Pergamon

Bioorganic & Medicinal Chemistry Letters 10 (2000) 1105–1108

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Diastereoisomers of an 'Arsenomethionine'-Based Structure from *Sargassum Lacerifolium*: The Formation of the Arsenic–Carbon Bond in Arsenic-Containing Natural Products

John S. Edmonds*

Department of Chemistry, De Montfort University, The Gateway, Leicester LE1 9BH, UK

Received 11 February 2000; accepted 14 March 2000

Abstract—Separation and ^{1}H NMR spectra of a pair of arsenic-containing diastereoisomers (**1a** and **1b**) isolated from a brown alga has provided support for their structures (proposed on the basis of NMR spectra of the unseparated mixture). The diastereoisomerism and analogies with nitrogen-containing algal lipids indicated that they were derived from an analogue of methionine in which the dimethylarsinoyl- group had replaced amino. Although *S*-adenosylmethionine is probably the source of methyl and 5′-deoxyribos-5′-yl groups in arsenic-containing natural products, the arsenic–carbon bonds in some compounds might be formed by a process in which arsenic replaces nitrogen in amino-acid synthesis.

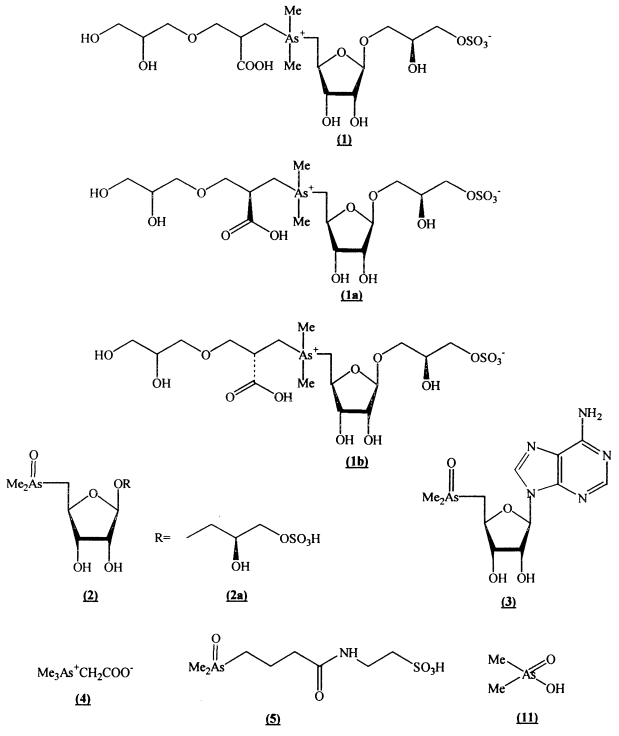
Compound **1** was isolated in low amount from the brown alga *Sargassum lacerifolium* where it represented about 0.1% of the total water-soluble arsenic present.[1] The ^{1}H and ^{13}C NMR spectra of **1** showed, through an apparent duplication of signals, that it probably consisted of a pair of unseparated diastereoisomers. The degree of separation of the duplicated signals suggested that the pair might be epimeric at the carbon bonded to the carboxyl group. There was no indication that the two further chiral carbon atoms in the molecule, or the ribose, were other than stereochemically homogeneous. I now report that separation of the pair of diastereoisomers (**1a** and **1b**) was effected by HPLC on an analytical ODS reversed phase column (Inertsil ODS-2, 4.6 mm i.d.× 250 mm, Gasukuro Kogyo, Japan; elution with 10 mM tetraethylammonium hydroxide, 4.5 mM malonic acid, 0.05% methanol at pH 6.8).[2] ICP–MS (Yokogawa PMS 2000) was used as an arsenic-specific detector to locate the two arsenic compounds. Buffer was removed from the separated arsenic compounds by passage through a Sephadex LH-20 column (850×26 mm), elution with 50% aq methanol. The ^{1}H NMR spectra of the two compounds[3] confirmed that the greatest differences, one from the other, were for the protons on carbon atoms next to the proposed epimeric centre. Thus, the chemical shifts[3] of corresponding methylene protons in **1a** and **1b** adjacent to C2″ differed by 0.02 to 0.04 ppm, whereas methylene protons adjacent to the other chiral carbon atoms in the molecule, and for the ribose moiety, did not differ between the two diastereoisomers by greater than 0.01 ppm. The suggested diastereoisomeric nature of **1** as originally reported was thus reinforced and interesting aspects of the biosynthesis of the compounds were raised.

Marine algae contain most of their arsenic (usually in the range of 2 to 40 mg/kg wet weight) in the form of dimethylarsinoylribosides **2**.[4] The biosynthesis of these compounds probably involves the transfer of two methyl groups from *S*-adenosylmethionine (AdoMet) to arsenic since this is the main biological donor of methyl groups and, in these methylation steps, probably follows the pathway suggested by Challenger[5] for the methylation of arsenic by the bread mould *Scopulariopsis brevicaulis*. AdoMet and its decarboxylated product can also act as donors of 3-amino-3-carboxypropyl and 3-aminopropyl groups, respectively; clearly each of the three alkyl groups, of the sulfonium ion is easily attacked by nucleophiles. Hence, it would not be surprising if AdoMet were also the donor to arsenic of the third such group, namely the 5′-deoxyadenosin-5′-yl.[6] Support for this proposal was provided by the isolation[7] of dimethylarsinoyladenosine **3**, a predicted intermediate,[8] from extracts of giant clam kidney (a rich source of dimethylarsinoylribosides **2**, possibly as a result of symbiotic algae in clam tissues[9]). Arsenobetaine **4**, the major form of arsenic in marine animals,[10] might come from breakdown of these ribosides, with the two carbons of the carboxymethyl group of arsenobetaine **4** being

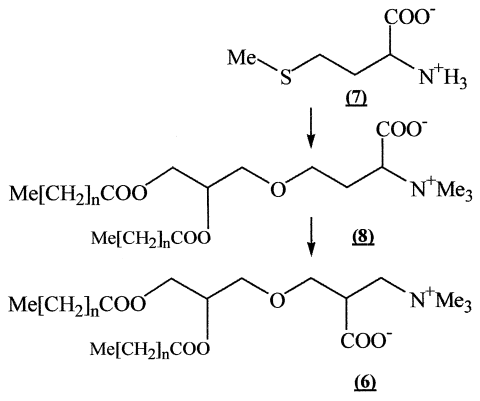
*Fax: +44-116-257-7135; e-mail: jedmonds@dmu.ac.uk

0960-894X/00/$ - see front matter
PII: S0960-894X(00)00176-1

derived from the 4 and 5 carbons of the ribose.[8] The site or sites of this breakdown and the reduction and further methylation of arsenic are rather ill-defined, although anaerobic sediments have been suggested.[8]



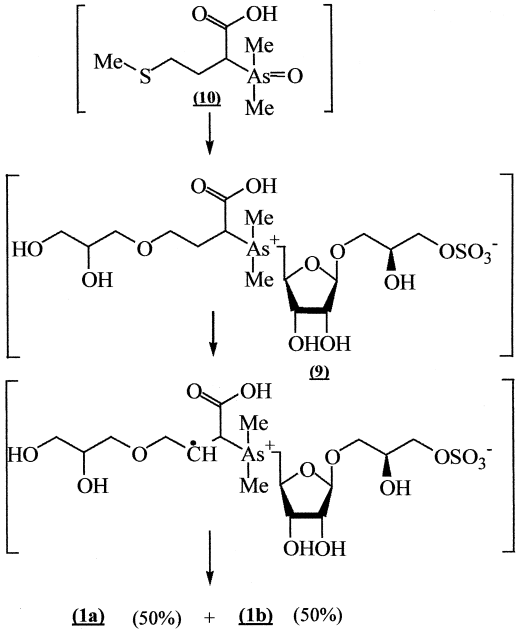
Compounds **1a** and **1b** from *Sargassum*, however, contain a substituent on arsenic that seems unrelated to any that could be provided by AdoMet. This is also true of compound **5** isolated from *Tridacna* kidney,[11] which has been identified by NMR spectroscopy and its structure confirmed by synthesis.[10] So the origin of these substituents can now be considered. Compounds **1a** and **1b** are similar to the polar portion of lipid **6** an important membrane component of several species of algae,[12] except that where **6** has the trimethylammonio group, **1a** and **1b** have an arsonio group that carries two methyls and a 5′-deoxyribos-5′-yl group as substituents. This last group is related to the major dimethylarsinoylribose derivative **2a** of *S. lacerifolium*. The similarities between compounds **1a** and **1b** and **6** suggest that they might originate by parallel biosynthetic processes.



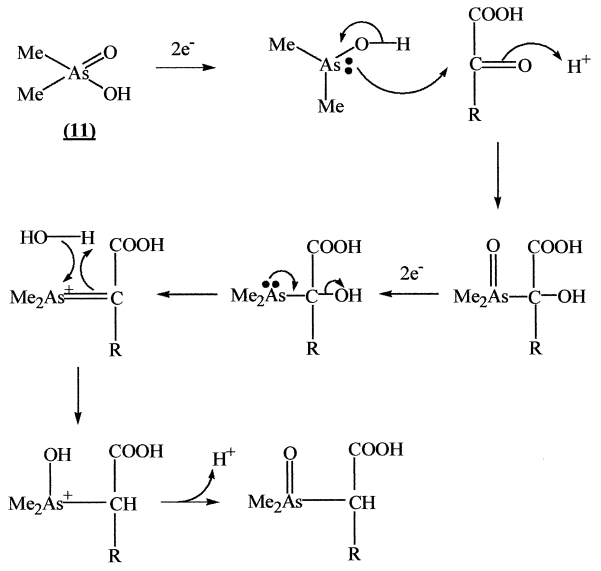
Scheme 1.

Studies[13,14] on **6** have determined that the C_4 unit, including its carboxy group and nitrogen, is derived from methionine, **7**, as are the methyl groups, presumably through AdoMet (Scheme 1). The initially formed homoserine derivative, **8**, which is also the basis of a membrane lipid in some algae, is converted into **6** by transfer of the carboxy group from C-2 to C-3.[14] Such exchanges are typically mediated by enzymes that use coenzyme B_{12} and involve the abstraction of the hydrogen atom with its lone electron from the substrate leaving a substrate radical.[15] The R group (carboxy in this case) then moves to the atom that bears the lone electron. This mechanism would involve the intermediate-CH·–CH(COOH)- radical which has the potential to epimerize. No stereochemical information has been published for the nitrogen-containing lipid **6**, but the existence of the diastereoisomeric pair of arsenic compounds **1a** and **1b** strongly suggests that they have arisen through a radical mechanism and thus have a common precursor which contains $X–CH_2–CH_2–CH(COOH)–As(O)Me_2$.

It is necessary then, when proposing a biosynthetic route for **1a** and **1b**, to consider how a compound containing such a component might have arisen, and how X might be replaced by glycerol. Replacement of X by glycerol would be simplest if X were Me-S- (Scheme 2) because then adding the adenosyl group will convert the thioether into a sulfonio group and hence convert $-CH_2–CH_2–CH(COOH)–As(O)Me_2$ into an alkylating agent. This could alkylate glycerol to obtain the product found as is indicated in Scheme 2. Thus a direct parallel with the biosynthesis of lipid **6** is suggested; nevertheless, there are other ways in which the same result might be achieved, such as acylating the hydroxy group of a homoserine analogue. As to the origin of a compound containing-CH_2-CH_2–CH(COOH)–$As(O)Me_2$, Scheme 3 shows a possible method. It is based on the analogy



Scheme 2.



Scheme 3.

noted above between nitrogen and arsenic compounds. All biological formation of C–As bonds appears to be alkylation of As(III), and just as amino acids can be formed by transamination, effectively alkylation of ammonia using a 2-oxo acid, so Scheme 3 shows how a 2-oxo acid might similarly act to alkylate arsenic. Finally, after reduction of the $Me_2As(O)$- group it could be given a 5′-deoxyadenosin-5′-yl group from AdoMet. Such steps may be correct even if they do not occur in this order.

Additional support for the idea of an arsenic-containing group replacing the ammonium ion or amino group in amination or transamination reactions is provided by consideration of compound **5** isolated from *Tridacna* kidney. In this case it was proposed[11] that the 4-carbon part bound to arsenic was derived from AdoMet in a manner analogous to the biosynthesis of spermine and spermidine. However, an unusual reductive deamination would be required before conjugation with taurine. In the light of the consideration of the biosynthesis of **1a** and **1b** above, it is feasible that **5** is formed by 'arsenylation' (again the dimethylarsinoyl-group seems most likely) of 2-oxo glutarate to form an arsenic equivalent of glutamate, which would decarboxylate and conjugate with taurine to give **5** (Scheme 3, $R = [CH_2]_2COOH$). It is possible then that an arsenic species might replace ammonium in this core pathway for the biosynthesis of amino acids.

It is interesting to speculate that a similar 'arsenylation' of pyruvate (paralleling the biosynthesis of alanine — Scheme 3, R = Me) or glyoxylate (Scheme 3, R = H) could give rise to arsenobetaine **4**. If a dimethylated arsenic group was involved further reduction and methylation at the arsenic atom would be required, and, for pyruvate, a demethylation of the 2-carbon, or, possibly, a shift of the methyl group from carbon to arsenic would also be necessary. The 'arsenylation' of 2-oxo acids, paralleling their amination, would naturally produce arsenic compounds analogous to α-amino acids and the similarity of arsenobetaine to glycine betaine would appear less coincidental.

In the above cases it is assumed that it is the ubiquitous metabolite of arsenic dimethylarsinic acid **11** that is involved in 'arsenylation' of 2-oxo acids. A possible mechanism for this reaction (Scheme 3) would involve attack by the lone pair of the reduced dimethylated arsenic on the carbonyl carbon (this reduced species is a postulated intermediate in the Challenger[5] pathway). A further reduction of arsenic (again analogous to the Challenger pathway) and the lone pair on arsenic could displace hydroxide ion. Addition of water could then lead to the compounds discussed (or to a precursor in the case of arsenobetaine).

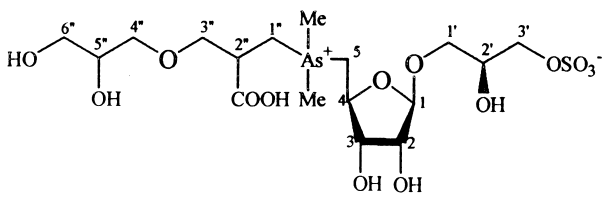
The concentrations of **1a** and **1b** and of **5** in *Sargassum* and *Tridacna*, respectively, are very low. Compounds **1a** and **1b** together made up only 2.5×10^{-7} of the dry weight of *Sargassum* (approximately 0.04 μM in the live alga; 0.1% of total arsenic), although more of the compounds might have been present, acylated with long-chain fatty acids, in lipid material that was not examined.[1] Compound **5** contributed about 8.5×10^{-6} of the dry weight of *Tridacna* kidney (about 5 μM in live kidneys; 0.2% of total arsenic).[11] The question arises as to whether compounds at this concentration represent directed biosyntheses or accidental involvement of arsenic in other processes. In other words, is the presence of these compounds merely a measure of the fidelity of amination processes and the concentration of arsenic in the cell? In contrast, arsenobetaine **4** is a major metabolite of arsenic contributing from about 5 to 500×10^{-6} of the dry weight of most marine animals (0.015 to 1.5 mM in living tissues) and usually accounting for virtually all arsenic present.[10] Although attempts have been made to explain the origin of this compound,[4,10] its biogenesis remains unresolved, and it is in this context that the processes outlined here are currently being investigated.

Acknowledgements

I thank Drs K. A. Francesconi, H. B. F. Dixon and Y. Shibata for their contributions to the work reported here.

References and Notes

1. Francesconi, K. A.; Edmonds, J. S.; Stick, R. V.; Skelton, B. W.; White, A. H. *J. Chem. Soc., Perkin Trans. 1*, **1991**, 2707.
2. Shibata, Y.; Morita, M. *Anal. Sci.* **1989**, *5*, 107.
3. ^{1}H NMR spectra



(500 MHz, D_2O) Fast diastereoisomer (eluting first from ODS column). δ 1.92, 1.96, 2s, Me_2As^+; 2.59, dd, $J_{1'1''}$ 13.7 Hz, $J_{1'',2''}$ 4.4 Hz, H1″; 2.64, dd, $J_{5,5}$ 14.3 Hz, $J_{4,5}$ 10.2 Hz, H5; 2.70, dd, $J_{1'',1''}$ 13.7 Hz, $J_{1'',2''}$ 10.9 Hz, H1″; 2.82, dd, $J_{5,5}$ 14.3 Hz, $J_{4,5}$ 2.8 Hz, H5; 2.93, m, H2″; 3.49, dd, $J_{4'',4''}$ 10.9Hz, $J_{4'',5''}$ 6.5Hz, H4″; 3.53, dd, $J_{6'',6''}$ 11.8 Hz, $J_{5'',6''}$ 6.2 Hz, H6″; 3.57–363, m, H1′, 4″, 6″; 3.74, m, H3″, 3″; 3.80, dd, $J_{1',1'}$ 10.8Hz, $J_{1',2'}$ 4.3 Hz, H1′; 3.87, m, H5″; 4.01–409, m, H2′, 3′, 3′; 4.12, d, $J_{2,3}$ 3.9 Hz, H2; 4.19.25, m, H3, 4; 5.00, s, H1. Slow diastereoisomer. δ 1.91, 1.96, 2s, Me_2As^+; 2.55, dd, $J_{1'',1''}$ 13.7 Hz, $J_{1'',2''}$ 4.2 Hz, H1″; 2.63, dd, $J_{5,5}$ 14.0 Hz, $J_{4,5}$ 11.0 Hz, H5; 2.72, dd, $J_{1'',1''}$ 13.7 Hz, $J_{1'',2''}$ 11.1 Hz, H1″; 2.81, dd, $J_{5,5}$ 14.0 Hz, $J_{4,5}$ 2.2 Hz, H5; 2.92, m, H2″; 3.50, dd, $J_{4'',4''}$ 10.6 Hz, $J_{4'',5''}$ 6.7 Hz, H4″; 3.54, dd, $J_{6'',6''}$ 11.6 Hz, $J_{5'',6''}$ 6.3 Hz, H6″; 3.57, dd, $J_{4'',4''}$ 10.6 Hz, $J_{4'',5''}$ 3.9 Hz, H4″; 3.60, dd, $J_{6'',6''}$ 11.6 Hz, $J_{5''6''}$ 3.5 Hz, H6″; 3.62, dd, $J_{1',1'}$ 10.6 Hz, $J_{1',2'}$ 4.4 Hz, H1′; 3.72, dd, $J_{3'',3''}$ 9.8 Hz, $J_{2'',3''}$ 4.5 Hz, H3″; 3.77, dd, $J_{3'',3''}$ 9.8 Hz, $J_{2'',3''}$ 5.8 Hz, H3″; 3.80, dd, $J_{1',1'}$ 10.6 Hz, $J_{1',2'}$ 4.5 Hz, H1′; 3.87, m, H5″; 4.01.09, m, H2′, 3′, 3′; 4.12, d, $J_{2,3}$ 3.9Hz, H2; 4.19.25, m, H3, 4; 5.00, s, H1.

4. Edmonds, J. S.; Francesconi, K. A.; Stick, R. V. *Nat. Prod. Rep*. **1993**, *10*, 421.
5. Challenger, F. *Adv. Enzymol*. **1951**, *12*, 429.
6. Edmonds, J. S.; Francesconi, K. A. *J. Chem. Soc., Perkin Trans. 1*, **1983**, 2375.
7. Francesconi, K. A.; Stick, R. V.; Edmonds, J. S. *J. Chem. Soc., Chem. Commun*., **1991**, 928.
8. Edmonds, J. S.; Francesconi, K. A. *Experientia* **1987**, *43*, 553.
9. Edmonds, J. S.; Francesconi, K. A.; Healy, P. C.; White, A. H. *J. Chem. Soc., Perkin Trans. 1*, **1982**, 2989.
10. Francesconi, K. A.; Edmonds, J. S. *Adv. Inorg. Chem*. **1996**, *44*, 147.
11. Francesconi, K. A.; Edmonds, J. S.; Stick, R. V. *J. Chem. Soc., Perkin Trans. 1*, **1992**, 1349.
12. Eichenberger, W. *Plant Physiol. Biochem*. **1993**, *31*, 213.
13. Sato, N. *Plant Physiol*. **1988**, *86*, 931.
14. Vogel, G.; Eichenberger, W. *Plant Cell Physiol*. **1992**, *33*, 427.
15. Halpern, J. *Science* **1985**, *227*, 869.