Correspondence to: Professor B. K. Nicholson
Chemistry Department
University of Waikato
Private Bag 3105
Hamilton
New Zealand.
Fax 0064 7 838 4219
e-mail: b.nicholson@waikato.ac.nz

Communication

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.

[*] 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.

Ehrlich introduced in 1910 the compound 3-amino-4-hydroxyphenylarsenic(I) [Salvarsan, arsphenamine, 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]. The compound was synthesised by Ehrlich by reaction of 3-nitro-4-hydroxyphenylarsonic acid with dithionite [3]. This simultaneously reduced both the $-NO_2$ group to $-NH_2$, and the As(V) to As(I) to give a material of stoichiometry [3-H_2N-4-HO-C_6H_3As] (Scheme 1). The product was isolated as the hydrated hydrochloride salt of empirical formula [3-H_2N-4-$HOC_6H_3As.HCl.H_2O$]. 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 [4]. A two-step process involving initial reduction of the NO_2 group by sodium dithionite, followed by reduction of the As(V) with hypophosphorous acid was subsequently shown by Christiansen to lead to sulfur-free material (Scheme 1) [5].

By analogy with azo-compounds, Ehrlich assigned structure **1** to the free base of Salvarsan. Although it is now recognised that As=As bonds are only found in sterically crowded molecules [6], the As=As form is repeatedly cited, with text-books 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 polymeric versions [8], but without strong supporting data.

We now report the first definitive evidence for composition of Salvarsan, based on electrospray mass spectrometric data.

Salvarsan was synthesised using a modification of the two-step synthesis first described by Christiansen [5]. The compound was isolated as the hydrochloride, giving consistent yields of a pale-yellow material that analysed for the mono-hydrate, corresponding exactly to Ehrlich's original reports [9]. We also isolated the mixed $H_2PO_2^-/H_2PO_3^-$ salt (from ^{31}P NMR) by direct precipitation from the reaction mixture.

^1H and ^{13}C NMR data were recorded for Salvarsan, but these were not particularly useful for structural assignment. Much more indicative were electrospray mass spectra. A typical spectrum is shown in Figure 1. Although there are many peaks, these can be readily assigned (Table 1). Three clear series can be identified. The first of these is the major one and consists of $[(RAs)_n + H]^+$ peaks, where R = 3-H_2N-4-HO-C_6H_3-, for n = 3-8, but with n = 3 and 5 being clearly dominant. Another lesser series comprises ions generated by the loss of R^- from $(RAs)_n$, i.e. $[R_{n-1}As_n]^+$ for n = 4-8. There are also very minor peaks assignable to species $[(RAs)_n + O + H]^+$, arising from mild oxidation presumably involving insertion of an oxygen atom into the As_n ring. There is a significant peak $[R_2As]^+$ which is a mass-spectrometer fragmentation product of the cyclic polyarsines (see below), and peaks $[RAs(OH)_2 + H]^+$ and $[RAs(OH)]^+$ from complete oxidative cleavage of the As-As bonds.

The small peak at m/z 459 is a doubly-charged ion (from the spacing in the high resolution isotope 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 of the evidence indicates that Salvarsan in solution consists of cyclic species $(RAs)_n$, with n= 3 (**2**) and n = 5 (**3**) the preferred size.*

We are confident that the ESMS data provides an accurate profile of the components of Salvarsan in solution, and that the species observed are not artefacts generated in the mass spectrometer for the following reasons.

(i) In many other areas it has been demonstrated that ESMS provides an accurate guide to solution speciation, since the chemical ionisation process is mild and does not lead to significant fragmentation [10]. Ions present in solution are transferred directly into the source. This is why ES ionisation is so much more informative than EI, FAB, or MALDI ionisation in speciation studies.

(ii) The overall ESMS profile does not change significantly with varying skimmer cone voltage (20 -100 V), with varying pH, nor with the different electrospray desolvation mechanisms on the two different instruments used. If the smaller rings were being produced by fragmentation 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 by MS/MS experiments the observed processes involved elimination of R, R_2As and R_3As fragments, with no sign at all of conversion of larger rings to smaller ones.

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

Since it has been established that other organo-arsenic(I) compounds such as $(PhAs)_n$ and $(MeAs)_n$ can form rings [n = 5,6 for $(PhAs)_n$, and n = 5 together with a polymeric ladder form for $(MeAs)_n$] [1] it is pertinent to ask why defining the structure of Salvarsan has proved so difficult. This is due to its obdurate physical features – the material is a mixture of compounds, it does not crystallise in a form suitable for X-ray analysis, it is not volatile so cannot be examined by traditional mass spectrometric methods, it is readily oxidised in solution (particularly at neutral or higher pH) [11], and has strong H-bonding functional groups which will have confused cryoscopic mass measurements (Salvarsan forms gels in aqueous solutions at critical concentrations/pH [5, 15]).

We have attempted to separate Salvarsan into its constituent ring species using HPLC experiments. However these have so far been unsuccessful, presumably because of the high reactivity and H-bonding properties of the compound.

The $(RAs)_n$ rings formed by Salvarsan are of interest, especially the trimer since an As_3 ring does not seem to have been confirmed before, although both $(RP)_3$ and $(RSb)_3$ examples are known [12]. 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 characterised examples for R = phenyl [1,7]. (We note that West *et al.* found a peak corresponding to $[(PhAs)_3]^+$ in the EI mass spectrum of $(PhAs)_n$ but they concluded it was a fragment ion rather than a true component of the mixture [13]).

It has been accepted that Salvarsan is the administered form in clinical use but is not the active form *in vivo*. Rather oxidation was needed to activate it [14]. This was assumed to give rise to what was originally formulated as the oxide (RAs=O), but is undoubtedly $RAs(OH)_2$ (**4**) or a condensation product thereof [15]. In essence,

Salvarsan $(RAs)_n$ appears to serve as a slow release source of $RAs(OH)_2$. This oxidation process was confirmed by our ESMS 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 as the commonly written form **1**.

Experimental Section

Electrospray 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 into the spectrometer via a Rheodyne injector with a 10 mL sample loop. A flow rate of 0.02 mL min^{-1} and a source temperature of 60°C was used, and nitrogen was used as both a nebulising and drying gas. For the LCQ the sample was directly injected into the spectrometer at 5 μL min^{-1} via a syringe pump. The capillary temperature was set at 100°C, nitrogen was used as drying gas and argon for collisionally-induced fragmentation. NMR data were recorded on a Bruker Avance 300 spectrometer. 3-Nitro-4-hydroxyphenylarsonic acid ('Roxarsone') was purchased from Aldrich.

Preparation of 3-amino-4-hydroxyphenylarsonic acid (adapted from Fargher [16])

3-Nitro-4-hydroxyphenylarsonic acid (13.1g, 0.05 mol) was dissolved in aqueous NaOH solution (1 mol L^{-1}, 100 mL) and cooled to 0ºC 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 colour changed from orange to pale yellow, concentrated aqueous HCl (12 mL) was added. This mixture was held at <0ºC until all frothing ceased and the product precipitated from solution. This was filtered and washed twice with ice-cold water to give crude 3-amino-4-hydroxyphenylarsonic acid (6.50 g, 56%) as a cream-coloured solid. This was dried under vacuum for 24 hours.

Purification of 3-amino-4-hydroxyphenylarsonic acid (Adapted from Christiansen [5])

The crude product (6.0 g) obtained above was dissolved in dilute aqueous HCl (H_2O 25 mL, conc HCl 2 mL). This was stirred with decolourising charcoal for 15 minutes.

The mixture was filtered and to the filtrate was added 25% sodium acetate solution until the solution was no longer acidic to Congo Red. The solution was cooled to 4°C for 20 minutes while the product precipitated. This 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. Elemental analysis: found C 30.76, H 3.41, N 6.32%; calcd for $C_6H_8AsNO_4$ C 30.90, H 3.43, N 6.01%. NMR (d^6-DMSO) ^{1}H: δ 6.97, 6.81; ^{13}C δ 114.5, 115.0, 119.0, 122.6, 137.8, 148.3. ESMS (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 $[3M-H]^-$.

Preparation of Salvarsan [3-amino-4-hydroxyphenylarsenic(I)] (adapted from Christiansen [5])

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). Aqueous KI solution (1 mL of 3%) was added and the solution was heated gradually to 55ºC and held between 55 and 60º for 90 minutes. The mixture was cooled to 10ºC and poured with vigorous stirring into cold HCl / H_2O (1:1 mixture, 164 mL). Salvarsan precipitated as a pale yellow gelatinous solid which was collected and dried under vacuum as the hydrochloride. Elemental analysis: found C 30.66, H 3.70, N 5.62%; calcd for $C_6H_6AsNO.HCl.H_2O$: C 30.37, H 3.80, N 5.91%.

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 characterised as a mixed hypophosphite/phosphite salt of 3-amino-4-hydroxyphenylarsenic(I). Elemental analysis: found C 27.50, H 3.92, N 5.14%; calcd

for $C_6H_6AsNO.H_2PO_2$ H_2O. C 26.97, H 4.12, N 5.24%; ^{31}P NMR (D_2O) δ 9.7 (t, $^1J_{H-P}$ 536 Hz) $H_2PO_2^-$, δ 4.1 (d, $^1J_{H-P}$ 648 Hz) HPO_3^{2-} (*ca* 5:1).

References

[1] O. M. N. Dhubhghaill and P. J. Sadler, *Structure and Bonding*, **1991**, *78*, 129.

[2] A. Lykknes and L. Kvittingen, *J. Chem., Ed.,* **2003**, *80*, 497; A. S. Levanson, *J. Chem. Ed.*, **1977**, *54*, 98.

[3] P. Ehrlich and A. Bertheim, *Berichte*, **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**, *23*, 28.

[4] 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 and F. L. Pyman, *J. Chem. Soc*., **1920**, *117*, 370.

[5] W. G. Christiansen, *J. Am. Chem. Soc*., **1920**, *42*, 2402.

[6] B. Twamley, C. D. Sofield, M. M. Olmstead and P. P. Power, *J. Am. Chem. Soc.*, **1999**, *121*, 3357.

[7] L. R. Smith and J. L. Mills*, J. Organometal. Chem*., **1975**, *84*, 1; T. Klapotke, Biol. Metals, (1 (1988) 69; Ch. Elschenbroich and A. Salzer, "*Organometallics: a Concise Introduction*", VCH, Weinheim, 2nd Ed., 1992, p158.

[8] M. Ya. Kraft and E. B. Agracheva, *Dokl. Akad. Nauk. SSSR*, **1955**, *100*, 279 (CAN 50:8273); M. Ya. Kraft and I. A. Bashchuk, *Dokl. Akad. Nauk. SSSR*, **1949**, *65*, 509 (CAN 45:16367).

[9] P. Ehrlich and A. Bertheim, US Patent 986148 (1911), (CAN 5:10870).

[10] R. B. Cole, Electrospray Ionisation Mass Spectrometry, Wiley, New York, **1997**; W. Henderson, B. K. Nicholson and L. J. McCaffrey, Polyhedron, **1998**,

17, 4291; R. Colton, A. D'Agostino and J. C. Traeger, Mass Spectrom. Rev., **1995**, 14, 79; N. B. Cech and C. G. Enke, *Mass Spec. Rev.,* **2001**, *20*, 362.

[11] C. Voegtlin and H. W. Smith, *J. Pharm. Expt. Therap.,* **1920**, *16*, 199; R. P. Hogan and H. Eagle, *J. Pharm. Expt. Therap*., **1944**, *80*, 93.

[12] C. Frenzel and E. Hey-Hawkins, *Phosphorus, Sulfur, Silicon and Relat. Elem.,* **1998**, *143*, 1; H. J. Breunig, R. Rosler and E. Lork, *Organometallics*, **1998**, *17*, 5594.

[13] P. S. Elmes, S. Middleton and B. O. West, *Aust. J. Chem*., **1970**, *23*, 1559

[14] C. Voegtlin, *Physiological Rev*., **1925**, *5*, 63; Eagle and Doak, *Pharm. Rev.,* **1951**, 107.

[15] G. O. Doak, H. G. Steinman and H. Eagle, *J. Amer. Chem. Soc*., **1941**, *63*, 99; P. V. Ioannou, *Appl. Organometal. Chem*., **2000**, *14*, 261.

[16] R. G. Fargher, *J. Chem. Soc*., **1919**, *115*, 982.

CAPTION TO FIGURE 1

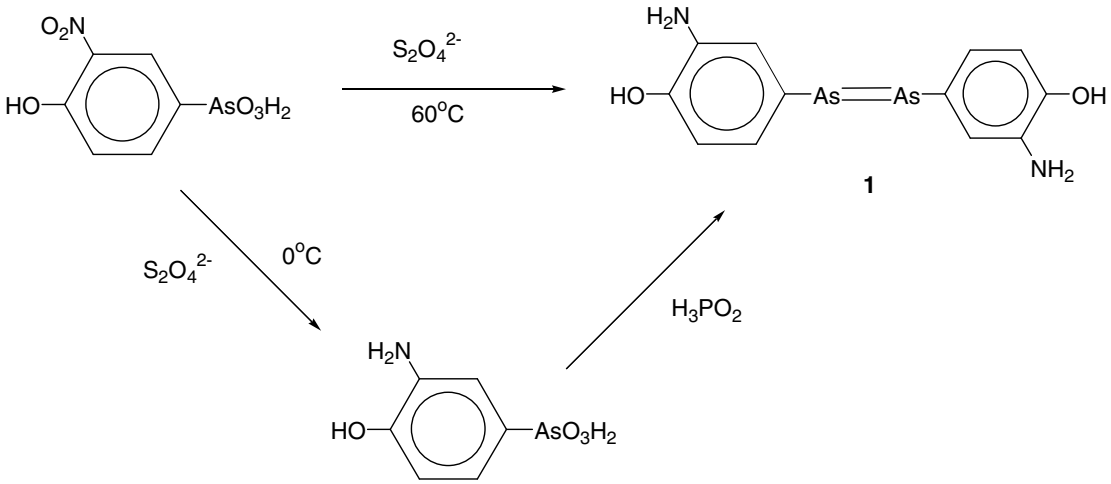
The electrospray mass spectrum of Salvarsan in H_2O. Peaks are annotated with *m/z* values, and the ions derived from the different cycloarsines, $(RAs)_n$, are as indicated. Assignment of other ions is given in Table 1

Table 1

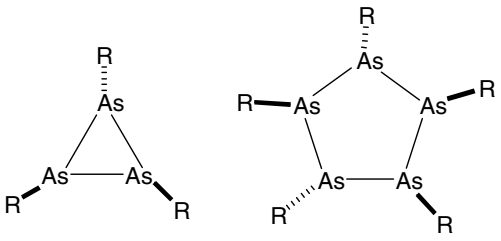
Electrospray mass spectral data for an aqueous solution of Salvarsan.

(R = 3-NH_2-4-HO-C_6H_3-) (c.f. Figure 1).

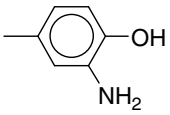
m/z	relative int. (%)	assignment	*m/z*	relative int. (%)	assignment
1465	2	$[(RAs)_8+H]^+$	807	10	$[(RAs)_5-R^-]^+$
1356	2	$[(RAs)_8-R^-]^+$	749	5	$[(RAs)_4O+H]^+$
1282	4	$[(RAs)_7+H]^+$	733	43	$[(RAs)_4+H]^+$
1173	2	$[(RAs)_7-R^-]^+$	624	10	$[(RAs)_4-R^-]^+$
1115	3	$[(RAs)_6O+H]^+$	550	100	$[(RAs)_3+H]^+$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

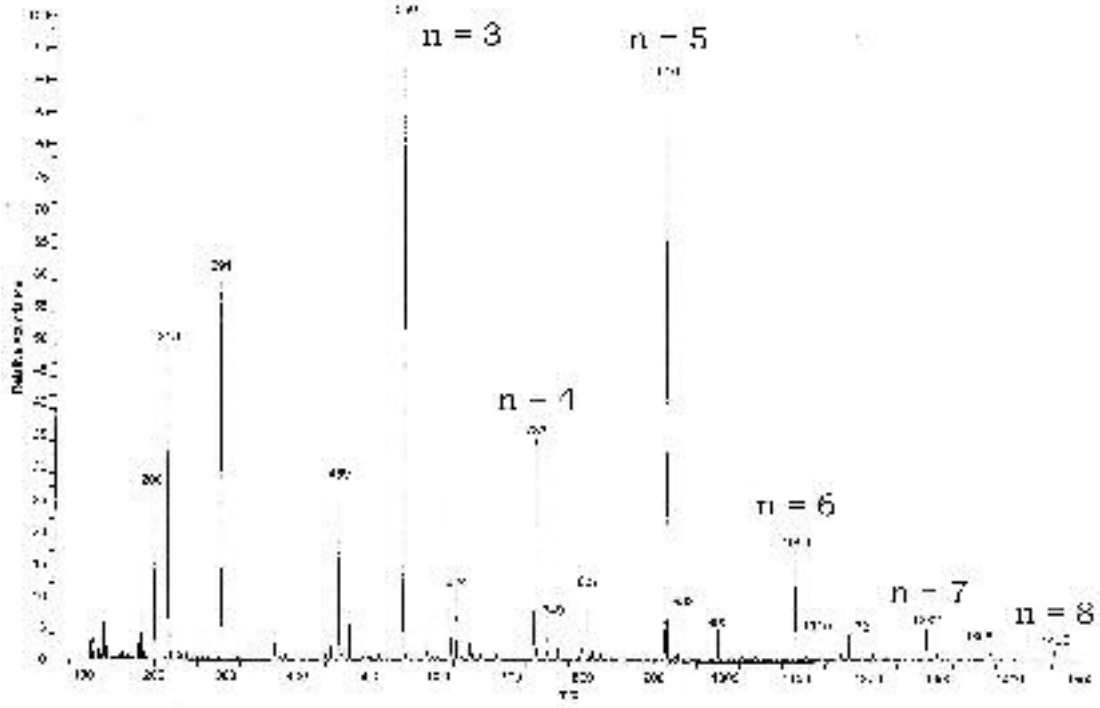

O2N
HO
AsO3H2
S2O4 2-
60°C
H2N
HO
As
As
OH
NH2
1
S2O4 2-
0°C
H3PO2
H2N
HO
AsO3H2

Scheme 1


R
As
As
As
R
R
R
As
As
As
R
R
As
As
R

2 **3**

R = 
OH
NH2


n = 3
n – 5
n – 1
n = 6
n – 7
n = 8