Glycolonitrile Oligomerization: Structure of Isolated Oxazolines, Potential Heterocycles on the Early Earth

Gustaf Arrhenius,[1a] Kim K. Baldridge,[1b] Sarah Richards-Gross,[1c] and Jay S. Siegel*,[1c]

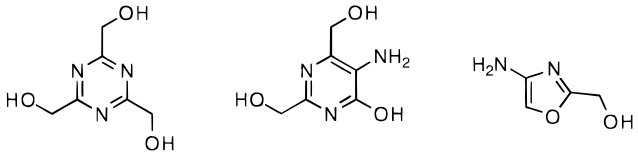
Department of Chemistry, University of California, San Diego, La Jolla California 92093-0358, Scripps Institution of Oceanography, University of California, San Diego, California 92093, and the San Diego Supercomputer Center, P.O. Box 85608, San Diego, California 92138-9784

Received November 22, 1996[⊗]

A study of glycolonitrile polymerization has led to the isolation and characterization of two 2,5-dihydro-4-aminooxazoles, **4** and **5**. Previous reports have misassigned these structures as *s*-triazines or pyrimidines. X-ray diffraction analysis of crystals of **4** and an acetylated oxazole derivative of **5** (**6**) confirm the proposed structures. Ab initio computations are used to assess the relative thermodynamic stability of three trimer isomers (an *s*-triazine, an aminohydroxypyrimidine, and an aminooxazoline), and the results indicate that **4** is a novel kinetic product. Mechanistic considerations rationalize kinetic oxazole formation over the more customary triazine or pyrimidine trimers.

Glycolonitrile, the archetypal cyanohydrin, forms spontaneously ($K = 10^6$ M^{-1}) from formaldehyde and hydrogen cyanide even under aqueous conditions and polymerizes at pH > 7. Given the preponderance of geochemical data pointing to formaldehyde as a primary single carbon source in the primordial atmosphere,[2] the presence of even small quantities of hydrogen cyanide[3] would presage the formation of glycolonitrile and descendant oligomers. Such oligomer mixtures serve as rich storehouses of multicarbon molecules, the cyclic isomers of which stand as principal candidates for the oldest heterocyclic compounds.[4] Despite the reasonable probability that these mixtures were among the first organic polymers, almost nothing definitive is known about the composition or structure of their tractable fractions.

Over the past 60 years, several groups have followed the base-catalyzed oligomerization of glycolonitrile and each has proposed a structure for an isolated crystalline material on the basis of mechanistic and elemental analysis. In the first report, Jacobsen opted for the symmetric 2,4,6-tris(hydroxymethyl)-*s*-triazine (**1**) as a tentative assignment.[5] Later, Lake and Londergan reinvestigated the reaction, and from spot tests, elemental analysis, and derivative formation they deduced the structure to be 5-amino-4-hydroxy-2,6-bis(hydroxymethyl)pyrimidine (**2**).[6] A further reinvestigation by Arrhenius and co-workers considered mass spectral data of the acetyl derivative and found it to come from a dimer not a trimer of glycolonitrile.[4] They proposed the dimer to be 4-amino-2-(hydroxymethyl)-1,3-oxazole (**3**); their underivatized product yielded an elemental analysis consistent with two molecules of glycolonitrile and one water, which they suggested to be a water of crystallization. Given the crystalline nature of the material and the wide variety of structural proposals, we decided to pursue a direct structural elucidation by X-ray diffraction methods.



1 **2** **3**

Results and Discussion

Polymerization of glycolonitrile initiated at 3 °C and maintained at −15 °C and pH 7.5 forms a crystalline material suitable for X-ray diffraction analysis. Data for the crystals were collected at 169 K and solved in the monoclinic space group $P2_1/n$ with four molecules per unit cell. The molecules turned out to be trimers, but not of the triazine or pyrimidine motif; rather, they were oxazolines resulting from the 1,4-addition of the alcohol portion of the third glycolonitrile across the aminooxazole (**4**). In the crystal, the molecule uses the free amino group, hydroxymethyl group, imine nitrogen, and ether oxygen to form hydrogen bonds (Figure 1). The ribbon network can be described by four Etter motifs,[7] an internal 1,4 hydrogen-bonded cycle **S**(5), a bimolecular cycle $\mathbf{R}_2^2(14)$, a tetramolecular cycle $\mathbf{R}_4^4(14)$, and an antiparallel double chain **C**(6). Within the *yz* plane these ribbons display nitrile fingers which create an intersplined (zipper) arrangement, consistent with minimum energy packing of shapes and dipoles. The molecular structure of **4** is asymmetric but the crystal is racemic; a center of symmetry relates the two strands of the ribbon. The specific bond lengths betray an aminooxazoline structure as the favored tautomer, and no unusual bond lengths or angles are observed.

In contrast to the lower temperature study, at pH 8 and 3° C glycolonitrile solutions become orange-brown and viscous in a few days with the eventual formation of black solid. Mass spectral analysis of the orange-brown mixture indicates that oligomers are present, with large amounts of dimer and trimer. Careful observation of the

⊗ Abstract published in *Advance ACS Abstracts,* July 15, 1997.

(1) (a) Scripps Institution of Oceanography. (b) San Diego Supercomputer Center. (c) University of California, San Diego.

(2) Pinto, J. P.; Gladston, G. R.; Yung, Y. L. *Science* **1980**, *210*, 183.

(3) Zahnle, K. J. *Geophys. Res.* **1986**, *91*, 2819.

(4) Arrhenius, T.; Arrhenius, G.; Paplawsky, W. J. *Origin Life* **1994**, *24*, 1.

(5) (a) Jacobsen, R. A. Chemical Deptartment, E.I. du Pont de Nemours & Company, private communication, 1936. (b) In a different context this structure had been proposed by Vakhatova. See: Vakhatova, G. M. *Khim. Farm. Zh.* **1981**, *15*, 50 (*Chem. Abstr.* **1982**, *96*, 20069).

(6) Lake, D. B.; Londergan, T. E. *J. Org. Chem.* **1954**, *19*, 2004.

(7) Bernstein, J.; Etter, M. C.; Leiserowitz, L. In *Structure Correlation*; Dunitz, J. D., Bürgi, H.-B., Eds.; VCH: Weinheim, 1994; Chapter 11.

S0022-3263(96)02185-8 CCC: $14.00

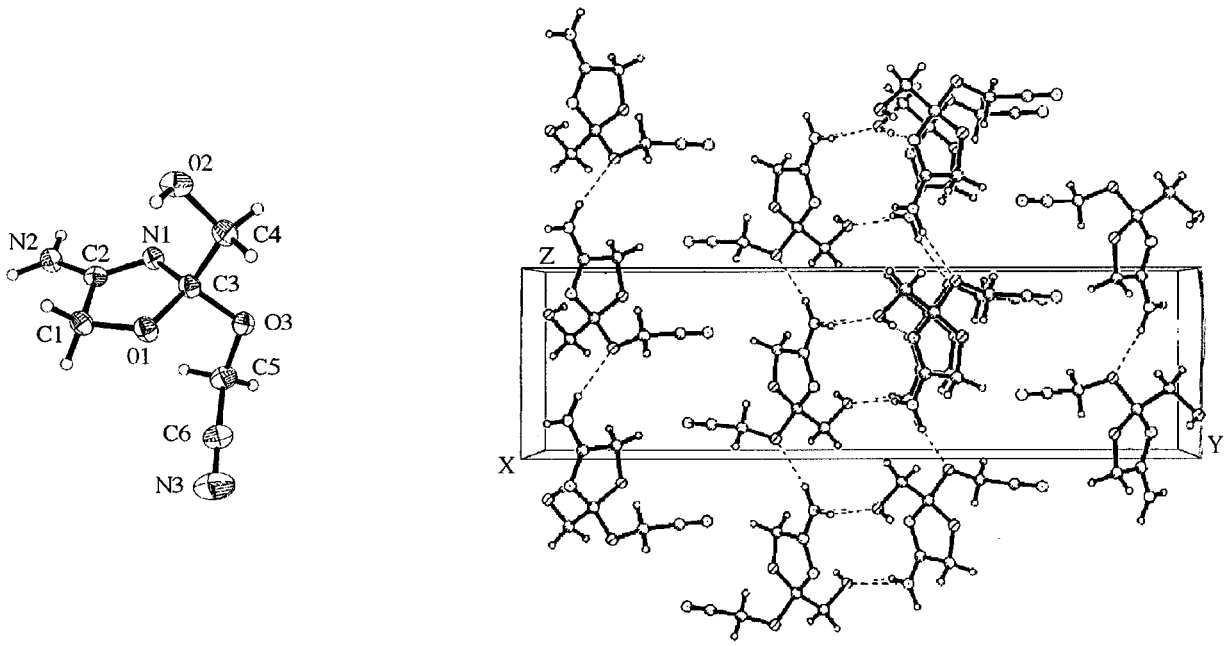


Figure 1. X-ray structure of **4**. Packing diagram shows the hydrogen-bonded network.

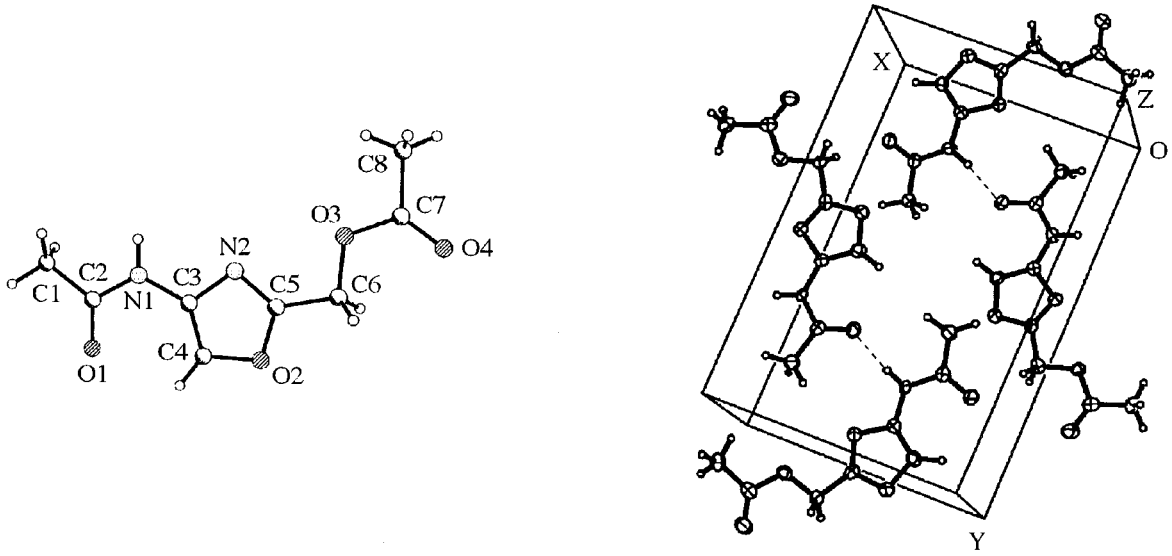
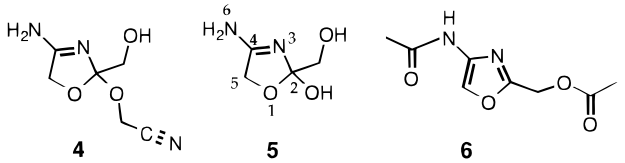


Figure 2. X-ray structure of **6**. Packing diagram shows the hydrogen-bonded chains.

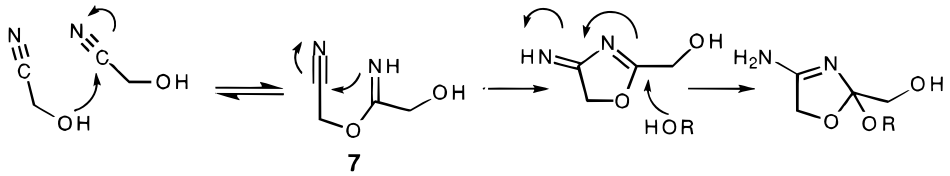
oligomerization process reveals the formation of a white microcrystalline material around days two or three, but the crystals are too small for standard single-crystal crystallographic study. This material analyzes as a dimer hydrate (**5**).[4] Treatment of the hydrate crystals with methanol yields a fine white crystalline powder which has lost water.[4] By analogy to the structure found for **4**, we tentatively assigned the structure of **5** to the oxazoline formed from the 1,4 addition of water to the aminooxazole nucleus. Treatment of **5** with acetic anhydride produced a crystalline acetylated derivative (**6**).



Diffraction data were collected on crystals of **6** at 130 K, and the structure was solved in the orthorhombic space group $Pna2_1$ with four molecules in the unit cell. The molecular structure was determined to be *N*-acetyl-4-amino-2-(acetoxymethyl)-1,3-oxazole, consistent with dehydration of **5** and concomitant acetylation of the free hydroxy and amino groups. The bonding in the ring shows the expected short C(3)−C(4) and C(5)−N(2) "double" bond distances. In addition the exocyclic N(1)−C(3) distance is shortened, indicating appreciable imine character. The molecular packing in the *xy* plane shows a standard amide chain motif **C**(4) with alternate chains running antiparallel to one another (Figure 2).

The additional question of whether the solid state structure persists in solution can be answered by NMR spectroscopy. Even though the proton NMR spectrum should be relatively simple, a few details are definitive. Unlike the other proposed structures which are achiral and should display singlets for their methylene protons, the asymmetry in **4** requires the hydrogens of each methylene to be diastereotopic and give rise to an AB quartet. In dimethyl sulfoxide, the ^{1}H NMR spectrum can

Scheme 1



7

Table 1. X-ray Crystallographic Data for 4 and 6

	4	**6**
formula	$C_6H_9N_3O_3$	$C_8H_{10}N_2O_4$
space group	$P2_1/n$ (monoclinic)	$Pna2_1$ (orthorhombic)
a, Å	4.906(2)	9.389(2)
b, Å	23.694(10)	14.786(2)
c, Å	6.677(3)	6.510(10)
β, deg	91.33	
V	775.9(6)	904.7(3)
ρ(calcd), mg/m^3	1.465	1.455
Z	4	4
data/parameters	14	6.2
GOF	1.34	0.658
R (wR), %	4.5 (6.2)	3.1 (7.0)

be obtained and these AB quartets are clearly seen. The same AB quartet pattern is present in the ^{1}H NMR spectrum of **5**, which provides additional support for our assignment of the hydrate structure.

The formation of **4** likely occurs through the intermediacy of the imidate dimer (**7**). Nitriles with electronegative substituents in the α position form stable imidates with acidic alcohols under catalytically basic conditions.[8] For *tert*-butoxyacetonitrile (the *tert*-butyl ether of glycolonitrile), the equilibrium with methanol lies 91% of the way toward imidate formation. Thus, it is reasonable to assume that, to a significant extent, glycolonitrile exists as an imidate dimer at neutral to mildly basic pH. The imidate **7** can either cyclize to form the imino tautomer of the oxazole and then add another molecule of glycolonitrile, or it can first add to the nitrile of another glycolonitrile and then snap shut to form **4** (Scheme 1). Similar mechanistic possibilities account for the formation of **5**.

The apparent absence of an aromatic fraction from the polymerization is not surprising in light of these mechanistic considerations. Nonetheless, oxazoles and oxazolines bearing 4-amino substituents are quite rare, and the relationship between the aromatic 4-aminooxazole and these potentially prebiotic heterocycles is worth examination. The majority of 4-aminooxazoles reported are stable only as acyl derivatives,[9] of which compound **6** is an example. Early attempts to synthesize the aromatic 2-phenyl-4-aminooxazole from an acyl precursor or from the isocyanate led only to the open-ring form.[10] Similarly, an amino group in the 5-position is reported to greatly destabilize the oxazole ring relative to the ring-opened forms.[11] The first free 4-aminooxazoles which are isolable in underivatized form have recently been prepared and isolated by Lakhan and Singh.[12] These are substituted in the 5-position by an aromatic group. A synthesis and proposed mechanism have been published. R may be aryl, alkyl, or H. After the initial condensations and subsequent cyclization, water is eliminated, conferring aromaticity. This intermediate is strikingly similar to compound **5**. Acidity arguments suggest that **5** would more readily aromatize than the 5-aryloxazoles. The C-5 proton, α to an imine in **5**, is more acidic than the proton which is removed to aromatize the ring in the synthesis. The OH leaving group, however, produces a benzylic cation in the aryl substituted case. Without a similar stabilization in **5** or **4**, no aromatization is seen. In the acetylated derivative **6**, the aromatic oxazole is obtained because hydroxyl is converted to an excellent leaving group in the form of acetic acid and because the electron-donating ability of the amino group is reduced.

Although the structure of the trimer has been elucidated as **4**, the question still remains as to why not **1** and/or **2**. Often one might expect the thermodynamically most stable product to crystallize. Indeed, thermodynamic biases may well have been the rationale for the speculations of Jacobsen[5] and Lake and Londergan.[6] To this point, the ab initio structures and energies of **1**, **2**, and **4** were optimized using a split-valence (DH(2dp)) basis set, and energies were further evaluated by single-point computations including dynamic electron correlation (MP2/DH(2dp)). The geometrical parameters look quite reasonable for all structures; **4** in particular was compared with the experimental X-ray structure and no significant deviations were found. The relative energies (MP2/DH(2dp)) were **1** (11.5 kcal/mol), **2** (0.0 kcal/mol), and **4** (28.9 kcal/mol); thus, the pyrimidine is predicted to be by far the most stable isomer, and the oxazoline the least. Clearly, the kinetic formation of the oxazoline must be mechanistically determined. Thermolysis or extended degradation of **4** does not lead to any detectable quantities of **1** or **2**. Hydrolysis of **4** leads primarily to glycolamide and glycolic acid. No report to date of the structure of **1** or **2** has appeared, although **1** should, in principle, be accessible by reduction of trimethyl *s*-triazinetricarboxylate.

Conclusions. Glycolonitrile readily polymerizes at neutral to basic pH to form a colored tar of complex composition. Close observation of the polymerizations conducted below ambient temperatures leads to the isolation of two related crystalline materials, **4** and **5**, trimer of glycolonitrile and a dimer hydrate, respectively. The dimer hydrate can be further derivatized to the acetyl-protected dimer, **6**. The structures of **4** and **6** are shown crystallographically to contain the basic amino 5-membered heterocycle of oxazoline and oxazole, respectively. Aminooxazolines are a rare class of heterocycles, and of possible glycolonitrile trimers, **5** is by far not the thermodynamically most stable. Nonetheless, kinetic factors in the polymerization coupled with solubility characteristics allow the isolation of **5** during the early stages of polymerization. Hydrogen-bonding patterns in the crystal structure of **5** are interesting. The fundamental nature of the formaldehyde−cyanohydrin, glycolonitrile, and its ease of polymerization suggest that **5** could well have been among the first heterocycles. The instability of **5**, however, makes it a questionable candi-

(8) Schaeffer, F. C.; Peters, G. A. *J. Org. Chem.* **1961**, *26*, 412.
(9) Turchi, I. J. Oxazoles. In *Heterocyclic Compounds: Oxazoles*; Turchi, I. J., Ed.; John Wiley & Sons: New York, 1986; pp 88−90.
(10) Cornforth, J. W.; Cookson, E. *J. Chem. Soc.* **1952**, 1085.
(11) Jansen, A. B. A.; Szelke, M. *J. Chem. Soc.*, **1961**, 405.
(12) (a) Lakhan, R.; Singh, R. L. *J. Het. Chem.* **1988**, *25*, 1413. (b) Lakhan, R.; Singh, R. L. *Org. Prep. Proc. Int.* **1989**, *21*, 141.

date for a prebiotic nucleotide mimic. Hydrolysis of the polymeric tar yields essentially only glycolic acid and glycolamide.

Computational Details

The molecular structures have been determined at a variety of theoretical methods to determine self-consistency. Reported here are restricted Hartree−Fock[13] (RHF) calculations performed with the aid of analytically determined gradients and the search algorithms contained within GAMESS,[13,14] using the DZV(2d,p)[15] double-ζ valence basis set. This basis set includes 2 sets of six d polarization functions on all heavy atoms, and 1 set of p polarization functions on hydrogen atoms. The nature of each stationary point was uniquely characterized by analytically calculating and diagonalizing the matrix of energy second derivatives (Hessian) to determine the number of imaginary frequencies. Additional single-point energy calculations using Moller−Plesset theory[16] of order 2 with the same basis set were performed to determine the effects of dynamical correlation [MP2/DZV(2d,p)//DZV(2d,p)]. These latter calculations were performed using the GAUSSIAN94 suite of programs.[17]

Experimental Section

General Procedure for the Polymerization of Glycolonitrile. Method A. Freshly distilled glycolonitrile (5 mL) or distilled glycolonitrile which had been stabilized with concd H_2SO_4 was adjusted to a pH between 7.5 and 9 with 10% NaOH or similar bases. The solution was placed at approximately 3 °C. After a minimum of 2 days white crystals appeared in the solution. As time progressed, the surrounding solution became increasingly viscous and orange in color. The dimer hydrate **5** was isolated from the surrounding polymeric mixture by trituration in methanol/acetone mixtures. The insoluble white powder was then isolated by filtration or by repeated decantation (370 mg). The compound is stable for long periods if stored at low temperatures (−15 °C) and in the absence of moisture. Samples left on the bench top will discolor readily as polymerization begins.

4-Amino-2-hydroxy-2-hyroxymethyloxazolidine (5): mp 110−140 °C (dec); IR 3200 broad, 1670, 1590, 1450 cm^{-1}; ^1H NMR (DMSO-d_6) δ 6.7 (br), 4.35 (d, 1H, J = 13.5 Hz), 4.29 (d, 1H, J = 13.5 Hz), 3.98 (d, 1H, J = 11.5 Hz), 3.41 (d, 1H, J = 11.5 Hz); ^1H NMR (D_2O) δ 4.68 (d, 1H, J = 13 Hz), 4.61 (d, 1H, J = 13.5 Hz), 4.20 (d, 1H, J = 12 Hz), 3.74 (d, 1H, J = 12 Hz); ^{13}C NMR (DMSO-d_6) δ 166.8, 118.4, 68.8, 67.2; MS (EI) m/z = 114, 85, 57.

Method B. The glycolonitrile solution was prepared as for method A, although the pH was kept as close to 7.5 or lower as possible. The solution was placed at 3 °C overnight and then the temperature reduced to −15 °C. The material was isolated in the same way as in method A, with trituration in methanol/acetone mixtures. Although this method was expected to grow X-ray quality crystals of **5**, compound **4** was isolated instead. Crystals selected from the mixture before trituration were suitable for X-ray analysis.

4-Amino-2-(cyanomethoxy)-2-(hydroxymethyl)oxazolidine (4): mp 40 °C (dec); IR 3400, 3150, 2900, 2800, 1670, 1610, 1450 cm^{-1}; ^1H NMR (DMSO-d_6) δ 6.8 (br), 4.18 (s, 2H), 4.44 (d, 1H, J = 12.5 Hz), 4.24 (d, 1H, J = 12.5 Hz), 3.37 (d, 1H, J = 7.5 Hz), 3.31 (d, 1H, J = 7.5 Hz); ^{13}C NMR (DMSO-d_6) δ 167.2, 126.4, 118.5, 70.6, 65.1, 46.3; HRMS (FAB) m/z $(M + H)^+$ = 172.0730 (172.0776 calcd).

***N*-Acetyl-4-amino-2-(acetoxymethyl)oxazole (6).** **5** (130 mg; 1 mmol) was dissolved in a small quantity of DMF (2 mL) and stirred overnight in a large excess of pyridine (2 mL) and acetic anhydride (20 mmol). The volatiles were removed by on a Rotovap, and the dark, oily residue was chromatographed on silica with ethyl acetate/hexane mixtures (1:5 to 1:1) to give crude **6**; recrystallization from ethyl acetate/hexane (1:2) gave X-ray quality crystals: mp 116 °C; ^1H NMR ($CDCl_3$) δ 8.92 (s, NH), 8.05 (s, 1H), 5.05 (s, 2H), 2.13 (s, 3H), 2.07 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 170.1, 167.8, 156.0, 136.6, 127.2, 57.5, 22.8, 20.4; MS (EI) m/z = 198; MS (CI) m/z + H = 199.

X-ray Crystallography.[18] Crystallographic data were collected at 132 K on a crystal of **6** (0.55 × 0.45 × 0.30 mm) using Cu Kα radiation (λ = 1.541 78 Å). The crystal was mounted in the 132 K cold stream of a Syntex $P2_1$ diffractometer equipped with a locally-modified LT-1 low temperature apparatus. Only random fluctuations of <0.1% in intensities of two standard reflections were observed during the course of data collection. The structure was solved by direct methods (SHELXTL PLUS) in the non-centrosymmetric space group $Pna2_1$ and refined by full-matrix (based on F^2) least squares. An absorption correction (XABS2) was applied. Hydrogen atoms were added geometrically and refined using a riding model and isotropic Us equal to 1.2 times the equivalent isotropic U of the bonded atom. In the final cycles of refinement non-hydrogen atoms were refined with anisotropic thermal parameters. A final difference map was essentially featureless.

Crystallographic data were collected at 169 K on a crystal of **4** (0.2 × 0.45 × 0.45 mm) using Mo Kα radiation (λ = 0.710 73 Å). The crystal was mounted in the 169 K cold stream of a Siemans R3m diffractometer equipped with a locally-modified low-temperature apparatus. Only random fluctuations of <0.1% in intensities of two standard reflections were observed during the course of data collection. The structure was solved by direct methods (SHELXTL PLUS) in the non-centrosymmetric space group $P2_1/n$ and refined by full-matrix (based on F^2) least squares. An absorption correction (XABS2) was applied. Hydrogen atoms were added geometrically and refined using a riding model and isotropic Us equal to 1.2 times the equivalent isotropic U of the bonded atom. In the final cycles of refinement non-hydrogen atoms were refined with anisotropic thermal parameters. A final difference map was essentially featureless.

Acknowledgment. This work was supported by the National Science Foundation (CHE-9628565, ASC-8902827, and the VPW program (K.K.B.)) and the National Aeronautics and Space Administration (NAGW 1031 and NAGW 2881, the latter provided through NASA NSCORT for Exobiology). We thank Dr. Peter Gantzel and Dr. Marilyn Olmstead for their crystallographic expertise. A grant for supercomputer time was provided by the grand challenge program.

Supporting Information Available: Ortep figures and tables of X-ray data for compounds **4** and **6** (15 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

JO962185R

(13) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Jensen, J. H.; Koseki, S.; Gordon, M. S.; Nguyen, K. A.; Windus, T. L.; Albert, S. T. *QCPE Bull.* **1990**, *10*.

(14) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Elbert, S. T. *J. Comput. Chem.* **1993**, *14*, 1347.

(15) Dunning, T. H., Jr.; Hay, P. J. *Modern Theoretical Chemistry*; Plenum: New York, 1976; pp 1−28.

(16) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 3396.

(17) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; DeFrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, 1994.

(18) The author has deposited atomic coordinates for structures **4** and **6** with the Cambridge Crystallographic Data Centre. The coordinates can be obtained, on request, from the Director, Cambridge Crystallographic DataCentre, 12 Union Road, Cambridge, CB2 1EZ, U.K.