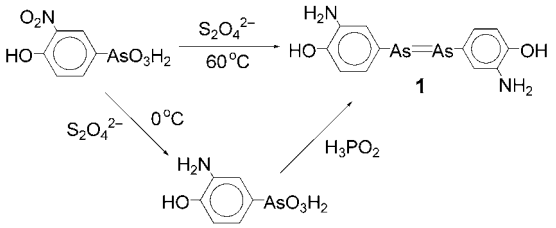


Scheme 1. Two routes to salvarsan (**1**). The initial route of Ehrlich and Bertheim involves simultaneous reduction of the nitro group and As^V (top arrow).[3] Christiansen's two-step process of sequential reductions is shown below. [4]

Molecular Structure

The Composition of Ehrlich's Salvarsan: Resolution of a Century-Old Debate**

Nicholas C. Lloyd, Hugh W. Morgan, Brian K. Nicholson, and Ron S. Ronimus*

In 1910, Ehrlich introduced the compound 3-amino-4-hydroxyphenylarsenic(I) (salvarsan, arsphenamine, or Ehrlich 606) as a remedy for syphilis, a disease caused by the spirochaete bacterium *Treponema pallidum.* His methodical search for a specific curative for an identified disease can be regarded as the introduction of targeted chemotherapy.

Despite its long history and importance, the actual structure of salvarsan is still debated.[1,2] Ehrlich synthesized the compound by the reaction of 3-nitro-4-hydroxyphenylarsonic acid with dithionite.[3] This simultaneously reduced the NO_2 group to NH_2 and As^V to As^I to give a material of formula 3-H_2N-4-HOC_6H_3As (Scheme 1). The product was isolated as the hydrated hydrochloride salt of formula 3-H_2N-4-$HOC_6H_3As{\cdot}HCl{\cdot}H_2O$. This synthesis was not always reproducible, and inevitably gave rise to sulfur-containing impurities, which may have accounted for the variable toxicity of different batches of salvarsan.[5] A two-step process that involved initial reduction of the NO_2 group by sodium dithionite followed by reduction of As^V with hypophosphorous acid was subsequently shown by Christiansen to lead to sulfur-free material (Scheme 1).[4]

By analogy with azo compounds, Ehrlich assigned structure **1** to the free base of salvarsan. Although As=As bonds are now known to be found only in sterically crowded molecules,[6] the As=As form is repeatedly cited, with textbooks and reviews still giving structure **1**.[1,2] Various suggestions for the true structure of salvarsan have been proposed, including large polycyclic molecules[7] and polymers,[8] but lack strong supporting data.

We now report the first definitive evidence for the composition of salvarsan based on electrospray ionization mass spectrometric data. Salvarsan was synthesized by using a modification of the two-step synthesis first described by Christiansen.[4] The compound was isolated as the hydrochloride, giving consistent yields of a pale-yellow material that was analyzed as the monohydrate, corresponding exactly to Ehrlich's original reports.[9] We also isolated the mixed $H_2PO_2^-/H_2PO_3^-$ salt form (^{31}P NMR spectroscopic data) by direct precipitation from the reaction mixture.

^{1}H and ^{13}C NMR spectroscopic data were recorded for salvarsan, but were not particularly useful for structural assignment; ESI MS proved much more practical for this purpose. A typical spectrum is shown in Figure 1. Although there are many peaks, they can be readily assigned. Three clear series can be identified (Table 1). The first of these is the major one and consists of $[(RAs)_n + H]^+$ peaks (R = 3-H_2N-4-HOC_6H_3) for n = 3–8. Peaks corresponding to $n = 3$ and $n = 5$ are clearly dominant in this series. A lesser series comprises ions generated by loss of R^- from $(RAs)_n$, that is, $[R_{n-1}As_n]^+$ for n = 4–8. There are also minor peaks, which can be assigned to species $[(RAs)_n + O + H]^+$ arising from mild oxidation, presumably through the insertion of an oxygen atom into the As_n ring. There is a significant peak ($[R_2As]^+$) which connotes a fragmentation product of the cyclic polyarsines (see below). Peaks corresponding to $[RAs(OH)_2 + H]^+$ and $[RAs(OH)]^+$ arise from complete oxidative cleavage of As–As bonds.

The small peak at m/z 459 is a doubly charged ion (as determined by the spacing in the high-resolution isotope

[*] N. C. Lloyd, Prof. H. W. Morgan, Prof. B. K. Nicholson, Dr. R. S. Ronimus
Departments of Chemistry and Biological Sciences
University of Waikato
Private Bag 3105, Hamilton (New Zealand)
Fax: (+64) 7-838-4219
E-mail: b.nicholson@waikato.ac.nz

[**] We thank the Marsden Fund, administered by the Royal Society of New Zealand, for financial support.

 DOI: 10.1002/ange.200461471

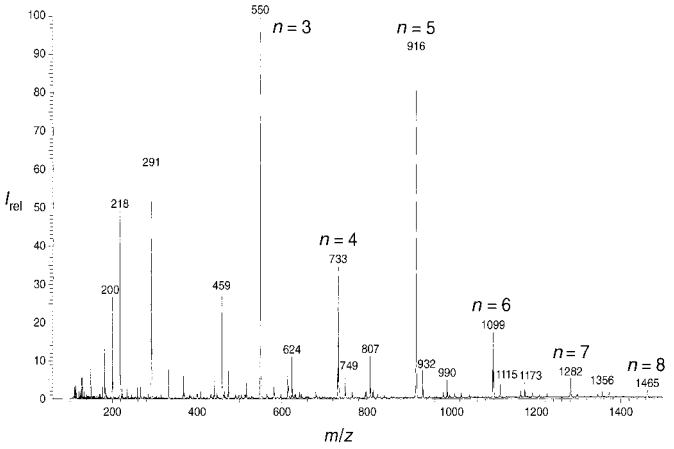


Figure 1. The ESI mass spectrum of salvarsan in H_2O. Ions derived from the different cycloarsines, $(RAs)_n$, are indicated as values of *n*. Assignment of other ions is given in Table 1.

Table 1: ESI mass spectral data for an aqueous solution of salvarsan.

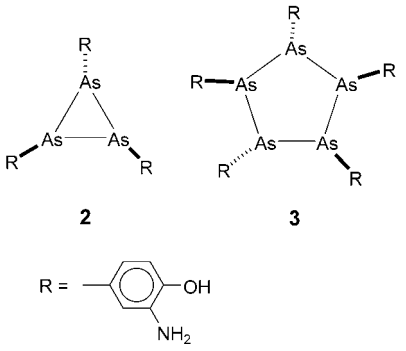
m/z	I_{rel} [%]	Assignment[a]
1465	2	$[(RAs)_8+H]^+$
1356	2	$[(RAs)_8-R^-]^+$
1282	4	$[(RAs)_7+H]^+$
1173	2	$[(RAs)_7-R^-]^+$
1115	3	$[(RAs)_6O+H]^+$
1099	18	$[(RAs)_6+H]^+$
990	3	$[(RAs)_6-R^-]^+$
932	5	$[(RAs)_5O+H]^+$
916	90	$[(RAs)_5+H]^+$
807	10	$[(RAs)_5-R^-]^+$
749	5	$[(RAs)_4O+H]^+$
733	43	$[(RAs)_4+H]^+$
624	10	$[(RAs)_4-R^-]^+$
550	100	$[(RAs)_3+H]^+$
458.5	30	$[(RAs)_6+2H]^{2+}$
291	61	$[R_2As]^+$
218	50	$[RAs(OH)_2+H]^+$
200	25	$[RAs(OH)]^+$

[a] R = 3-NH_2-4-HOC_6H_3-; see Figure 1

pattern) and can be attributed to $[(RAs)_5+2H]^{2+}$. Similarly, there is a small overlapping contribution to the $[(RAs)_3+H]^+$ peak at *m/z* 550 from doubly protonated $[(RAs)_6+2H]^{2+}$ ions. There were no features in any of the mass spectra of the type expected for higher-molecular-mass polymeric compounds.

All the evidence indicates that salvarsan in solution consists of cyclic species $(RAs)_n$, with $n=3$ (**2**) and $n=5$ (**3**) as the preferred sizes. We are confident that the ESI MS data provide an accurate profile of the components of salvarsan in solution, and that the species observed are not artifacts generated in the mass spectrometer, for the following reasons:

1) In many other areas it has been demonstrated that ESI MS provides an accurate guide to solution speciation, as the chemical ionization process is mild and does not lead to significant fragmentation.[10] Ions present in solution are transferred directly into the source. This is why ESI is so much more informative than EI, FAB, or MALDI in speciation studies.
2) The overall ESI MS profile does not change significantly with varying skimmer cone voltage (20–100 V), with varying pH, or with the different electrospray desolvation mechanisms of the two different instruments used. If the smaller rings were fragmentation products of larger molecules, then the distribution of peaks would change markedly under the different conditions. Furthermore, when fragmentation of the specific rings was deliberately induced in MS/MS experiments, the observed processes involved elimination of R, R_2As, and R_3As fragments, with no sign at all of conversion from larger into smaller rings.



Whereas the ESI MS results are believed to give a true qualitative reflection of the composition of salvarsan, quantitative allocation to different ring sizes is less certain, as the ion signal reflects ease of chemical ionization as well as relative abundance.[10] However, each molecule in the series is chemically ionized (protonated) on the same type of NH_2 group. Therefore, a moderate correlation between signal intensity and the amount present should exist. Certainly it appears that $(RAs)_3$ and $(RAs)_5$ are the predominant molecules in the mixture, with significant $(RAs)_4$ and $(RAs)_6$ present, and only minor representation from larger ring sizes.

It is pertinent to ask why elucidation of the structure of salvarsan has proved so difficult, as it has been established that other organoarsenic(I) compounds such as $(PhAs)_n$ and $(MeAs)_n$ can form rings ($n=5$ and 6 for $(PhAs)_n$, and $n=5$ together with a polymeric ladder form for $(MeAs)_n$).[1] This is a result of its resilient physical features—the material is a mixture of compounds; it does not crystallize in a form suitable for X-ray crystallographic analysis; it is not volatile, so it cannot be examined by traditional mass spectrometric methods; it is readily oxidized in solution (particularly at neutral or higher pH values[11]); and it has strong hydrogen bonding functional groups, which will complicate cryoscopic mass measurements (salvarsan forms gels in aqueous solutions at critical concentrations and pH values[4,12]).

We have attempted to separate salvarsan into its constituent ring species with HPLC experiments. However, these

have so far been unsuccessful, presumably because of the high reactivity and hydrogen bonding properties of the compound.

The $(RAs)_n$ rings formed by salvarsan are of interest, especially the trimer, as an As_3 ring does not seem to have been confirmed before, although examples of both $(RP)_3$ and $(RSb)_3$ are known.[13] It is striking that the trimer and pentamer appear to be the main forms when R = 4-hydroxy-3-aminophenyl, whereas the hexamer and pentamer are the only characterized examples for R = phenyl.[1,7] (We note that West and co-workers found a peak corresponding to $[(PhAs)_3]^+$ in the EI mass spectrum of $(PhAs)_n$, but concluded it to be a fragmentation ion rather than a true component of the mixture.[14])

It has been accepted that salvarsan in the clinically administered form differs from the active form in vivo. Oxidation is needed to activate the compound.[15] This was earlier assumed to give rise to what was originally formulated as the oxide (RAs = O), but is undoubtedly $RAs(OH)_2$ or a condensation product thereof.[12] In essence, salvarsan $(RAs)_n$ appears to serve as a slow-release source of $RAs(OH)_2$. This oxidation process was confirmed by our ESI MS results. Peaks at *m*/*z* 218 and 200 can be assigned to $[RAs(OH)_2+H]^+$ and $[RAs(OH)]^+$, respectively (Figure 1). As the sample was exposed to air, these signals increased in intensity at the expense of those from the various $[(RAs)_n+H]^+$ species, until after a few hours they completely dominated the spectrum.

In conclusion, we have demonstrated for the first time that salvarsan consists of small rings $(RAs)_n$, particularly **2** and **3**, rather than the commonly written form **1**. Although this assignment is based mainly on ESI mass spectrometry of the mixture because the properties of the compound have so far precluded separation into components, we are confident that this work has provided an accurate description of the dominant structures of salvarsan.

Experimental Section

ESI mass spectra were recorded on VG Platform and Finnegan LCQ mass spectrometers. For the former, the compounds were dissolved in the appropriate solvent and injected through a Rheodyne injector equipped with a 10-mL sample loop. A flow rate of 0.02 mL min^{-1} and a source temperature of 60 °C was used. Nitrogen was used as both the nebulizing and drying gas. For the LCQ instrument, the sample was directly injected at 5 μL min^{-1} with a syringe pump. The capillary temperature was set at 100 °C. Nitrogen was used as the drying gas, and argon for collisionally induced fragmentation. NMR spectroscopic data were recorded with a Bruker Avance 300 spectrometer. 3-Nitro-4-hydroxyphenylarsonic acid (roxarsone) was purchased from Aldrich.

Preparation of 3-amino-4-hydroxyphenylarsonic acid (adapted from the method of Fargher[16]): 3-Nitro-4-hydroxyphenylarsonic acid (13.1 g, 0.05 mol) was dissolved in a solution of aqueous NaOH (1M, 100 mL) and cooled to 0 °C in an ice/salt bath. Sodium dithionite ($Na_2S_2O_4$) (30.25 g) was added in one portion with vigorous stirring. The solution effervesced. As soon as the color changed from orange to pale yellow, concentrated aqueous HCl (12 mL) was added. This mixture was held at < 0 °C until all frothing ceased and the product precipitated from solution. The precipitate was filtered and washed twice with ice-cold water to give crude 3-amino-4-hydroxyphenylarsonic acid (6.50 g, 56 %) as a cream-colored solid. This was dried under vacuum for 24 h.

Purification of 3-amino-4-hydroxyphenylarsonic acid (adapted from the method of Christiansen[4]): The crude product (6.0 g) obtained above was dissolved in dilute aqueous HCl (H_2O (25 mL), concd HCl (2 mL)), and the solution was stirred with decolorizing charcoal for 15 min. The mixture was filtered, and a solution of sodium acetate (25 %) was added to the filtrate until the solution was no longer acidic to Congo Red. The solution was cooled to 4 °C for 20 min. The resulting precipitate was collected by filtration and dried under vacuum to yield 3-amino-4-hydroxyphenylarsonic acid (4.7 g, 78 %) as an off-white microcrystalline solid. ^{1}H NMR ([D$_6$]DMSO): δ = 6.97, 6.81 ppm; ^{13}C NMR: δ = 114.5, 115.0, 119.0, 122.6, 137.8, 148.3 ppm; ESI MS (H_2O): positive ion *m*/*z* 234 $[M+H]^+$, 449 $[2M-H_2O+H]^+$, 664 $[3M-2H_2O+H]^+$; negative ion *m*/*z* 232 $[M-H]^-$, 465 $[2M-H]^-$, 698 $[3M-H]^-$. Elemental analysis calcd for $C_6H_8AsNO_4$: C 30.90, H 3.43, N 6.01 %; found: C 30.76, H 3.41, N 6.32 %.

Preparation of salvarsan (3-amino-4-hydroxyphenylarsenic(I)) (adapted from the method of Christiansen[4]): All solvents and solutions were degassed and the reaction was carried out under nitrogen. 3-Amino-4-hydroxyphenylarsonic acid (2.3 g, 0.01 mol) was dissolved in a solution of hypophosphorous acid (14 mL, 50 %) in water (73 mL). An aqueous solution of KI (1 mL, 3 %) was added, and the solution was heated gradually to 55 °C and held between 55 and 60 °C for 90 min. The mixture was cooled to 10 °C and poured with vigorous stirring into a cold mixture of HCl/H_2O (1:1, 164 mL). Salvarsan precipitated as a pale yellow gelatinous solid, which was collected and dried under vacuum as the hydrochloride. Elemental analysis calcd for $C_6H_6AsNO{\cdot}HCl{\cdot}H_2O$: C 30.37, H 3.80, N 5.91 %; found: C 30.66, H 3.70, N 5.62 %. Alternatively, ethanol was added to the cooled reaction mixture after the heating step with vigorous stirring, until a precipitate formed. The solid was collected and dried as above, and was characterized as a mixed hypophosphite/phosphite salt of 3-amino-4-hydroxyphenylarsenic(I). ^{31}P NMR (D_2O): δ = 9.7 ppm (t, $^1J_{H\text{-}P}$ = 536 Hz) $H_2PO_2^-$, δ = 4.1 ppm (d, $^1J_{H\text{-}P}$ = 648 Hz) HPO_3^{2-} (≈5:1). Elemental analysis calcd for $C_6H_6AsNO{\cdot}H_2PO_2{\cdot}H_2O$: C 26.97, H 4.12, N 5.24 %; found: C 27.50, H 3.92, N 5.14 %.

Received: July 29, 2004
Revised: October 8, 2004
Published online: December 29, 2004

Keywords: arsenic · arsphenamine · mass spectrometry · salvarsan · structure elucidation

[1] O. M. N. Dhubhghaill, P. J. Sadler, *Struct. Bonding (Berlin)* **1991**, *78*, 129.
[2] A. Lykknes, L. Kvittingen, *J. Chem. Educ.* **2003**, *80*, 497; A. S. Levanson, *J. Chem. Educ.* **1977**, *54*, 98.
[3] P. Ehrlich, A. Bertheim, *Ber. Dtsch. Chem. Ges.* **1912**, *45*, 756. See also W. G. Christiansen, *J. Am. Chem. Soc.* **1921**, *43*, 2202; for a review, see: S. Riethmiller, *Bull. Hist. Chem.* **1999**, *24*, 28.
[4] W. G. Christiansen, *J. Am. Chem. Soc.* **1920**, *42*, 2402.
[5] W. G. Christiansen, *J. Am. Chem. Soc.* **1922**, *44*, 847; H. King, *J. Chem. Soc.* **1921**, *119*, 1107; H. King, *J. Chem. Soc.* **1921**, *119*, 1415; R. G. Fargher, F. L. Pyman, *J. Chem. Soc.* **1920**, *117*, 370.
[6] B. Twamley, C. D. Sofield, M. M. Olmstead, P. P. Power, *J. Am. Chem. Soc.* **1999**, *121*, 3357.
[7] L. R. Smith, J. L. Mills, *J. Organomet. Chem.* **1975**, *84*, 1; T. Klapotke, *Biol. Met.* **1988**, 69; C. Elschenbroich, A. Salzer, *Organometallics: a Concise Introduction*, *2nd ed.*, VCH, Weinheim, **1992**, p.158.
[8] M. Ya. Kraft, E. B. Agracheva, *Dokl. Akad. Nauk Az. SSR* **1955**, *100*, 279 (CAN 50:8273); M. Ya. Kraft, I. A. Bashchuk, *Dokl. Akad. Nauk Az. SSR* **1949**, *65*, 509 (CAN 45:16367).

[9] P. Ehrlich, A. Bertheim, US Patent 986148, **1911**, (CAN 5:10870).
[10] R. B. Cole, *Electrospray Ionization Mass Spectrometry*, Wiley, New York, **1997**; W. Henderson, B. K. Nicholson, L. J. McCaffrey, *Polyhedron* **1998**, *17*, 4291; R. Colton, A. D'Agostino, J. C. Traeger, *Mass Spectrom. Rev.* **1995**, *14*, 79; N. B. Cech, C. G. Enke, *Mass Spectrom. Rev.* **2001**, *20*, 362.
[11] C. Voegtlin, H. W. Smith, *J. Pharmacol. Exp. Ther.* **1920**, *16*, 199; R. P. Hogan, H. Eagle, *J. Pharmacol. Exp. Ther.* **1944**, *80*, 93.
[12] G. O. Doak, H. G. Steinman, H. Eagle, *J. Am. Chem. Soc.* **1941**, *63*, 99; P. V. Ioannou, *Appl. Organomet. Chem.* **2000**, *14*, 261.
[13] C. Frenzel, E. Hey-Hawkins, *Phosphorus Sulfur Silicon Relat. Elem.* **1998**, *143*, 1; H. J. Breunig, R. Rosler, E. Lork, *Organometallics* **1998**, *17*, 5594.
[14] P. S. Elmes, S. Middleton, B. O. West, *Aust. J. Chem.* **1970**, *23*, 1559.
[15] C. Voegtlin, *Physiol. Rev.* **1925**, *5*, 63; H. Eagle, G. O. Doak, *Pharm. Rev.* **1951**, 107.
[16] R. G. Fargher, *J. Chem. Soc.* **1919**, *115*, 982.

