Synthesis of Cytidine Ribonucleotides by Stepwise Assembly of the Heterocycle on a Sugar Phosphate

Abdul-Aziz Ingar,[a] Richard W. A. Luke,[b] Barry R. Hayter,[b] and John D. Sutherland*[a]

Although various syntheses of the nucleic acid bases exist and ribose is a product of the formose reaction, no prebiotically plausible methods for attaching pyrimidine bases to ribose to give nucleosides have been described. Kinetic and thermodynamic factors are thought to mitigate against such condensation reactions in aqueous solution. This inability to produce pyrimidine nucleosides and hence nucleotides is a major stumbling block of the "RNA World" hypothesis and has led to suggestions of alternative nucleic acids as evolutionary precursors to RNA. Here, we show that a process in which the base is assembled in stages on a sugar phosphate can produce cytidine nucleotides. The sequential action of cyanamide and cyanoacetylene on arabinose-3-phosphate produces cytidine-2′,3′-cyclophosphate and arabinocytidine-3′-phosphate.

KEYWORDS:

prebiotic chemistry · molecular evolution · nucleobases · nucleotides · RNA

Introduction

Chemical analysis of the structure of RNA suggests that it might have been produced prebiotically by polymerisation of activated nucleotides.[1] This conventional retrosynthetic view further reduces nucleotides to nucleosides and thence to ribose and the heterocyclic bases.[1a] Potentially prebiotic syntheses of ribose and the bases are cited as evidence for this prebiotic route of RNA synthesis.[2] However, the yields of some of these syntheses are very low and several of the steps do not appear particularly plausible.[3] In particular, no prebiotic condensation of uracil or cytosine with ribose to give nucleosides is known.[1b] Delocalisation of the N1 lone pair of these bases presumably presents a substantial kinetic barrier to glycosidation. Furthermore, the equilibrium constant for the reaction of ribose with uracil to form uridine has been calculated to be in the region of $10^{-3}\,\text{M}^{-1}$.[1c] An alternative route to pyrimidine nucleosides involving stepwise assembly of the nitrogenous base on the sugar has been demonstrated but a stereochemical impasse was encountered.[4] Sequential treatment of D-ribose with cyanamide and cyanoacetylene produced α-ribocytidine and not β-ribocytidine.

The lack of a route to the correctly configured nucleoside components of RNA poses a major challenge to the "RNA World" hypothesis[1a] and has resulted in a bleak outlook in the field summarised thus: "It is possible that some efficient prebiotic synthesis of the β-ribosides, or some method of separating the β-ribosides from closely related isomers, will be discovered, but there is no basis in organic chemistry for optimism."[1d]

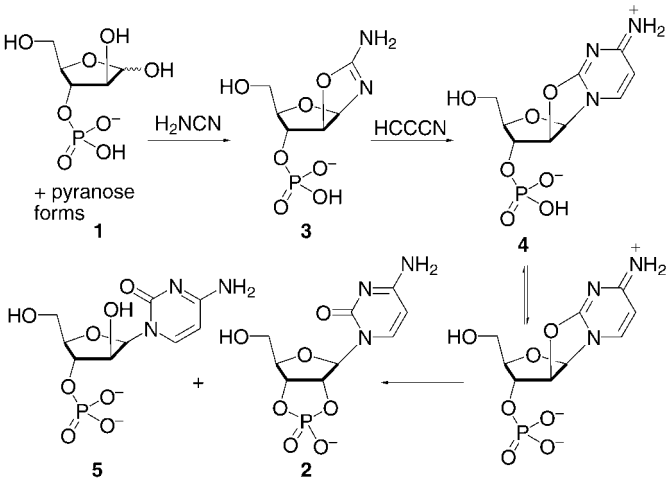
Results and Discussion

Prebiotically plausible routes to phosphorylated sugars have recently been demonstrated,[5] so we decided to investigate the possible conversion of such compounds into nucleotides. In addition to demonstrating that sugar phosphates can be produced by aldolisation of glycolaldehyde phosphate and formaldehyde,[5a] Eschenmoser and co-workers have also shown that sugars can be phosphorylated by using amidotriphosphate or diamidophosphate.[5b] In the pentose series, these chemistries produce 2,4-diphosphates and 2-phosphates (by hydrolysis of 1,2-cyclophosphates), respectively. However, we initially focussed on a pentose-3-phosphate, since we reasoned that arabinose-3-phosphate (**1**) has the potential to be elaborated to cytidine-2′,3′-cyclophosphate (**2**) by sequential assembly of the base and inversion of the 2′-stereochemistry (Scheme 1). Pentose-3-phosphates such as **1** could be derived from the corresponding 2-phosphates or 2,4-diphosphates by isomerisation/hydrolysis or by alternative routes. We reasoned that reaction of **1** with cyanamide should give the aminooxazoline **3**,

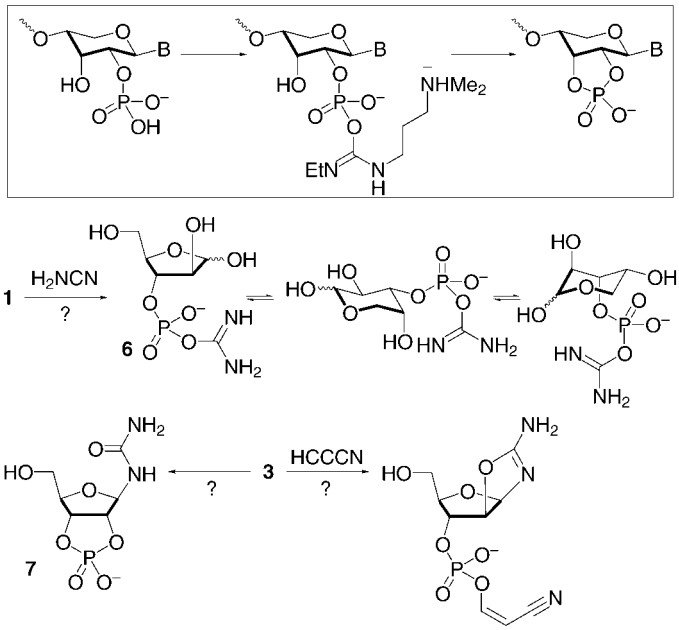
[a] *Prof. Dr. J. D. Sutherland, A.-A. Ingar*
Department of Chemistry
The University of Manchester
Oxford Road, Manchester M13 9PL (UK)
Fax: (+44) 161-275-4939
E-mail: john.sutherland@man.ac.uk

[b] *Dr. R. W. A. Luke, Dr. B. R. Hayter*
AstraZeneca, Mereside, Alderley Park
Macclesfield SK10 4TG (UK)

  DOI: 10.1002/cbic.200300554



Scheme 1. *Possible prebiotic conversion of arabinose-3′-phosphate (**1**) to cytidine-2′,3′-cyclophosphate (**2**).*



Scheme 2. *Potential side reactions in the conversion of arabinose-3′-phosphate (**1**) to anhydronucleotide **4**. Boxed: precedent for activated phosphates in the pyranosyl series to undergo intramolecular alcoholysis by vicinal hydroxy groups.[8] Treatment of 2′-phosphoryl-pyranosyl-RNA tetramers with ethyl-(3-dimethylaminopropyl)-carbodiimide gave 2′,3′-cyclophosphoryl-pRNA tetramers. Cyanamide, whilst a less powerful activating agent than a carbodiimide, has been used to effect the cyclisation of uridine-2′/3′-monophosphate to uridine-2′,3′-cyclophosphate in greater than 70% yield.[7] It was therefore possible that reaction of arabinose-3-phosphate (**1**) with cyanamide would lead to the activated derivative **6** and that the pyranose forms of **6** would undergo cyclisation to give arabinopyranose-2,3- or 3,4-cyclophosphates. The phosphate group of **3** has the potential either to displace the aminooxazoline to give a ribosylurea **7** or to react with cyanoacetylene to give a cyanovinyl adduct.*

which should react with cyanoacetylene to give the anhydronucleotide **4**. Compound **4** has previously been prepared by conventional chemical synthesis[6a] and shown to undergo competing hydrolysis and rearrangement to **5** and **2**, respectively.[6b] However, the success of the conversion of **1** to **4** was by no means guaranteed. Firstly, the steric and electronic effects of the 3-phosphate group of **1** on the formation of the aminooxazoline **3** could not be predicted. Secondly, it is known that phosphate monoesters react with cyanamide to give intermediates, which are hydrolysed or alcoholysed (in high yield if the attack is an intramolecular one by a vicinal hydroxy group) to urea or which react with additional cyanamide to give cyanoguanidine.[7] It was thus possible that such reactions would compete with aminooxazoline formation, since intramolecular alcoholysis by vicinal hydroxy groups would be possible in the pyranose forms of the phosphate–cyanamide adduct **6** (Scheme 2).[8] Thirdly, since inorganic phosphate is known to react with cyanoacetylene,[9] it was possible that the 3-phosphate of **3** might do so as well. Lastly, we recognised that the 3′-phosphate might displace the aminooxazoline function of **3** to give the ribosylurea **7**.

Conventional chemical synthesis was used to prepare **1**, which was found to exist predominantly in pyranose forms in solution.[10] Treatment of an aqueous solution of **1** (0.1 M) with two equivalents of cyanamide at pH 7.5 and 60 °C for 2.5 h resulted in the formation of the aminooxazoline **3** in high yield (> 70% as measured by 300-MHz ^{1}H NMR analysis; Figure 1a). A sample of **3** (0.05 M) was then allowed to react with five equivalents of cyanoacetylene at pH 5.5 and 50 °C for 2 h. Although **3** is probably *N*-protonated and therefore less reactive than **1** at this pH value,[11] these conditions were chosen to allow observation of any zwitterionic *O*2,2′-cyclocytidine-3′-phosphate (**4**) produced by rendering the reactive dianionic form of **4** unavailable. Compound **3** was converted into **4** in approximately 50% yield (Figure 1b). When the pH value of the reaction was subsequently raised to 7 by addition of aqueous sodium hydroxide, the dianionic form of **4** gradually transformed into **2** and **5**. After 10 days, **2** and **5** had been produced in a ratio of 1:4 and a combined yield of approximately 50% based on **4** (Figure 1c). ^{1}H NMR analysis also demonstrated the presence of three other pyrimidine nucleotide species that we have not identified as yet. The less favourable ratio of **2**:**5** is thought to be due in part to the nature of the counterions of **4** in this experiment. The previously reported conversion of **4** into **2** and **5** in a ratio of 2:1[6b] and an overall quantitative yield involved triethylammonium or tri-*n*-octylammonium counterions, in contrast to the sodium ions in our experiment. The effects of counterions and other variables on this reaction are currently being evaluated.

The fact that these conversions proceed readily in aqueous solution at near-neutral pH values suggests that this chemistry is prebiotically plausible.[12] Furthermore, the dominance of the chemistry shown in Scheme 1 over that shown in Scheme 2, despite the precedence for the latter, suggests that within the ensemble of arabinose-3-phosphate, cyanamide and cyanoacetylene, the reactions are predisposed to lead to cytidine nucleotides.[13] At present the sequence requires excesses of cyanamide and cyanoacetylene and suffers from a low yield in the final step. Both cyanamide and cyanoacetylene are presumed to be prebiotic compounds and can be produced in

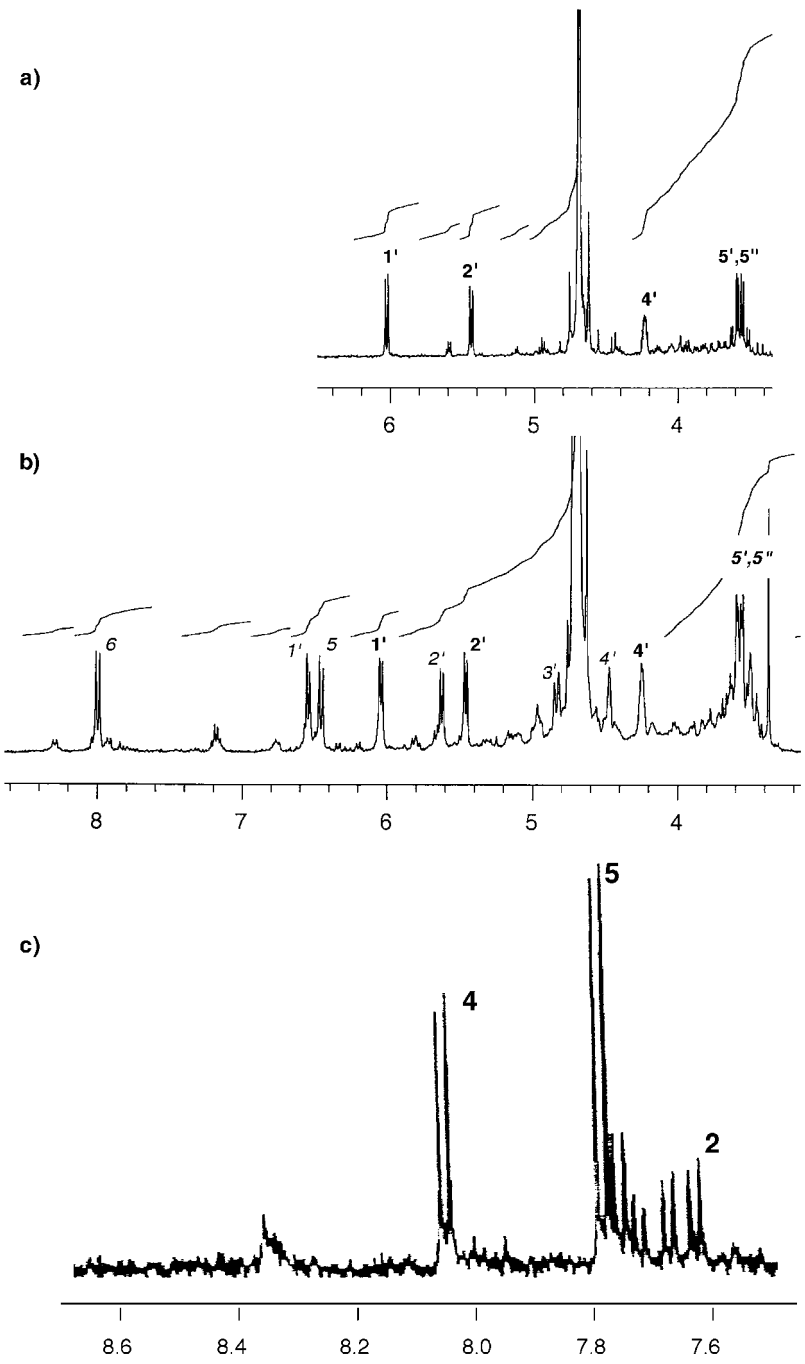


Figure 1. *^{1}H NMR analysis of the nucleotide assembly sequence (all spectra recorded in D_2O). a) 300 MHz spectrum of the crude reaction products from treatment of **1** with cyanamide. Peaks due to **3** are labelled with assignments in bold. b) 300 MHz spectrum showing the partial conversion of **3** into **4** after reaction with cyanoacetylene. Peaks due to **4** are labelled in italics, peaks due to residual **3** are labelled in bold and peaks due to both **3** and **5** are labelled in bold italics. c) Expansion of the 400 MHz spectrum showing conversion of **4** into **2**, **5** and three unidentified species presumed to be cytidine derivatives. The various doublets all correspond to H6 signals, with assignments labelled according to compound number. During the experiment, ^{1}H NMR analysis showed that the H6 peaks for **4** decreased whilst those for **2** and **5** increased in proportion with the signals assigned to the other protons of **2** and **5** (not shown).*

Experimental Section

Arabinose-3-phosphate (**1**) was prepared by conventional chemical synthesis (details to be reported elsewhere). Cyanamide was purchased from Aldrich and cyanoacetylene was prepared by dehydration of propiolamide according to the procedure of Eschenmoser et al.[15] The aminooxazoline **3** was purified by C18 reversed-phase HPLC (Waters Spherisorb ODS column) by isocratic elution with water. ^{1}H NMR data for **3** (300 MHz, D_2O): $\delta = 3.52$ (1H, dd, $J_{gem} = 12.8$, $J_{4',5'} = 4.7$ Hz; H5′), 3.60 (1H, dd, $J_{gem} = 12.8$, $J_{4',5''} = 3.5$ Hz; H5″), 4.23 (1H, m; H4′), 5.44 (1H, d, $J_{1''} = 5.6$ Hz; H2′), 6.02 (1H, d, $J_{1',2'} = 5.6$ Hz; H1′) ppm; the signal for H3′ was obscured by the residual HOD signal. Assignment of the ^{1}H NMR data for *O*2,2′-cyclocytidine-3′-phosphate (**4**), produced by reaction of **3** with cyanoacetylene, was made by comparison with data acquired on an authentic sample of **4** prepared by a slightly modified version of the literature method.[6a] ^{1}H NMR data for **4** (300 MHz, D_2O): $\delta = 3.51$ (1H, dd, $J_{gem} = 13.0$, $J_{4',5'} = 2.7$ Hz; H5′), 3.60 (1H, dd, $J_{gem} = 13.0$, $J_{4',5''} = 2.8$ Hz; H5″), 4.51, (1H, m; H4′), 4.88 (1H, br d, $J_{3',P} = 9.2$ Hz; H3′), 5.65 (1H, d, $J_{1',2'} = 6.0$ Hz; H2′), 6.48 (1H, d, $J_{5,6} = 7.4$ Hz; H5), 6.57 (1H, d, $J_{1',2'} = 6.0$ Hz; H1′), 8.02 (1H, d, $J_{5,6} = 7.4$ Hz; H6) ppm. Literature NMR chemical shift data for **4** (100 MHz, D_2O) for H2′ (6.24), H5 (7.09), H1′ (7.15) and H6 (8.61) are shifted approximately 0.6 ppm downfield relative to our data.[6a] Assignment of the ^{1}H NMR data for cytidine-2′,3′-cyclophosphate (**2**) produced when the pH value of the solution containing **4** was raised to 7 was made by comparison with literature data[16] and those acquired on a sample of **2** purchased from Sigma–Aldrich. ^{1}H NMR data for **2** (400 MHz, D_2O): $\delta = 3.85$ (1H, dd, $J_{gem} = 12.4$, $J_{4',5'} = 5.7$ Hz; H5′), 3.92 (1H, dd, $J_{gem} = 12.4$, $J_{4',5''} = 3.6$ Hz; H5″), 4.31 (1H, ddd, $J_{4',5'} = 5.7$, $J_{3',4'} = 5.4$, $J_{4',5''} = 3.6$ Hz; H4′), 4.95 (1H, ddd, $J_{3',P} = 12.1$, $J_{2',3'} = 6.7$, $J_{3',4'} = 5.4$ Hz; H3′), 5.14 (1H, ddd, $J_{2',3'} = 6.7$, $J_{2',P} = 6.5$, $J_{1',2'} = 2.7$ Hz; H2′), 5.85 (1H, d, $J_{1',2'} = 2.7$ Hz; H1′), 6.01 (1H, d, $J_{5,6} = 7.5$ Hz; H5), 7.67 (1H, d, $J_{5,6} = 7.5$ Hz; H6) ppm.

This work was funded by AstraZeneca and the Engineering and Physical Sciences Research Council (EPSRC) through the provision of a studentship to A.-A.I. Manchester University, the EPSRC, the Biotechnology and Biological Sciences Research Council, AstraZeneca, DSM Anti-Infectives and Pfizer have generously supported work in the group of J.D.S. We thank Drs. A. J. Lawrence and J. Lehbauer for helpful discussions and advice, and Professor A. Eschenmoser (Eidgenössische Technische Hochschule and Scripps) for generously providing many insightful comments.

reasonable yields.[14] If the yield of the conversion of **4** into **2** can be increased to that reported in the literature[6b] then the overall yield of **2** from **1** would be in excess of 20%; as it stands the overall yield is 3.5%.

[1] a) G. F. Joyce, *Nature* **1989**, *338*, 217–224; G. F. Joyce, *Nature* **2002**, *418*, 214–221; G. F. Joyce, L. E. Orgel in *The RNA World* (Eds.: R. F. Gesteland, J. F. Atkins), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, **1993**, pp. 1–25; b) L. E. Orgel, R. Lohrmann, *Acc. Chem. Res.* **1974**, *7*, 368–377; J. P. Dworkin, *Origins Life Evol. Biosphere* **1997**, *27*, 345–355; c) V. M. Kolb,

J. P. Dworkin, S. L. Miller, *J. Mol. Evol.* **1994**, *38*, 549 – 557; d) G. F. Joyce, A. W. Schwartz, S. L. Miller, L. E. Orgel, *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4398 – 4402.
[2] a) Ribose: T. Mizuno, A. H. Weiss, *Adv. Carbohydr. Chem. Biochem.* **1974**, *29*, 173 – 227; b) purines: W. D. Fuller, R. A. Sanchez, L. E. Orgel, *J. Mol. Evol.* **1972**, *1*, 249 – 257; W. D. Fuller, R. A. Sanchez, L. E. Orgel, *J. Mol. Biol.* **1972**, *67*, 25 – 33; R. A. Sanchez, J. P. Ferris, L. E. Orgel, *J. Mol. Biol.* **1967**, *30*, 223 – 253; c) pyrimidines: M. P. Robertson, S. L. Miller, *Nature* **1995**, *375*, 772 – 774; correction: M. P. Robertson, S. L. Miller, *Nature* **1995**, *377*, 257.
[3] R. Shapiro, *Origins Life Evol. Biosphere* **1988**, *18*, 71 – 85; R. Shapiro, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4396 – 4401.
[4] R. A. Sanchez, L. E. Orgel, *J. Mol. Biol.* **1970**, *47*, 531 – 543.
[5] a) D. Müller, S. Pitsch, A. Kittaka, E. Wagner, C. Wintner, A. Eschenmoser, *Helv. Chim. Acta* **1990**, *73*, 1410 – 1468; b) R. Krishnamurthy, S. Guntha, A. Eschenmoser, *Angew. Chem.* **2000**, *112*, 2369 – 2373; *Angew. Chem. Int. Ed.* **2000**, *39*, 2281 – 2285.
[6] a) J. Nagyvary, *J. Am. Chem. Soc.* **1969**, *91*, 5409 – 5410; b) C. M. Tapiero, J. Nagyvary, *Nature* **1971**, *231*, 42 – 43.
[7] R. Lohrmann, L. E. Orgel, *Science* **1968**, *161*, 64 – 66.
[8] M. Bolli, R. Micura, A. Eschenmoser, *Chem. Biol.* **1997**, *4*, 309 – 320.
[9] J. P. Ferris, G. Goldstein, D. J. Beaulieu, *J. Am. Chem. Soc.* **1970**, *92*, 6598 – 6603.
[10] The predominance of the pyranose forms of **1** over the furanose forms in solution was anticipated on the basis of the behaviour of the free sugar: K. N. Drew, J. Zajicek, J. Bondo, B. Bose, A. S. Serianni, *Carbohydr. Res.* **1998**, *307*, 199 – 209.
[11] D. H. Shannahoff, R. A. Sanchez, *J. Org. Chem.* **1973**, *38*, 593 – 598.
[12] Work investigating the potential origin of pentose-3-phosphates such as **1** is now underway in our laboratories. In the pentose-2-phosphate series, similar "anhydro" chemistry leading to β-ribonucleotides can be envisaged starting from xylose-2-phosphate derivatives. This possibility is also being experimentally investigated (for the behaviour of *O*2,3′-cyclocytidine, see: Y. Mizuno, T. Sasaki, *Tetrahedron Lett.* **1965**, 4579 – 4584).
[13] A. Eschenmoser, E. Loewenthal, *Chem. Soc. Rev.* **1992**, *21*, 1 – 16; J. D. Sutherland, J. N. Whitfield, *Tetrahedron* **1997**, *53*, 11493 – 11527.
[14] Cyanamide can be formed from ammonia and hydrogen cyanide by the action of ultraviolet or high-energy radiation: A. Schimpl, R. M. Lemmon, M. Calvin, *Science* **1965**, *147*, 149 – 150. Cyanoacetylene is the second most abundant nitrogen-containing product of electrical discharge through mixtures of N_2 and hydrocarbons such as methane: L. E. Orgel, *Origins Life Evol. Biosphere* **2002**, *32*, 279 – 281; R. A. Sanchez, J. P. Ferris, L. E. Orgel, *Science* **1966**, *154*, 784 – 785. The prebiotic abundance of these compounds cannot be known, so it is difficult to assess the significance of the excesses currently required by our reaction sequence.
[15] Y.-B. Xiang, S. Drenkard, K. Baumann, D. Hickey, A. Eschenmoser, *Helv. Chim. Acta* **1994**, *77*, 2209 – 2250.
[16] R. D. Lapper, I. C. P. Smith, *J. Am. Chem. Soc.* **1973**, *95*, 2880 – 2884.

Received: January 8, 2003 [F 554]

