DOI: 10.1002/ange.201102527

Bacteriocin Structure

The 3D Solution Structure of Thurincin H, a Bacteriocin with Four Sulfur to α-Carbon Crosslinks**

*Clarissa S. Sit, Marco J. van Belkum, Ryan T. McKay, Randy W. Worobo, and John C. Vederas**

Bacteriocins are a group of structurally diverse antimicrobial peptides produced by bacteria to target competing strains of bacteria within their immediate environment. The bacteriocin thurincin H, which is produced by *Bacillus thuringiensis* SF361, exhibits strong activity against a spectrum of *Bacillus* and *Listeria* spp., including the human pathogen *Listeria monocytogenes*.[1] Sequencing of the structural gene indicated that thurincin H is a 31 amino acid peptide with a predicted average molecular weight of 3147.61 Da.[1] However, the observed molecular weight of purified thurincin H was 3139.51 Da, eight Daltons less than predicted, thus suggesting the presence of unusual posttranslational modifications in the mature peptide.[1]

We now report the structural elucidation of thurincin H by using a combination of mass spectrometry and NMR spectroscopy techniques. High-resolution MALDI FTICR mass spectrometry showed that the eight Dalton mass difference was due to a loss of eight hydrogen atoms from the predicted molecular formula of the peptide. MS/MS sequencing performed on thurincin H confirmed the amino acid sequence of the peptide and indicated the positions of posttranslational modification (Figure 1). Specifically, four residues (Asn19, Thr22, Thr25, and Ser28) appear to be two mass units lighter than expected. This phenomenon has previously been observed with the MS/MS sequencing of Trn-α and Trn-β, the two components that constitute the bacteriocin thurincin CD.[2] Since Trn-α and Trn-β each feature three cysteine sulfur to α-carbon bridges at the positions of their modified residues, our MS/MS findings with thurincin H suggest that similar crosslinks exist between the four cysteine residues and four modified residues of the peptide.[2] The presence of *thnB*, a gene that encodes for an iron-sulfur oxidoreductase, in the biosynthetic gene cluster of thurincin H lends further support to this hypothesis.[1] ThnB is a member of the radical *S*-adenosylmethionine (SAM) superfamily of enzymes and shows homology to TrnC/TrnD and AlbA, the oxidoreductases thought to form the sulfur to α-carbon crosslinks in thuricin CD and subtilosin A, respectively.[2,3]

Figure 1. The sequence of thurincin H, as confirmed by MS/MS analysis. Modified residues that appear to have lost two mass units are highlighted in blue.

NMR spectroscopic studies were carried out in solution to determine the structure of thurincin H. [^{13}C,^{15}N]thurincin H was purified from *B. thuringiensis* SF361 grown in a 4:1 mixture of [^{13}C,^{15}N]Celtone-CN-rich media and unlabeled tryptic soy broth. A series of two- and three-dimensional NMR experiments were performed on the peptide to assign the chemical shifts of the majority of the atoms in thurincin H. Notably, the chemical shifts for the α-carbon atoms of Asn19, Thr22, Thr25, and Ser28 are 10 to 15 ppm downfield of the average values expected for unmodified Asn, Thr, and Ser residues in random coil peptides.[4] These downfield values are similar to the chemical shifts of the modified α-carbon atoms in thuricin CD and subtilosin A.[2,5] Examination of the TOCSY and ^{13}C-HSQC data indicated that there are no protons attached to the α-carbon atoms of the modified residues, which is consistent with the presence of sulfur to α-carbon linkages at positions 19, 22, 25, and 28 in thurincin H.

The connectivity of the cysteine residues to the modified residues was determined by analyzing NOE data, which show through-space interactions between atoms that are in proximity to each other. Long-range ^{1}H-^{1}H NOE interactions were observed between the β protons of Cys4 and the amide proton (HN) of Ser28, the β protons of Cys7 and the HN of Thr25, and the β protons of Cys13 and the HN proton of Asn19. Likewise, NOE interactions were seen between the α proton of Cys10 and the HN atom of Thr22, and one of the β protons of Cys10 and the HN proton of Thr22. Altogether, these NOE data indicate that sulfur to α-carbon thioether linkages connect Cys4 to Ser28, Cys7 to Thr25, Cys10 to Thr22, and Cys13 to Asn19.

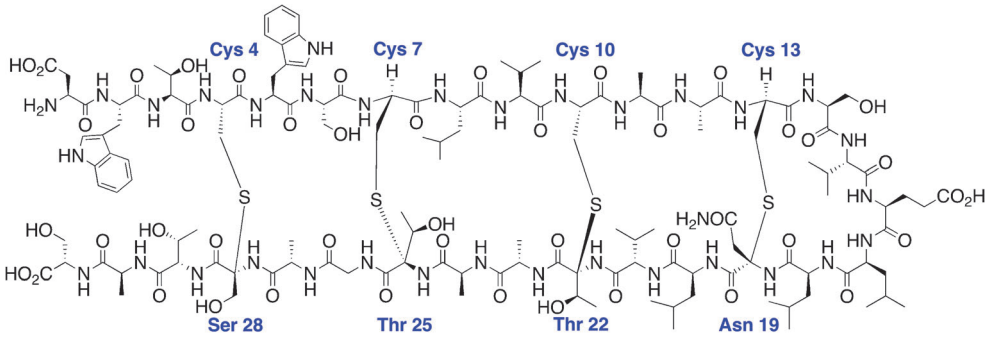
[*] C. S. Sit, M. J. van Belkum, Prof. J. C. Vederas
Department of Chemistry, University of Alberta
Edmonton, Alberta, T6G 2G2 (Canada)
E-mail: john.vederas@ualberta.ca
R. T. McKay
National High Field NMR Centre (NANUC)
University of Alberta
Edmonton, Alberta, T6G 2E1 (Canada)
Prof. R. W. Worobo
Department of Food Science
New York State Agricultural Experiment Station, Cornell University
630 W North St., Geneva, NY 14456 (USA)

[**] We thank Dr. Randy Whittal, Jing Zheng, and Bela Reiz for performing the mass spectrometry analysis. We thank Mark Miskolzie for advice regarding NMR experiments and data analysis. We thank Dr. Leah Martin-Visscher and Dr. Pascal Mercier for their assistance with CYANA. This research was supported by the Alberta Scholarship Programs (to C.S.S.), the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Research Chair in Bioorganic and Medicinal Chemistry, the Alberta Heritage Foundation for Medical Research (AHFMR), and the United States Department of Agriculture—National Integrated Food Safety Initiative (USDA-NIFSI) Grant no. 2008-51110-0688.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201102527.

Since there are four bridges, each of which can adopt one of two possible stereochemical conformations at the α-carbon atom, 16 distinct stereoisomers must be considered when determining the three-dimensional structure of thurincin H. Structure calculations for all 16 of these stereoisomers were carried out using the program CYANA 2.1.[6] Eight rounds of structure calculations were performed on each stereoisomer using the same NMR restraints and, following a protocol similar to that of Salvatella et al., the results were compared to determine which structure fit the NMR spectroscopic data best.[7]

Interestingly, the stereoisomer that gave the best match to the NOE data featured D configurations at Asn19 (α-S), Thr22 (α-S), Thr25 (α-S), and Ser28 (α-S; Scheme 1). Having



Scheme 1. The chemical structure of thurincin H.

the same stereochemistry at all four bridges (namely, the DDDD isomer) is a feature unique to thurincin H; the structures of thuricin CD and subtilosin A indicate that these peptides have LLD and LDD configurations, respectively.[5,8,9] The differences in the configurations may result from the mechanism of action of the radical SAM enzymes, which are thought to form these thioether linkages via a diradical intermediate.[5,9] Depending on the conformation of the active site of the enzyme, rapid bond rotation in the substrate could occur, thereby leading to a radical inversion prior to the formation of the sulfur to α-carbon bond.

The DDDD isomer was chosen as the representative structure of thurincin H instead of the other stereoisomers for several reasons. Firstly, the DDDD isomer was the only structure that did not generate any constraint violations in the CYANA calculations. All the other stereoisomers gave rise to structures with at least one distance, van der Waals, angle, or coupling constant violation. Secondly, the CYANA program incorporated the greatest number of assigned NOE interactions into the structure calculations for the DDDD isomer, with anywhere from 3 to 35 more NOE interactions used than for the other stereoisomers. The DDDD isomer also has, by far, the lowest average target function value of the 16 stereoisomers, which indicates that its structure most accurately reflects the NOE restraint data that formed the original basis of the structure calculations. The backbones of the 20 lowest energy conformers for the DDDD isomer superimpose quite well (see the Supporting Information), with a reasonably low root mean square deviation (rmsd) of (0.74 ± 0.17) Å for the backbone. The structural statistics of the DDDD isomer are summarized in Table 1.

Table 1: Structural statistics for thurincin H (DDDD isomer)

Structural statistics	Thurincin H
distance and angle restraints	
total cross-peak assignments	502
short ($\|i-j\| \leq 1$)	378
medium ($1 < \|i-j\| < 5$)	59
long ($\|i-j\| \geq 5$)	65
number of ϕ angles	22
average target function value	0.03
rmsd (Å) for residues 1–30	
backbone	0.74 ± 0.17
heavy atoms	1.44 ± 0.22

As shown in Figure 2, the three-dimensional structure of thurincin H features a helical backbone that is folded over and held in position by its four sulfur to α-carbon thioether bridges. Similar to thuricin CD and subtilosin A, most of the side chains of thurincin H point outwards (see the Supporting Information), thereby allowing the helical coils of the backbone to pack together more tightly along the central axis of the molecule.[5,8,9]

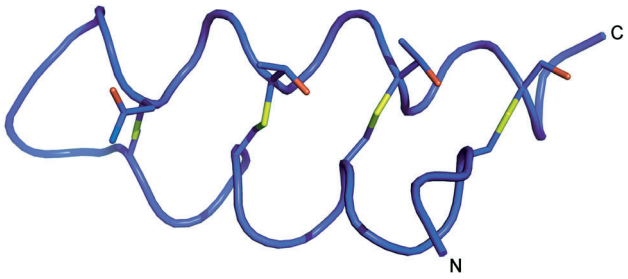


Figure 2. Schematic representation of the three-dimensional solution structure of thurincin H (DDDD isomer). The N and C termini are labeled in the structure.

Another unusual observation from the NMR spectroscopic characterization of thurincin H was that the protons of the Thr29 methyl group have a chemical shift of $\delta = 0.14$ ppm, which deviates significantly from the average expected value of $\delta = 1.2$ ppm for threonine Hγ protons.[4] Closer examination of the 20 lowest energy conformers reveals that the methyl group of Thr29 spends a predominate amount of time being held within 5 Å of the face of the Trp5 indole ring (Figure 3). This suggests that the Hγ protons experience significant diamagnetic anisotropy from the electron cloud of the indole ring, thus providing a rationale for the resultant

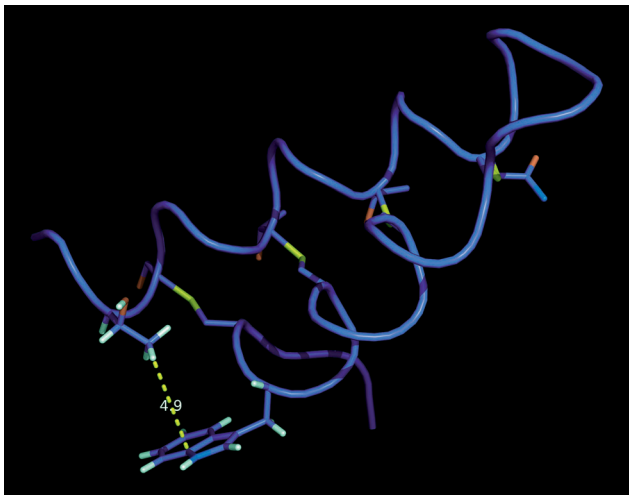


Figure 3. Schematic representation of thurincin H, showing the interaction of Thr29 with Trp5. A distance of 4.9 Å was measured between the two side chains.

upfield chemical shift. This structural feature involving Thr29 and Trp5 represents yet another interaction between two residues widely separated in sequence, thereby reinforcing the tightly packed nature of the peptide's structure.

The electrostatic surface potential of thurincin H is characterized by a net anionic charge (Figure 4A). The regions of negative charge localize at the C-terminal carboxylate and aspartic acid on one end of the molecule, as well as at the glutamic acid extending outwards from the other end of the molecule. Aside from the charged residues, the other hydrophilic residues cluster together on one face of thurincin H, while the hydrophobic residues form prominent patches over the remaining surface of the peptide (Figure 4B). If thurincin H operates by disrupting bacterial cell membranes, similar to the mechanism of action proposed for subtilosin A, then its amphipathic nature would suggest that it can form pores in the membranes of its target strains.[5,10]

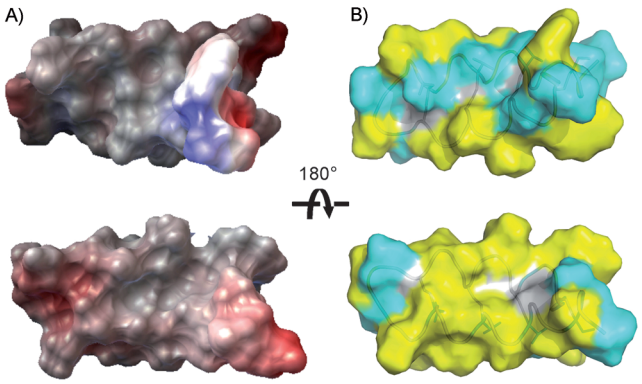


Figure 4. A) Electrostatic surface potential of thurincin H, where blue indicates positive charge and red indicates negative charge. B) Surface hydrophobicity of thurincin H, where yellow represents hydrophobic residues and cyan represents hydrophilic residues.

The structural elucidation of thurincin H is significant not only because it describes the first example of a peptide with four sulfur to α-carbon bridges, but also because it may represent the structure of multiple peptides reported in the literature. From our MALDI MS and FTICR analyses, we found the exact monoisotopic mass of thurincin H to be 3137.36 Da. By comparison, the monoisotopic masses reported for thuricin S and cerein MRX1 are 3137.61 Da and 3137.93 Da, respectively.[11,12] Edman or MS/MS sequencing indicated that both of these peptides have similar, if not identical, N-terminal sequences to thurincin H.[11,12] Gray et al. reported the average molecular weight of thuricin 17 to be 3162 Da. However, a smaller signal at 3139 Da, which is 23 Da or one sodium ion lighter than 3162 Da, can be observed in the MALDI-QTOF spectrum of the peptide.[13] If 3139 Da represents the average molecular weight of the parent $[M+H]^+$ ion, then its monoisotopic mass would be calculated as 3137 Da. Coincidentally, the open reading frame prediction for the thuricin 17 gene gives a predicted peptide sequence identical to that of thurincin H.[14] Likewise, bacthuricin F4 has a homologous N-terminal (DWTXWSXL) sequence as well as physical and biological properties that are highly similar to thuricin 17, and a molecular mass of 3160.05 Da, which also happens to be 23 Da heavier than 3137 Da.[15,16] Although further FTICR-MS/MS analysis would be needed to confirm their amino acid sequences, it is highly probable that thuricin S, cerein MRX1, thuricin 17, and bacthuricin F4 all have the same structure as thurincin H. As such, it is interesting to find that several distinct strains of *Bacillus thuringiensis* produce the same peptide, thus underscoring the biological and ecological importance of this molecule. Our future studies on thurincin H will focus on elucidating its mechanism of action through structure–activity relationship studies and on identifying its cellular target through NMR binding studies.

Received: April 12, 2011
Revised: June 14, 2011
Published online: July 22, 2011

Keywords: bacteriocins · mass spectrometry · NMR spectroscopy · peptides · structure elucidation

[1] H. Lee, J. J. Churey, R. W. Worobo, *FEMS Microbiol. Lett.* **2009**, *299*, 205–213.
[2] M. C. Rea, C. S. Sit, E. Clayton, P. M. O'Connor, R. M. Whittal, J. Zheng, J. C. Vederas, R. P. Ross, C. Hill, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 9352–9357.
[3] G. L. Zheng, L. Z. Yan, J. C. Vederas, P. Zuber, *J. Bacteriol.* **1999**, *181*, 7346–7355.
[4] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR* **1995**, *5*, 67–81.
[5] K. E. Kawulka, T. Sprules, C. M. Diaper, R. M. Whittal, R. T. McKay, P. Mercier, P. Zuber, J. C. Vederas, *Biochemistry* **2004**, *43*, 3385–3395.
[6] P. Guntert, C. Mumenthaler, K. Wuthrich, *J. Mol. Biol.* **1997**, *273*, 283–298.
[7] X. Salvatella, J. M. Caba, F. Albericio, E. Giralt, *J. Org. Chem.* **2003**, *68*, 211–215.

[8] K. Kawulka, T. Sprules, R. T. McKay, P. Mercier, C. M. Diaper, P. Zuber, J. C. Vederas, *J. Am. Chem. Soc.* **2003**, *125*, 4726 – 4727.
[9] C. S. Sit, R. T. McKay, C. Hill, R. P. Ross, J. C. Vederas, *J. Am. Chem. Soc.* **2011**, *133*, 7680 – 7683.
[10] S. Thennarasu, D. K. Lee, A. Poon, K. E. Kawulka, J. C. Vederas, A. Ramamoorthy, *Chem. Phys. Lipids* **2005**, *137*, 38 – 51.
[11] S. Chehimi, F. Delalande, S. Sable, A. R. Hajlaoui, A. Van Dorsselaer, F. Limam, A. M. Pons, *Can. J. Microbiol.* **2007**, *53*, 284 – 290.
[12] S. Sebei, T. Zendo, A. Boudabous, J. Nakayama, K. Sonomoto, *J. Appl. Microbiol.* **2007**, *103*, 1621 – 1631.
[13] E. J. Gray, K. D. Lee, A. M. Souleimanov, M. R. Di Falco, X. Zhou, A. Ly, T. C. Charles, B. T. Driscoll, D. L. Smith, *J. Appl. Microbiol.* **2006**, *100*, 545 – 554.
[14] K. D. Lee, E. J. Gray, F. Mabood, W. J. Jung, T. Charles, S. R. D. Clark, A. Ly, A. Souleimanov, X. M. Zhou, D. L. Smith, *Planta* **2009**, *229*, 747 – 755.
[15] F. Kamoun, H. Mejdoub, H. Aouissaoui, J. Reinbolt, A. Hammami, S. Jaoua, *J. Appl. Microbiol.* **2005**, *98*, 881 – 888.
[16] W. J. Jung, F. Mabood, A. Souleimanov, X. M. Zhou, S. Jaoua, F. Kamoun, D. L. Smith, *J. Microbiol. Biotechnol.* **2008**, *18*, 1836 – 1840.

