19841 | 1147.64805 | 11 |
| 27 | 2747.24133 | 991.54694 | 10 |
| 28 | 2818.27845 | 877.50401 | 9 |
| 29 | 2932.32137 | 806.4669 | 8 |
| 30 | 3003.35849 | 692.42397 | 7 |
| 31 | 3117.40141 | 621.38685 | 6 |
| 32 | 3230.48548 | 507.34393 | 5 |
| 33 | 3301.52259 | 394.25986 | 4 |
| 34 | 3376.58482 | 323.22275 | 3 |
| 35 | 3477.63249 | 248.16053 | 2 |
| 36 | 3605.72746 | 147.11285 | 1 |
| | 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 | b 1 72.04444 2 171.11285 3 270.18126 4 399.22386 5 513.26678 6 600.29881 7 737.35772 8 851.40065 9 922.43776 10 979.45923 11 1076.51199 12 1147.5491 13 1310.61243 14 1413.62162 15 1512.69003 16 1569.71149 17 1732.77482 18 1835.78401 19 1892.80547 20 2006.8484 21 2120.89133 22 2177.91279 23 2276.9812 24 2376.04962 25 2477.0973 26 2633.19841 27 2747.24133 28 2818.27845 29 2932.32137 30 3003.35849 <td< th=""><th>b y 1 72.04444 3623.73802 2 171.11285 3552.70091 3 270.18126 3453.63249 4 399.22386 3354.56408 5 513.26678 3225.52149 6 600.29881 3111.47856 7 737.35772 3024.44653 8 851.40065 2887.38762 9 922.43776 2773.34469 10 979.45923 2702.30758 11 1076.51199 2645.28612 12 1147.5491 2548.2335 13 1310.61243 2477.19624 14 1413.62162 2314.13291 15 1512.69003 2211.12372 16 1569.71149 2112.05531 17 1732.77482 2055.03385 18 1835.78401 1891.97052 19 1892.80547 1788.96133 20 2006.8484 1731.93987 21 2120.89133 1617.89694 22<</th></td<> | b y 1 72.04444 3623.73802 2 171.11285 3552.70091 3 270.18126 3453.63249 4 399.22386 3354.56408 5 513.26678 3225.52149 6 600.29881 3111.47856 7 737.35772 3024.44653 8 851.40065 2887.38762 9 922.43776 2773.34469 10 979.45923 2702.30758 11 1076.51199 2645.28612 12 1147.5491 2548.2335 13 1310.61243 2477.19624 14 1413.62162 2314.13291 15 1512.69003 2211.12372 16 1569.71149 2112.05531 17 1732.77482 2055.03385 18 1835.78401 1891.97052 19 1892.80547 1788.96133 20 2006.8484 1731.93987 21 2120.89133 1617.89694 22< |

Table S7 - Theoretical mass fragments of peptide $C2_{28-63}A^{61}d4$

| Table S8 - Theoretical Mass fragm | nents of peptide C228-63A ⁶¹ d3 |
|-----------------------------------|--|
|-----------------------------------|--|

| | | b | у | |
|---|----|------------|------------|----|
| Α | 1 | 72.04444 | 3622.73174 | 36 |
| V | 2 | 171.11285 | 3551.69463 | 35 |
| V | 3 | 270.18126 | 3452.62621 | 34 |
| Е | 4 | 399.22386 | 3353.5578 | 33 |
| Ν | 5 | 513.26678 | 3224.51521 | 32 |
| S | 6 | 600.29881 | 3110.47228 | 31 |
| Н | 7 | 737.35772 | 3023.44025 | 30 |
| N | 8 | 851.40065 | 2886.38134 | 29 |
| Α | 9 | 922.43776 | 2772.33841 | 28 |
| G | 10 | 979.45923 | 2701.3013 | 27 |
| Р | 11 | 1076.51199 | 2644.27984 | 26 |
| Α | 12 | 1147.5491 | 2547.22707 | 25 |
| Y | 13 | 1310.61243 | 2476.18996 | 24 |
| С | 14 | 1413.62162 | 2313.12663 | 23 |
| V | 15 | 1512.69003 | 2210.11744 | 22 |
| G | 16 | 1569.71149 | 2111.04903 | 21 |
| Y | 17 | 1732.77482 | 2054.02757 | 20 |
| С | 18 | 1835.78401 | 1890.96424 | 19 |
| G | 19 | 1892.80547 | 1787.95505 | 18 |
| Ν | 20 | 2006.8484 | 1730.93359 | 17 |
| Ν | 21 | 2120.89133 | 1616.89066 | 16 |
| G | 22 | 2177.91279 | 1502.84774 | 15 |
| V | 23 | 2276.9812 | 1445.82627 | 14 |
| V | 24 | 2376.04962 | 1346.75786 | 13 |
| Т | 25 | 2477.0973 | 1247.68945 | 12 |
| R | 26 | 2633.19841 | 1146.64177 | 11 |
| Ν | 27 | 2747.24133 | 990.54066 | 10 |
| Α | 28 | 2818.27845 | 876.49773 | 9 |
| Ν | 29 | 2932.32137 | 805.46062 | 8 |
| Α | 30 | 3003.35849 | 691.41769 | 7 |
| Ν | 31 | 3117.40141 | 620.38057 | 6 |
| L | 32 | 3230.48548 | 506.33765 | 5 |
| Α | 33 | 3301.52259 | 393.25358 | 4 |
| Α | 34 | 3375.57854 | 322.21647 | 3 |
| Т | 35 | 3476.62621 | 248.16053 | 2 |
| K | 36 | 3604.72118 | 147.11285 | 1 |

Supplementary references

1. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. **25**, 1189-1191