ORGANIC LETTERS

2012
Vol. 14, No. 22
5720–5723

Total Synthesis, Structure, and Oral Absorption of a Thiazole Cyclic Peptide, Sanguinamide A

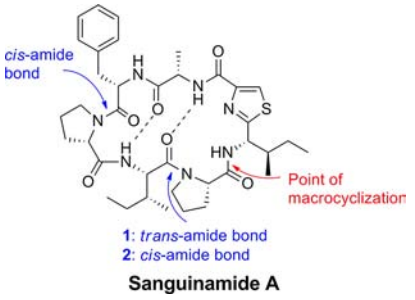
Daniel S. Nielsen, Huy N. Hoang, Rink-Jan Lohman, Frederik Diness,*,† and David P. Fairlie*

Division of Chemistry and Structural Biology, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Qld 4072, Australia

d.fairlie@imb.uq.edu.au; fdi@farma.ku.dk

Received October 3, 2012

ABSTRACT



The first total synthesis and three-dimensional solution structure are reported for sanguinamide A, a thiazole-containing cyclic peptide from the sea slug *H. sanguineus*. Solution phase fragment synthesis, solid phase fragment assembly, and solution macrocyclization were combined to give (1) in 10% yield. Spectral properties were identical for the natural product, requiring revision of its structure from (2) to (1). Intramolecular transannular hydrogen bonds help to bury polar atoms, which enables oral absorption from the gut.

Hexabranchus sanguineus (Spanish dancer) is one of the largest and most colorful nudibranchs[1] (sea slugs), marine invertebrates conspicuous throughout Indian and Pacific ocean coral reefs for their spectacular red or yellow colors and undulating shapes. Many bioactive peptides with oxazole/oxazoline or thiazole/thiazoline heterocycles incorporated in the peptide backbone have been isolated from *H. sanguineus*.[2−4] We have suggested that the many peptides featuring such *azoles*[2−7] be named *azotides*,[6] and our interest has been on how these kinds of heterocycles can act as conformational constraints to influence molecular shape[7] and biological activities[8] of cyclic peptides more generally. Limited quantities of azotides are accessible through isolation from organisms, prompting a need for synthetic routes in order to permit studies of their chemical structures and properties.

Sanguinamide A is a novel thiazole-containing cyclic peptide isolated as a minor component in extracts from *H. sanguineus*.[3] Unlike the cyclic octapeptide analogue sanguinamide B, cyclo-[Val-Ala(Thz)-Leu-Pro(Thz)-Ox-Pro], that features all L-amino acids, two thiazoles,

† Present address: Department of Drug Design and Pharmacology, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark.

(1) Coleman, N. *Nudibranchs of the South Pacific* **1989**, 7.

(2) (a) Pawlik, J. R.; Kernan, M. R.; Molinski, T. F.; Harper, M. K.; Faulkner, D. J. *J. Exp. Marine Biol. Ecol.* **1988**, *119*, 99. (b) Chattopadhyay, S. K.; Pattenden, G. *J. Chem. Soc., Perkin Trans. 1* **2000**, 2429.

(3) Dalisay, D. S.; Rogers, E. W.; Edison, A. S.; Molinski, T. F. *J. Nat. Prod.* **2009**, *72*, 732.

(4) Singh, E. K.; Ramsey, D. M.; McAlpine, S. R. *Org. Lett.* **2012**, *14*, 1198.

(5) (a) Jin, Z. *Nat. Prod. Rep.* **2011**, *28*, 1143. (b) Houssen, W. E.; Jaspars, M. *ChemBioChem* **2010**, *11*, 1803.

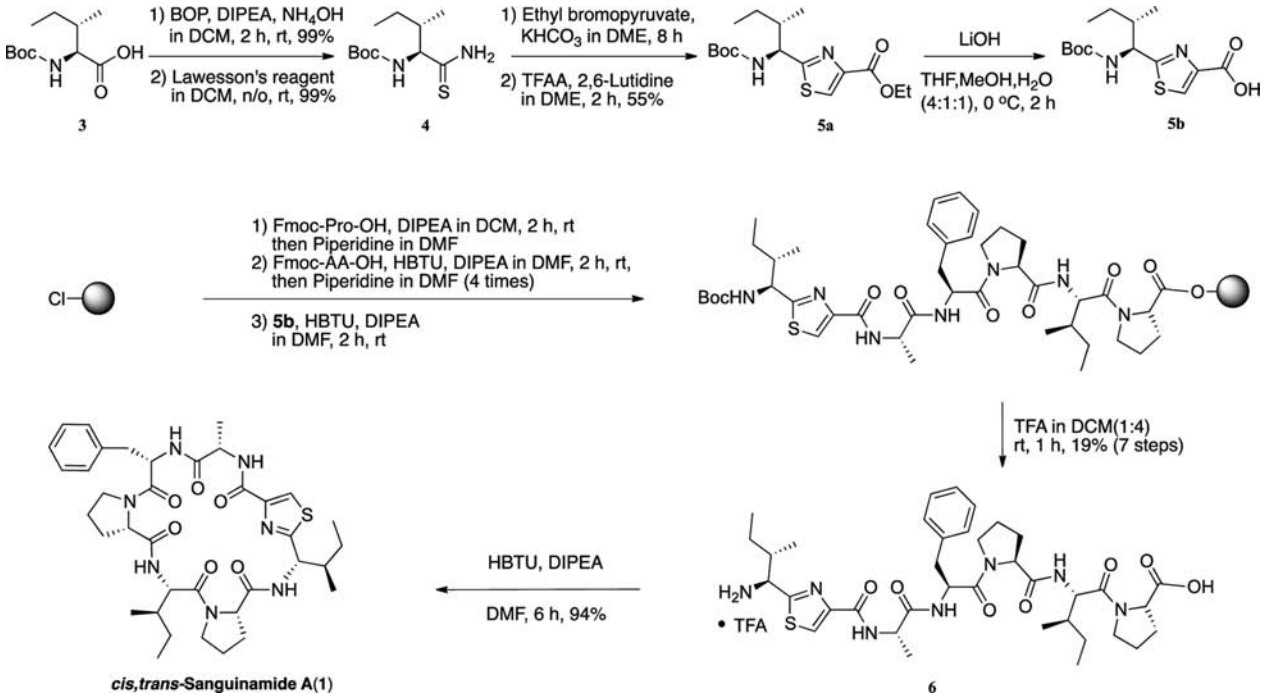
(6) Diness, F.; Nielsen, D. S.; Fairlie, D. P. *J. Org. Chem.* **2011**, *76*, 9845.

(7) (a) Abbenante, G.; Fairlie, D. P.; Gahan, L. R.; Hanson, G. R.; Pierens, G.; van den Brenk, A. L. *J. Am. Chem. Soc.* **1996**, *118*, 10384. (b) Singh, Y.; Sokolenko, N.; Kelso, M. J.; Gahan, L. R.; Abbenante, G.; Fairlie, D. P. *J. Am. Chem. Soc.* **2001**, *123*, 333. (c) Singh, Y.; Stoermer, M. J.; Lucke, A.; Glenn, M. P.; Fairlie, D. P. *Org. Lett.* **2002**, *4*, 3367. (d) Singh, Y.; Stoermer, M. J.; Lucke, A.; Guthrie, T.; Fairlie, D. P. *J. Am. Chem. Soc.* **2005**, *127*, 6563.

(8) (a) Fairlie, D. P.; Abbenante, G.; March, D. *Curr. Med. Chem.* **1995**, *2*, 672. (b) McGeary, R. P.; Fairlie, D. P. *Curr. Opin. Drug Discuss. Dev.* **1998**, *1*, 208.

10.1021/ol3027347
Published on Web 11/06/2012

Scheme 1. Total synthesis of *cis,trans*-Sanguinamide A (**1**)



oxazole, and all trans amide bonds,[4] sanguinamide A is a heptapeptide derivative cyclo-[Ile(Thz)-Ala-Phe-Pro-Ile-Pro] with all L-amino acids, two prolines, phenylalanine, alanine, isoleucine and an isoleucine-thiazole dipeptide surrogate. Two *cis*-amide bonds were reported in this natural product,[3] as shown in **2**. We wished to synthesize sanguinamide A to enable determination of its structure and to identify how the heterocycles might constrain the cyclic peptide shape. Here, we report a facile total synthesis of *cis,trans*-sanguinamide A (**1**) in 10% overall yield and show extensive NMR data that support this as also being the identity of the reported natural product sanguinamide A, rather than the *cis,cis*-sanguinamide A structure **2** attributed previously.

The synthesis was accomplished by combining solution and solid phase synthesis (Scheme 1). Boc-L-isoleucine **3** was converted to thioamide **4**,[7] followed by a modified Hantzsch procedure,[7,9] to give the thiazole-containing dipeptide surrogate **5a** (overall yield 54%). The linear peptide **6** was assembled via solid phase synthesis on a polystyrene resin using a 2-chloro-trityl linker. Macrocyclization was effected between Pro6 and the isoleucine-thiazole, since **5b** was Boc-protected and required acid for N-deprotection with simultaneous cleavage of the linker. Fmoc-L-proline was therefore linked to the solid support for subsequent peptide couplings, including coupling of building block **5b**, mediated by HBTU/DIPEA. This gave polymer-bound Boc-protected linear precursor (Boc-Ile-(Thz)-Ala-Phe-Pro-Ile-Pro), which was deprotected and cleaved from resin with TFA/DCM (1:4). The crude peptide was purified on an RP-18 column and obtained as the TFA-salt **6**. Cyclization was achieved in solution with HBTU, HATU, or BOP at 1 mM concentrations in DMF. LC-MS confirmed rapid, exclusive formation of cyclic product **1**, the highest yield (94%) obtained with use of HBTU. Sufficient quantities of pure synthetic sanguinamide A (**1**) enabled acquisition of multiple 1D and 2D NMR spectra (1D ^{1}H, COSY, TOCSY, NOESY, 1D ^{13}C, HSQC, HMBC) in $CDCl_3$ and d_6-DMSO (Supporting Information (SI)).

NMR spectral data for synthetic sanguinamide A (**1**) matched well with published data[3] for the slightly impure natural product, except for the β-carbon ^{13}C chemical shift of Pro6 previously assigned[3] as 30.5 ppm. A detailed study of the $^{1}H-^{13}C$ HSQC spectrum (Figure 1) showed a clear signal at 25.88 ppm for synthetic **1**, with unambiguous single-bond correlations to the β-protons of Pro6 (Figure 1A). Examination of the published[3] $^{1}H-^{13}C$ HSQC spectrum (Figure 1B) for the natural product assigned as **2** revealed a signal at 30.5 ppm previously thought to couple with the two β-protons of Pro6 at 1.83 and 2.71 ppm. Instead, the published spectrum (Figure 1B) displayed the same correlative $^{1}H-^{13}C$ resonances as the synthetic compound **1** (Figure 1A).

The above HSQC data for both synthetic and isolated sanguinamide A are thus consistent with **1** rather than **2**. The availability of more compound here enabled reinvestigation of the macrocycle structure, leading to this

(9) (a) Brendenkamp, M. W.; Holzapfel, C. W.; van Zyl, W. *Synth. Commun.* **1990**, *20*, 2235. (b) Bredenkamp, M. W.; Holzapfel, C. W.; Snyman, R. M.; van Zyl, W. J. *Synth. Commun.* **1992**, *22*, 3029.

(10) (a) Zimmerman, S. S.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 408. (b) Dorman, D. E.; Bovey, F. A. *J. Org. Chem.* **1973**, *38*, 2379. (c) Wüthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1984**, *180*, 715.

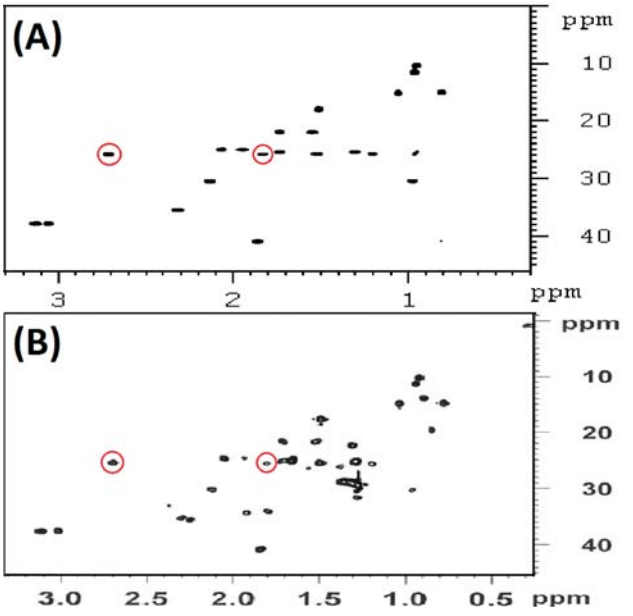


Figure 1. HSQC spectrum of aliphatic region for (A) synthetic *cis, trans*-sanguinamide A (**1**) and (B) isolated "*cis,cis*-sanguinamide A" (**2**) reported in reference 3. Cross-peaks for differently assigned β-protons and carbon of Pro1 circled red.

reassignment. Second, peptide bonds to the proline tertiary nitrogen can adopt a *cis*- or *trans*-conformation,[10a] with *trans*-amide bonds generally preferred over a smaller percentage of *cis*-amide. These geometric isomers can be distinguished by ^{13}C NMR chemical shifts, which differ between C-β and C-γ ($\Delta\delta_{\beta\gamma} = \delta_\beta - \delta_\gamma$), a *cis* configuration having a larger $\Delta\delta_{\beta\gamma}$ compared to a *trans* configuration.[10b] Third, NOE ^1H NMR correlations also reflect a shorter distance in the *cis* isomer between α(Xaa)-α(Pro) protons giving a strong NOE, whereas the *trans* configuration has a shorter distance between α(Xaa)-δ(Pro) protons (Figure 2).[10c] The $\Delta\delta_{\beta\gamma}$ for Pro4 and Pro6 (Figure 2) in the natural product were measured at 8.54 and 0.86 ppm, respectively. These values clearly indicated a *cis* amide for Phe3-Pro4 and a *trans* amide for Ile5-Pro6 (Figure 2). NOEs (SI Figure S9) also strongly supported *cis*- and *trans*-conformations for Phe3-Pro4 and Ile5-Pro6 amides, respectively. All these NMR spectral parameters for synthetic compound **1** are thus identical to those for the natural product and preclude two *cis*-amide bonds as required for structure **2**. Thus, the geometry can be reassigned to *cis,trans-*, rather than *cis,cis*-, sanguinamide A, requiring revision of the natural product structure from the reported structure **2** to the structure **1** found here.

Variable temperature ^1H NMR experiments revealed that two amide-NH protons had low temperature coefficients in d_6-DMSO ($\Delta\partial/T = +0.5$ ppb/K, Ala2; -1.5 ppb/K, Ile5), consistent with being protected from solvent likely due to intramolecular hydrogen bonds (SI Figure S10A). This was supported by H−D exchange experiments in d_6-DMSO containing D_2O, which showed that Ala2 and

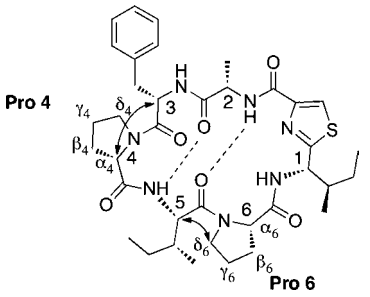


Figure 2. Sanguinamide A with residue numbers (α-, β-, γ-, δ-positions labeled for Pro4, Pro6). Double headed arrows indicate NOE correlations between α- and δ-protons that define *cis*- versus *trans*- amides. Two intramolecular hydrogen bonds between Ala2 and Ile5 are shown by dashed lines.

Ile5 exchanged more slowly than Phe3 (SI Figure S10B). Together, these data support the presence of two intramolecular hydrogen bonds (Figure 2), which direct Ala2 and Ile5 polar atoms to the interior of the cycle while restricting adjacent side chain locations.

The solution structure for **1** was determined in d_6-DMSO at 298 K using NOESY 2D ^1H NMR spectra, calculated from 36 NOE distance restraints, 4 backbone ϕ-dihedral angle restraints derived from $^3J_{NH-CH\alpha}$, one *cis*-amide between Phe3-Pro4, and without any hydrogen bond restraints. Structures were calculated in XPLOR-NIH[11a] using a dynamic simulated annealing protocol in a geometric force field and energy minimized using the CHARMm force field.[11b] The 19 lowest energy structures (Figure 3A) for **1** had no distance (≥0.2 Å) or dihedral angle (≥2°) violations and were quite rigid, convergent structures (ave. pairwise backbone RMSD 0.01 Å). The structure for **1** supported observations made in the VT NMR and H−D exchange experiments, with reciprocal Ala2 NH...OC Ile5 and Ile5 NH...OC Ala2 amide hydrogen bonds forming an antiparallel β-sheet connected by a hairpin turn centered at Phe3-Pro4 (Figure 3A). Pro6 and Ile1(Thz) form an α-turn at the other end of **1**. Side chains from Pro4-Ile5-Pro6-Ile1 create a contiguous hydrophobic surface along one side of the molecule (Figure 3B), which shields polar atoms in Ile5 and Ile1 from solvent. The opposite side of the molecule exposes amide protons from Phe3 and, to some extent, Ala2 making them more accessible to solvent.

Since the cyclic peptide structure forces amides to the interior of the molecule through two hydrogen bonds, and shields others from water through hydrophobic side chains, we were interested in whether **1** would be absorbed into the bloodstream after oral administration to rats. Although Lipinski's rule-of-five[12] is violated by three parameters (MW 721, HBA 13, CLogP 5.4) with high

(11) (a) Brunger, A. T. *X-PLOR Manual*, version 3.1; Yale University Press: New Haven, CT, 1992. (b) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 18.

(12) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3.

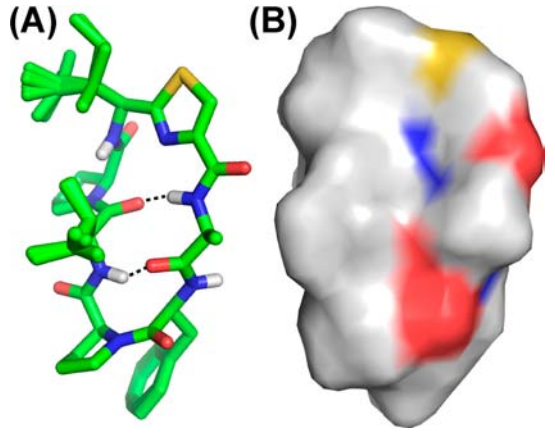


Figure 3. (A) Backbone superimposition of the 19 lowest energy structures for **1** in d_6-DMSO. Nonpolar hydrogens omitted for clarity; hydrogen bonds indicated by dashed lines. (B) Surface of **1** showing hydrophobic surface (gray) impregnated by polar groups (nitrogen, blue; oxygen, red). Ile1 and Ile5 amides are shielded from solvent, while Thz-S (nonpolar sulfur, yellow), Thz-N, Ile1-CO, Ala2-NH (slightly), Phe3-NH, Phe3-CO, Pro4-CO, and Pro6-CO are solvent exposed. Ala2 and Ile5 polar atoms are all hydrogen bonded.

polar surface area (tPSA 169 Å^2), compound **1** was found in sera after oral administration at 10 mg/kg to Wistar rats (F 7 ± 4%, Cmax 40 nM, Tmax 60 min), despite rapid clearance (70 mL/min) and a short half-life ($t_{1/2}$ 23 min). This oral availability is higher than that for most peptides of this size[13] and is attributed to (a) replacement of three amide NH protons by heterocycles, (b) shielding of some polar groups from water by hydrophobic side chains, and (c) two intramolecular hydrogen bonds that force polar atoms to the interior of the macrocycle. It is likely that a contiguous hydrophobic surface is important for membrane permeation, with water-exposed polar groups mostly on one surface of the molecule.

In summary, a total synthesis has been described here for *cis,trans*-sanguinamide A (**1**), featuring a *cis*-amide at Phe3-Pro4. ^{1}H, ^{13}C, HSQC, and NOESY spectra establish unambiguously that the synthetic product is identical to a previously reported[3] natural product, assigned at that time to *cis,cis*-sanguinamide A (**2**) with two *cis*-amide bonds, one at Phe3-Pro4 and one at Ile5-Pro6. The structure of the natural product must now be revised to *cis,trans*-sanguinamide A (**1**). ^{1}H NMR studies support two intramolecular hydrogen bonds that, together with hydrophobic side chains, protect residues Ile1, Ala2, and Ile5 from water solvation. These structural features were predicted to favor a degree of membrane permeability, which was verified by oral absorption from the gut in rats. This finding adds to our knowledge of how cyclic peptides can promote oral bioavailability of peptides that normally do not survive oral delivery *in vivo*. It may help in the design of orally bioavailable compounds that defy rule-of-five parameters normally used to guide medicinal chemists in the development of orally bioavailable drugs.

Acknowledgment. We acknowledge the Carlsberg Foundation (Denmark) for a postdoctoral fellowship (to F.D.); the Australian Research Council for a Federation Fellowship (to D.F.) and the National Health and Medical Research Council for a Senior Principal Research Fellowship (to D.F); and both agencies for Grants (FF0668733, DP1096290).

Supporting Information Available. Detailed experimental procedures, NMR, HPLC, and characterization data of synthetic compounds, and solution structure data for **1**. This material is available free of charge via the Internet at http://pubs.acs.org.

The authors declare no competing financial interest.

(13) (a) Paulettia, G. M.; Gangwara, S; Siahaana, T. J.; Aubéd, J; Borchardta, R. T. *Adv. Drug Delivery Rev.* **1997**, *27*, 235. (b) Hamman, J. H.; Enslin, G. M.; Kotzé, A. F. *BioDrugs* **2005**, *19*, 165.