JOC Article

On the Aromatic Character of the Heterocyclic Bases of DNA and RNA[#]

Michał K. Cyrański,[†] Mirosław Gilski,[‡,§] Mariusz Jaskólski,[‡,§] and Tadeusz Marek Krygowski*,[†]

Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland, Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznań, Poland, and Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland

tmkryg@chem.uw.edu.pl

Received June 4, 2003

Studies based on ab initio optimized geometries (at B3LYP/6-311+G** and MP2/6-311+G** levels) and on experimental structures retrieved from the Cambridge Structural Database (CSD) reveal that the nucleobases constituting DNA and RNA differ significantly in their aromatic character, as shown by the geometry-based index of aromaticity HOMA that ranges from 0.466 for thymine to 0.917 for adenine, based on B3LYP/6-311+G** calculations, and 0.495−0.926, respectively, if based on the MP2/6-311+G** level. Aromaticity of the bases decreases markedly with an increase of the number of double-bond C═X (X = N, O) substituents at the rings. H-bonds involving C═O groups in Watson−Crick pairs cause an increase of the aromatic character of the rings.

Introduction

Aromaticity is a theoretical concept of immense practical importance and wide use in organic chemistry and related fields.[1] According to the database of the Institute for Science Information in Philadelphia,[2,3] the terms aromaticity and/or aromatic (properties, character) have been used in over 60 000 scientific papers as a key word, in a title, or in the abstracts in the last two dacades. Although the term *aromaticity* is neither precisely defined nor has a unique interpretation,[4] the aromatic character[5−9] is most frequently associated with (mostly planar) cyclic π-electron systems exhibiting the following: (i) increased stability, initially described by the low chemical reactivity of benzene (and related systems), and later in a more quantitative way by resonance energy[10−12] [nowadays, the aromatic stabilization energies determined either theoretically or experimentally (from heats of formation) by isodesmic[13,14] or preferably homodesmotic[15,16] schemes of reactions are widely considered (e.g., refs 4 and 6) as a much more reliable source of information; (ii) averaging of bond lengths to values that are intermediate between those characterizing typical single and double bonds;[6,17,18] (iii) π-electron ring current formation when the molecule is exposed to external magnetic field, a phenomenon that is associated with an anisotropy

[#] Dedicated to our friend and teacher Prof. Maciej Wiewiórowski on the occasion of his 85th birthday.

* Address correspondence to this author. FAX: (+48) 22 8222 892.
[†] University of Warsaw.
[‡] A. Mickiewicz University.
[§] Polish Academy of Sciences.

(1) Krygowski, T. M.; Cyrański, M. K.; Czarnocki, Z.; Häfelinger, G.; Katritzky, A. R. *Tetrahedron*, **2000**, *56*, 1783−1796 and references therein.
(2) Institute for Scientific Information, Philadelphia (1981−1999). Retrieved in February 2000.
(3) Krygowski, T. M.; Cyrański, M. K. *Chem. Rev.* **2001**, *101*, 1385−1420.
(4) (a) Cyrański, M. K.; Krygowski, T. M.; Schleyer, P. v. R.; Katritzky, A. R. *Chem Listy* **2000**, *94*, 788−789. (b) Cyrański, M. K.; Krygowski, T. M.; Katritzky, A. R.; Schleyer, P. v. R. *J. Org. Chem.* **2002**, *67*, 1333−1338.
(5) Schleyer, P. v. R.; Jiao, H. *Pure Appl. Chem.* **1996**, *68*, 209−218.
(6) Minkin, V. I.; Glukhovtsev, M. N.; Simkin, B. Ya. *Aromaticity and Antiaromaticity−Electronic and Structural Aspects*; Wiley: New York, 1994.
(7) Bergmann, E. D.; Pullman, B., Eds. *Aromaticity, Pseudoaromaticity and Antiaromaticity*; Proceedings of an International Symposium held in Jerusalem 1970, Israel Academy of Sciences and Humanities, Jerusalem, 1971.
(8) (a) Watson, M. D.; Fechtenkötter, A.; Müllen, K. *Chem. Rev.* **2001**, *101*, 1267−1300. (b) Katritzky, A. R.; Jug, K.; Oniciu, D. C. *Chem. Rev.* **2001**, *101*, 1421−145. (c) Schaad, L. J.; Hess, B. A., Jr. *Chem. Rev.* **2001**, *101*, 1465−1476. (d) Shaik, S.; Shurki, A.; Danovich, D.; Hiberty, P. C. *Chem. Rev.* **2001**, *101*, 1501−1540. (e) Minkin, V. I.; Minyaev, R. M. *Chem. Rev.* **2001**, *101*, 1247−1266. (f) Wiberg, K. B. *Chem. Rev.* **2001**, *101*, 1317−1332. (g) Williams, R. V. *Chem. Rev.* **2001**, *101*, 1185−1204. (h) Bühl, M.; Hirsch, A. *Chem. Rev.* **2001**, *101*, 1153−1184. (i) Mitchell, R. H. *Chem. Rev.* **2001**, *101*, 1301−1316. (j) Bleeke, J. R. *Chem. Rev.* **2001**, *101*, 1205−1228. (k) King, R. B. *Chem. Rev.* **2001**, *101*, 1119−1152.
(9) Sadlej-Sosnowska, N. *J. Org. Chem.* **2001**, *66*, 8737−8743.
(10) Pauling, L.; Sherman, J. *J. Chem. Phys.* **1933**, *1*, 606.
(11) Kistiakowski, B.; Ruhoff, J. R.; Smith, H. A.; Vaughan, W. E. *J. Am. Chem. Soc.* **1936**, *58*, 146−153.
(12) (a) Wheland, G. W. *The Theory of Resonance and its Application to Organic Chemistry*; J. Wiley: New York, 1944. (b) Wheland, G. W. *The Theory of Resonance and its Application to Organic Chemistry*; J. Wiley: New York, 1944.
(13) Hehre, W. J.; Ditchfield, D.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4786−4795.
(14) George, P.; Trachtman, M.; Bock, C. W.; Brett A. M. *J. Chem. Soc., Perkin Trans. 2* **1976**, 1222−1227.
(15) Hehre, W. J.; McIver, R. T.; Pople, J. A.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1974**, *96*, 7162−7163.
(16) George, P.; Trachtman, M.; Brett, A. M.; Bock, C. W. *J. Chem. Soc., Perkin Trans. 2* **1977**, 1036−1047.
(17) Julg, A.; François, Ph. *Theor. Chim. Acta* **1967**, *7*, 249.
(18) Kruszewski, J.; Krygowski, T. M. *Tetrahedron Lett.* **1972**, 3839−3842.

10.1021/jo034760e CCC: $25.00 © 2003 American Chemical Society
Published on Web 10/09/2003

of magnetic susceptibility, increase of exaltation of the magnetic susceptibility, and ^{1}H NMR chemical shifts;[19−26] and (iv) a non-ground-state property that is associated with a typical reactivity of these systems in which the π-electron structure tends to be retained.[27] It is best exemplified by the tendency of aromatic systems to undergo substitution reactions in contrast to their olefinic analogues, which react by addition.

The last criterion often has been a subject of heated dispute.[5−7] However, the meaning of the term in organic synthesis cannot be neglected and hence the compounds which fulfill all (i)−(iv) criteria are expected to be fully aromatic.[1]

Among the aromatic systems, of particular importance, at least from the point of view of biology and bioactive systems, are the basic constituents of DNA and RNA: adenine, cytosine, guanine, thymine, and uracyl. The electronic properties of nucleobases have been the subject of numerous studies, e.g., ref 28, as they are of fundamental importance for the unique character of these molecular systems, i.e., for their ability to form H-bonded pairs and to assemble into π-electron stacks. To the best of our knowledge, however, the concept of aromaticity, so seminal in organic chemistry, has not been applied to nucleobases in a systematic manner, although recently it has been used in a study on the structure of DNA.[29]

The purpose of this report is to investigate how the basic components of nucleic acids differ in their aromatic character, what are the main sources of the differences, to what extent their aromatic character depends on intramolecular factors, as substituent effects, and finally to what extent the aromatic character depends on the intermolecular interactions, most importantly on H-bonding within canonical Watson−Crick pairs. To do so we have analyzed the molecular geometries of these bases retrieved from the CSD[30] as well as optimized by quantum chemical methods.[31] In addition to the canonical forms, also unusual tautomeric forms as well as nucleobases involved in Watson−Crick pairs were considered.

Methodology

To study the aromatic character of the bases as well as their individual fragments (local aromatic character) we have undertaken two approaches. By use of ab initio calculations at the DFT B3LYP/6-311+G** level of theory, we have optimized the geometry of all basic constituents of DNA and RNA, as well as of some Watson−Crick pairs. All species corresponded to the minima on the potential energy surface at the B3LYP /6-311+G** level, with no imaginary frequencies. For comparison we have also optimized the nucleobases at the MP2/6-311+G** level. They also corresponded to the minima at the potential energy surface (at the MP2/6-311+G** level). Unfortunately we were unable to compute the Watson−Crick pairs at this level of theory due to problems with the energy convergence and approaching of a system to another geometric arrangement of lower energy. On the other hand, the CSD has been searched for relevant X-ray structures. The geometry-based index of aromaticity HOMA[32] was then calculated for the heterocyclic systems and for their fragments by using the geometry of the experimental and theoretical models. HOMA is the geometry-based index of aromaticity. It is defined as follows:

$$\mathrm{HOMA} = 1 - \alpha/n \sum (d_{\mathrm{opt}} - d_{\mathrm{i}})^2 \qquad (1)$$

where n is the number of bonds taken into account, α is a normalization constant (fixed to give HOMA = 0 for a model nonaromatic system and HOMA = 1 for the system with all bonds equal to the optimal value; α equal to 257.7 and 93.52 for CC and CN bonds, respectively), d_{opt} stands for optimal bond length (1.388 and 1.334 Å for CC and CN bonds, respectively),[32] which is assumed to be realized when full delocalization of π-electrons occurs, and d_{i} are running bond lengths in the ring. The aromatic character of many systems with use of the HOMA model has recently been reported.[1,3] Typical values are as follows: HOMA for benzene,[33] naphthalene[33] and anthracene[33] is 0.991, 0.811, and 0.718, respectively, and for pyridine,[32c] pyrrole,[4] tiophene,[4] and furan[4] HOMA is 0.998, 0.876, 0.891, and 0.298, respectively. For porhyrin[43] HOMA = 0.666 whereas for its magnesium salt[43] and dianion[43] HOMA = 0.671 and 0.448, respectively. Another aromaticity index, NICS,[34] is a magnetic one that is purely theoretical, and can only be used for estimation of local aromatic character. It is defined as a negative value of the absolute magnetic shielding computed at ring centers. Negative value of NICS denote aromaticity, positive one denote antiaromaticity. The NICS was calculated by using the GIAO method at the HF/6-31+G* level of theory. The choice of the quantitative definitions of aromaticity was dictated by the general expectation[4b,35,36] that among many easily accessible quantitative definitions of aromaticity[6] these two models give the most accurate estimates of stabilization energy due to cyclic π-electron delocalization.[35]

Results and Discussion

Ab Initio Geometries. Both MP2/6-311+G** and B3LYP/6-311+G** based data lead to a similar picture. The diferences of relative energies between the related

(19) Flygare, W. H. *Chem. Rev.* **1974**, *74*, 653−687.
(20) Bothner-By, A. A.; Pople, J. A. *Rev. Phys. Chem.* **1965**, *16*, 43.
(21) London, F. *J. Phys. Radium* **1937**, *8*, 397.
(22) Kutzelnigg, W. *Isr. J. Chem.* **1980**, *19*, 193.
(23) Woliński, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.
(24) Dauben, H. J.; Wilson, J. D.; Laity, J. L. *J. Am. Chem. Soc.* **1968**, *90*, 811−813.
(25) Selwood, P. W. *Magnetochemistry*; Interscience: New York, 1956.
(26) Haberdietzl, H. *Angew. Chem.* **1966**, *78*, 277−288.
(27) (a) Breslow, R. *Angew. Chem.* **1968**, *80*, 573−578. (b) Lloyd, D.; Marshall, D. R. *Angew. Chem., Int. Ed. Engl.* **1972**, *11*, 404−408.
(28) (a) For an excellent review see: Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, *99*, 3247−3276 and papers cited therein. (b) Šponer, J.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 3161−3164. (c) Kabelac, M.; Kratochvil, M.; Šponer, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **2000**, *17*, 1077−1086.
(29) Box, V. G. S.; Jean-Mary, F. *J. Mol. Model.* **2001**, *7*, 334−342.
(30) (a) Allen, F. H.; Davies, J. E.; Galloy, J. J.; Johnson, O.; Kennard, O.; McRae, E. M.; Mitchell, G. F.; Smith, J. M.; Watson, D. G. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 187. (b) Allen, F. H.; Kennard, O. *Chem. Design Automation News* **1993**, *8*, 31−37.
(31) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheesman, J. R.; Keith, T.; Peterson, 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.; Gonzales, C.; Pople, J. A. *Gaussian 98*; Gaussian Inc.: Pittsburgh, PA, 1998.
(32) (a) Krygowski, T. M. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 70−78. (b) Krygowski, T. M.; Cyrański, M. *Tetrahedron* **1996**, *52*, 1713−1722. (c) Krygowski, T. M.; Cyrański, M. *Tetrahedron* **1996**, *52*, 10255−10264.
(33) Cyrański, M. K.; Stępień, B. T.; Krygowski, T. M. *Tetrahedron* **2000**, *56*, 9663−9667.
(34) Schleyer, P. v. R.; Maerker, C.; Dransfeld, A.; Jiao, H.; Hommes, N. J. R. v. E. *J. Am. Chem. Soc.* **1996**, *118*, 6317−6318.
(35) Schleyer, P. v. R. *Chem. Rev.* **2001**, *101*, 1115−1118.
(36) (a) Stępień, B. T.; Cyrański, M. K.; Krygowski, T. M. *Chem. Phys. Lett.* **2001**, *350*, 537−542. (b) Stępień, B. T.; Krygowski, T. M.; Cyrański, M. K. *J. Org. Chem.* **2002**, *67*, 5987−5992.

TABLE 1. Aromaticity Indices HOMA[32] and NICS[34] for the Whole Systems (total values) and for the Six-Membered (6) and Five-Membered (5) Fragments Based on B3LYP/6-311+G Optimized Geometries[a]**

				HOMA			NICS [NICS(1)]	
no.	base	rel energy	*N*	total	(6)	(5)	(6)	(5)
1	**adenine**	**0.0**	**0**	**0.917**	**0.976**	**0.860**	**−7.2 (−7.4)**	**−11.9 (−12.1)**
				(0.926)	**(0.965)**	**(0.884)**	**[−8.9 (−9.1)]**	**[−9.8 (−10.0)]**
2	adenine	+12.4	1	0.762	0.687	0.898	−1.4 (−1.5)	−12.8 (−13.0)
		(+12.5)		(0.781)	(0.705)	(0.916)	[−3.1 (−3.3)]	[−10.1 (−10.4)]
3	**guanine**	**0.0**	**1**	**0.745**	**0.682**	**0.871**	**−3.1 (−3.0)**	**−12.4 (−12.7)**
				(0.751)	**(0.664)**	**(0.903)**	**[−3.8 (−3.7)]**	**[−9.7 (−10.0)]**
4	guanine	+1.3	0	0.910	0.974	0.839	−6.9 (−7.0)	−11.1 (−11.5)
		(+0.1)		(0.920)	(0.960)	(0.871)	[−7.8 (−8.0)]	[−9.2 (−9.5)]
5	guanine	+13.4	2	0.665	0.572	0.841	−1.8 (−1.4)	−12.7 (−12.8)
		(+15.2)		(0.664)	(0.537)	(0.883)	[−1.5 (−1.1)]	[−9.4 (−9.5)]
6	**cytosine**	**0.0**	**1**	**0.672**	**0.672**		**−1.3 (−1.4)**	
				(0.698)	**(0.698)**		**[−3.2 (−3.4)]**	
7	cytosine	+2.2	2	0.506	0.506		−0.6 (−0.6)	
		(+1.8)		(0.531)	(0.531)		[−1.2 (−1.4)]	
8	**thymine**	**0.0**	**2**	**0.466**	**0.466**		**−1.5 (−1.5)**	
				(0.495)	**(0.495)**		**[−2.2 (−2.5)]**	
9	thymine	+13.1	1	0.703	0.703		−2.7 (−2.9)	
		(+11.4)		(0.724)	(0.724)		[−4.1 (−4.3)]	
10	thymine	+18.7	1	0.350	0.350		−1.7 (−1.7)	
		(+16.2)		(0.438)	(0.438)		[−3.6 (−3.7)]	
11	**uracyl**	**0.0**	**2**	**0.503**	**0.503**		**−1.2 (−1.2)**	
				(0.516)	**(0.516)**		**[−1.9 (−2.2)]**	
12	uracyl	+12.2	1	0.715	0.715		−2.1 (−2.2)	
		(+10.9)		(0.727)	(0.727)		[−3.8 (−3.8)]	
13	uracyl	+19.4	1	0.411	0.411		−1.1 (−1.1)	
		(+17.1)		(0.465)	(0.465)		[−3.1 (−3.2)]	

[a] NICS(1) stands for the NICS value calculated 1 Å above the system plane [in square brackets]. *N* is the number of C=X (X = O, N) groups attached to the ring. Relative energies are given in kcal/mol. The values based on the MP2/6-311+G** level are given in parentheses.

isomers at these two levels are usually very small (ca. 1.5 kcal/mol). Also, the NICS and NICS(1) values usually do not deviate more than 0.2 ppm. The largest discrepancies can be found in the molecular geometries. However, still the analysis based on the geometry-based descriptor of the extent of cyclic π-electron delocalization (HOMA) is consistent in both cases. The most obvious conclusion from a survey of the results based on ab initio geometries is that the aromatic character of the nucleobases as complete molecules varies in a wide range (0.466−0.917 units of HOMA, based on the B3LYP/6-311+G** level), as shown in Table 1. The MP2/6-311+G**-based data show similar range: 0.495−0.926. The data in bold in Table 1 are for the most stable tautomers. The energies given for all other tautomers correspond to the lowering of stability relative to the most stable tautomer. Chart 1 presents the labeling system and the bond lengths (Å) for all optimized nucleobases, as well as for some tautomers. The ring fragments do not deviate from planarity in a way to affect signifficantly the extent of cyclic π-electron delocalization.[37]

An inspection of the most stable tautomers (and those very close energetically, within 2 kcal/mol) indicates that aromaticity decreases in a regular way with the increase of the number of C=X (X = O, N) groups that are attached to the ring. Figure 1 presents the HOMA values for the whole molecules plotted against the number of C=O/C=N groups, *n*, bonded to the ring system of the base. The scatter graph contains all 7 data points for the most stable tautomers in Table 1 (bold entries). Numbers in parentheses correspond to entry numbers in Table 1.

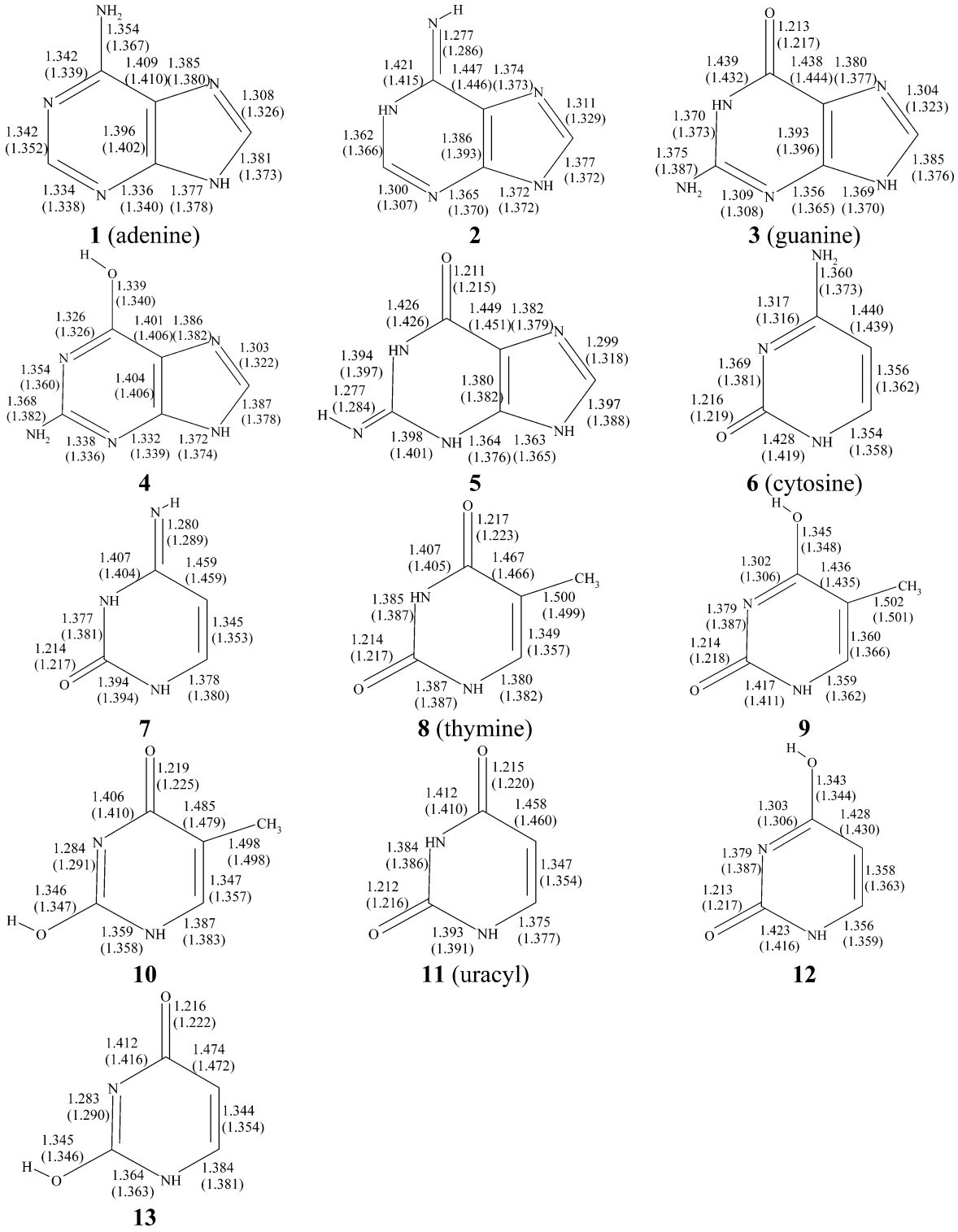
(37) Bodwell, G. J.; Bridson, J. N.; Cyrański, M. K.; Kennedy, J. W. J.; Krygowski, T. M.; Mannion, M. R.; Miller, D. O. *J. Org. Chem.*, **2003**, *68*, 2089−2098.

It is worth noting that HOMA values for MP2 geometry differ only sligthly but always in a direction to increase π-electron delocalization. For most stable conformers (**1** and **4**) the differences in HOMA estimated form B3LYP and MP2 are around 0.01. With a decrease of aromaticity this difference increases and for **7**, **11**, and **8** it is 0.024.

The local aromatic character changes in an even wider range, from HOMA = 0.976 (six-membered ring of adenine) to 0.466 (six-membered ring of thymine), and follows the same pattern as that for the whole molecules. The NICS values estimated for the individual rings are usually determined in two ways: in the center of the ring (denoted NICS) or one angstrom above the center (denoted as NICS(1)). Also these magnetic parameters exhibit a considerable spread, from NICS(1) = −10.1 for the five-membered ring of the unstable (ΔE = 12.4 kcal/mol) tautomer of adenine (**2**) to NICS(1) = −1.2 for the six-membered ring of the slightly less stable (ΔE = 2.2 kcal/mol) tautomer of cytosine (**7**). The NICS values estimated in the plane of the molecules are still more negative, by ca. 1−2 units. The more negative values of NICS for the five-membered rings are in line with a general observation for this index that the magnitude of NICS depends on the size of the ring.[1,34]

The correlation between the aromatic character, measured by HOMA, and the number of exocyclic double bonds (C=O and C=N) attached to the entire ring system of the nucleobase is also observed in the individual rings for which the NICS and HOMA parameters have been calculated (Table 1). As in the case of the whole molecules, the deviations from the linear trend for HOMA are observed only for the unstable tautomers of guanine (ΔE = 3.4 kcal/mol), thymine (ΔE = 18.7 kcal/mol), and uracyl (ΔE = 19.4 kcal/mol). In the case of NICS, the

CHART 1. Molecular Geometries of Adenine, Guanine, Cytosine, Thymine, and Uracyl Optimized atthe B3LYP/6-311+G Level of Theory[a]**



[a] The data optimized at MP2?6-311+G** level are given in parentheses.

trend is preserved but is blurred with larger scatter, and the deviations are different. Nevertheless, the two numerical characteristics of aromaticity are consistent, indicating dependence on the number of exocyclic C═O(N) bonds.

The variability of the aromatic character of the rings in the bases in question may be interpreted in terms of a dearomatizing effect of the exocyclic substituents (or of any other chemical links) that are of a double-bond character. A similar effect was used by Beak et al.[38,39]

(38) Beak, P.; Bonham, J.; Lee, J. T., Jr. *J. Am. Chem. Soc.* **1968**, *90*, 1569−1582.

(39) Beak, P.; Lee, J. T., Jr. *J. Org. Chem.* **1969**, *34*, 2125−2128.

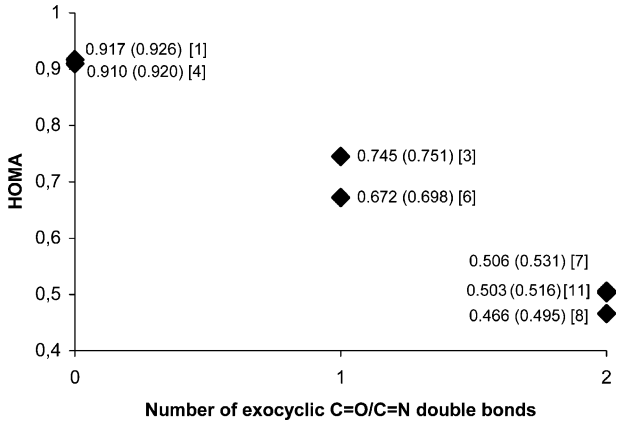


FIGURE 1. The dependence of the global aromaticity index HOMA on the number of C=O and C=N bonds attached to the ring systems. The values based on the MP2/6-311+G** level are given in parentheses.

and by others[40−42] for approximate estimation of the resonance energy from the difference in energy between the aromatic 2-methoxypyridine and nonaromatic *N*-methyl-2-pyridone. Recently this effect has been described in porphyrine derivatives[43] where the partially double CC bonds (1.394−1.400 Å) of the bridges between the pyrrole rings cause a substantial decrease of their aromatic character, as measured by the HOMA values ranging from 0.450 to 0.660, depending on the protonation state of the pyrrole nitrogen atoms. These values could be compared with those for the less perturbed pyrrole derivatives (for instance, single-bonded substituents at 2 and/or 5 positions) characterized by HOMA of 0.907−0.932. A similar observation was also made for pyrazole derivatives.[44]

The analysis of aromaticity of the bases involved in Watson−Crick pairs (Table 2) leads to a general conclusion that in all cases when the bases contain one (cytosine, guanine) or two (uracyl, thymine) carbonyl groups, their aromatic character increases as a consequence of H-bond formation. The double C=O bonds become longer resulting in higher aromatic character of the ring to which the C=O group is attached. The increase is substantial and ranges from 0.03 to 0.13 units of HOMA. The largest change is observed for guanine, both considered as a whole moiety and taking into account only the six-membered ring. In the case of adenine forming pairs with thymine or uracyl, its aromatic character is practically unchanged. It may be argued that any Lewis acid capable of similar interaction as the C=O groups would lead to analogous consequences. NICS (and also NICS(1)) exhibits a smaller trend for the CG pair and a rather opposite trend for AT and AU pairs.

TABLE 2. Aromaticity Indices HOMA[32] and NICS[34] for Watson−Crick Paired Nucleobases (total values) and for the Six-Membered (6) and Five-Membered (5) Fragments[a]

			HOMA			NICS [NICS(1)]	
no.	pair	base	total	(6)	(5)	(6)	(5)
14	AT	adenine	0.914	0.966	0.867	−6.5 [−8.2]	−11.8 [−9.7]
14	AT	thymine	0.533	0.533		−1.5 [−2.3]	
15	AU	adenine	0.914	0.965	0.867	−6.5 [−8.2]	−11.8 [−9.7]
15	AU	uracyl	0.566	0.566		−1.2 [−2.0]	
16	CG	guanine	0.815	0.810	0.853	−3.5 [−4.1]	−11.7 [−9.2]
16	CG	cytosine	0.715	0.715		−1.5 [−3.1]	

[a] NICS(1) stands for the NICS value calculated 1 Å above the system plane (in square brackets).

One of the great adventages of the HOMA model (and also other geometry-based models) is that it also can be applied to quantify the π-electron delocalization within selected fragments of the bases. For this purpose we have extracted from the ab initio optimized geometries those fragments, shown in Chart 3, that are involved in Watson−Crick base pairing. For those fragments the HOMA index was calculated, both in Watson−Crick paired states and in isolated forms. In this way a much more clear picture of the role of base pairing via H-bond formation is revealed. Chart 3 defines the fragments used in the analysis, whereas the numerical data are presented in Table 3.

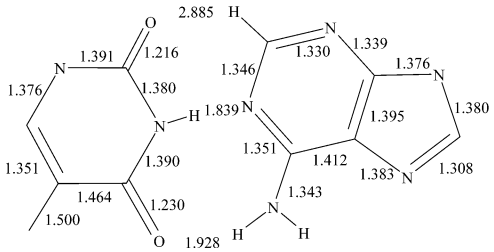
The above results clearly demonstrate that, with the exception of adenine, in all other cases base pairing markedly increases the delocalization (as shown by the Δ values), which is an important part of the concept of aromaticity of π-electron systems. The larger changes of delocalization within base fragments than within the whole molecules reflect the "resistance" of the other parts of the π-electron systems of the nucleobases.

Experimental Geometries. The experimentally determined geometries of the five nucleobases have been retrieved in a search of the Cambridge Structural Database.[30] The search excluded metal complexes and protonated forms of the nucleobases and was limited to error-free, disorder-free, and good-quality crystal structures. The number of hits was as follows: adenine, 40; uracyl, 35; cytosine, 32; thymine, 22; and guanine, 6. With use of the experimental bond distances, the HOMA values were calculated for each case individually. Next, the HOMA parameters for each nucleobase were determined as mean values of the individual numbers, and their standard deviations were calculated as well. A summary of the results, together with a comparison with the theoretically derived parameters, is presented in Table 4.

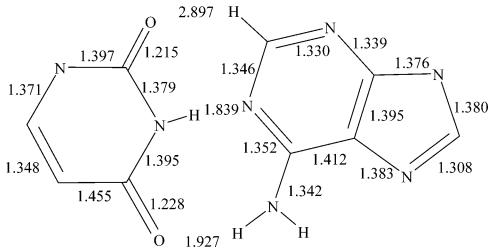
Except for thymine and uracyl, the agreement between the experimental and theoretically estimated HOMA values is quite good, but the experimental values are always higher (except the case of cytosine) and the agreement is better for the MP2/6-311+G** optimized systems. The guanine sample is too small for reliable statistical conclusions. The differences are most signifi-

(40) Cook, M. J.; Katritzky, A. R.; Linda, P.; Tack, R. D. *Chem. Commun.* **1971**, 510−511.
(41) Cook, M. J.; Katritzky, A. R.; Linda, P.; Tack, R. D. *J. Chem. Soc., Perkin Trans. 2* **1972**, 1295−1301.
(42) Cook, M. J.; Katritzky, A. R.; Linda, P.; Tack, R. D. *J. Chem. Soc., Perkin Trans. 2* **1973**, 1080−1086.
(43) Cyrański, M. K.; Krygowski, T. M.; Wisiorowski, M.; Hommes, N. J. R. v. E.; Schleyer, P. v. R. *Angew. Chem., Int. Ed.* **1998**, *37*, 177−180.
(44) Krygowski, T. M.; Anulewicz, R.; Cyrański, M. K.; Puchała, A.; Rasała, D. *Tetrahedron* **1998**, *54*, 12295−12300.

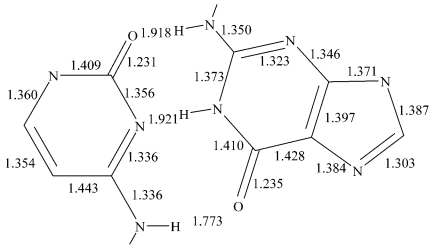
CHART 2. Molecular Geometries of the Nucleobases Involved in Watson−Crick Pairs Optimized at the B3LYP/6-311+G Level of Theory.**



14 (TA)

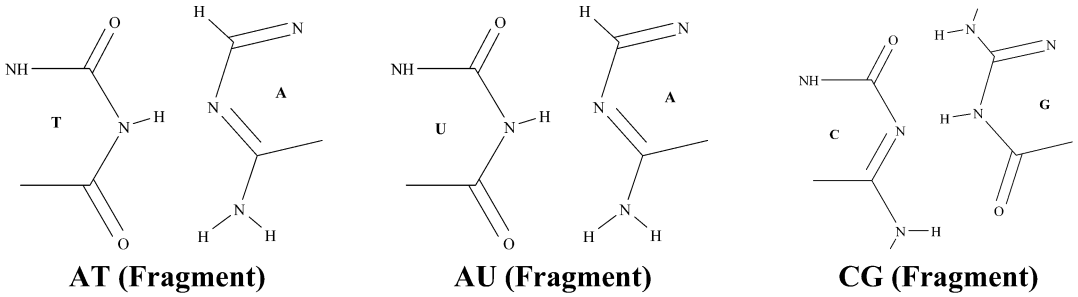


15 (UA)



16 (CG)

CHART 3



AT (Fragment) **AU (Fragment)** **CG (Fragment)**

TABLE 3. HOMA Values for π-Electron Fragments (defined in Chart 3) of Isolated (iso) and H-bond-Paired (pair) Nucleobases

pair	fragment	HOMA (iso)	HOMA (pair)	Δ
AT	A	0.969	0.952	−0.017
AT	T	0.347	0.429	0.082
AU	A	0.969	0.952	−0.017
AU	U	0.402	0.484	0.082
CG	G	0.536	0.723	0.187
CG	C	0.584	0.662	0.078

TABLE 4. HOMA Values from Experimental and Theoretically Optimized Molecular Geometries[a]

base (*n*)	$HOMA_{exp}$	SD	$HOMA_{theor}$	Δ
adenine (40)	0.966	0.016	0.917 (0.926)	0.049 (0.040)
guanine (6)	0.860	0.010	0.745 (0.751)	0.115 (0.109)
cytosine (33)	0.640	0.070	0.672 (0.698)	−0.032 (−0.058)
thymine (22)	0.631	0.050	0.466 (0.495)	0.165 (0.136)
uracyl (35)	0.670	0.050	0.503 (0.516)	0.167 (0.154)

[a] The numbers of experimental entries (*n*) are given in parentheses. The $HOMA_{exp}$ values are mean values calculated for each set of *n* structures. SD is the standard deviation of $HOMA_{exp}$. $HOMA_{theor}$ is calculated for B3LYP/6-311+G** optimized geometry (and for MP2/6-311+G** in parentheses). $\Delta = HOMA_{exp} - HOMA_{theor}$.

cant for those two bases which have two double-bond substituents at the ring. Both these substituents decrease delocalization[38−42] and in consequence the HOMA values. The differences may have their origin in the fact that the experimental structures retrieved from the CSD interact with each other by intermolecular interactions in the crystal lattice, and usually they are fragments of larger molecular systems, or in other words, the nucleobases have additional substituents. These substituents can be expected to significantly affect the aromaticity of the weakly aromatic thymine and uracyl moieties. It is a known fact that in nonaromatic or antiaromatic systems, the substituent effect may lead to a substantial increase of the aromatic character. It was found for example to increase by almost 1 unit of HOMA for fulvene or by 0.6 for heptafulvene derivatives.[36] The above differences and the large values of the standard deviations for the experimental HOMA values indicate that aromaticity of the less aromatic pyrimidine bases (cytosine, thymine, and uracyl) is sensitive to perturbations such as substituent effect. This is in line with the observation that these systems are also sensitive to much weaker perturbations, such as those caused by hydrogen bonding.

Conclusions

On the basis of the geometry model of aromaticity HOMA and magnetic index NICS we have analyzed the extent of the cyclic π-electron delocalization that operates in the nucleobases constituting DNA and RNA. The bases differ significantly in their aromatic character, as shown by the HOMA ranges: from 0.466 for thymine to 0.917 for adenine (at the B3LYP/6-311+G** level). Clearly the

main source of these differences is the number of double-bond C═X (X = N, O) substituents attached to the six-membered rings. Most importantly, this factor affects the cyclic π-electron delocalization of the bases when they are paired in the Watson−Crick pairs. The H-bonds involving C═O groups in the pairs cause an increase of the cyclic π-electron delocalization, which is clearly observed by an increased aromatic character of the rings.

Acknowledgment. M.K.C. gratefully acknowledges 120-501/68-BW-1522/19/2001 for financial support. He also gratefully acknowledges the Interdisciplinary Centre for Mathematical and Computational Modelling (Warsaw, Poland) for computational facilities. M.K.C. also acknowledges a stipend from the Polityka and PKN Orlen S.A. (2001/2002).

Supporting Information Available: Absolute electronic energies, optimized molecular geometries, and zero-point vibrational energies (ZPE) at B3LYP/6-311+G** and MP2/6-311+G**. This material is available free of charge via the Internet at http://pubs.acs.org.

JO034760E