JOC Article

# 5-Cyanoimino-4-oxomethylene-4,5-dihydroimidazole and Nitrosative Guanine Deamination. A Theoretical Study of Geometries, Electronic Structures, and N-Protonation†

Sundeep Rayat and Rainer Glaser*

*Department of Chemistry, University of Missouri-Columbia, Columbia, Missouri 65211*

*glaserr@missouri.edu*

*Received August 5, 2003*

The 5-cyanoimino-4-oxomethylene-4,5-dihydroimidazole **1** (R = H), its *N*1-derivatives **2** (R = Me) and **3** (R = MOM) and their cyano-*N* (**4**, **6**, **8**) and imino-*N* protonated (**5**, **7**, **9**) derivatives were studied with RHF, B3LYP, and MP2 theory. Solvation effects were estimated with the isodensity polarized continuum model (IPCM) at the MP2 level using the dielectric constant of water. Carbodiimide **10**, cyanamide **12**, *N*-cyanomethyleneimine **13**, and its protonated derivatives **14** and **15** were considered for comparison as well. Adequate theoretical treatment requires the inclusion of dispersion because of the presence of intramolecular van der Waals, charge-dipole, and dipole-dipole (including H-bonding) interactions. All conformers were considered for the MOM-substituted systems, and direct consequences on the preferred site of protonation were found. The vicinal push (oxomethylene)−pull (cyanoimino) pattern of the 5-cyanoimino-4-oxomethylene-4,5-dihydroimidazoles results in the electronic structure of aromatic imidazoles with 4-acylium and 5-cyanoamido groups. The gas-phase proton affinities of **1**−**3** are over 30 kcal/mol higher than that for *N*-cyanomethyleneimine **13**, and this result provides compelling evidence in support of the zwitterionic character of **1**−**3**. Protonation enhances the push−pull interaction; the OC charge is increased from about one-half in **1**−**3** to about two-thirds in the protonated systems. In the gas phase, cyano-*N* protonation is generally preferred but imino-*N* protonation can compete if the R-group contains a suitable heteroatom (hydrogen-bond acceptor, Lewis base). In polar solution, however, imino-*N* protonation is generally preferred. Solvation has a marked consequence on the propensity for protonation. Whereas protonation is fast and exergonic in the gas phase, it is endergonic in the polar condensed phase. It is an immediate consequence of this result that the direct observation of the cations **8** and **9** should be possible in the gas phase only.

## Introduction

It has been known since 1861 that nitrous acid deaminates the DNA bases,[1] but it was not until 100 years later that the toxicological effects of nitrous acid on nucleic acids were realized.[2] Subsequent studies showed that nitrosative deamination converts adenine to hypoxanthine,[3] cytosine to uracil,[4,5] and guanine to xanthine.[6] More recently, it was shown that the deamination of 2′-deoxyguanosine also yields 2′-deoxyoxanosine in large amounts (21.5%, Scheme 1).[7,8] Aside from nitrous acid, the bioregulatory agent nitric oxide[9] can cause DNA base deaminations in vitro,[10] and $HNO_2$[11] and NO[12] can induce mutations through the formation of interstrand DNA cross-links. The cross-links dG-to-dG and dG-to-dA have been isolated and spectrometrically identified (Scheme 1).[13,14] Nitrosative deamination and cross-linking of DNA

* To whom correspondence should be addressed. Telephone: (573) 882−0331. Facsimile: (573) 882−2754.

† Part 4 in the series "DNA Base Deamination." For part 3, see ref 24.

(1) (a) Strecker, A. *Ann.* **1861**, *118*, 151. (b) Kossel, A. *Ber.* **1885**, *18*, 1928. (c) Kossel, A.; Steudal, H. *Z. Physiol. Chem.* **1903**, *37*, 377.

(2) (a) Gierer, A.; Mundry, K. W. *Nature* **1958**, *182*, 1457−1458. (b) Schuster, H.; Schramm, G. *Z. Naturforsch.* **1958**, *13b*, 697. (c) Geiduschek, E. P. *Proc. Natl. Acad. Sci. U.S.A.* **1961**, *47*, 950−955. (d) Horn, E. E.; Herriott, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1962**, *48*, 1409−1416.

(3) (a) Nair, V.; Richardson, S. G. *Tetrahedron Lett.* **1979**, 1181−1184. (b) Nair, V.; Richardson, S. G. *J. Org. Chem.* **1980**, *45*, 3969−3974. (c) Nair, V.; Chamberlain, S. D. *Synthesis* **1984**, 401−403.

(4) Duncan, B. K.; Miller, J. H. *Nature* **1980**, *287*, 560−561.

(5) Viswanathan, A.; You, H. J.; Doetsch, P. W. *Science* **1999**, *284*, 159−162 and references therein.

(6) Schuster, V. H.; Wilhelm, R. C. *Biochim. Biophys. Acta* **1963**, *68*, 554−560.

(7) Suzuki, T.; Yamaoka, R.; Nishi, M.; Ide, H.; Makino, K. *J. Am. Chem. Soc.* **1996**, *118*, 2515−2516.

(8) Lucas, L. T.; Gatehouse, D.; Shuker, D. E. G. *J. Biol. Chem.* **1999**, *274*, 18319−18326.

(9) (a) Ignaro, L. J. *Annu. Rev. Pharmacol.* **1990**, *30*, 535. (b) Schuman, E. M.; Madison, D. V. *Science* **1991**, *254*, 1503.

(10) Wink, P. A.; Kasprzak, K. S.; Maragos, C. M.; Elespuru, R. K.; Misra, M.; Dunams, T. M.; Cebula, T. A.; Koch, W. H.; Andrews, A. W.; Allen, J. S.; Keefer, L. K. *Science* **1991**, *254*, 1001.

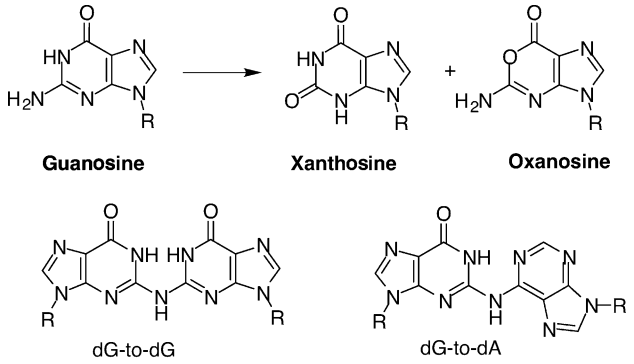
(11) (a) Geiduschek, E. P. *Biochemistry* **1961**, *47*, 950−955. (b) Geiduschek, E. P. *J. Mol. Biol.* **1962**, *4*, 467−487. (c) Becker, E. F., Jr.; Zimmerman, B. K.; Geiduschek, E. P. *J. Mol. Biol.* **1964**, *8*, 377−391. (d) Becker, E. F. *Biochim. Biophys. Acta* **1967**, *142*, 238−244.

(12) Burney, S.; Caulfield, J. L.; Niles, J. C.; Wishnok, J. S.; Tannenbaum, S. *Mutat. Res.* **1999**, *424*, 37−49.

(13) (a) Kirchner, J. J.: Hopkins, P. B. *J. Am. Chem. Soc.* **1991**, *113*, 4681−4682. (b) Kirchner, J. J.; Sigurdsson, S. T.; Hopkins, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 4021−4027. (c) Harwood: E. A.; Sigurdsson, S. T.; Edfelt, N. B. F.; Reid, B. R.; Hopkins, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 5081−5082.

10.1021/jo0351522 CCC: $25.00 
Published on Web 11/20/2003

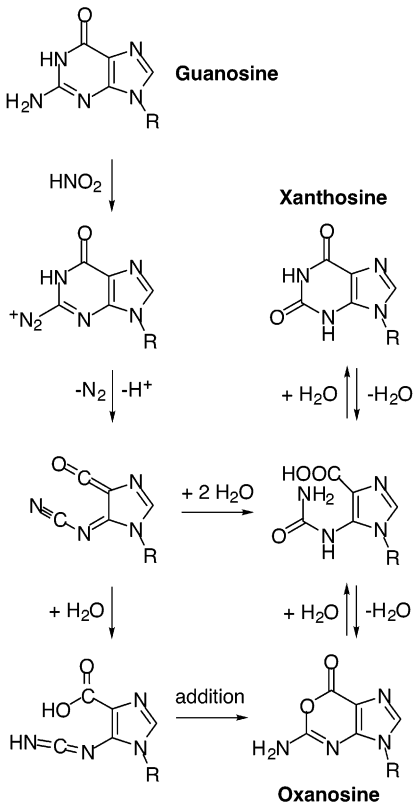
**SCHEME 1**




represent important mechanisms of genomic alteration.[15] Chronic inflammation is associated with the increased production of nitric oxide, and there seems to be a direct relationship between chronic inflammation and different types of cancers.[16] Because of the considerable dietary[17] and environmental[18] exposure of humans to the oxides of nitrogen, the deamination chemistry of DNA has attracted significant attention in recent years.

On the basis of the diazo chemistry of primary amines,[19] it was generally assumed that the deamination of nucleobases and cross-link formation proceed through the diazonium ions,[14,20,21] followed by $S_N1$ or $S_N2$ attack of water to form the $NH_2$/OH replacement product or of an amino group of a neighboring nucleobase to form a cross-link, respectively. Yet, the diazonium ions of the nucleobases have never been isolated or observed experimentally, their properties and stabilities are not known, and their reaction chemistry is not well understood. Our theoretical studies of the unimolecular dediazoniations of the DNA base diazonium ion revealed that these ions are much more prone to lose dinitrogen than the prototypical benzenediazonium ion.[22] Moreover, we found that the unimolecular dediazoniation of the guaninediazonium ion is accompanied by concomitant amide bond cleavage leading to a pyrimidine ring-opened cation (Scheme 2).[23−25] The results of our ab initio study provided a consistent explanation for the formation of all of the major products, xanthinosine, oxanosine and the cross-links, from a single common intermediate. Recently,

**SCHEME 2**



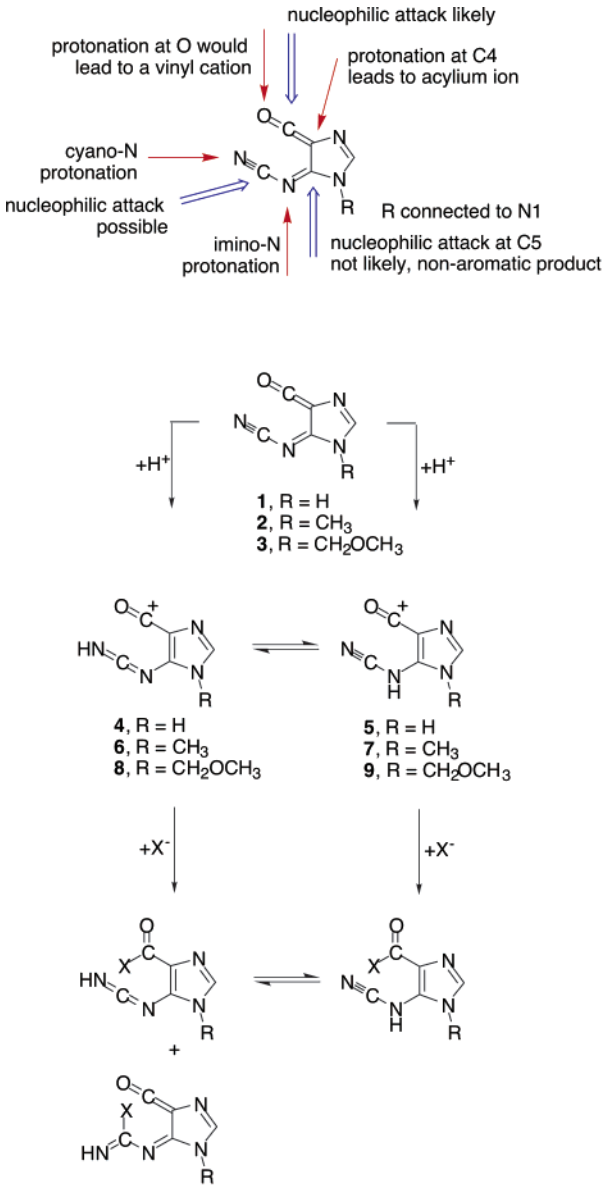

Suzuki et al. subjected guanosine and its methyl derivatives to nitrosative deamination, and it was shown that the exocyclic amino group of 2′-deoxyoxanosine arises from the imino nitrogen of 2′-deoxyguanosine.[26] This finding corroborates our mechanistic hypothesis.

The present Scheme 2 was modified in one significant way from the respective diagram we presented previously. We have now included the cyanoimine as the neutral key intermediate obtained by deprotonation of the product of the dediazoniation with pyrimidine ring-opening. Our more recent study of the dediazoniation of guaninediazonium ion in the presence of cytosine showed the formation of the cyanoimine as the result of a deprotonation-dediazoniation sequence. The present Scheme 2 reflects that a cyanoimine is the key product irrespective of the timing of deprotonation and dediazoniation. Protonation of the cyanoimine then also presents the first step in the addition chemistry leading to the products.

In this article we present the results of a theoretical study of *N*-protonation of the 5-cyanoiminodihydroimidazoles **1**−**3**, which are models of the postulated key intermediate in guanine deamination. The neutral compounds **1**−**3** and their protonated derivatives contain interesting and as yet entirely unexplored electronic structures. We have recently succeeded in the mass-spectroscopic generation and detection of protonated **1** (R = H) from various precursors that include guanosine and (the pyrimidine-ring opened) 5-cyanoamino-4-car-

(14) (a) Shapiro, R.; Dubelman, S.; Feinberg, A, M.; Crain, P. F.; McCloskey, J. A. *J. Am. Chem. Soc.* **1977**, *99*, 302−303. (b) Dubelman, S.; Shapiro, R. *Nucleic Acids Res.* **1977**, *4*, 1815−1827.
(15) Felley-Bosco, E. *Cancer Metastasis Rev.* **1998**, *17*, 25−37.
(16) Ohshima, H.; Bartsch, H. *Mutat. Res.* **1994**, *305*, 253−264.
(17) Wolff, E. A.; Wasserman, A. E. *Science* **1972**, *177*, 15.
(18) O'Sullivan, D. A. *Chem. Eng. News* **1970**, *48*, 38.
(19) Zollinger, H. *Diazo Chemistry I−Aromatic and Heteroaromatic Compounds*; VCH Verlagsgesellschaft: Weinheim, 1994.
(20) Bunton, C. A.; Wolfe, B. B. *J. Am. Chem. Soc.* **1974**, *96*, 7747−7752.
(21) Caulfield, J. L.; Wishnok, J. S.; Tannenbaum, S. *J. Biol. Chem.* **1998**, *273*, 12689−12695.
(22) (a) Glaser, R.; Horan, C. J. *J. Org. Chem.* **1995**, *60*, 7518−7528. (b) Glaser, R.; Horan, C. J.; Zollinger, H. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 2210−2213. (c) Glaser, R.; Horan, C. J.; Lewis, M.; Zollinger, H. *J. Org. Chem.* **1999**, *64*, 902−913.
(23) (a) Glaser, R.; Son, M.-S. *J. Am. Chem. Soc.* **1996**, *118*, 10942−10943. (b) Glaser, R.; Rayat, S.; Lewis, M.; Son, M.-S.; Meyer, S. *J. Am. Chem. Soc.* **1999**, *121*, 6108−6119.
(24) Glaser, R.; Lewis, M. *Org. Lett.* **1999**, *1*, 273−276.
(25) More recent studies have revealed that pyrimidine ring opening also may play an important role in the deaminations of cytosine (Rayat, S.; Glaser, R., unpublished results) and adenine (Hodgen, B.; Rayat, S.; Glaser, R. *Org. Lett.* **2003**, *5*, 4077−4080).
(26) (a) Suzuki, T.; Ide, H.; Yamada, M.; Endo, N.; Kanaori, K.; Tajima, K.; Morii, T.; Makino, K. *Nucleic Acids Res.* **2000**, *28*, 544−551. (b) Suzuki, T.; Kanaori, K.; Makino, K. *Nucleic Acids Symp. Ser.* **1997**, *37*, 313−314.

**SCHEME 3**

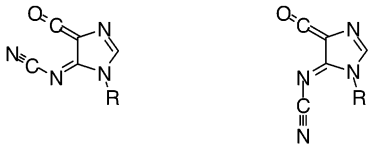



boxamide-imidazole (R = -$CH_2$-O-$CH_2$-$CH_2$-O-$CH_3$).[27] We compare the tautomers obtained by either protonation at the cyano-*N* or at the imino-*N* (Scheme 3).

In DNA, the R group is deoxyribose and computationally rather demanding. We replaced the sugar by an H atom, a methyl group, and by a methoxymethyl group. The methoxymethyl group is a simple model of the C1−O−C4 sugar fragment. The main structural issue concerns the regiochemistry of the protonation and we have determined cyano-*N* protonation preference energies to discuss this aspect. The other main structural issue concerns the stereochemical preference about the C5═N bond (Scheme 4). For **1**−**3** this is a configurational issue about the C5═N double bond and the labels *E* and *Z* could be used to describe the issue. For the cations **4**−**9**, however, the same stereochemical aspect is a conformational issue and the labels *E* and *Z* would seem less appropriate. Hence, we use the labels **a** and **b** to

(27) Qian, M.; Rayat, S.; Glaser, R. Manuscript in preparation.

**SCHEME 4**



***E*** ***Z***
**a** **b**

distinguish between these isomers, and the **a**-structure always features parallel-alignment of the ketene moiety and the NCN fragments at carbons 4 and 5. We have determined **a**-preference energies to quantify these stereochemical preferences. The structures of the neutral and cationic systems are analyzed and compared to the structures of prototypical systems to discriminate between various resonance forms (Scheme 5).

## Theoretical and Computational Methods

Potential energy surface explorations were performed with the programs Gaussian94 and Gaussian98[28] on an SGI PowerChallenge L minisupercomputer with 8 R12000 processors and two clusters of 8 ES40 and 8 ES45 Compaq alphaservers.[29] The structures were first determined by gradient optimization at the restricted Hartree−Fock level (RHF). Electron correlation effects on the structures were examined by performing structure optimizations again both with second-order Møller−Plesset perturbation theory[30] and with hybrid density functional theory at the B3LYP level.[31] All calculations were performed with the 6-31G* basis set. All geometry optimizations were followed by vibrational analyses to obtain the thermodynamical data and to ensure that the structures were indeed stationary. Density functional theory is a cost-effective method to account for parts of the electron correlation effects in a semiempirical fashion.[32] Density functional theory is frequently employed instead of the more computer-time-intensive perturbation theory, and the DFT results often are considered more reliable.[33] We will show substantial deviations between the DFT and MP2 results, and the analysis demonstrates that perturbation theory yields the more reliable data. As the present study progressed, we recognized more fully the significance of intramolecular van der Waals interactions and the inability of the B3LYP method to account for dispersion.[34]

(28) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.

(29) *Essentials of Computational Chemistry*; Cramer, C., Ed.; Wiley: New York, NY, 2002.

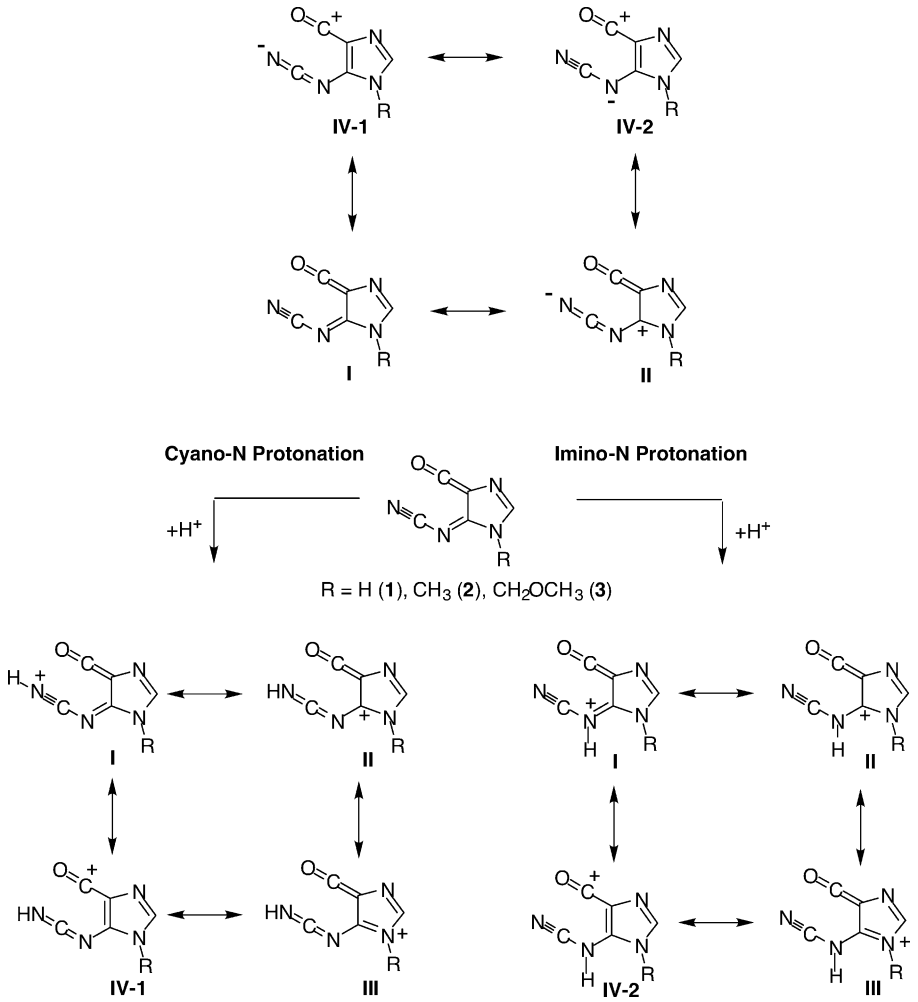
(30) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(31) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1998**, *37*, 785. (c) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.

(32) St-Amant, A. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publisher: New York, 1996; Vol. 7, p 217.

(33) (a) Florian, J.; Johnson, B. G. *J. Phys. Chem.* **1994**, *98*, 3681−3687. (b) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627. (c) Wang, J.; Eriksson, L. A.; Boyd, R. J.; Shi, Z.; Johnson, B. G. *J. Phys. Chem.* **1994**, *98*, 1844−1850.

**SCHEME 5**




The known problems with the description of hydrogen bonding at the B3LYP/6-31G* level also might contribute to the problem.[35] Total energies and thermodynamical data computed at the RHF/6-31G*, B3LYP/6-31G*, and MP2(full)/6-31G* levels are documented in Supporting Information. Models of the optimized structures are shown in the figures.

We computed relative energies and relative Gibbs free energies at all three theoretical levels, and these are given in Tables 1−3. Thermal energies were used as computed without any scaling. Unless otherwise noted, we discuss the MP2 derived $\Delta G_{rel}$ values. The computed proton affinities consider the thermal translational energy of the proton, and the entropy of the proton was determined with the Sackur Tetrode equation.[36] Gas-phase proton affinities contain a term that reflects the volume work associated with the protonation, and this term increases the proton affinity by $pV = RT = 0.59$ kcal/mol. Volume effects for protonations in condensed phase are practically null and our $\Delta G_{rel}$ values therefore do not contain any volume terms.

The computation of solvation effects remains challenging and requires the use of approximate theory. Solvent effects were determined with the isodensity surface polarized continuum model[37] (IPCM) for water ($\epsilon = 78.39$). These IPCM calculations were carried out at the MP2(full)/6-31G* level and were based on the gas-phase structures optimized at the MP2-(full)/6-31G* theoretical level. We used the thermal data computed for the gas phase to derive $\Delta G$ data. The enthalpy of hydration $\Delta H_{298}(H^+) = -262.2$ kcal/mol and the Gibbs free energy of hydration $\Delta G_{298}(H^+) = -262.5$ kcal/mol (gas phase to water transfer) were taken from the literature.[38] Total energies computed with solvation are given in Supporting Information (Table S1.3), and relative energies appear in Tables 1−3.

Electronic structures were analyzed by way of the natural population analysis[39] at the MP2(full)/6-31G* level in the gas phase (Table S4 ) and with the inclusion of solvation effects via the IPCM model (Table S5).

## Results and Discussion

**Parent 5-Cyanoimino-4-oxomethylene-4,5-dihydroimidazole (1) and *N*1-Methyl-5-cyanoimino-4-oxomethylene-4,5-dihydroimidazole (2).** The C5═N geometrical isomers of **1** and **2** are shown in Figure 1. There is a small **a**-preference of $\Delta G_{rel} = 1.2$ kcal/mol for

(34) (a) Kristyan, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175. (b) Hobza, P.; Sponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315−1325. (c) Papai, I.; Jancso, G. *J. Phys. Chem. A* **2000**, *104*, 2132.

(35) Lii, J.-H.; Ma, B.; Allinger, N. L. *J. Comput. Chem.* **1999**, *20*, 1593.

(36) Wedler, G. *Lehrbuch der Physikalischen Chemie*; Verlag Chemie: Weinheim, 1982; p 644.

(37) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. *J. Phys. Chem.* **1996**, *100*, 16098.

(38) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610.

(39) (a) Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. *NBO*, Version 3.1. (b) Glendening, E. D.; Weinhold, F. *J. Comput. Chem.* **1998**, *19*, 628.

**TABLE 1. Conformational a-Preference Energies about the C−N Bond**[a]

| | | RHF | | B3LYP | | MP2 | | IPCM | |
|---|---|---|---|---|---|---|---|---|---|
| R | tautomer | $\Delta E_{rel}$ | $\Delta G_{rel}$ | $\Delta E_{rel}$ | $\Delta G_{rel}$ | $\Delta E_{rel}$ | $\Delta G_{rel}$ | $\Delta E_{rel}$ | $\Delta G_{rel}$ |
| H | **1a−1b** | −0.69 | −0.73 | 0.35 | 0.26 | 1.34 | 1.16 | 0.25 | 0.07 |
| $CH_3$ | **2a−2b** | 2.87 | 3.06 | 2.96 | 2.92 | 4.15 | 4.23 | 3.08 | 3.16 |
| $CH_2OCH_3$ | **3a-Sg−3b-Sg** | 3.02 | 3.09 | 2.81 | 2.92 | 3.53 | 3.55 | −0.32 | −0.29 |
| $CH_2OCH_3$ | **3a-Sst−3b-Sst** | 1.93 | 2.69 | 1.55 | 2.41 | 1.78 | 2.82 | 0.99 | 2.03 |
| $CH_2OCH_3$ | **3a-Ag−3b-Ag** | 3.01 | 3.25 | 2.50 | 2.88 | 3.58 | 3.84 | 0.64 | 0.90 |
| H | **4a−4b** | −1.18 | −1.24 | −0.19 | −0.16 | −0.10 | −0.11 | −0.47 | −0.48 |
| $CH_3$ | **6a−6b** | 0.19 | 0.69 | 0.68 | 1.14 | 0.75 | 1.33 | 1.37 | 1.95 |
| $CH_2OCH_3$ | **8a-Sg−8b-Sg** | −3.41 | −2.68 | −1.97 | −1.40 | −2.82 | −2.15 | −4.59 | −3.92 |
| $CH_2OCH_3$ | **8a-Sst−8b-Sst** | −1.74 | −0.42 | −1.00 | −0.10 | −1.82 | −0.70 | −2.26 | −1.14 |
| $CH_2OCH_3$ | **8a-Ag−8b-Ag** | 1.08 | 1.63 | 1.22 | 1.12 | 1.23 | 1.29 | 3.42 | 3.48 |
| H | **5a−5b** | 1.01 | 0.94 | 0.83 | 0.74 | 2.15 | 2.01 | 2.43 | 2.30 |
| $CH_3$ | **7a−7b** | 4.28 | 3.38 | 3.57 | 3.01 | 4.17 | 4.53 | 0.30 | 0.66 |
| $CH_2OCH_3$ | **9a-Sg−9b-Sg** | 9.27 | 8.79 | 8.53 | 7.90 | 9.32 | 8.79 | 0.63 | 0.10 |
| $CH_2OCH_3$ | **9a-Sst−9b-Sst** | 7.52 | 7.51 | 7.45 | 7.33 | 7.35 | 7.36 | 6.11 | 6.12 |
| $CH_2OCH_3$ | **9a-Ag−9b-Ag** | 3.43 | 3.38 | 2.76 | 2.36 | 3.35 | 3.57 | 0.88 | 1.10 |

[a] Relative energies $\Delta E_{rel}$ and relative free energies $\Delta G_{rel}$ in kcal/mol.

**TABLE 2. Preference Energies for Cyano-*N* Protonation**[a]

| | | RHF | | B3LYP | | MP2 | | IPCM | |
|---|---|---|---|---|---|---|---|---|---|
| R | tautomer | $\Delta E_{rel}$ | $\Delta G_{rel}$ | $\Delta E_{rel}$ | $\Delta G_{rel}$ | $\Delta E_{rel}$ | $\Delta G_{rel}$ | $\Delta E_{rel}$ | $\Delta G_{rel}$ |
| H | **4a → 5a** | 4.97 | 6.92 | 12.28 | 13.48 | 4.93 | 5.65 | −8.74 | −8.02 |
| $CH_3$ | **6a → 7a** | 4.41 | 7.14 | 11.96 | 13.71 | 5.00 | 6.16 | −3.05 | −1.89 |
| $CH_2OCH_3$ | **8a-Sg → 9a-Sg** | −1.50 | 1.18 | 5.97 | 7.83 | −1.50 | 0.04 | −8.34 | −6.79 |
| $CH_2OCH_3$ | **8a-Sst → 9a-Sst** | −0.55 | 2.65 | 6.22 | 8.62 | −0.90 | 1.00 | −8.27 | −6.37 |
| $CH_2OCH_3$ | **8a-Ag → 9a-Ag** | 5.59 | 7.64 | 12.79 | 13.58 | 5.93 | 6.57 | −2.48 | −1.84 |
| H | **4b → 5b** | 7.17 | 9.11 | 13.30 | 14.38 | 7.18 | 7.77 | −5.83 | −5.24 |
| $CH_3$ | **6b → 7b** | 8.49 | 9.83 | 14.86 | 15.58 | 8.87 | 10.60 | −6.09 | −4.37 |
| $CH_2OCH_3$ | **8b-Sg → 9b-Sg** | 11.17 | 12.66 | 16.47 | 17.13 | 10.64 | 10.98 | −3.12 | −2.77 |
| $CH_2OCH_3$ | **8b-Sst → 9b-Sst** | 8.71 | 10.58 | 14.67 | 16.06 | 8.28 | 9.06 | 0.10 | 0.89 |
| $CH_2OCH_3$ | **8b-Ag → 9b-Ag** | 7.94 | 9.39 | 14.34 | 14.82 | 8.05 | 8.86 | −5.02 | −4.22 |
| model | **14 → 15** | 9.38 | 10.94 | 12.15 | 14.06 | 6.32 | 7.80 | −4.98 | −3.51 |

[a] Relative energies $\Delta E_{rel}$ and relative free energies $\Delta G_{rel}$ in kcal/mol.

**TABLE 3. Proton Affinities for the Most Stable Isomer**[a]

| | | RHF | | B3LYP | | MP2 | | IPCM |
|---|---|---|---|---|---|---|---|---|
| R | isomer | $\Delta E_{rel}$ | $\Delta G_{rel}$ | $\Delta E_{rel}$ | $\Delta G_{rel}$ | $\Delta E_{rel}$ | $\Delta G_{rel}$ | $\Delta G_{rel}$ |
| | | | | Cyano-*N* Protonation | | | | |
| H | **4b** | 219.7 | 206.7 | 223.3 | 210.0 | 217.6 | 204.0 | −16.7 |
| H | **4a** | | | | | | | −7.0 |
| $CH_3$ | **6a** | 221.5 | 209.2 | 225.6 | 212.8 | 219.1 | 206.0 | −14.7 |
| $CH_2OCH_3$ | **8b-Sg** | 228.5 | 215.4 | 231.1 | 217.5 | 225.8 | 212.0 | −15.0 |
| $CH_2OCH_3$ | **8a-Ag** | | | | | | | −14.3 |
| model system | **14** | 198.0 | 183.9 | 197.5 | 184.4 | 190.0 | 176.3 | −24.8 |
| | | | | Imine-*N* Protonation | | | | |
| H | **5a** | 214.2 | 199.3 | 210.5 | 196.1 | 211.2 | 197.0 | −9.3 |
| $CH_3$ | **7a** | 217.1 | 202.1 | 213.6 | 199.1 | 214.1 | 199.9 | −12.8 |
| $CH_3$ | **7b** | | | | | | | −9.1 |
| $CH_2OCH_3$ | **9a-Sg** | 223.6 | 208.4 | 220.3 | 205.34 | 220.9 | 206.2 | −11.9 |
| $CH_2OCH_3$ | **9a-Sst** | | | | | | | −9.5 |
| model system | **15** | 188.6 | 173.0 | 185.4 | 170.3 | 183.6 | 168.5 | −21.9 |

[a] Relative energies $\Delta E_{rel}$ and relative free energies $\Delta G_{rel}$ in kcal/mol.

isomer **1a** as compared to **1b**. Methyl substitution at N1 increases the **a**-preference to $\Delta G_{rel}$ = 4.2 kcal/mol, and this increase is consistent with some steric interference between the *N*1-methyl group and the NCN fragment in the **b**-structures. The distance between NCN-carbon and the methyl-C is 2.970 Å and well within vdW distance (vdW radius of C is 1.7 Å).

***N*1-Methoxymethyl-5-cyanoimino-4-oxomethylene-4,5-dihydroimidazole (3).** Several conformers can result by rotation about the glycosidic C−N bond and the C−O bond of the MOM group. With respect to the rotation about the glycosidic C−N bond, we considered *syn* (**S**) and *anti* (**A**) conformations (Scheme 6). The *syn* structures usually are synclinal rather than synperiplanar, and the *anti* structures are anticlinal rather than antiperiplanar. Both the **S** and **A** structures may be realized with various conformations with respect to the C−O bond of the MOM group, and these include two *gauche* conformations **g** and **g′** and one *s-trans* conformation **st** (Scheme 6). The C−O *gauche* conformations **g** and

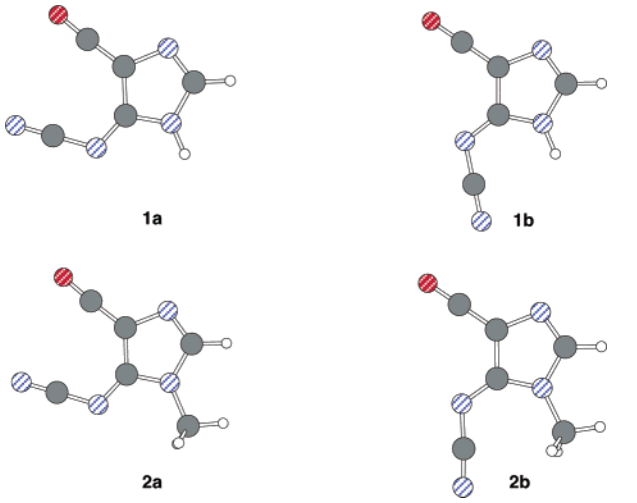
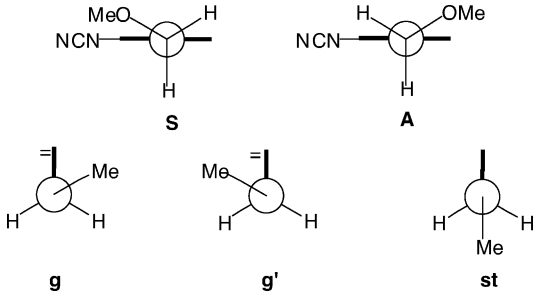



**FIGURE 1.** Structures of the cyanoimines **1** (R = H) and **2** (R = $CH_3$).

**SCHEME 6**




**g′** are not equivalent for synclinal and anticlinal C−N conformations. For the **S**-structure, there is a C−O *gauche* conformation that leaves the methyl group far away from the NCN fragment (**g**) and there is another *gauche* conformation where the methyl group would be in close proximity (**g′**). Similarly, for the **A**-structure, there is a C−O *gauche* conformation that leaves the methyl group far away from the imidazole-CH (**g**) and there is another *gauche* conformation where the methyl group would be placed close to the imidazole-CH (**g′**). None of these **g′** structures exist as minima on the potential energy surface. With the choice between **a** and **b**, **A** and **S**, and **g** and **st**, one may expect eight structures. While eight minima indeed exist on the RHF potential energy surface, only six minima persist on the potential energy surfaces computed with the correlated methods and these structures are shown in Figure 2.

The most stable structure is **3a-Ag** and relative energies of the other conformers and isomers are indicated in Figure 2. A theoretical model dependency occurs for the **3-Ast** structures. At the RHF level, both of the $C_s$-structures **3a-Ast** and **3b-Ast** are minima, whereas these structures are transition state structures at the B3LYP and MP2 level. Optimization at the correlated level without constraint to $C_s$ symmetry affords the **3a-Sst** and **3b-Sst** structures. Structures **3a-Sg** and **3a-Sst**, respectively, are $\Delta G_{rel}$ = 2.2 and 2.3 kcal/mol less stable than **3a-Ag**. For each pair with the same C−N and C−O conformations, the **a**-isomer is consistently preferred (Table 1) with preferences energies of 2.8−3.8 kcal/mol.

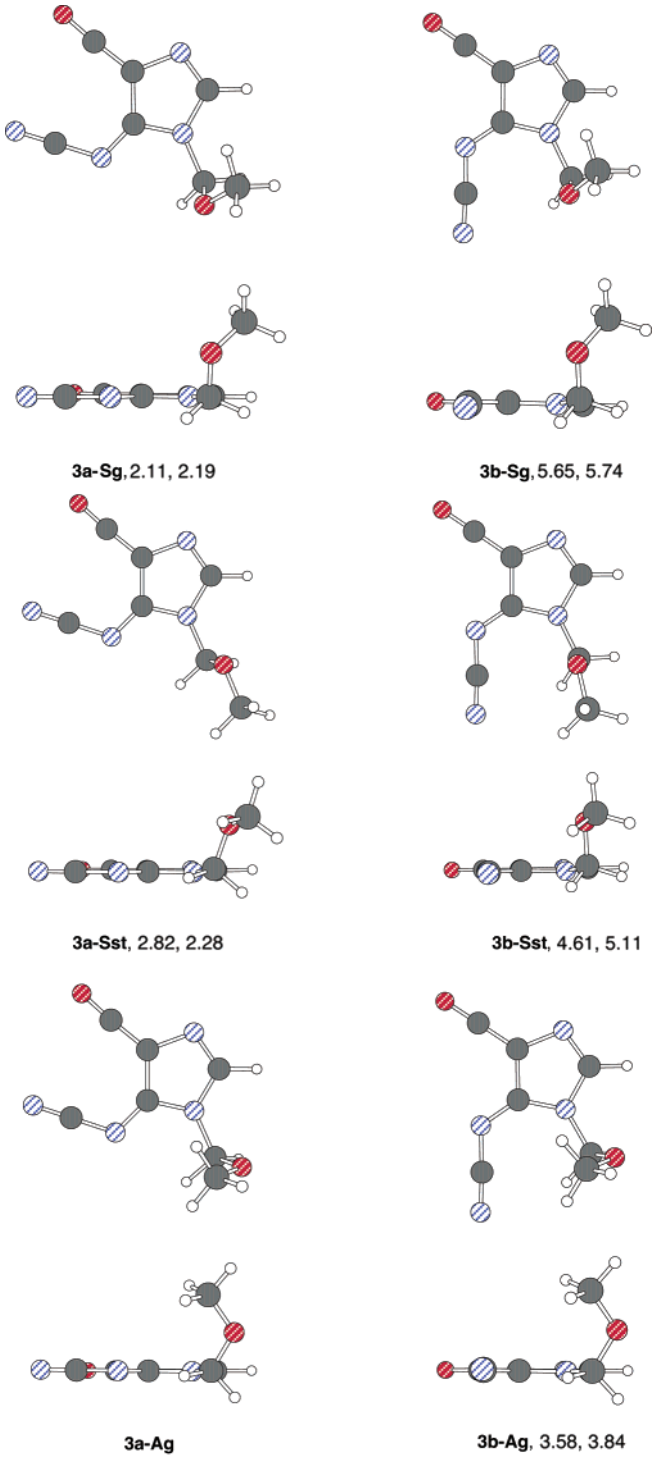



**FIGURE 2.** Structures of the cyanoimine **3** (R = -$CH_2$-O-$CH_3$). Structures with **S** conformations about the glycosidic C−N bond are shown in the top row, and structures with **A** conformations are shown on the bottom. Relatives energies $E_{rel}$ and relative Gibbs free energies $G_{rel}$ values are given in kcal/mol.

**Protonation of the Parent 5-Cyanoimino-4-oxomethylene-4,5-dihydroimidazole (1) and *N*1-Methyl-5-cyanoimino-4-oxomethylene-4,5-dihydroimidazole (2).** The optimized structures of the cyano-*N* protonated structures **4a** and **4b** and of the imino-*N* protonated structures **5a** and **5b** are shown in Figure 3.

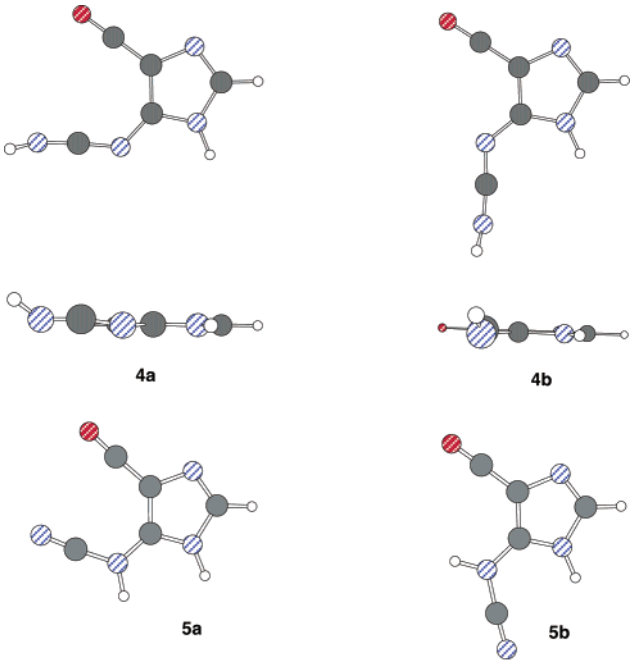



**FIGURE 3.** Structures of the cations **4** and **5** (R = H).

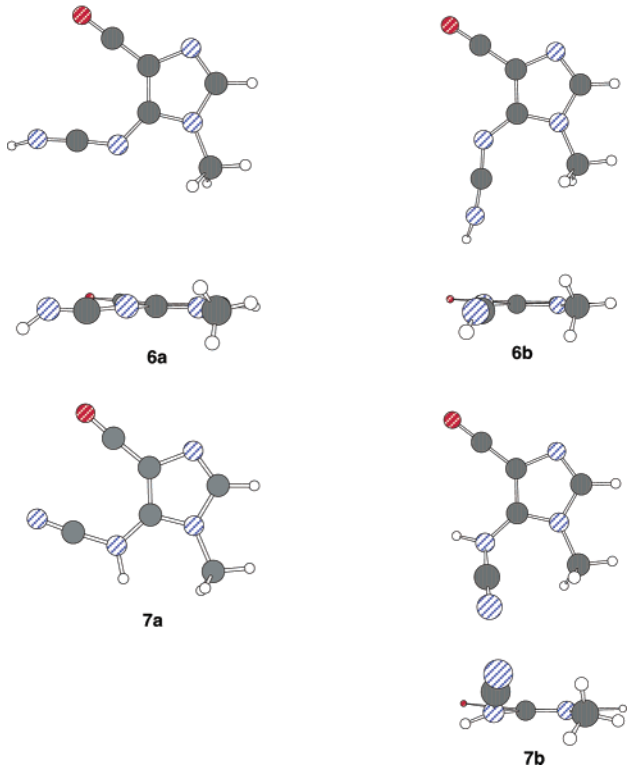



**FIGURE 4.** Structures of the cations **6** and **7** (R = $CH_3$).

The structures of **5** are planar at every level, but a significant structural effect of electron correlation occurs for **4**. Although the RHF structures of **4** are planar, the correlated levels indicate nonplanar minima in which protonation of the cyano-*N* results in a distinctly nonlinear H−N−C moiety.

The respective protonated derivatives **6** and **7** of the *N*1-methylcyanoimine **2** are shown in Figure 4. Although all structures of **5** are planar at all levels, **7b** stands out in that it is predicted to be planar at RHF and B3LYP while the MP2 level shows a small preference of 1.2 kcal/mol for a nonplanar structure. As with **4**, the respective structures of **6** are predicted to be nonplanar at the correlated level. Although the RHF level still prefers a $C_s$ structure for **6a**, a slight preference of 0.5 kcal/mol for the nonplanar structure **6b** even occurs at the RHF level.

The small **a**-preference of $\Delta G_{rel}$ = 1.2 kcal/mol for **1a** over **1b** vanishes upon cyano-*N* protonation and **4a** and **4b** are virtually isoenergetic with a slightly negative **a**-preference $\Delta G_{rel}$ = −0.11 kcal/mol. More significant is the much larger reduction of the **a**-preference from $\Delta G_{rel}$ = 4.2 kcal/mol for **2** to 1.3 kcal/mol for **6** upon cyano-*N* protonation.

Imino-*N* protonation maintains the preference for the **a**-isomers. The **a**-preference of $\Delta G_{rel}$ = 2.0 kcal/mol for **5** remains positive and becomes a bit larger than for **1**, and the **a**-preference of $\Delta G_{rel}$ = 4.5 kcal/mol for **7** remains significant.

The preference energies for cyano-*N* over imino-*N* protonation are listed in Table 2. There is a clear and pronounced preference in favor of cyano-*N* protonation. This preference is consistently higher for the **b**-isomers as compared to the **a**-isomers and varies by 2−5 kcal/mol.

The preference energies for cyano-*N* protonation of **1** are 5.7 and 7.8 kcal/mol for the **a**- and **b**-structures, respectively, and the respective values for **2** are 6.2 and 10.6 kcal/mol, respectively.

The preference energies for cyano-*N* protonation over imino-*N* protonation show a major and significant theoretical level dependency. One might expect that there is relatively close agreement between the correlated methods and the correlated data should deviate from the RHF data to some degree. However, it is found that the data computed at the correlated levels B3LYP and MP2 differ greatly.

**Protonation of *N*1-Methoxymethyl-5-cyanoimino-4-oxomethylene-4,5-dihydroimidazole (3).** As with the neutral cyanoimine **3**, rotational isomerism in the MOM group complicates the potential energy surface of the cyano-*N* and imino-*N* protonated derivatives **8** and **9**. With the choice between **a** and **b**, **A** and **S,** and **g** and **st**, we expected up to 8 structures each for **8** and **9**. We located six minima for **8** and seven minima for **9**, and their structures are shown in Figures 5 and 6. In several instances the **Ast** and **Sst** structures merge and exist as separate minima only for **9b**. The **Sg** and **Sst** structures are shown in Figures 5 and 6, respectively, and the **Ag** and **Ast** structures are shown in Figures 5-2 and 6-2 in the Supporting Information. As before, we show the cyano-*N* protonated structures on top and the imino-*N* protonated systems on the bottom in each figure. The most stable structures are **8b-Sg** for cyano-*N* protonation and **9a-Sg** for imino-*N* protonation, respectively, and relative energies of the other conformers and isomers are indicated in the Figures 5 and 6.

**Cyano-*N* Protonation and Electrostatic Stabilization of b-Structures.** There is a major difference between the MOM-substituted cations and the H- or methyl-substituted systems: The **8b-S** structures may be stabilized by through-space electrostatic stabilization between the MOM-O atom and the positively charged

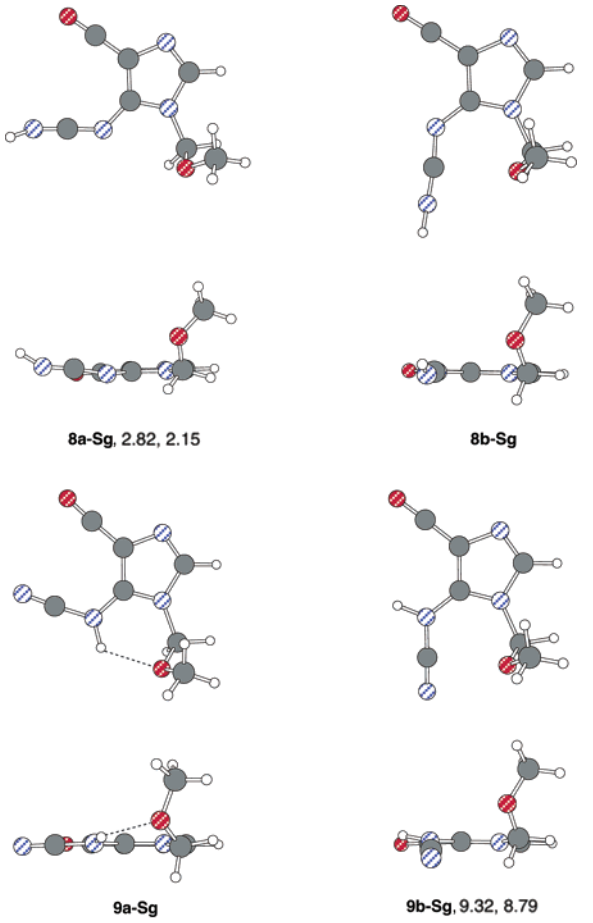



**FIGURE 5.** Structures of cations **8** and **9** (R = MOM) with synclinal (**S**) conformations with respect to glycosidic C−N bond and *gauche* (**g**) C−O bond conformation. Structures with (**A**) conformations are given in Supporting Information.

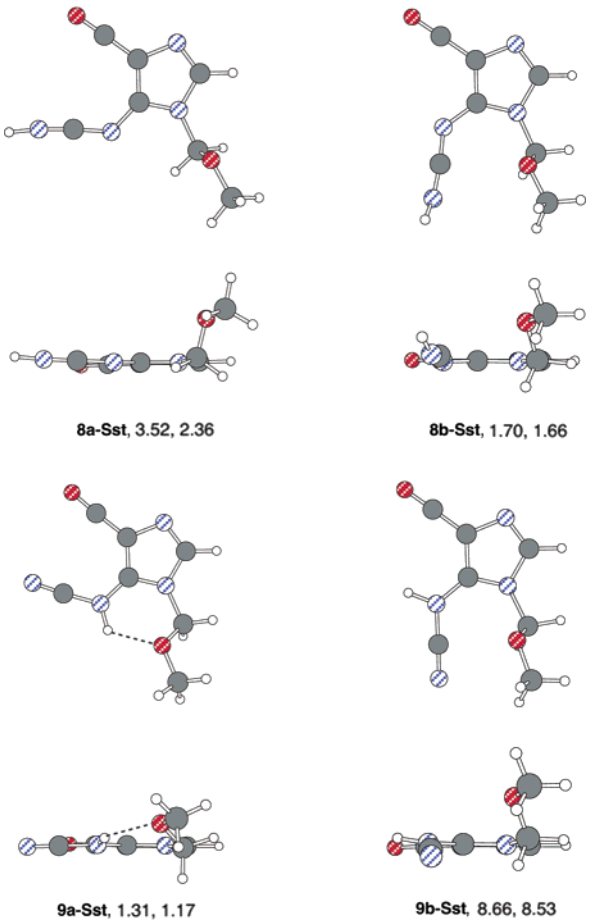



**FIGURE 6.** Structures of the cations **8** and **9** (R = MOM) with synclinal (**S**) conformation with respect to the glycosidic C−N bond and *s-trans* (**st**) C−O bond conformation. Structures with (**A**) conformations are given in Supporting Information.

protonated cyanoimine fragment. In the most stable structure, **8b-Sg**, the MOM-O atom is positioned well to stabilize the positive charge of the protonated cyanoimine fragment. This electrostatic interaction is best in the C−O *gauche* conformation, which orients the O lone pairs toward the cyanoimine moiety. In **8b-Sst**, only one lone pair is oriented favorably, and this structure is $\Delta G_{rel}$ = 1.7 kcal/mol less stable. Complete loss of this interaction in **8b-Ag** leaves this structure $\Delta G_{rel}$ = 2.6 kcal/mol less stable than **8b-Sg**.

Cyano-*N* protonation of **1** and **2** gives nearly isoenergetic C5−N isomers **4a** and **4b** and **6a** and **6b**, respectively. Cyano-*N* protonation of **3**, however, leads to a negative **a**-preference of $\Delta G_{rel}$ = −2.2 kcal/mol; **8b-Sg** is preferred over **8a-Sg**. This preference for the **b**-structure decreases as the O bridging of the protonated cyanoimine moiety is diminished: The **a**-preference energies for the pair **8a** and **8b** increases from −2.2 (**Sg**) to −0.7 (**Sst**) to +1.3 (**Ag**).

**Imino-*N* Protonation and Stabilization of a-Structures by Hydrogen Bonding.** There also is one major difference between the MOM-substituted cations and the H- or methyl-substituted systems: The **9a-S** structures may be stabilized by hydrogen bonding between the MOM-O atom and the H atom at the imino-*N* atom. Hydrogen bonding occurs in **9a-Sg** and **9a-Sst**, and these are the most stable structures of **9a**. **9a-Sst** is only

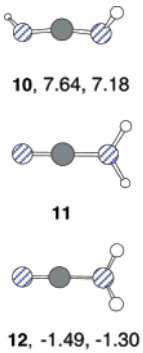



**FIGURE 7.** Structures of the parent carbodiimide **10**, of the planar transition state structure of cyanoamine **11**, and of the pyramidal minimum structures of cyanoamine **12**.

$\Delta G_{rel}$ = 1.2 kcal/mol less stable then **9a-Sg**. The absence of hydrogen bonding in **9a-Ag** leads to a high $\Delta G_{rel}$ value of 5.7 kcal/mol.

For each pair with the same C−N and C−O conformations, the **9a** isomer is consistently preferred over **9b** (Table 1), but the preference energies now greatly depend on the N1−C conformation. Imino-*N* protonation of **3-A** maintains the preference with an **a**-preference of $\Delta G_{rel}$ = 3.6 kcal/mol; the **A** structures of **9a** do not feature any additional hydrogen bonding. This value is in the range

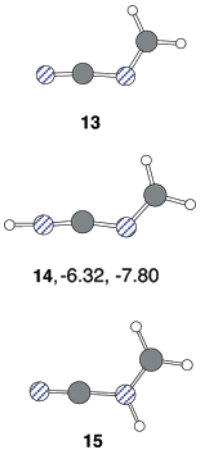



**FIGURE 8.** Structures of the cyanoimine of formaldehyde, **13**, and of the cations **14** and **15** obtained by cyano-*N* and imino-*N* protonation.

of the **a**-preferences of $\Delta G_{rel} = 2.0-4.5$ kcal/mol for the parent and the *N*1-methyl compounds **5** and **7**. With the possibility for hydrogen bonding, the **a**-preference drastically increases to $\Delta G_{rel} = 8.8$ (**Sg** structures) or $\Delta G_{rel} = 7.4$ kcal/mol (**Sst** structures). Hence, the strength of the hydrogen bonding is about 3.8−5.2 kcal/mol.

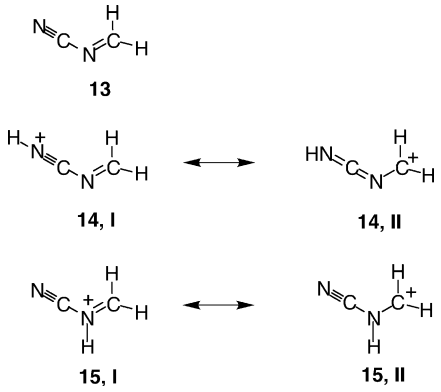
**Preference for Cyano-*N* Protonation.** We have seen that the preference for cyano-*N* protonation of **1** and **2** is in excess of 5.5 kcal/mol and that this preference consistently is higher for the **b**-isomers (Table 2). The same preference is found for the **8b** and **9b** pairs, and the $\Delta G_{rel}$ preference energies of 11.0 (**Sg**) > 9.1 (**Sst**) > 8.9 (**Ag**) again reflect the discussed extra electrostatic stabilization of the **8b** structures. Once the hydrogen bonding is possible in the imino-*N* protonated **9a** structure, then the preference for cyano-*N* protonation is greatly diminished but persists nevertheless. The preference for cyano-*N* protonation (**8a** versus **9a**) is reduced to 0.0 (**Sg**) and 1.0 (**Sst**) kcal/mol when hydrogen bonding occurs and remains 6.6 kcal/mol and "normal" in the absence of such hydrogen bonding (**Ag**).

This result points up the need to include R-groups at N1 that closely mimic a sugar moiety. In the absence of the sugar or the sugar model, an overwhelming preference for cyano-*N* protonation occurs and only with the availability of the O atom in the R-group does imino-*N* protonation become competitive.

**Structural Properties of 5-Cyanoimino-4-oxomethylene-4,5-dihydroimidazoles and Structural Relaxation upon Protonation.** The analysis of the structural effects of protonation allows for insights into the electronic structure of the cationic intermediates and provides information about the relative importance of the resonance forms shown in Scheme 5. Important structural parameters are collected in Tables S2.1−S2.3 in Supporting Information.

The consideration of two simple model systems is helpful. One comparison involves the parent carbodiimide **10** and its cyanoamine (aka cyanamide) tautomer **12**. The properties of these molecules should manifest themselves in the protonated cyanoimines if resonance forms **III** and **IV** (and to a lesser degree **II**) dominate. The other models involve the parent cyanoimine **13**, the cyanoimine of

**SCHEME 7**




formaldehyde (Scheme 7), and its cyano-*N* and imino-*N* protonated derivatives **14** and **15**, respectively.

**NH═C═NH and N≡C−NH$_2$.**[40] The optimized structures are shown in Figure 7. The carbodiimide **10** is $C_2$ symmetric,[41] and it is important for the following discussion to realize that the N−H bonds are almost orthogonal with respect to each other. Cyanoamine prefers a $C_s$-symmetric structure **12** with a pyramidal amino group,[42] and $C2_v$-**11** is the transition state for *N*-inversion. The carbodiimide is $\Delta G_{rel} = 7.2$ kcal/mol less stable than the cyanoamine. A typical carbodiimide C═N bond is 1.23 Å long, while the N≡C and C−N bond lengths in a cyanoamine are 1.18 and 1.35 Å, respectively.

**Protonation of *N*-Cyanomethyleneimine.** The optimized structures of **13** and of the cyano-*N* and imino-*N* protonated derivatives **14** and **15** are shown in Figure 8. There is a clear preference of 7.8 kcal/mol for cyano-*N* protonation (Table 2). Structure **13** shows that the N≡C bond length in the cyano*imine* is 1.18 Å and the same as in the cyano*amine* **12**, whereas the C−N bond in the cyano*imine* is 1.36 Å and again almost the same as in the cyano*amine* (1.35 Å). Structure **14** shows that resonance form **I** is best; **14** is a protonated cyanoimine rather than a carbodiimidylmethanium ion. The nitrilium ion **14** features N≡C and C−N bond lengths that *both are shorter* than those in **13**. Similarly, the cyanoiminium ion resonance form **I** best describes ion **15**, and this protonation occurs with hardly any structural relaxation.

**Neutral and Protonated 5-Cyanoimino-4,5-dihydroimidazoles.** The data in Tables S2.1−S2.3 show that the cyanoimine fragments in **1**−**3** differ only marginally (less than 0.02 Å) from the structure of the model cyanoimine **13**. The structures of the cyano-*N* protonated systems reveal several interesting features. The cyano-CN bond in **1**−**3** does *not* lengthen to any significant extent upon cyano-*N* protonation! This CN bond remains shorter than 1.2 Å in all isomers of **4**, **6**, and **8**. The slight shortening observed in the cyano-*N* protonation of **13** to **14** does not occur, but there certainly is no significant lengthening of that bond either. The short terminal CN bond length alone would suggest that the nitrilium

(40) For a recent theoretical study with comparison to experiment, see: Tordini, F.; Bencini, A.; Bruschi, M.; Dr Gioia, L.; Zampella, G.; Fantucci, P. *J. Phys. Chem A* **2003**, *107*, 1188 and references therein.

(41) Glaser, R.; Lewis, M.; Wu, Z. *J. Phys. Chem. A* **2002**, *106*, 7950 and references therein.

(42) This type of structure occurs in the gas phase and in the solid state: Denner, L.; Lugar, P.; Buschmann, J. *Acta Crystallogr. C* **1988**, *44*, 1979 and references therein.

resonance form **I** is important (Scheme 5). However, the cyano-*N* protonation has a major effect on the C−N and C═N bond lengths, and the data in Tables S2.1−S2.3 show a drastic shortening of the former and a lengthening of the latter, respectively. Overall, the HNCN fragment in **4**, **6**, and **8** is best described as an unsymmetrical carbodiimide with an average CN bond length of 1.226 Å and an average CN bond difference of 0.062 Å. The major consequences of imino-*N* protonation are lengthenings of the C−N and C═N bonds by about 0.02 and 0.06 Å, respectively. Clearly, both cyano-*N* and imino-*N* protonation lead to some delocalization of positive charge via resonance forms **II**−**IV**.

**Proton Affinities.** Computed proton affinities are summarized in Table 3. Surprisingly, the proton affinities increase in going from the methyl to the methoxymethyl systems. Even more surprising was the finding of the large increase of the proton affinities of all the dihydroimidazole systems as compared to the parent system *N*-cyanomethyleneimine **13**. The proton affinities of the MOM-substituted system are about 35 kcal/mol higher than those for **13**. The origin of this large effect was recognized by application of electron density analysis.

**Electron Density Analysis of *N*1-Methoxymethyl- 5-cyanoimino-4-oxomethylene-4,5-dihydroimidazole (3) and Its Relaxation Upon Protonation.** Electron density analysis was carried out for the most stable isomers of the neutral systems **3a-Ag** and **3b-Ag**, the most stable cyano-*N* protonated systems **8a-Ag** and **8b-Sg**, and the most stable imino-*N* protonated ions **9a-Sg** and **9b-Sst**. Results are provided in a Table in Supporting Information.

In **3**, there is about half of a negative charge located on the NCN fragment while half of a positive charge is localized on the CO fragment. The electron density distribution of the NCN fragment is of the type (− + −). hence, the aromatic resonance forms **IV-1** and **IV-2** (Scheme 5) are as important for protonated **3** as is **I**. The C4−C5 bond is highly polarized consistent with the donor−acceptor substitution. Note that the positive C5 charge always is larger in magnitude than the negative C4 charge; some contribution is made by **II**. The recognition of the importance of **IV-1** and **IV-2** reveals the resemblance of the "dihydroimidazole" and imidazole itself. The MOM group is a donor. The N1−C2−N3 fragment is of the type (− + −) with large and almost equal negative charges on N1 and N3.

Protonation essentially neutralizes the NCN fragment; that is, about one-half of the positive charge of the proton stays with the H-atom and the other half charge is delocalized onto the NCN fragment (about 0.1), the CO (about 0.13−0.16), the imidazole ring (about 0.2), and the MOM group (about 0.07). Resonance form **III** clearly is not important; while diminished somewhat, the N1 and N3 charges all remain rather negative. The CO group carries *two-thirds* of the positive charge after protonation, and this finding highlights the dominance of **IV-1** or **IV-2**, respectively. We chose the labels of the resonance forms in Scheme 5 to emphasize that the cyano-*N* and the imino-*N* protonated structures **IV-1** and **IV-2** are so similar to the resonance forms **IV-1** and **IV-2** of **3**, respectively. As with neutral **3**, we find the resonance forms with an aromatic imidazole to be a major contributor. Contributions from **II** remain essentially unchanged upon protonation.

The discussion up to now applied to **3** and its protonated derivatives irrespective of the site of the protonation. The effects of the protonation site are entirely localized on the NCN fragment, and we characterize these effects using the group charges of both CN groups. The CN groups charges are about the same in each of the cyano-*N* protonated systems, and this finding corroborates well the carbodiimide nature of these ions. The CN group charges also support the electrostatic argument made to explain the stability of **8b-Sg**. The CN-group charge in **8b-Sg** is higher than in **8a-Ag**, and this charge localization increases any electrostatic attraction with the proximate MOM group. The other important intramolecular interaction, the hydrogen-bonding interaction in **9a-Sg**, manifests itself in the more negative MOM O-charge and the increased positive charge on the H-atom of the cyanoimine functionality; both features strengthen the hydrogen bond.

**Solvent Effects.** The theoretical prediction of solvent effects remains challenging and approximate. Studies of the effects of specific solvation and of the effects of the permanent anisotropic environment in DNA are desirable, but they remain conceptually difficult and computationally demanding. Here, we employ a polarized continuum model that should provide a first indication of solvation effects on the present chemistry in homogeneous, polar solution. The conformers of the MOM-substituted systems differ significantly in their shapes, and the choice of the cavity determination might thus be critical. We employed the isodensity polarized continuum model (IPCM) as it seems most suited to account for these shape changes.

The IPCM data in Table 1 shows that the **a**-preferences are largely retained in aqueous solution. Aside from this general correlation, however, the solvation effects do not show any systematic trends. While most **a**-preferences decrease, some increase. Perhaps most puzzling are the greatly different solvation effects on the **a**-preference energies of the (**S**)-conformers of **9**.

The IPCM model predicts a consistent and significant solvent effect on the site preference for protonation: Solvation favors imino-*N* protonation over cyano-*N* protonation (Table 2). For the isolated molecules, imino-*N* protonation is preferred only in those cases where an internal hydrogen bonding opportunity is realized. In contrast, solvation favors imino-*N* protonation over cyano-*N* protonation for all systems by 7−14 kcal/mol.

In the gas phase, protonation of **1**−**3** occurs quickly and the protonated species are abundant. The gas-phase proton affinity of water is 167.9 kcal/mol,[43] and proton transfer from hydronium ion to **1**−**3** in the gas phase thus also is highly exothermic. The IPCM calculations suggest, however, that solvation causes the proton affinities $\Delta G_{298}$ of **1**−**3** to become slightly negative, that is, that the protonations of **1**−**3** are slow and endergonic processes. While the protonations of **1**−**3** are highly exothermic processes in the gas phase, this does not seem to be the case in solution because such a proton transfer would replace the well-solvated hydronium ion by the much less

(43) Collyer, S. M.; McMahon, T. B. *J. Phys. Chem.* **1983**, *87*, 909.

well-solvated ions **8** and **9**. The population analysis explains and fully supports this conclusion.

The population analyses indicate increased bond polarities in the polar solvent as compared to those in the gas phase, while all of the above-discussed characteristic features of intramolecular charge distribution persist. The neutral molecules **1**−**3** remain highly zwitterionic "imidazoles" and the protonated derivatives feature almost neutral HNCN fragments and extreme positive charge dispersal. The comparatively high solvation energy for the zwitterionic proton acceptors and the comparatively low solvation energy for the highly delocalized cations explain the reduced propensity for protonation of **1**−**3** in condensed phase.

## Conclusions

The results show that intramolecular van der Waals, charge-dipole, and dipole-dipole (including hydrogen bonding) interactions all are important factors contributing to the isomer preferences. Therefore, the adequate theoretical treatment requires correlated methods that include dispersion.

The vicinal push (oxomethylene)−pull (cyanoimino) pattern of the 5-cyanoimino-4-oxomethylene-4,5-dihydroimidazoles results in the electronic structure of aromatic imidazoles with 4-acylium and 5-cyanoamido groups. The gas-phase proton affinities of **1**−**3** are much higher than those for *N*-cyanomethyleneimine **13**, and this result provides compelling evidence in support of the zwitterionic character of **1**−**3**.

Protonation enhances this push−pull interaction; the OC-charge is increased from about one-half in **1**−**3** to about two-thirds upon protonation. This also means that cyano-*N* protonation yields ions that are best described as unsymmetrical carbodiimides rather than localized nitrilium ions. Similarly, imino-*N* protonation yields cyanoamines rather than cyanoiminium ions.

In the gas phase, cyano-*N* protonation is generally preferred but imino-*N* protonation can compete if the R-group contains a suitable hydrogen-bond acceptor. Clearly, the use of the MOM group as the smallest possible mimic of a sugar moiety is critical if the computational study is to support experimental work on guanosine.

The characteristic features of intramolecular charge distributions of **1**−**3** and their protonated derivatives persist when solvation effects are considered by way of a continuum model. However, solvation does have a marked consequence on the site of and the propensity for protonation. Whereas cyano-*N* protonation is preferred in gas phase, imino-*N* protonation is preferred in polar condensed phase. Whereas protonation is fast and exergonic in the gas phase, it is endergonic in polar condensed phase. It is an immediate consequence of this computational result that the direct observation of the cations **8** and **9** should be possible in the gas phase only.

The 5-cyanoimino-4-oxomethylene-4,5-dihydroimidazoles **1**−**3** all prefer that CN conformation in which the ketene and the cyanoimine arms are parallel with each other; the **a**-isomer is favored over the **b**-isomer. Cyano-*N* protonation of **1** and **2** produce isoenergetic **a**- and **b**-structures. For **3**, however, the most stable products of cyano-*N* protonation are **8b-Sg** and **8b-Sst** because of through-space electrostatic interaction between the MOM-O atom and the positively charged protonated cyanoimine fragment. In the absence of this kind of stabilization, as in **8b-Ag**, the MOM-substituted **a**-isomer becomes preferred just as in the cases of **1** and **2**. Imino-*N* protonation of **1** and **2** does not produce much variation in the **a**-preference energies. A significant **a**-preference for imino-*N* protonation of **3** occurs as a result of the stabilization of the **9a-S** structures by hydrogen bonding involving the MOM-O atom. These trends in the conformational preferences largely persist in polar solvents, although solvent effects on relative preferences may vary significantly and are not easily traced.

**Acknowledgment.** This work was supported by grant NIGMS R01GM61027 from the National Institutes of Health. We thank the University of Missouri-Columbia research computing services for computer time and Drs. G. Springer and L. Sanders for their assistance.

**Supporting Information Available:** A discussion of nomenclature, two figures of the structures of cations **8** and **9** (R = MOM) with anticlinal (**A**) conformations with respect to glycosidic C−N bond and *gauche* (**g**) C−O bond conformation (Figure 5-2) or *s-trans* (**st**) C−O bond conformation (Figure 6-2), tables documenting total energies and thermodynamic data (Tables S1.1−S1.3), structures (Tables S2.1−S2.3) computed at the levels RHF/6-31G*, B3LYP/6-31G*, and MP2(full)/6-31G*, a table with the dipole moments in the gas phase and in solution (Table S3), two tables with computed charges in the gas phase (Table S4) and in solution (Table S5), and Cartesian coordinates of all MP2(full)/6-31G* structures. This material is available free of charge via the Internet at http://pubs.acs.org.

JO0351522