Isolation of a new antibacterial peptide achromosin from Streptomyces achromogenes subsp. achromogenes based on genome mining

Note

Isolation of a new antibacterial peptide achromosin from *Streptomyces achromogenes* subsp. *achromogenes* based on genome mining

Authors: Issara Kaweewan[1], Mayumi Ohnishi-Kameyama[2], Shinya Kodani[1,3,4]*

Affiliations: [1]Graduate School of Integrated Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529 Japan; [2] Food Research Institute, National Agriculture and Food Research Organization (NARO), 2–1-12 Kan-nondai, Tsukuba, Ibaraki 305-8642 Japan; [3]College of Agriculture, Academic Institute, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529 Japan; [4]Graduate School of Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529 Japan

*To whom correspondence should be addressed: Shinya Kodani, College of Agriculture, Academic Institute, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529 Japan, Tel/Fax; +81(54)238-5008, E-mail; kodani.shinya@shizuoka.ac.jp

Running title: Isolation of a new antibacterial peptide achromosin

Keywords: *Streptomyces achromogenes* subsp. *achromogenes*; lasso peptide; MS/MS analysis; biosynthetic genes

Lasso peptides are a class of ribosomally biosynthesized and post-translationally modified peptides with a common motif of knot structure in the molecule.[1] The amino group of the N-terminal amino acid forms a peptide bond with γ-carboxyl group of Asp or Glu in the 8th or the 9th position from the N-terminus, resulting in formation of a macrolactam ring. The macrolactam ring looks like a loop of a "lasso" with a tail of the C-terminal linear peptide that normally locates through the ring. Regarding lasso peptides, a wide variety of biological activities such as anti-HIV,[2] antimycobacterial,[3] endothelin type B receptor antagonist,[4] and prolyl endopeptidase inhibition[5] were reported. In addition, lasso peptides normally show a stable property against proteolytic, thermal, and chemical degradation, which makes lasso peptides attractive in terms of practical application as pharmaceutical reagents.

Lasso peptides derived from actinobacteria have been classified into three main classes on the basis of their *N*-terminal residues and the number of disulfide bridges.[1] The class I lasso peptides include siamycins I and II,[2] aborycin,[6] and sviceucin,[7] which have an internal peptide linkage between γ-carboxyl residue of Asp9 (9th amino acid residue from the N-terminus) and the amino residue of Cys1. These peptides commonly have additional two disulfide bridges between Cys1 and Cys13, and Cys7 and Cys19. The class II lasso peptides include anantin,[8] lariatins,[3] propeptin,[5] RES-701-1,[4] SRO15-2005,[9] and sungsanpin.[10] These peptides have an internal peptide linkage between γ-carboxyl residue of Asp8 or Asp9 and the amino residue of Gly1 without any disulfide bonds. The class III lasso peptide includes only one peptide named

BI-32169.[11] The peptide BI-32169 has an internal peptide linkage between γ-carboxyl residue of Asp9 and the amino residue of Gly1 with one disulfide bond between Cys6 and Cys19.

The lasso peptide microcin J25 was isolated from *Escherichia coli*, which is regarded as the archetype of lasso peptides.[12] Its biosynthetic gene cluster consists of four genes including a precursor peptide coding gene: gene A (*mcjA*), two maturation enzymes including gene B (*mcjB,* cleavage of leader peptide) and gene C (*mcjC*, formation of macrolactam ring) and an ATP-binding cassette transporter coding gene: gene D (*mcjD*).[13] The protein McjC was reported to form the macrolactam ring, and the function of the protein McjB was assigned to cleave off the leader peptide from precursor peptide by in vitro experiments.[14] Normally lasso peptide biosynthetic genes in proteobacteria have a corresponding set of the genes, although the transporter gene is optional.[1] In actinobacteria, lasso peptide biosynthetic genes consist of a similar gene set, except that a maturation enzyme gene B have split-B genes (gene B1 and gene B2).[1] By genome mining, biosynthetic genes of a lasso peptide sviceucin was found on the genome of *Streptomyces sviceus*, and the lasso peptide was isolated and structure-determined by heterologous expression.[7] The lasso peptide SRO15-2005 was identified by matrix-assisted laser desorption/ionization-time-of-flight tandem mass spectrometry (MALDI-TOF-MS/MS) from the extract of *Streptomyces roseosporus*, based on genome sequence data.[9] Based on genome mining, a new lasso peptide chaxapeptin was also isolated as a lung cancer invasion inhibitor from *Streptomyces leeuwenhoekii*.[15] These results prompted us to find a

57 new lasso peptide from streptomycetes using genome sequence data. By genome search
58 approach, we found new lasso peptide biosynthetic genes on the genome sequence of
59 *Streptomyces achromogenes* subsp. *achromogenes*.[16] The new antibacterial peptide was
60 isolated by chromatographic separation from the culture of *S. achromogenes subsp.*
61 *achromogenes*. Here we describe isolation and structure determination of a new antibacterial
62 peptide named achromosin.

63 In the genome sequence of *Streptomyces achromogenes* subsp. *achromogenes*,[16] lasso
64 peptide modification enzyme coding genes (gene C named *acrC*: WP_063755122.1, *acrB2*:
65 WP_037654156.1, *acrB1*: WP_037654159.1, shown in Fig 1a) were found by blastp
66 similarity search. Since the lasso precursor peptide coding gene was not annotated, we
67 searched for the lasso precursor peptide coding gene in the close region to the modification
68 enzyme coding genes. Upstream of the gene *acrC* (WP_063755122.1), a new putative
69 precursor peptide coding gene for new peptide named achromosin (126 base pairs, 42 amino
70 acids, Fig. 1b) similar to chaxapeptin[15] was found from position 72827 to 72952 bp in the
71 genome sequence (GenBank accession number: NZ_JODT01000002.1). On the upstream of
72 9 residues of the precursor peptide coding region (72827-72952), Shine-Dalgarno sequence
73 (AGGAGGA) was present. As shown in Fig. 1b, the expected peptide achromosin was
74 deduced to have the amino acid sequence of GIGSQTWDTIWLWD (monoisotopic molecular
75 weight: 1676.7 Da), after cleaving off the leader peptide at the same position as chaxapeptin

(arrow in Fig. 1b). The expected monoisotopic molecular weight of achromosin was calculated to be 1658.7 Da considering the loss of 18 Da, resulting in macrolactam formation of lasso peptide biosynthesis. The preliminary chemical investigation of *S. achromogenes* subsp. *achromogenes* NBRC12735^T indicated that the expected peptide was present in the methanol extract of aerial hyphae and spore cells by HPLC and ESI-MS (Data not shown). Thus, cultivation of *S. achromogenes* subsp. *achromogenes* was performed using 5 L of ISP2 agar media to obtain enough amount of the peptide for structure determination. After 7 days of cultivation, cells of spore and aerial hyphae were harvested by steel spatula. The cells were extracted with double volume of methanol (MeOH), followed by centrifugation. After condensation using rotary evaporation, the extract was subjected to open column chromatography using hydrophobic resin (CHP-20P), eluted with 10%, 60%, and 100% MeOH. The expected peptide achromosin was detected in 100% MeOH fraction by HPLC (Fig. S1) and ESI-MS analysis (Fig. S2). The ESI- MS analysis of the peptide gave an ion peak at *m/z* 1659.7 for $[M+H]^+$. The 100% MeOH fraction was repeatedly subjected to HPLC purification to give pure achromosin.

The molecular formula of achromosin was established to be $C_{79}H_{106}N_{18}O_{22}$ by accurate mass analysis using the ESI-FTICR mass spectrometer ($[M+2H]^{2+}$ was observed at *m/z* 830.3941 corresponding to $C_{79}H_{108}N_{18}O_{22}$ whose calculated value was 830.3937). The amino acid content analysis was performed on achromosin following the reported method.[17]

The amino acid content analysis on achromosin afforded the relative molar ratios of the constituent amino acids (2 moles each of Asp, Gly, Ile, and Thr, and 1 mole each of Glu, Leu, and Ser), as shown in Fig. S3. NMR analysis using dimethyl sulfoxide-d_6 as a solvent was not possible due to ambiguous broad peaks in the NMR spectrum. To obtain peptide sequence, MALDI-TOF MS/MS analysis on achromosin was accomplished. As a result, the product ions from achromosin at *m/z* 1659 were of *b*-series peptides, *b*8-*b*13 (Fig. 2a), which indicated that the sequence of TIWLWD was the C-terminus tail sequence. Macrolactam ring structure was reported not to give fragment ions,[9] thus we proposed the structure of achromosin to be shown in Fig. 2a, based on the amino acid sequence of precursor peptide gene. To confirm the amino acid sequence in the macrolactam ring, C-terminal peptide bonds of tryptophans were cleaved by BNPS-skatole. After BNPS-skatole reaction, the cleaved achromosin (BNPS-achromosin) was purified by HPLC separation. ESI-TOF-MS analysis on BNPS-achromosin gave an ion peak at *m/z* 1291.5 for $[M+H]^+$ (Fig. S4). The molecular formula of BNPS-achromosin was clarified to be $C_{58}H_{78}N_{14}O_{20}$ by the accurate mass analysis. That is, $[M+2H]^{2+}$ was observed at *m/z* 646.2832 corresponding to $C_{58}H_{80}N_{14}O_{20}$ whose calculated value was 646.2831. By the reaction of BNPS-skatole the Trp residue in a peptide is oxidized and transformed to 3-oxindole with a spirolactone, which increases the molecular weight due to the addition of two oxygens by 32 Da. As shown Fig. 2b, the MALDI-TOF MS/MS of the cleaved achromosin gave the sequence of the peptide

with one N-terminus and two C-terminal ends. The product ions of *b1*, *b2*, and *b3* supported the sequence of DTIW* and *b4* ion especially indicated that Trp at C-terminus was oxidized (indicated with asterisk, Fig. 2b). The product ions of *y2* to *y7* supported the sequence of GIGSQTW* (Fig. 2b). Above all, the structure of achromosin was proposed to be a peptide with the sequence of GIGSQTWDTIWLWD having one macrolactam ring which was formed by peptide bond between amino residue of 1st Gly and γ-carboxyl residue of 8th Asp (Fig. 2a). The structure of achromosin did not include cysteine residues, and it was classified into class II lasso peptide.

The antimicrobial activity of achromosin was tested using a paper disk agar diffusion assay against microorganisms (Bacterial strains including *Escherichia coli, Pseudomonas aeruginosa, Serratia marcescens, Bacillus subtilis, Staphylococcus aureus, Micrococcus luteus, Streptomyces antibioticus*; Yeast strains including *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kloeckera apiculata*; fungi strains including *Aspergillus niger*, *Aspergillus oryzae, Mucor hiemalis*). At the dosage of 10 μg per disk, achromosin showed an inhibitory zone of 11 mm diameter against *M. luteus* (Fig. S5). On the other hand, achromosin did not show any inhibitory activity against the other testing microorganisms at the same dosage.

Biosynthetic genes clusters of lasso peptides of actinobacteria have been identified for lasso peptides including lariatin,[18] SRO15-2005,[9] lassomycin,[19] sviceucin,[7] chaxapeptin,[15] and

streptomonomicin.[20] The biosynthetic gene cluster of chaxapeptin consisted of 4 genes including *cptA*, *cptC*, *cptB1*, and *cptB2*.[15] Interestingly, the gene cluster of chaxapeptin lacked of transporter gene that often exists in the lasso peptide biosynthetic gene cluster. The gene *cptA* encoded chaxapeptin precursor peptide, and the three genes including *cptC*, *cptB1*, and *cptB2* were proposed to be involved in macrolactam formation and leader peptide cleavage. The amino acid sequence of precursor peptide gene *acrA* which was found on the genome of *S. achromogenes* subsp. *achromogenes*[16] showed high similarity with that of *cptA* (46% identity, 68% positive matches). By reference to chaxapeptin biosynthetic genes, we assigned the biosynthetic gene cluster for achromosin, which have 4 genes, *acrA* (annoted in this study, 42 aa), *acrC* (WP_063755122.1, 616 aa), *acrB2* (WP_037654156.1, 150 aa), and *acrB1* (WP_037654159.1, 95 aa) in this order with all the same direction (Fig. 1a). Interestingly, there was no transport protein coding genes near the gene cluster. The lack of transport gene was also reported in the chaxapeptin gene cluster.[15] Based on the similarity of each gene, we proposed the functions of the genes as shown in Fig. 1a. The gene *acrA* encoded the precursor of achromosin and the genes including *acrC*, *acrB1*, and *acrB2* were proposed to be modification enzymes to give the mature lasso peptide. The gene *acrC* encoded putative asparagine synthase possibly responsible for formation of the Gly1-Glu8 amide bond, which showed high similarity to *cptC* by using a BLAST homology search (37% identity, 51% positive matches). The amino acid sequence of *acrB2* showed high similarity

to that of *cptB2* by using a BLAST homology search (55% identity, 69% positive matches), and the amino acid sequence of *acrB1* showed high similarity to that of *cptB1* by using a BLAST homology search (40% identity, 54% positive matches). Above all, the biosynthetic genes of achromosin showed the similarity to those of chaxapeptin.

So far, no similar peptide has been found by the blastp search, which indicates the novelty of achromosin. As shown in Fig. 1b, the amino acid sequence of core peptide is different even from that of chaxapeptin, the closest lasso peptide. The lasso peptide in class II were reported to have a wide variety of biological activity such as antimycobacterial,[3] endothelin type B receptor antagonist,[4] and prolyl endopeptidase inhibition[5]. In this paper, the antimicrobial activity was tested on achromosin, and the further other bioactivity tests may lead to another activity of achromosin. In addition, the biosynthetic genes of achromosin was identified from the genome of *S. achromogenes* subsp. *achromogenes*, indicating that the gene cluster functioned properly enough to produce achromosin in *S. achromogenes* subsp. *achromogenes*. This information will lead to genetic engineering using the gene cluster to create mutated lasso peptide based on achromosin by heterologous expression. More potent antibacterial peptides may be produced based on the discovery of achromosin by the further genetic engineering experiments.

Conflict of interest

171 The authors declare no conflict of interest.

172 **Acknowledgments**

173 This study was supported by the Japan Society for the Promotion of Science by Grants-in-

174 aids (grant number 25350964). We thank Ms. Tomoko Satoh (Bruker Daltonics) for her

175 technical assistance in the MS analysis.

176

References

1. Li, Y.; Zirah, S.; Rebuffat, S., *Lasso Peptides: Bacterial Strategies to Make and Maintain Bioactive Entangled Scaffolds*. Springer: 2015.

2. Detlefsen, D. J.; Hill, S. E.; Volk, K. J.; Klohr, S. E.; Tsunakawa, M.; Furumai, T.; Lin, P. F.; Nishio, M.; Kawano, K.; Oki, T.; et al., Siamycins I and II, new anti-HIV-1 peptides: II. Sequence analysis and structure determination of siamycin I. *J Antibiot (Tokyo)* **1995,** 48, 1515-7.

3. Iwatsuki, M.; Tomoda, H.; Uchida, R.; Gouda, H.; Hirono, S.; Omura, S., Lariatins, antimycobacterial peptides produced by *Rhodococcus* sp. K01-B0171, have a lasso structure. *J Am Chem Soc* **2006,** 128, 7486-91.

4. Morishita, Y.; Chiba, S.; Tsukuda, E.; Tanaka, T.; Ogawa, T.; Yamasaki, M.; Yoshida, M.; Kawamoto, I.; Matsuda, Y., RES-701-1, a novel and selective endothelin type B receptor antagonist produced by *Streptomyces* sp. RE-701. I. Characterization of producing strain, fermentation, isolation, physico-chemical and biological properties. *J Antibiot (Tokyo)* **1994,** 47, 269-75.

5. Kimura, K.; Kanou, F.; Takahashi, H.; Esumi, Y.; Uramoto, M.; Yoshihama, M., Propeptin, a new inhibitor of prolyl endopeptidase produced by *Microbispora*. I. Fermentation, isolation and biological properties. *J Antibiot (Tokyo)* **1997,** 50, 373-8.

6. Potterat, O.; Stephan, H.; MetzgerJ.W.; Gnau, V.; Zähner, H.; Jung, G., Aborycin A -

196 Tricyclic 21-Peptide Antibiotic Isolated from *Streptomyces griseoflavus*. *Liebigs Ann.* **1994,**
197 1994, 741-743.
198 7. Li, Y.; Ducasse, R.; Zirah, S.; Blond, A.; Goulard, C.; Lescop, E.; Giraud, C.; Hartke,
199 A.; Guittet, E.; Pernodet, J. L.; Rebuffat, S., Characterization of sviceucin from *streptomyces*
200 provides insight into enzyme exchangeability and disulfide bond formation in lasso peptides.
201 *ACS Chem Biol* **2015,** 10, 2641-9.
202 8. Weber, W.; Fischli, W.; Hochuli, E.; Kupfer, E.; Weibel, E. K., Anantin-a peptide
203 antagonist of the atrial natriuretic factor (ANF). I. Producing organism, fermentation, isolation
204 and biological activity. *J Antibiot (Tokyo)* **1991,** 44, 164-71.
205 9. Kersten, R. D.; Yang, Y. L.; Xu, Y.; Cimermancic, P.; Nam, S. J.; Fenical, W.;
206 Fischbach, M. A.; Moore, B. S.; Dorrestein, P. C., A mass spectrometry-guided genome
207 mining approach for natural product peptidogenomics. *Nat Chem Biol* **2011,** 7, 794-802.
208 10. Um, S.; Kim, Y. J.; Kwon, H.; Wen, H.; Kim, S. H.; Kwon, H. C.; Park, S.; Shin, J.;
209 Oh, D. C., Sungsanpin, a lasso peptide from a deep-sea streptomycete. *J Nat Prod* **2013,** 76,
210 873-9.
211 11. Potterat, O.; Wagner, K.; Gemmecker, G.; Mack, J.; Puder, C.; Vettermann, R.;
212 Streicher, R., BI-32169, a bicyclic 19-peptide with strong glucagon receptor antagonist
213 activity from *Streptomyces* sp. *J Nat Prod* **2004,** 67, 1528-31.
214 12. Salomon, R. A.; Farias, R. N., Microcin 25, a novel antimicrobial peptide produced

215 by Escherichia coli. *J Bacteriol* **1992,** 174, 7428-35.

216 13. Solbiati, J. O.; Ciaccio, M.; Farias, R. N.; Salomon, R. A., Genetic analysis of

217 plasmid determinants for microcin J25 production and immunity. *J Bacteriol* **1996,** 178, 3661-

218 3.

219 14. Yan, K. P.; Li, Y.; Zirah, S.; Goulard, C.; Knappe, T. A.; Marahiel, M. A.; Rebuffat,

220 S., Dissecting the maturation steps of the lasso peptide microcin J25 in vitro. *Chembiochem*

221 **2012,** 13, 1046-52.

222 15. Elsayed, S. S.; Trusch, F.; Deng, H.; Raab, A.; Prokes, I.; Busarakam, K.; Asenjo, J.

223 A.; Andrews, B. A.; van West, P.; Bull, A. T.; Goodfellow, M.; Yi, Y.; Ebel, R.; Jaspars, M.;

224 Rateb, M. E., Chaxapeptin, a lasso peptide from extremotolerant *Streptomyces leeuwenhoekii*

225 strain C58 from the hyperarid atacama desert. *J Org Chem* **2015,** 80, 10252-60.

226 16. Ju, K. S.; Gao, J.; Doroghazi, J. R.; Wang, K. K.; Thibodeaux, C. J.; Li, S.; Metzger,

227 E.; Fudala, J.; Su, J.; Zhang, J. K.; Lee, J.; Cioni, J. P.; Evans, B. S.; Hirota, R.; Labeda, D. P.;

228 van der Donk, W. A.; Metcalf, W. W., Discovery of phosphonic acid natural products by

229 mining the genomes of 10,000 actinomycetes. *Proc Natl Acad Sci U S A* **2015,** 112, 12175-80.

230 17. Heinrikson, R. L.; Meredith, S. C., Amino acid analysis by reverse-phase high-

231 performance liquid chromatography: precolumn derivatization with phenylisothiocyanate.

232 *Anal Biochem* **1984,** 136, 65-74.

233 18. Inokoshi, J.; Matsuhama, M.; Miyake, M.; Ikeda, H.; Tomoda, H., Molecular cloning

of the gene cluster for lariatin biosynthesis of *Rhodococcus jostii* K01-B0171. *Appl Microbiol Biotechnol* **2012,** 95, 451-60.

19. Gavrish, E.; Sit, C. S.; Cao, S.; Kandror, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; Torrey, H.; Akopian, T.; Mueller, A.; Epstein, S.; Goldberg, A.; Clardy, J.; Lewis, K., Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. *Chem Biol* **2014,** 21, 509-18.

20. Metelev, M.; Tietz, J. I.; Melby, J. O.; Blair, P. M.; Zhu, L.; Livnat, I.; Severinov, K.; Mitchell, D. A., Structure, bioactivity, and resistance mechanism of streptomonomicin, an unusual lasso peptide from an understudied halophilic actinomycete. *Chem Biol* **2015,** 22, 241-50.

247 Figure legends

248

249 Fig. 1. a) Alignment of amino acid sequences of achromosin and chaxapeptin precursor peptide
250 genes (Letters with underline: leader peptide, bold letter: conserved amino acid, arrow: cleavage
251 position), b) Gene cluster for biosynthesis of achromosin including 4 genes (*acrA*: structural gene,
252 and modification genes: *acrC*, *acrB1*, and *acrB2*)

253 Fig. 2. MALDI-TOF-MS/MS analyses of achromosin (a) and BNPS-skatole treated
254 achromosin (b) (The oxidized Trp residue is marked by an asterisk.) a) MALDI-TOF-MS/MS
255 analysis of achromosin, b) MALDI-TOF-MS/MS analysis of BNPS-skatole cleaved
256 achromosin (oxidization was indicated with asterisk)

257

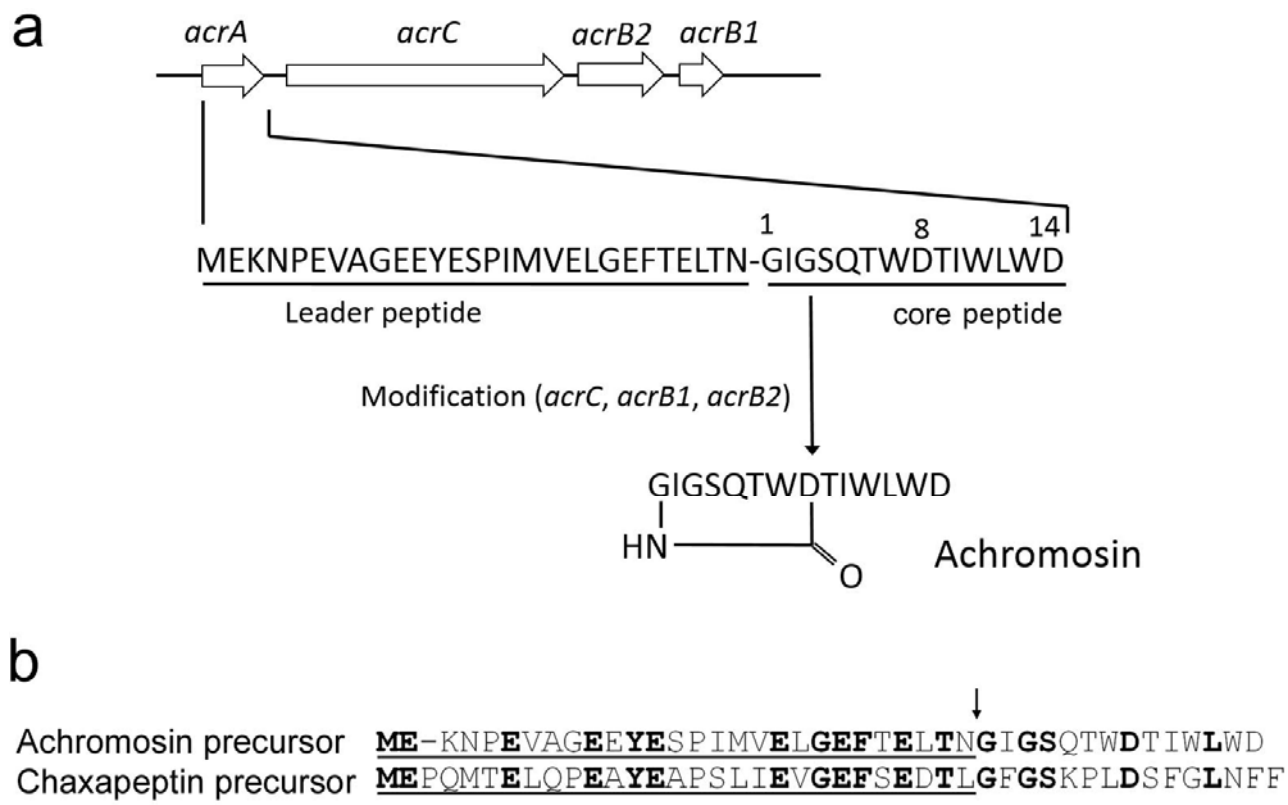

Figure 1. Issara et al.

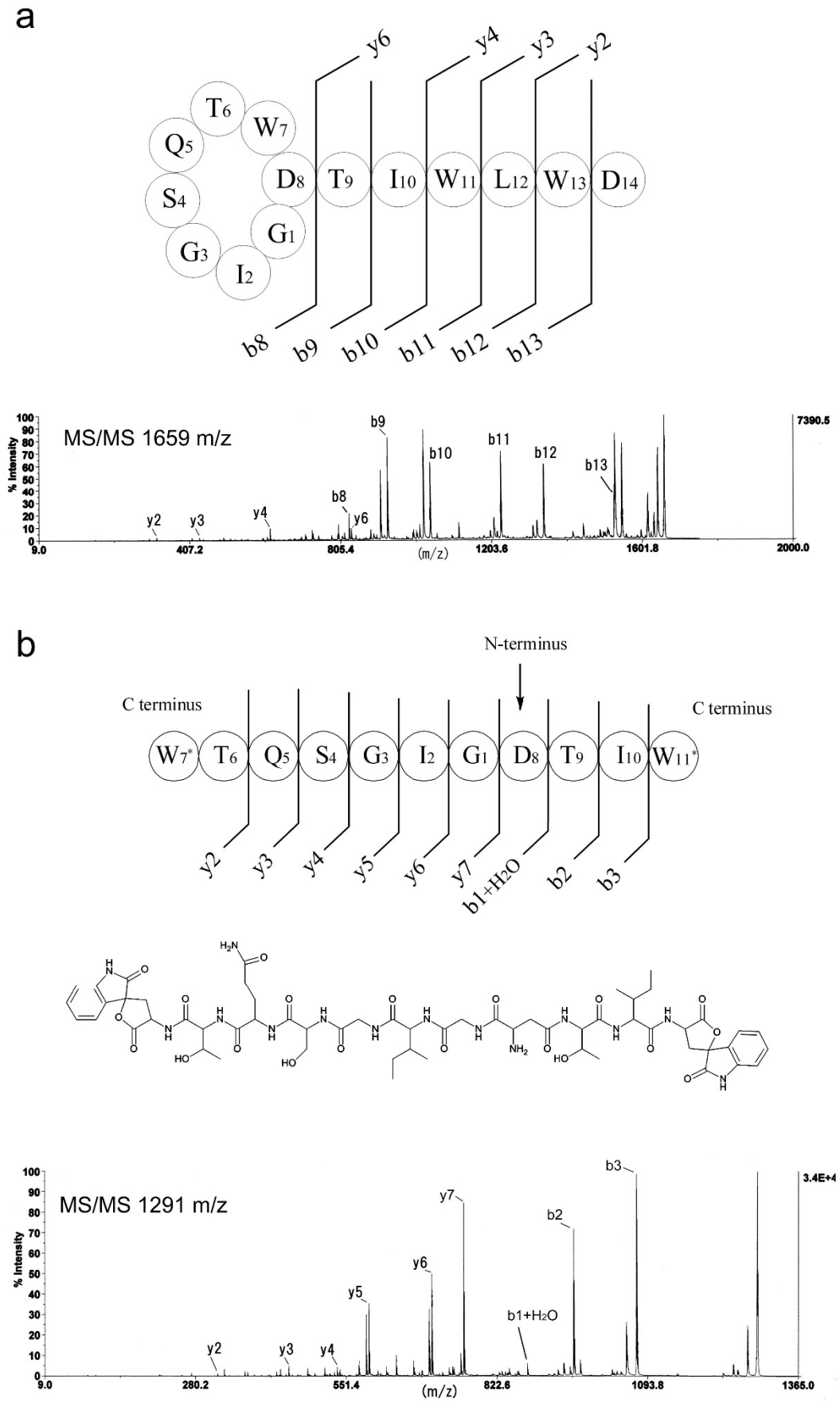

Figure 2. Issara Kaweewan et al.