VIP Protein Structures Very Important Paper

Deutsche Ausgabe: DOI: 10.1002/ange.201602719
Internationale Ausgabe: DOI: 10.1002/anie.201602719

Radiation Damage and Racemic Protein Crystallography Reveal the Unique Structure of the GASA/Snakin Protein Superfamily

Ho Yeung, Christopher J. Squire, Yuliana Yosaatmadja, Santosh Panjikar, Gemma López, Antonio Molina, Edward N. Baker, Paul W. R. Harris, and Margaret A. Brimble**

Abstract: *Proteins from the GASA/snakin superfamily are common in plant proteomes and have diverse functions, including hormonal crosstalk, development, and defense. One 63-residue member of this family, snakin-1, an antimicrobial protein from potatoes, has previously been chemically synthesized in a fully active form. Herein the 1.5 Å structure of snakin-1, determined by a novel combination of racemic protein crystallization and radiation-damage-induced phasing (RIP), is reported. Racemic crystals of snakin-1 and quasi-racemic crystals incorporating an unnatural 4-iodophenylalanine residue were prepared from chemically synthesized* D- *and* L-*proteins. Breakage of the C–I bonds in the quasi-racemic crystals facilitated structure determination by RIP. The crystal structure reveals a unique protein fold with six disulfide crosslinks, presenting a distinct electrostatic surface that may target the protein to microbial cell surfaces.*

Proteins of the GASA/snakin superfamily are found in many plant species, such as tomato,[1] *Arabidopsis*,[2] petunia,[3] rice,[4] strawberry,[5] and potato,[6] where they are involved in a diverse range of functions including hormonal crosstalk, development, and defense.[7] The snakins are a subset of this superfamily comprising proteins that have antimicrobial properties. To date four distinct snakin proteins have been identified: snakin-1,[6a] snakin-2,[6b] CaSN,[8] and snakin-Z (Figure 1).[9] All members of this group are produced by genes encoding an N-terminal signal peptide, a variable region, and a C-terminal circa 60 residue GASA domain with 12 conserved cysteines.[7]

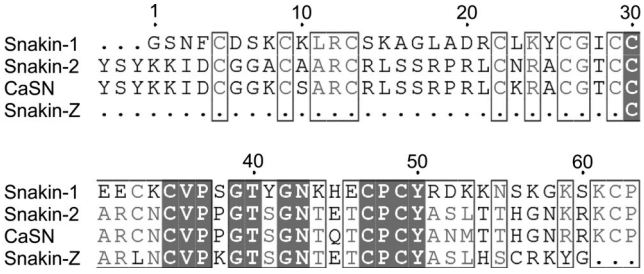


Figure 1. Sequence alignment of the snakin proteins snakin-1, snakin-2, CaSn (putative), and snakin-Z.

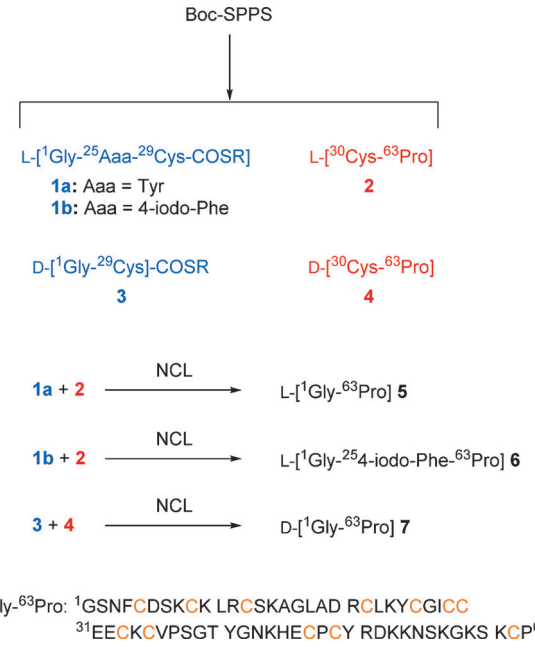
We previously reported the total chemical synthesis of fully functional snakin-1, by native chemical ligation, and revealed the connectivity of three of the six total disulfide bonds by mass spectrometry.[6a,10] The crosslinks identified were different to those suggested by ab initio structure prediction and the protein structure remained unknown.[11] Herein we report the crystal structure of snakin-1 enabled by total chemical protein synthesis to unambiguously resolve the disulfide connectivity and to inform investigations of its antibacterial mechanism of action.

Initial attempts to crystallize synthetic L-snakin-1 from 480 different conditions were unsuccessful. We therefore turned to racemic protein crystallography in which the mirror image D-protein is prepared by chemical synthesis.[12] Racemic protein crystallography is a newly developed approach that benefits protein crystallization by accessing achiral space groups.[13]

Synthesis of the unnatural mirror image D-snakin-1 was completed by *tert*-butoxycarbonyl (Boc) chemistry solid-phase peptide synthesis (SPPS) and native chemical ligation (NCL) using a similar strategy to that used for the synthesis of L-snakin-1 (see Supporting Information).[10] Initial crystallization trials on a racemic mixture of synthetic D- and L-snakin-1 quickly identified more than 10 conditions producing crystals. Several of these crystals were tested at the Australian Synchrotron and gave high quality diffraction data at resolutions up to 1.6 Å, and in centrosymmetric space groups $P\bar{1}$, *Pbcn*, and $P2_1/c$, indicative of true racemic crystals. However, attempts to solve the structure by molecular replacement or ab initio phasing methods were unsuccessful (see the Supporting Information).

We next sought to obtain experimental phases by producing quasi-racemic crystals containing an iodophenyl-

[*] Dr. P. W. R. Harris, Prof. Dr. M. A. Brimble
School of Chemical Sciences, The University of Auckland
23 Symonds St, Auckland Central, 1010 (New Zealand)
E-mail: m.brimble@auckland.ac.nz
H. Yeung, Dr. C. J. Squire, Y. Yosaatmadja, Prof. Dr. E. N. Baker
School of Biological Sciences, The University of Auckland
3A Symonds St, Auckland Central, 1010 (New Zealand)
E-mail: c.squire@auckland.ac.nz
Dr. C. J. Squire, Prof. Dr. E. N. Baker, Dr. P. W. R. Harris, Prof. Dr. M. A. Brimble
Maurice Wilkins Centre for Molecular Biodiscovery
Thomas Building Level 2
3A Symonds St, Auckland Central, 1010 (New Zealand)
Dr. S. Panjikar
Australian Synchrotron
800 Blackburn Road, Clayton, Victoria 3168 (Australia)
G. López, Prof. Dr. A. Molina
Centro de Biotecnología y Genómica de Plantas (UPM-INIA)
Universidad Politécnica de Madrid (UPM)
Campus Montegancedo, M-40 (Km 38)
28223-Pozuelo de Alarcón, Madrid (Spain)
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under http://dx.doi.org/10.1002/anie.201602719.



Scheme 1. Synthetic strategy for the preparation of L-snakin-1 (**5**), iodo-L-snakin-1 (**6**), and D-snakin-1 (**7**).

alanine modification of the L-protein (Scheme 1). Incorporation of heavy atoms, such as Se, Br, or I, provides a means for structure determination using anomalous scattering while retaining the improved likelihood of crystallization of true racemic crystals.[13b]

Iodo-L-snakin-1 was produced with Boc-L-4-iodophenylalanine using the same strategy as used for the synthesis of native L-snakin-1.[10] To obtain a useful anomalous signal, the scattering atom must be well ordered in the crystal. For this reason, a single modification site at ^{25}Tyr was chosen based on secondary structure predictions that suggested this residue was internal. Mass spectrometry of the purified peptide thioester **1b** confirmed iodine incorporation. Native chemical ligation and subsequent oxidative folding afforded a purified product with a mass of 7031.3 ± 0.6 Da. Crystallization screening with the quasi-racemic mixture (iodo-L-snakin-1 + D-snakin-1) again produced crystals in multiple conditions and that diffracted to resolutions of up to 1.5 Å. Single-wavelength anomalous dispersion (SAD) data collection was attempted but the crystals were extremely sensitive to radiation damage and deteriorated extensively before the high data redundancy required for SAD phasing was achieved.

To mitigate the high redundancy required for SAD phasing, we attempted multiple-wavelength anomalous dispersion (MAD) phasing, used successfully to solve the structure of quasi-racemic snow flea antifreeze protein (sfAFP) crystals containing selenium atoms.[13b] This method requires an anomalous scattering atom with an absorption edge readily accessible by X-rays, such as Se or Br. We chose bromine since there are numerous examples of soaking bromide ions into crystals for successful MAD phasing.[14] Quasi-racemic iodo-snakin-1 crystals were soaked briefly in cryo-protectant containing bromide prior to flash-cooling for X-ray diffraction.[15] Multiple MAD data sets were collected, but no anomalous dispersion signal was observed. The data again showed signs of significant radiation damage and suggested radiation damage-induced phasing (RIP) as a means of structure solution.[16]

RIP exploits specific radiation damage occurring between two data sets (before and after radiation damage) to obtain phase information; disulfide bonds and carboxyl groups in protein crystals are particularly susceptible to radiation damage.[16,17] The presence of these chemical groups prior to radiation damage and their loss later in data collection can then provide an exploitable difference, in a manner similar to having a native protein and heavy atom derivative. Retrospective analysis of our snakin-1 structures showed that the C–I bond of the iodophenylalanine residues was particularly susceptible to radiation damage, suggesting iodophenylalanine incorporation combined with RIP as a useful general approach in experimental phasing.

Application of RIP methods to our "MAD" data found multiple sites of specific radiation damage, allowing phasing and automated tracing of the complete backbones of both the L- and D-proteins in the electron density. A final structure of quasi-racemic snakin-1 was obtained with R_{cryst}/R_{free} values of 0.18/0.22 at 1.5 Å in space group $P1$ (see Table S1 in the Supporting Information). Molecular replacement of the iodo-L-snakin-1 model into a true racemic snakin-1 crystal data set produced a final model of the racemate at a resolution of 1.6 Å with R_{cryst}/R_{free} values of 0.24/0.28 in the centrosymmetric space group $P21/c$, equivalent to values of 0.16/0.19 for non-centrosymmetric space groups (Table S3).[18] The coordinates of the final models have been uploaded to the Protein Data Bank with PDB codes 5e5t, 5e5y, and 5e5q.

The structure of snakin-1 comprises two short helices (α1 and α2) forming a helix–turn–helix domain, two large loops held in place by three disulfide bonds, and a short helical region located between these loops (Figure 2A). The short helical region contains a three-residue 3_{10} helix (η1) and a five residue α-helix (α3) divided by a single cysteine that tethers this portion of the backbone to the end of the helix–turn–helix domain (Figure 2B). The second larger loop and the face of the helix–turn–helix domain adjacent to it form a highly positively charged electrostatic surface (Figure 2C). Although there are superficial similarities to the previously predicted structure,[11] the overall topologies are distinct, reflected by a backbone root-mean-square deviation (RMSD) of 5.2 Å, and only two of the disulfides are consistent between the models.

This structure shows some similarity to those of the plant thionin proteins and α-helical hairpin protein classes of cysteine-rich antimicrobial peptides, which also contain a helix–turn–helix domain held together by disulfide bonds (Figure 2D). Structural alignment of the α-helical hairpin protein EcAMP1 with the helix–turn–helix domain of snakin-1 reveals that the helices and two disulfide bonds are arranged identically (Figure S11). The helix–turn–helix domain of the membrane permeabilizing thionin PpTH is also arranged similarly, and this may indicate that the disulfide-bridged helix–turn–helix (dbHTH) domain is an evolutionarily conserved motif.[19]

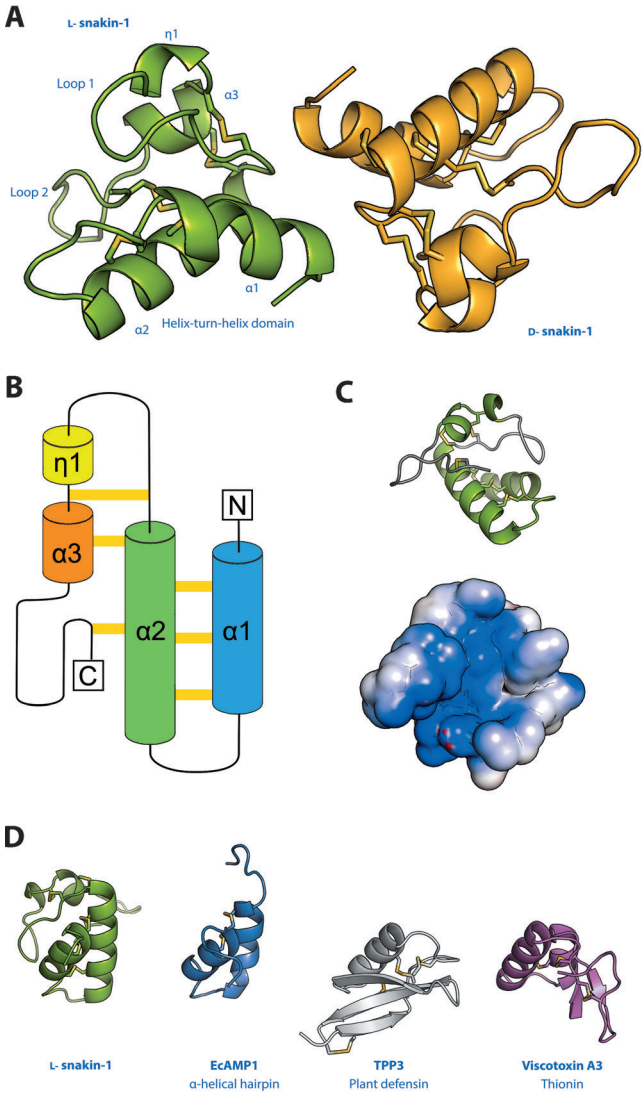


Figure 2. A) Crystal structures of L-snakin-1 (green) and D-snakin-1 (orange) with structural elements labelled. B) Topology diagram showing the backbone connectivity of snakin-1. Disulfides are shown in yellow. C) Ribbon representation of snakin-1 (top) and surface electrostatics of snakin-1 in the same orientation (bottom), showing a highly positively charged cleft (blue). D) Structures of the cysteine-rich antimicrobial peptides L-snakin-1 (green), EcAMP1 (α-helical hairpin; blue), TPP3 (plant defensin; silver), and viscotoxin A3 (thionin; magenta).

In addition to the dbHTH domain, the two rigidly held loops in snakin-1 are presumably a feature shared by the snakin-1 class, as they are held in place by half of the highly conserved cysteine residues characteristic of the GASA/snakin family. That these loops help form a large positive electrostatic surface suggests that they may be important in the targeting of snakin-1 to its site of action, potentially negatively charged bacterial membrane surfaces. In support of a lipid-membrane-targeting mechanism against pathogen species is the distribution of hydrophobic residues in snakin-1, which shows that the protein is amphiphilic, like many known membrane-disrupting antimicrobial peptides (Figure S12).

Previous studies of L-snakin-1 have shown it to be active against *Plectosphaerella cucumerina BBM* and *Fusarium oxysporum f. sp. conglutinans* 699,[6a,20] prompting us to investigate the antimicrobial properties of D-snakin-1 towards these fungal species. Remarkably, the D-protein shows full activity (Figure S13), suggesting that snakin-1 does not act through interaction with any chiral species. It may instead act by mechanical disruption of target membranes similarly to other antimicrobial peptides whose D-enantiomers are also fully active,[21] although an early study of snakin-1 showed no disruption of anionic phospholipid micelles.[22]

Our novel snakin-1 structure will facilitate the investigation of the structure–activity relationship of this protein. For example, a truncated construct of the thionin PpTH comprising only the dbHTH domain retains full activity,[19] thus a similar minimum functional unit may exist for snakin-1. Armed with the structure and disulfide arrangement of snakin-1, we can now prepare truncated analogues and assess their activity. We can also prepare analogues replacing the disulfide bonds with stable isosteres to determine whether the 12 cysteine residues have a functional role.

We have experimentally determined the first structure of a GASA/snakin protein, snakin-1, by X-ray crystallography using total chemical synthesis, racemic and quasi-racemic protein crystallography, and radiation damage-induced phasing methods. This novel protein structure contains a helix–turn–helix domain, an additional helical section comprising a short α-helix and a 3_{10}-helix, two rigidly held loops which may have a functional role, and six disulfide bonds that "glue" the whole structure together. We have unambiguously determined the disulfide connectivity of the snakin-1 protein molecule and shown that similarities exist with the α-helical hairpin and the thionin classes of cysteine-rich antimicrobial peptides within the helix–turn–helix domain. The three-dimensional structure will inform the design of modifications and truncations aiming to produce new antimicrobial peptides.

Acknowledgements

This work was supported by the Marsden Fund (New Zealand; UOA1115) and the Maurice Wilkins Centre for Molecular Biodiscovery. Research was undertaken on the MX1 and MX2 beamlines at the Australian Synchrotron, Victoria, Australia. We thank Dr. T. T. Caradoc-Davies at the Australian Synchrotron for help with X-ray diffraction data collection. We thank Dr. R. Bunker at the Friedrich Miescher Institute for Biomedical Research, University of Basel, and Prof. Dr. S. B. H. Kent at the Department of Biochemistry and Molecular Biology, Institute for Biophysical Dynamic, University of Chicago for useful discussions.

Keywords: peptides · protein structures · racemic protein crystallography · solid-phase synthesis

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 7930–7933
Angew. Chem. **2016**, *128*, 8062–8065

[1] L. Shi, R. T. Gast, M. Gopalraj, N. E. Olszewski, *Plant J.* **1992**, *2*, 153–159.

[2] M. Herzog, A.-M. Dorne, F. Grellet, *Plant Mol. Biol.* **1995**, *27*, 743 – 752.
[3] G. Ben-Nissan, D. Weiss, *Plant Mol. Biol.* **1996**, *32*, 1067 – 1074.
[4] T. Furukawa, N. Sakaguchi, H. Shimada, *Genes Genet. Syst.* **2006**, *81*, 171 – 180.
[5] E. Moyano-Cañete, M. L. Bellido, N. García-Caparrós, L. Medina-Puche, F. Amil-Ruiz, J. A. González-Reyes, J. L. Caballero, J. Muñoz-Blanco, R. Blanco-Portales, *Plant Cell Physiol.* **2013**, *54*, 218 – 236.
[6] a) A. Segura, M. Moreno, F. Madueño, A. Molina, F. García-Olmedo, *Mol. Plant-Microbe Interact.* **1999**, *12*, 16 – 23; b) M. Berrocal-Lobo, A. Segura, M. Moreno, G. López, F. García-Olmedo, A. Molina, *Plant Physiol.* **2002**, *128*, 951 – 961.
[7] V. Nahirñak, N. I. Almasia, H. E. Hopp, C. Vazquez-Rovere, *Plant Signaling Behav.* **2012**, *7*, 1004 – 1008.
[8] Z. Mao, J. Zheng, Y. Wang, G. Chen, Y. Yang, D. Feng, B. Xie, *Phytoparasitica* **2011**, *39*, 151 – 164.
[9] F. Daneshmand, H. Zare-Zardini, L. Ebrahimi, *Nat. Prod. Res.* **2013**, *27*, 2292 – 2296.
[10] P. W. R. Harris, S.-H. Yang, A. Molina, G. López, M. Middleditch, M. A. Brimble, *Chem. Eur. J.* **2014**, *20*, 5102 – 5110.
[11] W. F. Porto, O. L. Franco, *Peptides* **2013**, *44*, 163 – 167.
[12] T. O. Yeates, S. B. H. Kent, *Annu. Rev. Biophys.* **2012**, *41*, 41 – 61.
[13] a) C. K. Wang, G. J. King, S. E. Northfield, P. G. Ojeda, D. J. Craik, *Angew. Chem. Int. Ed.* **2014**, *53*, 11236 – 11241; *Angew. Chem.* **2014**, *126*, 11418 – 11423; b) B. L. Pentelute, Z. P. Gates, V. Tereshko, J. L. Dashnau, J. M. Vanderkooi, A. A. Kossiakoff, S. B. Kent, *J. Am. Chem. Soc.* **2008**, *130*, 9695 – 9701; c) K. Mandal, B. L. Pentelute, V. Tereshko, A. A. Kossiakoff, S. B. Kent, *J. Am. Chem. Soc.* **2009**, *131*, 1362 – 1363; d) J. R. Banigan, K. Mandal, M. R. Sawaya, V. Thammavongsa, A. Hendrickx, O. Schneewind, T. O. Yeates, S. B. Kent, *Protein Sci.* **2010**, *19*, 1840 – 1849; e) R. Okamoto, K. Mandal, M. R. Sawaya, Y. Kajihara, T. O. Yeates, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2014**, *53*, 5194 – 5198; *Angew. Chem.* **2014**, *126*, 5294 – 5298.
[14] M. Dauter, Z. Dauter in *Macromolecular Crystallography Protocols*, *Vol. 364* (Ed.: S. Doublié), Humana, Totowa, NJ, **2007**, pp. 149 – 158.
[15] Z. Dauter, M. Dauter, K. R. Rajashankar, *Acta Crystallogr. Sect. D* **2000**, *56*, 232 – 237.
[16] R. B. G. Ravelli, H.-K. S. Leiros, B. Pan, M. Caffrey, S. McSweeney, *Structure* **2003**, *11*, 217 – 224.
[17] a) M. Weik, R. B. G. Ravelli, G. Kryger, S. McSweeney, M. L. Raves, M. Harel, P. Gros, I. Silman, J. Kroon, J. L. Sussman, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 623 – 628; b) R. B. G. Ravelli, S. M. McSweeney, *Structure* **2000**, *8*, 315 – 328; c) M. H. Nanao, G. M. Sheldrick, R. B. G. Ravelli, *Acta Crystallogr. Sect. D* **2005**, *61*, 1227 – 1237.
[18] V. Luzzati, *Acta Crystallogr.* **1952**, *5*, 802 – 810.
[19] M. Vila-Perelló, A. Sánchez-Vallet, F. García-Olmedo, A. Molina, D. Andreu, *J. Biol. Chem.* **2005**, *280*, 1661 – 1668.
[20] N. Kovalskaya, R. W. Hammond, *Protein Expression Purif.* **2009**, *63*, 12 – 17.
[21] D. Wade, A. Boman, B. Wåhlin, C. M. Drain, D. Andreu, H. G. Boman, R. B. Merrifield, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 4761 – 4765.
[22] J. M. M. Caaveiro, A. Molina, J. M. González-Mañas, P. Rodríguez-Palenzuela, F. García-Olmedo, F. M. Goñi, *FEBS Lett.* **1997**, *410*, 338 – 342.

Received: March 17, 2016
Published online: May 4, 2016

