ORGANIC LETTERS

2000
Vol. 2, No. 17
2721−2723

Structure of (−)-Neodysidenin from *Dysidea herbacea*. Implications for Biosynthesis of 5,5,5-Trichloroleucine Peptides

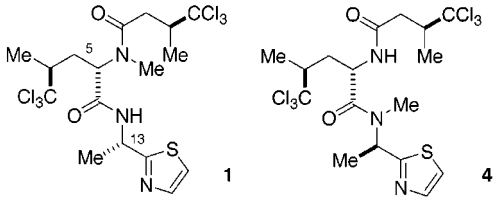
John B. MacMillan, Eugene K. Trousdale, and Tadeusz F. Molinski*

Department of Chemistry, University of California, One Shields Avenue, Davis, California 95616

tfmolinski@ucdavis.edu

Received July 12, 2000

ABSTRACT



Neodysidenin was isolated from the marine sponge *Dysidea herbacea* (Keller 1889) collected on the Great Barrier Reef. The complete configuration was obtained from a combination of methods, including capillary electrophoresis of Marfey's derivatives. Neodysidenin belongs to the L-series of trichloroleucine peptides, and the configuration of the *N*-methyl thiazolyl alanine residue (13*R*) is opposite to that of dysidenin.

Dysidenin (**1**), first discovered by Kazlauskas *et al*.,[1] is a member of a family of chloroleucine peptides produced by a symbiotic assemblage: the marine sponge *Dysidea herbacea* (Keller 1889) and the cyanobacterium *Oscillatoria spongeliae*.[2] Dysidenin is a potent inhibitor of the iodide−sodium co-transporter in bovine thyroid.[3] The ichthyotoxic 5-epimer of **1**, isodysidenin (**2**), was isolated from a sample of *D. herbacea* from New Guinea.[4] Many other examples of trichloromethyl derivatives such as diketopiperazine dimers,[5] pyrrolidinones,[6] and simple *N*-acyl trichloroleucines have been found in *D. herbacea*.[7] It is known that the peptide composition and content of *D. herbacea* is highly variable with geographic location.

The remarkable amino acid *N*-methyl 5,5,5-trichloroleucine (**3a**) is the most common structural motif in *Dysidea* peptides. It appears that introduction of Cl in **1** may involve chlorination of leucine, or an *N*-acyl derivative thereof. The 4,4,4-trichloroisovaleric acid side chain may be a product of catabolism of **3a** or **3b**.[8] We now report the complete structure of (−)-neodysidenin (**4**), a constitutional isomer of **1** containing a non-*N*-methylated trichloroleucine residue. This discovery has implications for the order of reactions involved in the biosynthesis of **1** and **3** and, ultimately, the origin of **3a** and **3b**.

(1) Kazlauskas, R.; Lidgard, R. O.; Wells, R. J.; Vetter, W. *Tetrahedron Lett*. **1977**, 3183−3186.

(2) Hinde, R.; Pironet, F.; Borowitzka, M. A. *Mar. Biol*. **1994**, *119*, 99−104.

(3) (a) Vroye, L.; Beauwens, R.; Van Sande, J.; Daloze, D.; Braekman, J. C.; Golstein, P. E. *Belg. Pfluegers Arch*. **1998**, *435*(*2*), 259−266. (b) Van Sande, J.; Deneubourg, F.; Beauwens, R.; Braekman, J. C.; Daloze, D.; Dumont, J. E. *Mol. Pharmacol*. **1990**, *37*(*4*), 583−9.

(4) (a) Charles, C.; Braekman, J. C.; Daloze, D.; Tursch, B. *Tetrahedron Lett*. **1978**, 1519−1520. Note the configuration of **2** reported in this paper is incorrect. For the correction of configuration (2*S*,5*R*,7*S*,13*S*), see: (b) Biskupiak, J. E.; Ireland, C. M. *Tetrahedron Lett*. **1984**, *25*, 2935−2936 and ref 12b for a comprehensive review of configurational assignments in this family of peptides.

(5) Fu, X.; Ferreira, M. L. G.; Schmitz, F. J.; Kelly-Borges, M. *J. Nat. Prod*. **1998**, *61*, 1226−1231 and references within.

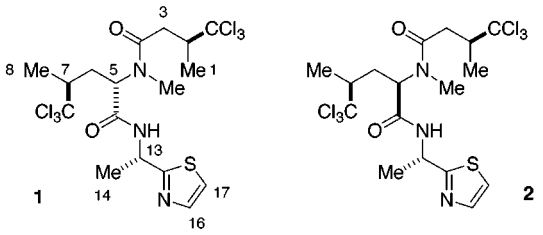
(6) Hofheinz, W.; Oberhänsli, W. E. *Helv. Chim. Acta* **1977**, *60*, 660−669.

(7) (a) MacMillan, J. B.; Molinski, T. F. *J. Nat. Prod*. **2000**, *63*, 155−157. (b) Lee, G. M.; Molinski, T. F. *Tetrahedron Lett*. **1992**, *33*, 7671−7674.

(8) (a) Sitachitta, N.; Rossi, J.; Roberts, M. A.; Gerwick, W. H.; Fletcher, M. D.; Willis, C. L. *J. Am. Chem. Soc*. **1998**, *120*, 7131−7132. (b) Hartung, J. *Angew. Chem., Int. Ed*. **1999**, *38*, 1209−1211.

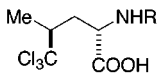
10.1021/ol006326u CCC: $19.00
Published on Web 08/03/2000

(−)-Neodysidenin (**4**) was isolated as a minor constituent of the Indo-Pacific sponge *D. herbacea*,[9] along with the major components **1**, diketopiperazine (−)-**5**,[10] and dysidin.[6]

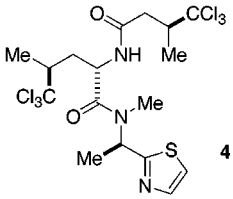


1

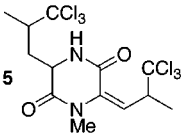
2



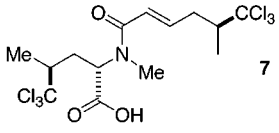
3a R = Me
3b R = H



4



5



7

HRFABMS data ([M + H]$^+$, *m*/*z* 543.9706, Δmmu 1.4) provided the formula $C_{17}H_{23}Cl_6N_3O_2S$ for **4** which is isomeric with that of **1** and **2**. The ^{1}H (Table 1) and ^{13}C NMR data of **4** revealed the presence of two mutually coupled protons belonging to a 2-substituted thiazole (δ 7.34, d, *J* = 3.3 Hz, H-17; 7.74, d, *J* = 3.3 Hz, H-16), but significant differences in ^{1}H and ^{13}C chemical shifts clearly pointed to a compound distinct from **1** or **2**. In particular, the *N*-Me signal (δ 2.94, s) appeared significantly upfield from that of **1** (δ 3.26, s).

Table 1. ^{1}H NMR Data for (−)-Neodysidenin (**4**)[a]

#	δ ($CDCl_3$, 300 MHz)	δ (1:1 C_6D_6/$CDCl_3$, 300 MHz)
1	1.37 (d, 3H, *J* = 6.7 Hz)	1.27 (d, 3H, *J* = 6.4 Hz)
2	3.22 (m, 1H, *J* = 10.2, 6.7, 3.2 Hz)	3.23 (m, *J* = 9.6, 6.7, 2.7 Hz, 1H)
3a	2.32 (m, 1H)	2.02 (dd, 1H, *J* = 14.9, 9.6 Hz)
3b	3.09 (dd, 1H, *J* = 15.3, 2.4)	2.89 (dd, 1H, 14.9, 2.7 Hz)
5	5.06 (ddd, 1H, *J* = 11.1, 8.4, 2.4 Hz)	5.04 (m, 1H, *J* = 9.8, 2.4 Hz)
6a	1.67[b]	1.68 (m, 1H, *J* = 10.1, 7.7 Hz)
6b	2.38 (m, 1H)	2.37 (m, 1H)
7	2.60 (m, 1H)	2.6[b]
8	1.48 (d, 3H, *J* = 6.7 Hz)	1.44 (d, 3H, *J* = 6.7 Hz)
13	6.17 (q, 1H, *J* = 7.1 Hz)	6.13 (q, 1H, *J* = 7.0 Hz)
14	1.67 (d, 3H, *J* = 7.1 Hz)	1.48 (d, 3H, *J* = 7.1 Hz)
16	7.74 (d, 1H, *J* = 3.3 Hz)	7.54 (d, 1H, *J* = 3.0 Hz)
17	7.34 (d, 1H, *J* = 3.3 Hz)	6.88 (d, 1H, *J* = 3.3 Hz)
N-Me	2.94 (s, 3H)	2.64 (s, 3H)
NH	6.70 (bd, 1H, *J* = 8.4 Hz)	

[a] The ^{1}H NMR spectrum of **4** shows a 7:1 ratio of rotamers about the *N*-methylamide bond. Assignments are those of the major rotamer only. Numbering scheme follows that of ref 1. All assignments were made from COSY and decoupling experiments. [b] Obscured by overlap.

^{1}H−^{1}H COSY, long-range COSY, and NOESY (t_m = 600 mS) experiments ($CDCl_3$)[11] revealed that the NH signal was contiguously coupled to signals of the trichloroleucine residue. Conversely, the thiazole-modified alanine signal H-13 appeared as a clean quartet (δ 6.17, q, 1H, *J* = 7.1 Hz) which required placement of the *N*-Me group at the α-nitrogen of the thiazolyl alanine residue. Thus, **4** is related to **1** and **2** by transposition of the *N*-Me group to the thiazolyl-alanine NH and constitutes a rare example of a non-*N*-methylated trichloroleucine peptide.

Independent investigations[6,12] showed that the configuration at each trichloroisopropyl group in **1** and related compounds is invariably *S*; however, C-5 can belong to either the *S* or *R* series as is shown in **1** and **2**, respectively. Optical rotation measurements of dysidenin analogues show that the sign of $[\alpha]_D$ is usually correlated with C-5 stereochemistry.[12b] Both **1** and **4** are levorotatory ($[\alpha]_D$ −98°[1] and −52.1°, respectively, in $CHCl_3$), but the CD spectra of the two compound showed Cotton effects of opposite sign at the $n-\pi^*$ transition maximum for amide bonds [**1**, λ 227 (Δε −5.4); **4**, 227 (+1.4)].

In addition, **4** exhibited a more negative Cotton effect at the thiazole chromophore [λ 245 (Δε −4.5)] which may reflect different configurations at C-5 or C-13 or simply the consequence of different rotamer populations arising from *N*-methylation at a different amide group.

To resolve the ambiguity of the chiroptical data, we turned to degradative methods to assign the configuration of **4**

(9) Isolation of trichloroleucine compounds follows earlier published methods (e.g., Lee, G. M.; Molinski, T. F. *Tetrahedron Lett*. **1992**, *33*, 7671−7674). Briefly, *Dysidea herbacea* (91-063) was frozen at the time of collection (Norman Reef, Great Barrier Reef, February, 1991), lyophilized, and extracted exhaustively with MeOH. The MeOH extract was progressively diluted with water and solvent-partitioned against *n*-hexane, CCl_4, $CHCl_3$, and *n*-BuOH. Chromatography of the $CHCl_3$ extract (silica then Florisil, EtOAc/*n*-hexane gradients) followed by HPLC of (Dynamax, silica, 2:3 EtOAc/*n*-hexane, then 7:3 MTBE/*n*-hexane) gave (−)-**1** and the new compound (−)-neodysidenin (**4**, 0.015% dry wt, >90% pure by ^{1}H NMR) in a ratio of 8:1. $[\alpha]_D$ −52.1° (*c* 0.165, $CHCl_3$). UV (MeOH) 241 nm (ε 4300). ^{1}H NMR, see Table 1. ^{13}C NMR ($CDCl_3$) δ 16.3 q, 16.5 q, 16.8 q, 29.8 q 36.9 t, 40.3 t, 47.5 d, 50.6 d, 51.71 d, 51.74 d, 104.9 s, 105.3 s, 119.9 d, 142.6 d, 169.4 s, 170.4 s, 171.2 s. HRFABMS *m*/*z* 543.9706 [M + H]$^+$, calcd for $C_{17}H_{24}Cl_6N_3O_2S$ 543.9720. Dysidin (ref 6) and (−)-**5** (ref 10) were also isolated from this sample.

(10) (a) Unson, M. D.; Faulkner, D. J. *Experientia* **1993**, *49*, 349−353. (b) Dumdei, E. J.; Simpson, J. S.; Garson, M. J.; Byriel, K. A.; Kennard, C. H. L. *Aust. J. Chem*. **1997**, *50*, 139−144.

(11) Overlap of some signals was resolved by COSY in 1:1 C_6D_6:$CDCl_3$.

(12) (a) Carmely, S.; Gebreyesus, T.; Kashman, Y.; Skelton, B. W.; White, A. H.; Yosief, T. *Aust. J. Chem*. **1990**, *43*, 1882−1888. (b) Unson, M. D.; Rose, C. B.; Faulkner, D. J.; Brinen, L. S.; Steiner, J. R.; Clardy, J. *J. Org. Chem*. **1993**, *58*, 6336−6343. (c) Biskupiak, J. E.; Ireland, C. M. *Tetrahedron Lett*. **1984**, *25*, 2935−2936. (d) Trousdale, E. K.; Taylor, S. W.; Parkin, S.; Hope, H.; Molinski, T. F. *Nat. Prod. Lett*. **1998**, 61−66 and ref 10b.

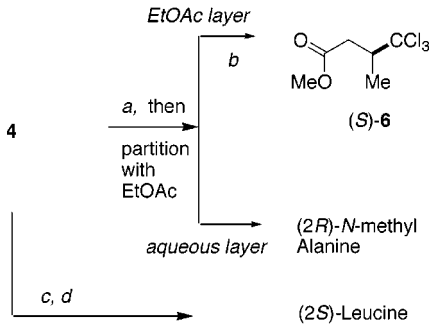
(13) It is interesting to note that, under the same hydrolysis conditions, herbacic acid (**7**, see ref 7a) returned starting material, essentially unchanged, as did other *N*-methyl trichloroleucine derivatives **1** (Molinski, T. F.; Taylor, S. W., unpublished) and dysideathiazole (see ref 12b).

(14) Marfey, P. *Carlsberg. Res. Commun*. **1984**, *49*, 591−596. L-leucine derivative, rt 18.9 min; D-leucine, 23.1 min. See Supporting Information for HPLC experimental conditions.

(15) Acid hydrolysis of *N*-thiazole-modified peptides, without prior oxidative degradation of the heterocyclic ring, results in racemization at the α-carbon (ref 4b).

(16) Tran, A. D.; Blanc, T.; Leopold, E. J. *J. Chromatogr*. **1990**, *516*, 241−249. A 100 mM aqueous borate buffer (pH 8.5) containing 200 mM SDS, 50m × 40 cm capillary, 20 kV. (*S*)-*N*-methylalanine Marfey's derivative, rt 11.2 min; (*R*)-derivative, 11.5 min. Using HPLC, we were unable to obtain satisfactory resolution of these two diastereomers from reagent artifacts. Use of different HPLC gradient elution profiles or attempts to remove interference by destroying excess reagent (NH_3 aq or H_2S aq) were unsuccessful. MECC, a variant of capillary electrophoresis, gave excellent separation, removal of neutral artifacts, and short retention times.

Scheme 1[a]



a (i) O_3, MeOH, -78°C, (ii) H_2O_2, HCOOH, 40° C (iii) 6M HCl, 100°, 10 h; *b*, CH_2N_2, Et_2O-MeOH, 0° C; *c*, 6M HCl, 110°, 10 h, *d*, (i) Zn-AcOH, 60° C, (ii) H_2S aq.

(Scheme 1). Determination of stereochemistry of **4** was made difficult by the limited amount of sample available (~1 mg) which necessitated careful handling and deployment of several sensitive techniques. A sample of **4** (ca. 0.4 mg) was subjected to hydrolysis (6 M HCl, 110 °C, 16 h).[13] The water-soluble fraction was concentrated and subjected to reductive dechlorination (Zn−AcOH, 60 °C), followed by precipitation of Zn salts (H_2S, aq) to obtain a solution containing leucine that was shown to be 2*S* by Marfey's method.[14] Thus, **4** has the (5*S*) configuration and is correlated with dysidenin (**1**) instead of isodysidenin (**2**).

The configuration at C-13 was determined as follows. A second sample of **4** was first ozonolyzed (O_3, MeOH, −78 °C, 15 min), followed by treatment with performic acid (HCOOH, H_2O_2, 50 °C), hydrolysis (6 M HCl, 100 °C, 10 h),[15] and partitioning with EtOAc. Derivatization of the aqueous fraction with Marfey's reagent and analysis of the mixture by micellar electrokinetic capillary chromatography (MECC)[16] revealed the presence of (*R*)-*N*-methylalanine. This is the first report of a peptide with a (13*R*) configuration in the "dysidenin family".

Treatment of the EtOAc-soluble fraction from the hydrolysis with diazomethane followed by chiral GCMS analysis (α-permethylated cyclodextrin capillary column) gave predominantly (*S*)-methyl 4,4,4-trichloromethyl-3-methylbutanoate [(3*S*)-**6**] with a retention time of 11.8 min.[17] Authentic (3*S*)-**6**[18] and (3*R*)-**6** gave retention times of 11.8 and 12.1 min, respectively. Thus, the C-2 configuration of **4** is *S* and it is highly likely the C-7 configuration is the same on the basis of literature precedents.[6,12] The complete configuration of **4** is, therefore, (2*S*,5*S*,7*S*,13*R*).

Non-*N*-methylated trichloroleucine, as found in **4**, is rare; almost all acylated trichloroleucine peptides reported to date are *N*-methylated. It has been proposed that chlorination of leucine occurs prior to *N*-methylation and *N*-acylation as the simplest trichloroleucine derivative (herbacic acid, **7**)[7a] is *N*-methylated and several diketopiperzines of **2** have one or both nitrogens methylated. This finding for **4** clearly shows that *N*-methylation is not an obligatory event prior to peptide chain extension in derivatives of **3**, which leads to the expectation that the enzymes responsible for production of **1**, **2**, and **4** may not be tightly coupled to the *N*-methylase activity. Consequently, it is possible that simple *N*-acylated analogues of leucine may serve as surrogate substrates for the putative "halogenase" [8b] that carries out the remarkable stereospecific substitution of three hydrogens for three chlorines at the unactivated pro-4*S* methyl group in the side chain of leucine.

Acknowledgment. This work was supported by the Office of Research, University of California, Davis. We are grateful to Professor Carlito Lebrilla and Hyun Joo An (University of California, Davis) for assistance with CE measurements and to the Great Barrier Reef Marine Park Authority for research permits.

Supporting Information Available: 1D and 2D spectra of **4**, CD spectra of **1** and **4**, experimental for hydrolyses of **4**, Marfey's analysis, chiral GCMS of **6**, and MECC. This information is available free of charge via the Internet at http://pubs.acs.org.

OL006326U

(17) GCMS of the sample revealed partial racemization of **6** under hydrolysis conditions (6 M HCl, 100 °C, 10 h, ~12:1 ratio of *S*:*R*). We surmise that the mechanism of epimerization at C-3 of **6** is reversible thermal elimination of HCl from the trichloroisopropyl group. During a second hydrolysis of **4**, under more harsh conditions (12 h, 110 °C), the racemization of **6** appeared to be complete.

(18) Brantley, S. E.; Molinski, T. F. *Org. Lett.* **1999**, *1*, 2165−2167.