RZ 3872 (#Z1205-001) 05/14/2012
Physical Sciences 5 pages

Research Report

A combined atomic force microscopy and computational approach for structural elucidation of breitfussin A and B, highly modified halogenated dipeptides from the Arctic hydrozoan *Thuiaria breitfussi*

K. Hanssen[1], B. Schuler[2], A.Williams[3], T.B. Demissie[4], E. Hansen[1], J. Andersen[1], J. Svenson[1], K. Blinov[5], M. Repisky[4], F. Mohn[2], G. Meyer[2], J.-S. Svendsen[1], K. Ruud[4], M. Elyashberg[6], L. Gross[2], M. Jaspars[1], and J. Isaksson[1]

[1]Centre for Research-based Innovation on Marine Bioactivities and Drug Discovery (MABCENT), University of Tromsø, 9037 Tromsø, Norway

[2]IBM Research - Zurich, 8803 Rüschlikon, Switzerland

[3]Royal Society of Chemistry, 904 Tamaras Circle, Wake Forest, NC, 27587, USA

[4]Centre for Theoretical and Computational Chemistry (CTCC), University of Tromsø, 9037 Tromsø, Norway

[5]Molecule Apps LLC, Moscow, Russian Federation

[6]Advanced Chemistry Development, Moscow Department, 6 Akademik Bakulev Street, Moscow 117513, Russian Federation

This is the accepted version of the following article: *Angewandte Chemie Int'l Edition*, vol. 51, no. 49, pp 12238–12241 (December 3, 2012) (DOI: 10.1002/anie.201203960), which has been published in final form at
http://onlinelibrary.wiley.com/doi/10.1002/anie.201203960/abstract

A combined atomic force microscopy and computational approach for structural elucidation of breitfussin A and B, highly modified halogenated dipeptides from the Arctic hydrozoan *Thuiaria breitfussi*

Kine Ø. Hanssen[†], Bruno Schuler[†], Antony Williams, Taye B. Demissie, Espen Hansen, Jeanette Andersen, Johan Svenson, Kirill Blinov, Michal Repisky, Fabian Mohn, Gerhard Meyer, John-Sigurd Svendsen, Kenneth Ruud, Mikhail Elyashberg, Leo Gross, Marcel Jaspars[†], Johan Isaksson*[†]*

[*] K. Hansen, Prof. Dr. M. Jaspars, Dr. J. Isaksson, Dr. J. Svenson, Dr. E. Hansen, Dr. J. Andersen, Prof. Dr. J. S. Svendsen, Centre for Research-based Innovation on Marine Bioactivities and Drug Discovery (MABCENT), University of Tromsø, N-9037 Tromsø, Norway
E-mail: m.jaspars@abdn.ac.uk, johan.isaksson@uit.no
Homepage: http://www0.nfh.uit.no/mabcent/

B. Schuler, Dr. L. Gross, F. Mohn, Dr. G. Meyer, IBM Research - Zurich, 8803 Rüschlikon, Switzerland

A. Williams, Royal Society of Chemistry, 904 Tamaras Circle, Wake Forest, NC, 27587, USA

T. B. Demisse, Dr. M. Repisky, Prof. Dr. K. Ruud
Centre for Theoretical and Computational Chemistry (CTCC), University of Tromsø, N-9037 Tromsø, Norway

K. Blinov, Molecule Apps LLC, Moscow, Russian Federation

Prof. Dr. M. Elyashberg, Advanced Chemistry Development, Moscow Department, 6 Akademik Bakulev Street, Moscow 117513, Russian Federation

[**] Marbank is acknowledged for collection and taxonomy of *thuiaria breitfussi*. Financial support from EU projects ARTIST (Contract No. 243421), HERODOT (Contract No. 214954) and the ERC advanced grant CEMAS, as well as Int.PhD.P.P. of the Foundation for Polish Science co-financed from E.R.D.F. "Grants for Innovation" and Research Council of Norway by grant no. 179568/V30 and 177558/V00, are gratefully acknowledged.

[†] Shared authorships

Supporting information for this article is available from the author.

The use of atomic force microscopy (AFM) with atomic resolution shows great potential for the structural characterization of planar, proton-poor compounds.[1] Structural elucidation of such compounds by spectroscopic methods alone is often problematic resulting in a high proportion of structural corrections.[2] Currently, AFM has limited ability to identify element type and consequently functional groups. Additional computational techniques such as computer-aided structure elucidation (CASE) and the calculation of ^{13}C NMR shifts using electronic structure calculations (DFT) may assist in this respect. We show here the combined use of spectroscopic methods, AFM, CASE and DFT to solve the structures of breitfussins A and B, highly modified halogenated dipeptides, which could not be solved using either method alone, and thus present a novel route in which organic structure analysis may progress in the future.

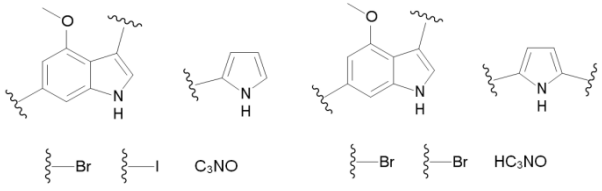
The subject of this study was the Arctic hydrozoan *Thuiaria breitfussi* (Family Sertulariidae). The few publications on the chemistry of this family show the presence of sterols[3] polyhalogenated monoterpenes[4] and anthracenone derivatives.[5] Arctic marine environments support highly diverse and dense populations of marine invertebrates.[6, 7] The adaptation of these organisms to low-temperature habitats include protection against ice formation and changes in enzyme structure to enable them to maintain metabolic rates comparable to those living in warmer conditions. It has been proposed that the ecological stress in cold water is similar to that in warmer waters and that these organisms therefore are expected to produce a range of secondary metabolites. A diverse range of natural products has been found in cold-adapted marine invertebrates and microorganisms, but there appear to be no clear cold-water structural types.[8] Recent work on Arctic invertebrates has yielded novel structures[9] and analogues of known compounds.[10, 11]

The sample of *Thuiaria breitfussi* was collected from Bjørnøya (Bear island) in 2007 and extracted with MeOH/CH_2Cl_2. A series of liquid-liquid extractions yielded 6.2 mg of purified breitfussin A (**1**) and 4.0 mg breitfussin B (**2**) (Figure 1) as the major compounds, using mass-guided HPLC fractionation.

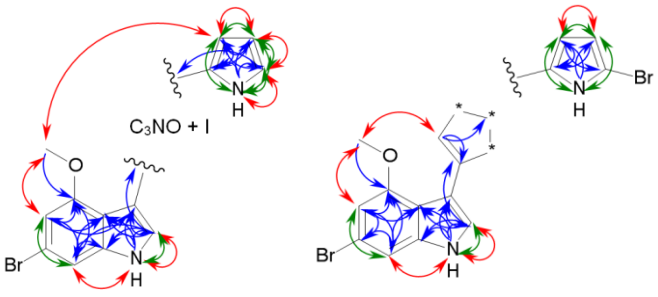
Mass spectrometry. High-resolution mass spectrometry gave the molecular formula $C_{16}H_{11}N_3O_2BrI$ for (**1**) and $C_{16}H_{11}N_3O_2Br_2$ for (**2**), both with concomitant 12 double-bond equivalents. Fragmentation analysis of **1** revealed fragments corresponding to the loss of -I (MS^2), -CH_3O (MS^3), -Br (MS^3) and -CH_3 (MS^3).

NMR spectroscopy. The ratio of heavy atoms to protons (*ca* 2 : 1) indicated that structure determination by spectroscopic methods would be challenging.[12] ^1H NMR in DMSO-d_6 displayed 2 heteroprotons between δ_H 11.5 and 12.5, 6 aromatic ^1H singlets between δ_H 6.0 and 8.0 and a MeO group between δ_H 3.7 and 4.0 for both **1** and **2** (Figures S3 and S8). The ^{13}C NMR spectrum showed 1 methyl carbon, 6 methines and 9 quaternary carbons (Figures S7 and S12). The 1D and 2D NMR data are summarized in Table S2.

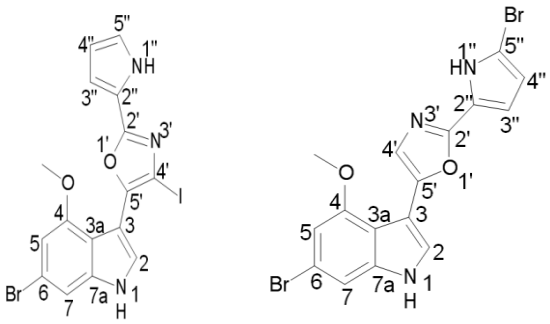
In methanol-d_4, coupling constants were consistent with a 2-substituted pyrrole moiety, which in the case of **2** was also substituted in the 5 position. Thus, **2** was not simply the result of Br replacing I in **1**, but the halogenation pattern was also different. Both molecules further had two aromatic protons in a *meta* relationship (Figures S5, S6, S10 and S11). In addition to the pyrrole, manual structure determination derived two structural elements (Figures 1a and 1b) - a probable 3,4,6-substituted indole and a MeO group, which could be connected to the 4 position of the indole on the basis of an NOE correlation with one of the *meta* coupled singlets (H5) but not to the other (H7) (Figures 1c, 1d, S6 and S11, and Table S2). By incorporating Br/I, this accounted for 10 double bond equivalents, in addition to which the remaining C_3NO in the molecular formula must account for an additional 2 double-bond equivalents. A MS fragment C_9H_8BrNO (223.9707 *m/z*) corresponds to a brominated 4-methoxy-indole. A second fragment $C_7H_6N_2O$ (135.0551 *m/z*) corresponds to the remaining silent C_3NO atoms plus the characteristic pyrrole (C_4H_4N) observed by NMR (Figure S1). A $^3J_{CH}$ from the indole H2 to one of the remaining non-assigned atoms C5' connects this grouping to the C3 of the indole and thus the Br must be attached at C6, giving the substructures presented in Figures 1c and 1d. The remaining grouping could not be unambiguously elucidated from spectroscopic data alone.



(a) (b)



(c) (d)

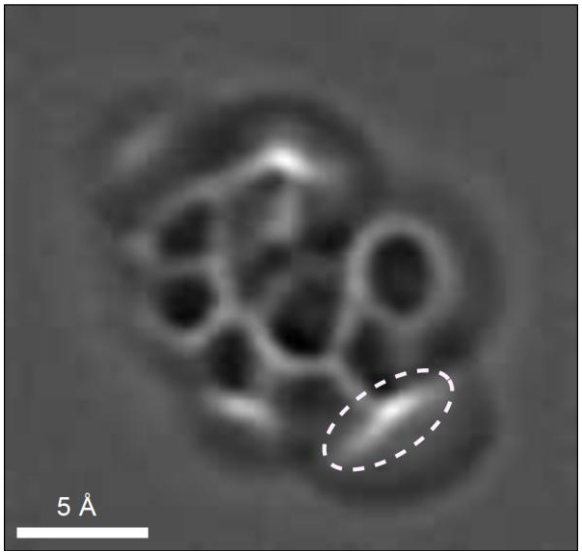


(e) (f)

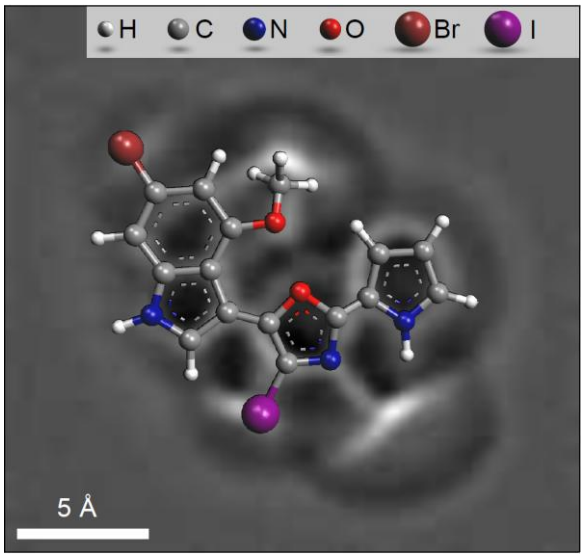
Figure 1. Substructures initially derived by manual analysis of the spectroscopic data for a) **1** and b) **2**. MS fragmentation and NMR correlations (J_{HH}, green; J_{HC}, blue; *ROE*, red) was used to connect all substructures except a C_3NO grouping plus c) I for **1** and d) H for **2** (* = C, N or O). The final suggested structures of e) Breitfussin A (**1**) and f) Breitfussin B (**2**)

The presence of the H4' in the remaining grouping in **2** added some further insight into this moiety. The long-range carbon couplings to C5' (weak) and C2' (strong) indicated that the former is a $^2J_{CH}$ and the latter a $^3J_{CH}$ in a planar ring structure (Figure S9). A strong ROE cross peak between H4' and the MeO protons fixes the position of H4'. The above observations suggested an oxazole substructure, where the exact substitution pattern with respect to the nitrogen and oxygen atoms could not be directly assessed.

AFM. Additional information to support a structural proposal was obtained using AFM with CO functionalized tips[13] on individual molecules of **1** adsorbed on Cu(111) (Figure 2). Atomically resolved AFM measurements revealed a bicyclic system including a 5-membered ring and two additional connected rings where the symmetry could not be resolved unambiguously as a consequence of relaxation due to the attached adatoms. Thus, the indole fragment could be readily confirmed. In addition, the connection points of the rings at 3,5' and 2',2" could be determined. Br and I, which are imaged as elongated single spots probably due to their anisotropic electron distribution, are connected at 6 and 4' or *vice versa* to the ring system. Furthermore we could allocate the major spot in the upper part of Figure 2a as the bulky MeO moiety. Consequently, the structure illustrated in Figure 2b is consistent with the MS, NMR and AFM experimental data. Note that I is imaged with increased contrast compared to the Br, which can be explained by the additional filled shell and the corresponding larger atomic radius of I compared to Br. The prominent spot encircled in Figure 2a is not attributed to an intrinsic molecular feature as it disappeared after the molecule was moved with the tip and it did not appear consistently on every individual molecule that was scanned (Figure S17).



(a)



(b)

Figure 2. (a) Low-pass and Laplace filtered AFM image of **1** on Cu(111) with CO tip. The white encircled region marks a non-intrinsic molecule feature (see AFM Experimental Details and Supplementary Information). (b) AFM image of **1** overlaid with the proposed structure

CASE. ACD/Structure Elucidator was utilized by enumerating all possible structures given chemical shifts and coupling restraints from NMR (Table S3). The resulting structures were ranked on the basis of chemical shift prediction methods.[14] Two iterations of CASE analysis were performed as detailed in the Supplementary information. The first run included several manual NMR-derived constraints but did not yield structures that agreed with the AFM images. A second run was performed with less manually introduced constraints and the resulting highest ranked structure was identical to **1** (Figures S13-16).

DFT calculations. The only remaining possible ambiguity is the practically unlikely but theoretically possible switch of the oxazole nitrogen and oxygen positions, which cannot be *directly* assessed by any of the above methods. Although the CASE analysis ranks the final suggested structures higher than any other of the six possible configurations of this ring for **1** and four for **2** (Scheme S1), there is an intrinsic risk of bias towards known structures due to database coverage. In order to rule this out, relativistic four-component DFT chemical shift calculations[15, 16] were performed on all the theoretically possible reshuffled configurations of the oxazole ring. The calculated chemical shifts all show the least average error for the proposed structures compared to any other configuration for both ^{13}C and ^{1}H shifts and for both **1** and **2** (Tables S4 to S7). Thus, both DFT calculations and the database approach are in agreement with respect to the silent atoms in the oxazole.

Structurally, the breitfussins comprise a novel molecular framework, with the combination of an indole, oxazole and a pyrrole. The diazonamides, isolated from an ascidian of the genus *Diazona*, contain an indole-oxazole moiety, halogenated on the 2 position of the indole and the 4 position of the oxazole.[17] Simpler compounds containing the indole-oxazole grouping have been isolated from a red algae[18, 19] and bacteria.[20, 21] All structures contained few protons, making unequivocal structure determination difficult using spectroscopic means, and thus the structure clarification was in this case achieved using x-ray crystallography. The structure of almazole D[18] was proposed using NMR data in conjunction with an x-ray structure of an analogue, but was later corrected by synthesis.[22] The combination of oxazole joined to pyrrole is even rarer, the only reported instance being the polychlorinated phorbazoles isolated from a sponge of the genus *Phorbas*, and solved using x-ray crystallography.[23] The chlorination is variable, but occurs on the 2, 3 and 4 positions of the pyrrole and the 4 position of the oxazole. The halogenation in breitfussin A is also unusual, with an iodinated oxazole not having been reported previously. The origin of the structure is most likely the dipeptide Pro-Trp having undergone the formation of the oxazole from the peptide bond, oxidation of the Pro to pyrrole and decoration of the molecule with MeO, Br and I (Figure S18).

This study presents the first structure isolated from *Thuiria breitfussi* from the underinvestigated class Hydrozoa. The Breitfussins represent novel halogenated modified dipeptides forming a new structural class containing a combination of indole-oxazole-pyrrole. Given the limited quantity isolated, X-ray crystallography was not possible and structure determination using spectroscopic data was unable to propose a single unequivocal structure consistent with all the data. The structure was therefore supported by a combination of AFM, CASE and DFT calculations, none of which were able to propose a unique solution individually. Remarkably, AFM could be used to determine all the connection positions of the cyclic systems as well as those of the substituent groups: MeO, Br, and I - information that is difficult to obtain with other techniques. Moreover, different AFM contrast above halogen atoms and the MeO group indicate chemical sensitivity within individual molecules. Using such powerful tools on limited quantities of bioactive natural products with complex structures will become more common as scientists begin to access unusual taxa from extreme locations presenting unique chemistry.

Experimental Section

Breitfussin A (**1**) (5-(6-bromo-4-methoxy-1H-indol-3-yl)-4-iodo-2-(1H-pyrrol-2-yl)oxazole) ^{1}H NMR (600 MHz, DMSO-d6) δ 11.95 (m, 1H), 11.78 (d, *J* = 2.8 Hz, 1H), 7.65 (d, *J* = 2.7 Hz, 1H), 7.28 (d, *J* = 1.5 Hz, 1H), 6.96 (m, 1H), 6.73 (d, *J* = 1.5 Hz, 1H), 6.68 (m, 1H),

6.20 (m, 1H), 3.77 (s, 3H). ^{13}C NMR (151 MHz, dmso) δ 157.37, 154.09, 146.86, 138.33, 127.74, 122.75, 119.66, 115.69, 115.48, 110.41, 110.01, 108.39, 104.75, 101.19, 84.42, 56.17. HRESIMS *m/z* 483.9163 $[M + H]^+$ (calcd for $C_{16}H_{12}N_3O_2BrI$ 483.9158).

Breitfussin B (**2**) (2-(5-bromo-1H-pyrrol-2-yl)-5-(6-bromo-4-methoxy-1H-indol-3-yl)oxazole) ^{1}H NMR (600 MHz, DMSO-d6) δ 12.54 (t, *J* = 2.5 Hz, 1H), 11.73 (d, *J* = 2.7 Hz, 1H), 7.77 (d, *J* = 2.6 Hz, 1H), 7.35 (s, 1H), 7.25 (d, *J* = 1.5 Hz, 1H), 6.73 (d, *J* = 1.5 Hz, 1H), 6.71 (dd, *J* = 3.7, 2.6 Hz, 1H), 6.26 (dd, *J* = 3.7, 2.3 Hz, 1H), 3.95 (s, 3H). ^{13}C NMR (151 MHz, dmso) δ 154.09, 153.42, 146.46, 138.94, 124.07, 123.16, 122.27, 115.79, 113.14, 112.12, 111.20, 108.47, 104.47, 104.39, 102.40, 56.03. HRESIMS *m/z* 435.9301 $[M + H]^+$ (calcd for $C_{16}H_{12}N_3O_2Br_2$ 435.9296)

[1] L. Gross, F. Mohn, N. Moll, G. Meyer, R. Ebel, W. M. Abdel-Mageed, M. Jaspars, *Nature Chem.* **2010**, *2*, 821.
[2] M. Elyashberg, A. J. Williams, K. Blinov, *Nat. Prod. Rep.* **2010**, *27*, 1296.
[3] E. Fattorusso, V. Lanzotti, S. Magno, E. Novellino, *Biochem. Syst. Ecol.* **1985**, *13*, 167.
[4] L. De Napoli, E. Fattorusso, S. Magno, L. Mayol, *Biochem. Syst. Ecol.* **1984**, *12*, 321.
[5] C. Pathirana, R. J. Andersen, J. L. C. Wright, *Can. J. Chem.* **1990**, *68*, 394.
[6] B. I. Sirenko, in *Arctic*, **2003**, pp. 49.
[7] B. Gullriksen, R. Palerud, T. Brattegard, J. Sneli, in *Distribution of marine benthic macro-organisms at Svalbard (including Bear Island) and Jan Mayen*, Research Report for DN 1999-4. Directorate for Nature Management, **1999**.
[8] M. D. Lebar, J. L. Heimbegner, B. J. Baker, *Nat. Prod. Rep.* **2007**, *24*, 774.
[9] M. Tadesse, M. B. Strøm, J. Svenson, M. Jaspars, B. F. Milne, V. Tørfoss, J. H. Andersen, E. Hansen, K. Stensvåg, T. Haug, *Org. Lett.* **2010**, *12*, 4752.
[10] M. Tadesse, J. N. Tabudravu, M. Jaspars, M. B. Strøm, E. Hansen, J. H. Andersen, P. E. Kristiansen, T. Haug, *J. Nat. Prod.* **2011**, *74*, 837.
[11] C. A. Volk, M. Köck, *Org. Lett.* **2003**, *5*, 3567.
[12] T. Amagata, M. Editors-in-Chief: Lew, L. Hung-Wen, in *Comprehensive Natural Products II*, Elsevier, Oxford, **2010**, pp. 581.
[13] L. Gross, F. Mohn, N. Moll, P. Liljeroth, G. Meyer, *Science* **2009**, *325*, 1110.
[14] M. E. Elyashberg, A. Williams, G. E. Martin, *Prog. Nucl. Magn. Reson. Spectrosc.* **2008**, *53*.
[15] M. Repiský, S. Komorovský, O. L. Malkina, V. G. Malkin, *Chem. Phys.* **2009**, *356*, 236.
[16] S. Komorovsky, M. Repisky, O. L. Malkina, V. G. Malkin, *J. Chem. Phys.* **2010**, *132*, 154101.
[17] N. Lindquist, W. Fenical, G. D. Van Duyne, J. Clardy, *J. Am. Chem. Soc.* **1991**, *113*, 2303.
[18] I. N'Diaye, G. Guella, I. Mancini, F. Pietra, *Tetrahedron Lett.* **1996**, *37*, 3049.
[19] S. Takahashi, T. Matsunaga, C. Hasegawa, H. Saito, D. Fujita, F. Kiuchi, Y. Tsuda, *Chem. Pharm. Bull.* **1998**, *46*, 1527.
[20] G. R. Pettit, J. C. Knight, D. L. Herald, R. Davenport, R. K. Pettit, B. E. Tucker, J. M. Schmidt, *J. Nat. Prod.* **2002**, *65*, 1793.
[21] K. Umehara, K. Yoshida, M. Okamoto, M. Iwami, H. Tanaka, M. Kohsaka, H. Imanaka, *J. Antibiot.* **1984**, *37*, 1153.
[22] F. Miyake, M. Hashimoto, S. Tonsiengsom, K. Yakushijin, D. A. Horne, *Tetrahedron* **2010**, *66*, 4888.
[23] A. Rudi, Z. Stein, S. Green, I. Goldberg, Y. Kashman, Y. Benayahu, M. Schleyer, *Tetrahedron Lett.* **1994**, *35*, 2589.