Molecular Structure of 9*H*-Adenine Tautomer: Gas-Phase Electron Diffraction and Quantum-Chemical Studies

Natalja Vogt,*,† Olga V. Dorofeeva,†,‡ Victor A. Sipachev,‡ and Anatolii N. Rykov‡

Chemieinformationssysteme, Universität Ulm, 89069 Ulm, Germany, and Department of Chemistry, Moscow State University, Moscow 119992, Russia

Received: June 18, 2009; Revised Manuscript Received: September 22, 2009

Molecular geometry of 9*H*-adenine tautomer was calculated by MP2 method using several basis sets (up to cc-pVQZ). According to the results of all quantum-chemical calculations, the molecule has an essentially planar heavy-atom skeleton and a quasi-planar amino group. Since the bond lengths of adenine are of similar magnitude, the structural problem could not be solved by the gas-phase electron diffraction (GED) method alone. Therefore the differences between similar bond lengths derived from ab initio geometry and rotational constants from microwave (MW) spectroscopic study (Brown, R. D.; et al. *Chem. Phys. Lett.* **1989**, *156*, 61) were used as supplementary data. To bring the data of the different experimental methods to the same basis (equilibrium structure), GED internuclear distances r_a and MW rotational constants $B_0^{(i)}$ (i = A, B, C) were corrected for vibrational effects. Harmonic and anharmonic corrections were estimated using quadratic and cubic force constants from MP2/cc-pVTZ calculations. Anharmonic corrections to r_a distances were calculated using improved theoretical approximation.

The molecular structure of 9*H*-adenine is determined experimentally for the first time. Since the GED intensities are not sensitive to hydrogen positions, and small deviations of skeleton cannot be determined with appropriate uncertainty, the molecular configuration of adenine was asumed to be planar (C_s symmetry) in the GED analysis. The main equilibrium structural parameters determined from GED data supplemented by rotational constants and results of MP2/cc-pVTZ calculations are the following (bond lengths in angstroms and bond angles in degrees with 3σ in parentheses): r_e(C2−N1) = 1.344(3), r_e(C2−N3) = 1.330(3), r_e(C4−N3) = 1.333(3), r_e(C4−C5) = 1.401(3), r_e(C5−C6) = 1.409(3), r_e(C6−N1) = 1.332(3), r_e(C5−N7) = 1.380(4), r_e(C8−N7) = 1.319(3), r_e(C8−N9) = 1.371(4), r_e(C4−N9) = 1.377(4), r_e(C6−N10) = 1.357(4), $\angle_e$(N1−C2−N3) = 129.0(1), $\angle_e$(C2−N3−C4) = 111.0(1), $\angle_e$(N3−C4−C5) = 127.2(1), $\angle_e$(C4−C5−N7) = 111.9(2), $\angle_e$(C5−N7−C8) = 103.4(2), and $\angle_e$(C5−C6−N10) = 121.9(2). The determined experimental bond lengths of adenine are in good agreement with those from MP2 calculations and with experimental bond lengths of pyrimidine and 1*H*-imidazole (except for the C−C double bond in imidazole). Being close to typical aromatic internuclear distances, the obtained C−C and C−N bond lengths indicate the aromatic nature of this molecule. The calculated aromaticity indexes (GIAO-MP2/cc-pVTZ) confirm this statement.

Introduction

Being responsible for transfer and storage of genetic information, nucleic acid (NA) bases (adenine, thymine, guanine, cytosine, and uracil) are subjects of particular scientific interest. The most stable canonical tautomers are dominantly present in DNA and RNA, whereas the rare tautomeric forms may be responsible for spontaneous mutations.[1,2]

The molecular structure and tautomerism of NA bases were topics of numerous quantum-chemical studies at various levels of theory (see review in ref 3). On the contrary, the experimental data are relatively rare and incomplete, although they are necessary to estimate the quality of the theoretical results.

The experimental geometries of isolated NA bases have been determined by gas-phase electron diffraction (GED) and microwave spectroscopy (MW) methods for uracil,[4−6] thymine,[7,8] and guanine.[6] After recent success with thymine[8] we decided to investigate adenine, which is even more problematic for GED analysis because it has more similar bond lengths compared to thymine and uracil. It should be noted that quadrupole splitting by five nitrogen atoms makes the structural study of adenine by the MW method extremely difficult.

Quantum-chemical calculations of the structural parameters of adenine tautomers have been carried out at different levels of theory using single-confirgurational and multiconfigurational and coupled-cluster approximations (HF/6-31G,[9] B3LYP/cc-pVDZ,[10] BP86/QZ4P,[11] RI-MP2/TZVPP,[12] MP2/aug-cc-pVDZ,[13] CASSCF(4/4)/6-31G,[9] MCSCF(12,10)/DZP,[14] and CCSD/aug-cc-pVDZ,[13] etc.[15]). Nevertheless, since the differences between the calculated geometries (up to 0.03 Å, 1°, and 8° for bond lengths, valence angles, and dihedral angles, respectively[13]) are larger than the typical GED uncertainties, experimental determination of the structure and high-level quantum-chemical calculations with more extended basis sets are desirable.

Tautomerism of adenine has been studied theoretically in different approximations. The BP86/QZ4P,[11] MP2/6-311++G(d,p),[16] RI-MP2/TZVPP,[12] and MCQDPT2/MCSCF[14] calculations showed that the canonical form (9H; see Figure 1) corresponds to the global minimum of the potential energy surface, while the 7H and 3H tautomers (and perhaps some

* To whom correspondence should be addressed.
† Universität Ulm.
‡ Moscow State University.

10.1021/jp905755u CCC: $40.75
Published on Web 10/29/2009

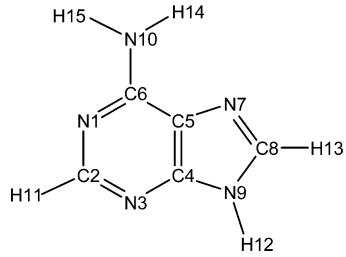

Figure 1. Structure of 9*H*-adenine tautomer with atom numbering.

others[11]) corresponding to the lowest local minima are considerably less stable (both by about 7−8 kcal/mol) being separated from the 9H tautomer by very high barriers (63 kcal/mol for 9H → 3H transformation[14]). In comparison with the gas phase, in the aqueous environment, the coexistence of the 3H and 7H tautomers with the dominant canonical form was predicted.[12]

The experimental data concerning tautomerization are ambiguous. Only the canonical tautomer was observed in the matrix-isolated[17,18] and gas-phase spectra,[19] and it was found to be highly dominant in the supersonic beam MW experiment,[20] whereas the coexistence of the less favorable 7H tautomer was proved in the matrix-isolated[21] and gas-phase (at 280 °C)[22] IR studies. The high-energy tautomers could be induced for any reason (for example, by the presence of water, metal cations,[23] etc.).

In the present study, the GED analysis has been carried out for the canonical 9H tautomer. The 7H or other tautomers could not be distinguished by the GED method because they differ only by hydrogen positions. However, these tautomers are not expected to be present in noticeable amounts under our experimental conditions.

Being the main method for the structure determination of free molecules,[24] GED cannot solve the structural problem completely if the bond lengths in the molecule are of similar magnitude. In this case, a support of GED analysis by data from other methods (MW spectroscopy and quantum chemistry) is mandatory. High-level quantum-chemical calculations (up to MP2/cc-pVQZ) have been carried out in order to derive structural differences for close bond lengths and to estimate the harmonic and anharmonic vibrational corrections to thermal-average experimental internuclear distances and the root-mean-square (rms) vibrational amplitudes.

Quantum-Chemical Calculations. Quantum-chemical calculations were carried out by means of the GAUSSIAN03 program package[25] running under Linux operating system on a cluster of quad-core Intel Xeon processors (University of Ulm, Group of Chemical Information Systems). The molecular structure of 9*H*-adenine tautomer was optimized at the B3LYP (cc-pVTZ and aug-cc-pVTZ basis sets) and MP2 (cc-pVTZ, aug-cc-pVTZ, and cc-pVQZ basis sets) levels of theory. The optimized geometries are presented in Table 1. The equilibrium conformations were confirmed by frequency calculations (see Table 2). The anharmonic (cubic) force constants were calculated in the B3LYP/cc-pVTZ and MP2/cc-pVTZ approximations. For very time-consuming numerical calculations of these constants the new version of the Gaussian program G03, Revision E.01, which allowed parallelization of such calculations, was required.

GED Experiment. The adenine sample was purchased from Sigma-Aldrich Co. (purity >99%) and used without further purification. The presence of impurities and decomposition products of the sample in the gas phase in the region of the sublimation temperature 220(20) °C was checked by mass-spectrometry (MS) using chemical ionization (CI) method

TABLE 1: Equilibrium Structure of 9*H*-Adenine Tautomer Calculated by Quantum-Chemical Methods (Bond Lengths in Angstroms; Angles in Degrees)

	B3LYP		MP2		
parameter	cc-pVTZ	aug-cc-pVTZ	cc-pVTZ	aug-cc-pVTZ	cc-pVQZ
C2−N1	1.339	1.338	1.348	1.348	1.345
C2−N3	1.331	1.331	1.334	1.334	1.331
C4−N3	1.334	1.333	1.337	1.337	1.334
C4−C5	1.393	1.393	1.396	1.397	1.394
C5−C6	1.406	1.406	1.405	1.405	1.403
C6−N1	1.339	1.339	1.335	1.336	1.333
C5−N7	1.382	1.382	1.376	1.376	1.374
C8−N7	1.305	1.305	1.322	1.323	1.320
C8−N9	1.376	1.376	1.368	1.368	1.365
C4−N9	1.374	1.374	1.374	1.374	1.372
C6−N10	1.351	1.350	1.358	1.357	1.354
C2−H11	1.084	1.083	1.082	1.083	1.081
N9−H12	1.005	1.006	1.006	1.007	1.006
C8−H13	1.078	1.078	1.077	1.078	1.077
N10−H14	1.004	1.003	1.005	1.006	1.004
N10−H15	1.004	1.004	1.005	1.006	1.004
N1−C2−N3	128.6	128.4	128.8	128.6	128.6
C2−N3−C4	111.4	111.5	110.9	111.0	111.1
N3−C4−C5	126.7	126.6	127.0	127.0	126.9
C4−C5−N7	111.2	111.2	111.9	111.8	111.8
C5−N7−C8	104.2	104.3	103.3	103.5	103.5
C5−C6−N10	122.4	122.4	121.8	121.9	121.9
N3−C2−H11	116.1	116.1	116.0	116.1	116.1
C4−N9−H12	125.9	125.9	125.9	125.9	125.9
N7−C8−H13	125.2	125.2	124.8	124.8	124.8
C6−N10−H14	119.6	120.3	117.1	118.0	117.8
C6−N10−H15	118.5	119.1	116.2	117.1	116.9
C4−C5−C6-N10	−178.8	−179.5	−177.4	−177.8	−177.7
C5−C6−N10-H14	−8.9	−3.2	−18.0	−15.3	−15.8
N1−C6−N10-H15	8.4	2.8	17.1	14.8	15.2

TABLE 2: Comparison of Experimental and Theoretical Frequencies (cm^{-1}) of 9*H*-Adenine[a]

mode	B3LYP/aug-cc-pVTZ (I)	MP2/aug-cc-pVTZ (II)	MP2/cc- pVTZ (III)	MP2/cc-pVTZ, transition state	expt IR MI (Ar) [b]	expt IR gas [c]	I [d]	II[d]	III[d]
39	3736 (59)	3743 (64)	3749 (57)	3791 (72)	3565 (84)	3552	4.80	4.99	5.16
38	3646 (83)	3658 (109)	3672 (110)	3672 (108)	3506 (135)	3501	3.99	4.34	4.73
37	3604 (95)	3603 (90)	3610 (89)	3641 (129)	3448 (110)	3434	4.52	4.50	4.70
36	3238 (0.1)	3277 (0.2)	3284 (0.1)	3284 (0.1)	3057 (6)	3061	5.92	7.20	7.43
35	3164 (18)	3214 (12)	3219 (17)	3218 (17)	3041 (3)		4.04	5.69	5.85
34	1655 (612)	1680 (486)	1684 (453)	1688 (542)	1633 (447)	1625	1.35	2.88	3.12
33	1631 (111)	1644 (100)	1651 (92)	1649 (94)	1612 (219)		1.18	1.99	2.42
32	1609 (19)	1613 (2)	1613 (7)	1611 (13)	1599 (49)		0.63	0.88	0.88
31	1514 (4)	1514 (47)	1520 (39)	1521 (32)	1482 (11)		2.16	2.16	2.56
30	1502 (89)	1488 (33)	1497 (31)	1496 (32)	1474 (71)	1468	1.90	0.95	1.56
29	1434 (20)	1453 (86)	1459 (73)	1461 (67)	1419 (49)	1415	1.06	2.40	2.82
28	1413 (14)	1430 (44)	1436 (44)	1437 (47)	1389 (45)		1.73	2.95	3.38
27	1365 (25)	1410 (28)	1414 (29)	1413 (32)	1345 (21)	1346	1.49	4.83	5.13
26	1355 (42)	1364 (18)	1368 (15)	1371 (13)	1328 (40)	1326	2.03	2.71	3.01
25	1323 (66)	1345 (43)	1349 (48)	1349 (47)	1290 (68)	1280	2.56	4.26	4.57
24	1268 (28)	1258 (26)	1263 (29)	1262 (29)	1240 (28)		2.26	1.45	1.85
23	1242 (15)	1248 (11)	1255 (10)	1247 (11)	1229 (13)	1234	1.06	1.55	2.12
22	1144 (23)	1146 (15)	1149 (14)	1149 (17)	1127 (6)	1126	1.51	1.69	1.95
21	1079 (19)	1096 (16)	1099 (13)	1099 (14)	1061 (13)	1065	1.70	3.30	3.58
20	1009 (6)	1025 (11)	1032 (11)	1009 (7)	1005 (9)	1018	0.40	1.99	2.69
19	985 (3)	966 (5)	969 (5)	970 (5)	958 (3)	957	2.82	0.84	1.15
18	945 (16)	932 (11)	935 (10)	936 (11)	927 (13)	926	1.94	0.54	0.86
17	899 (12)	893 (11)	896 (11)	895 (15)	887 (8)		1.35	0.68	1.01
16	864 (6)	855 (10)	856 (10)	853 (10)	848 (6)	847	1.89	0.83	0.94
15	829 (11)	813 (5)	812 (5)	810 (4)	802 (9)	801	3.37	1.37	1.25
14	726 (3)	723 (2)	725 (3)	725 (2)	717 (5)		1.26	0.84	1.12
13	696 (0.4)	681 (2)	682 (4)	682 (1)	678 (2)		2.65	0.44	0.59
12	673 (7)	670 (3)	672 (4)	672 (3)	655 (6)	650	2.75	2.29	2.60
11	618 (1)	614 (1)	616 (1)	614 (2)	610 (5)	600	1.31	0.66	0.98
10	583 (53)	591 (60)	592 (44)	591 (50)	566 (46)	563	3.00	4.42	4.59
9	543 (4)	544 (57)	535 (70)	541 (25)					
8	531 (4)	529 (10)	531 (15)	528 (40)					
7	530 (58)	521 (10)	522 (12)	528 (3)	513 (92)	515	3.31	1.56	1.75
6	520 (4)	498 (2)	490 (4)	514 (5)	503 (4)	506	3.38	0.99	2.58
5	302 (1)	300 (9)	302 (5)	302 (1)					
4	277 (11)	270 (14)	270 (12)	270 (11)	276 (12)	270	0.36	2.17	2.17
3	66 (167)	336 (241)	391 (285)	−283 (182)	242 (66)	244			
2	221 (0.3)	217 (13)	219 (9)	220 (2)	214 (15)	208	3.27	1.40	2.34
1	169 (29)	166 (6)	167 (7)	168 (15)		162			
						av	*2.32*	*2.40*	*2.75*

[a] Intensities in kilometers per mole are given in parentheses. [b] Matrix isolation infrared spectra (Ar).[17] [c] Infrared spectra in the gas phase.[19] [d] Deviations of B3LYP/aug-cc-pVTZ (I), MP2/aug-cc-pVTZ (II), or MP2/cc-pVTZ (III) theoretical frequencies from experimental values[17] in percent.

on a Finnigan SSQ-7000 mass spectrometer. The analysis showed (see mass spectrum in the Supporting Information, Figure S1) that the evaporated adenine was the main component and that the amount of volatile impurities did not exceed ~1%. Such an amount of impurities and/or decomposition products is not significant for GED analysis. The decomposition products were not detected also by a gas chromatography−mass-spectrometric method at 220 °C during 1 h. The GED experiment was carried out at the Moscow State University on the EG-100 M apparatus using the new r^3 sector from brass produced with high precision in the workshop of the University of Ulm. The experimental conditions were the following: accelerating voltage of ca. 57 kV, electron beam current of 2.4 μA, and vacuum of 2×10^{-5} mm Hg. The sample was evaporated from a stainless steel cell by heating to 211(15) °C to produce a vapor pressure sufficient for measurements. The diffraction patterns were recorded on photographic films (MACO EM-FILM EMS) at long (LD) and short (SD) nozzle-to-film distances of 362.16 and 193.82 mm, respectively. The wavelength of electron beam λ was calibrated using CCl_4 standard with structural parameters taken from ref 26. The stability of λ between two experiments was less than 0.1%. The diffraction patterns were scanned at the University of Ulm using an EPSON PERFECTION V750 PRO commercial scanner in the 16 bit/300 dpi scanning mode. The calibration of the scanner was carried out using a gray scale for Kodak ektachrome films. Data of scanning were transformed into intensity curves $I(s)$ using the UNEX program by the method described elsewhere.[27] The intensity curves for each nozzle-to-film distance could be very well adjusted to each other by shift and scaling; i.e., they were found to be compatible. The averaged experimental intensities $I(s)$ were obtained over the ranges $s = 3.75-18.0$ Å^{-1} (LD) and $s = 8.25-31.5$ Å^{-1} (SD) in steps of 0.125 Å^{-1}. They are presented together with background lines in Figures S2 and S3 and Table S1 of the Supporting Information.

Structural Analysis of the GED Data with Use of Rotational Constants and the Results of ab Initio Calculations. To carry out a joint analysis of data obtained by different experimental methods and quantum-chemical calculations correctly, it is necessary to bring them to the same basis. Therefore the GED thermal-average bond distances $r_{ij,a}$ and MW rotational constants[20] for the ground vibrational state $B_0^{(i)}$ (i = A, B, C) were converted to the parameters of the

TABLE 3: Total Vibrational Corrections $\Delta(r_{ij,e} - r_{ij,a})$ to the Experimental Internuclear Distances $r_{ij,a}$ (Å) Calculated in Different Approximations Using MP2/cc-pVTZ and B3LYP/cc-pVTZ Force Constants (Quadratic and Cubic)

bond	I,[a] MP2	II,[b] B3LYP	II,[b] MP2	III,[b,c] MP2	II,[b] MP2 scaled [d]
C2−N1	−0.0076	−0.0046	−0.0069	−0.0081	−0.0070
C2−N3	−0.0072	−0.0062	−0.0068	−0.0075	−0.0069
C4−N3	−0.0066	−0.0089	−0.0068	−0.0061	−0.0071
C4−C5	−0.0013	−0.0032	−0.0019	−0.0022	−0.0014
C5−C6	−0.0002	−0.0132	−0.0003	0.0002	−0.0004
C6−N1	−0.0042	−0.0111	−0.0035	−0.0041	−0.0038
C5−N7	−0.0054	−0.0049	−0.0033	−0.0073	−0.0025
C8−N7	−0.0049	−0.0049	−0.0043	−0.0053	−0.0043
C8−N9	−0.0103	−0.0128	−0.0122	−0.0090	−0.0122
C4−N9	−0.0096	−0.0124	−0.0115	−0.0082	−0.0117
C6−N10	0.0013	0.0263	−0.0017	−0.0017	0.0038
C2−H11	−0.0159	−0.0166	−0.0158	−0.0158	−0.0156
N9−H12	−0.0150	−0.0150	−0.0150	−0.0150	−0.0149
C8−H13	−0.0165	−0.0156	−0.0165	−0.0165	−0.0174
N10−H14	−0.0439	−0.0057	−0.0451	−0.0451	−0.0454
N10−H15	−0.0453	−0.0063	−0.0465	−0.0465	−0.0469

[a] Previously used (old) approximation (see, for example, refs 8 and 36). [b] Improved approximation (see text and eq 1). [c] Using the Gauss constrained motion theorem (see text). [d] MP2/cc-pVTZ force field was scaled so that experimental frequencies ν_3, ν_{10}, ν_{21}, ν_{25}, $\nu_{27} - \nu_{29}$, ν_{31}, $\nu_{33} - \nu_{39}$ (see Table 2)[17] were reproduced exactly (see text).

TABLE 4: Experimental Rotational Constants (MW) in Megahertz and Those from GED and MP2 Geometries of 9*H*-Adenine Tautomer

$B^{(i)}$	$B_0^{(i)}$(MW) [a,b]	$(B_e^{(i)} - B_0^{(i)})$ [c]	$B_e^{(i)}$(MW)	$B_e^{(i)}$(GED) [d]	$B_e^{(i)}$(MP2/cc-pVTZ) [d]
A	2371.8730(40)	15.4229	2387.2959	2386.31 (0.04%)	2373.1399 (0.6%)
B	1573.3565(8)	1.0226	1574.3791	1574.02 (0.02%)	1582.4657 (0.5%)
C	946.2576(4)	2.5482	948.8058	948.43 (0.04%)	949.7835 (0.1%)

[a] Uncertainties are given in parentheses. [b] Ref 20. [c] Anharmonic corrections calculated from MP2/cc-pVTZ cubic force field. [d] Deviations from $B_e^{(i)}$(MW) are given in parentheses.

TABLE 5: Molecular Structure of 9*H*-Adenine Tautomer Obtained from GED Data Supplemented by Experimental Rotational Constants from Reference 20 and Results of MP2/cc-pVTZ Calculations [a]

bonds	r_e	r_a	r_g	angles	$\angle_e$	dihedral angles	τ_e
C2−N1	1.344(3)[b]	1.351(3)	1.353(3)	N1−C2−N3	129.0(1)[f]	C4−C5−C6−N10	180.0[j]
C2−N3	1.330(3)[b]	1.337(3)	1.339(3)	C2−N3−C4	111.0(1)[f]	C5−C6−N10−H14	0.0[j]
C4−N3	1.333(3)[b]	1.341(3)	1.342(3)	N3−C4−C5	127.2(1)[f]	N1−C6−N10−H15	0.0[j]
C4−C5	1.401(3)[c]	1.402(3)	1.404(3)	C4−C5−N7	111.9(2)[g]		
C5−C6	1.409(3)[c]	1.410(3)	1.411(3)	C5−N7−C8	103.4(2)[g]		
C6−N1	1.332(3)[b]	1.336(3)	1.338(3)	C5−C6−N10	121.9(2)[f]		
C5−N7	1.380(4)[d]	1.382(4)	1.384(4)	N3−C2−H11	116.0[h]		
C8−N7	1.319(3)[b]	1.323(3)	1.325(3)	C4−N9−H12	125.9[h]		
C8−N9	1.371(4)[d]	1.384(4)	1.386(4)	N7−C8−H13	124.8[h]		
C4−N9	1.377(4)[d]	1.388(4)	1.391(4)	C6−N10−H14	117.1[h]		
C6−N10	1.357(4)[d]	1.353(4)	1.355(4)	C6−N10−H15	116.2[h]		
C2−H11	1.075(5)[e]	1.091(5)	1.096(5)	C4−C5−C6	115.2(1)[i]		
N9−H12	1.000(5)[e]	1.015(5)	1.020(5)	C5−C6−N1	119.6(3)[i]		
C8−H13	1.071(5)[e]	1.088(5)	1.094(5)	C6−N1−C2	118.0(3)[i]		
N10−H14	0.999(5)[e]	1.044(5)	1.048(5)	N7−C8−N9	113.5(6)[i]		
N10−H15	0.999(5)[e]	1.045(5)	1.050(5)	C4−N9−C8	107.0(7)[i]		
				C5−C4−N9	104.1(6)[i]		
				N1−C6−N10	118.5(3)[i]		

$R_{f,tot} = 3.2$ ($R_{f,LD} = 2.1$, $R_{f,SD} = 4.3$)[k]

[a] Equilibrium (r_e) and thermal-average (r_a, r_g) bond lengths are in angstroms, bond angles ($\angle_e$) and torsional angles (τ_e) are in degrees. Vibrational corrections calculated in approximation II (see Table 3). Values in parentheses are three times the estimated standard errors. [b,c,d,e,f,g] Parameters with the same superscript were refined in one group. Differences between parameters in each group were fixed at the values from MP2/cc-pVTZ calculations. [h] Fixed at the values from MP2/cc-pVTZ calculations. [i] Dependent parameter. [j] Assumed (see text). [k] Total disagreement factor $R_{f,tot}$, R_f factors for long ($R_{f,LD}$) and short ($R_{f,SD}$) nozzle-to-film distances are given in percent.

equilibrium structure, $r_{ij,e}$ and $B_e^{(i)}$ values, respectively. The transformations were performed taking into account nonlinear kinematic effects at the level of the first-order perturbation theory.[28,29] The harmonic ($\delta_{ij,h1}$ and $\delta_{h1}^{(i)}$, including also centrifugal distortions due to vibrations[30]) and anharmonic ($\delta_{ij,anh1}$ and $\delta_{anh1}^{(i)}$[29]) corrections were calculated using MP2/cc-pVTZ quadratic and cubic force fields, respectively.

To include corrections for "local" centrifugal distortions caused by intramolecular motions (Bartell's "bond on a block" problem[31]), calculations were performed in internal coordinates.

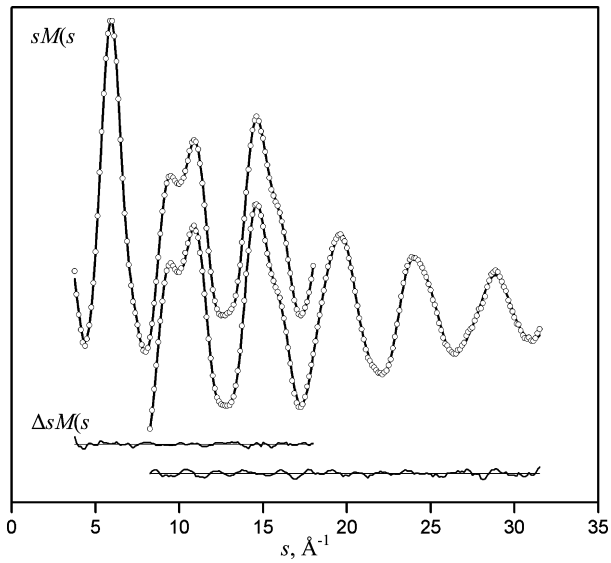

Figure 2. Experimental (open circles) and theoretical (solid line) molecular intensities $sM(s)$ and the difference curve.

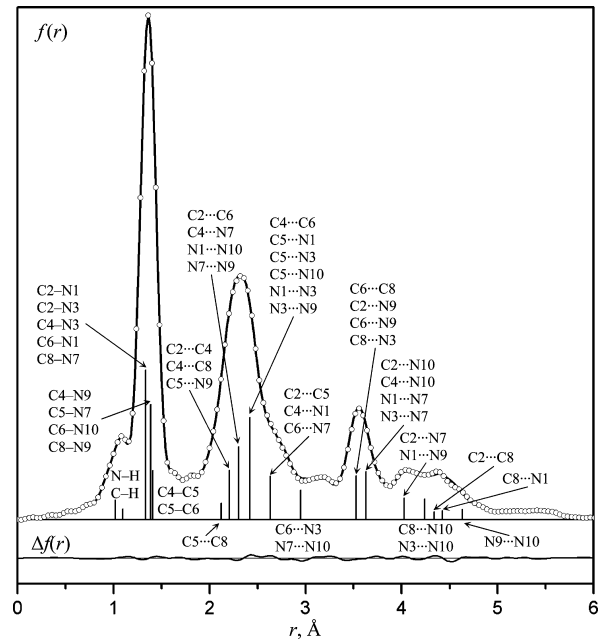

Figure 3. Experimental (open circles) and theoretical (solid line) radial distribution curves $f(r)$ with the difference curve. The distance distribution is indicated by vertical bars.

Cubic force constants in Cartesian coordinates obtained in quantum-chemical calculations were transformed into cubic force constants in internal coordinates as

$$B^{\dagger}(F_{\mathrm{I}}^{3}B)B = F_{\mathrm{C}}^{3} - \partial B^{\dagger}/\partial x\, F_{\mathrm{I}}^{2}B - B^{\dagger}F_{\mathrm{I}}^{2}\,\partial B/\partial x - \partial B^{\dagger}/\partial x\, F_{\mathrm{I}}^{2}B \quad (1)$$

F^3 and F^2 denote tables of cubic and quadratic force constants, respectively; indices "I" and "C" correspond to internal and Cartesian coordinates; B is the transformation matrix between the internal and Cartesian coordinates, $q = Bx$; and the dagger denotes transposition. The second and fourth terms on the right-hand side of this equation differ by the order of indices of the $\partial B^{\dagger}/\partial x$ tensor; see ref 32 for details. Local centrifugal distortions were calculated as described in ref 30. Lastly, centrifugal distortions caused by molecule rotations as a whole were calculated as recommended by Iwasaki and Hedberg,[33] and the transition from internal to Cartesian coordinates for calculations of amplitudes and corrections was performed following ref 28. Since the molecule is cyclic, and the corresponding system of internal coordinates inevitably contains many dependencies, we used the Gauss constrained motion theorem[34] to take this circumstance into account. The total corrections $(r_{ij,\mathrm{e}} - r_{ij,\mathrm{a}})$ and $(B_{\mathrm{e}}^{(k)} - B_0^{(k)})$ calculated by means of the SHRINK-07 program[28,35] are listed in Tables 3 and 4, respectively.

The GED molecular model of 9*H*-adenine tautomer was described by the following 30 parameters: 16 bond lengths C2−N1, C2−N3, C4−N3, C4−C5, C5−C6, C6−N1, C5−N7, C8−N7, C8−N9, C4−N9, C6−N10, C2−H11, N9−H12, C8−H13, N10−H14, and N10−H15; 11 bond angles N1−C2−N3, C2−N3−C4, N3−C4−C5, C4−C5−N7, C5−N7−C8, C5−C6−N10, N3−C2−H11, C4−N9−H12, N7−C8−H13, C6−N10−H14, and C6−N10−H15; and three torsional angles C4−C5−C6−N10, C5−C6−N10−H14, and N1−C6−N10−H15. Since the GED intensities are hardly sensitive to the positions of hydrogens, the bond angles associated with these atoms were fixed at the values from MP2/cc-pVTZ calculations. The torsional angles were assumed as discussed below. In final refinement, 22 geometrical parameters were combined into the following six groups: (1) both C−C bond lengths; (2) all C−N of six-membered ring and C8−N7, i.e., the C−N bond lengths which are shorter than the single N6−C10 bond of ∼1.35 Å; (3) C6−N10 and all C−N which are longer than C6−N10, i.e., all C−N of the five-membered ring except C8−N7; (4) all N−H and C−H bond lengths; (5) three bond angles describing the six-membered ring; and (6) C5−C6−N10 and two bond angles of the five-membered ring (see Table 5). The differences between parameters within each group were fixed at the values derived from the results of MP2/cc-pVTZ calculation.

The rms vibrational amplitudes $u_{ij,\mathrm{h1}}$ were calculated using the quadratic MP2/cc-pVTZ force field and were used as starting values in the structural analysis. The amplitudes for hydrogen containing distances were not refined, whereas the other amplitudes were refined in groups (see Table S2 of the Supporting Information) constraining their differences at theoretical values.

The fitting of GED intensities and MW rotational constants was performed by means of the UNEX program.[27] The weights of the rotational constants relative to the GED intensities were optimized so that to reach the maximal accuracy of determination of the GED component,[37] i.e., so that the GED *R*-factor remained at the minimal value, which was reached without the use of MW rotational constants. The *R*-factor was obtained to be 3.2%. The differences between the MW rotational constants and those calculated from the GED molecular geometry did not exceed 0.04% (see Table 4). The reduced molecular intensity $sM(s)$ and radial distribution curves $f(r)$ with their theoretical analogues for the final molecular model are shown in Figures 2 and 3, respectively.

Results, Discussion, and Conclusions

The vibrational spectra of the 9*H*-adenine tautomer calculated by B3LYP and MP2 methods using different basis sets (see Table 2) are very similar, except for the out-of-plane NH_2 frequencies ν_3, which are extremely different (66, 336, and 391 cm^{-1} in the B3LYP/aug-cc-pVTZ, MP2/aug-cc-pVTZ, and MP2/cc-pVTZ approximations, respectively). Most theoretical frequencies are relatively close to the experimental ones[17] even

TABLE 6: Comparison of Bond Lengths (Å) of 9*H*-Adenine Tautomer with Those of Similar Molecules

	9*H*-adenine		pyrimidine			1*H*-imidazole
	MP2/cc-pVQZ [b]	GED + MW[b]	MW [c]	ED + LCNMR + MW [d,e]	GED [d]	MW [f]
bond[a]	r_e	r_e	r_s	r_α	r_α	r_s
C2−N1	1.345	1.344(3)				
C2−N3	1.331	1.330(3)				
C2−N	1.338[g]	1.337(3)[g]	1.337(2)	1.328(7)	1.35(2)	
C6−N1	1.333	1.332(3)				
C4−N3	1.334 [h]	1.333(3) [h]	1.332(3)	1.350(7)	1.32(2)	
C4−C5	1.394	1.401(3)	1.393(2)	1.393(3)		1.3638
C5−C6	1.403	1.409(3)				
C5−N7	1.374	1.380(4)				1.3822
C8−N7	1.320	1.319(3)				1.3135
C8−N9	1.365	1.371(4)				1.3643
C4−N9	1.372	1.377(4)				1.3774

[a] Atom numbering is given for adenine. [b] Present work. Rotational constants from ref 20. [c] Ref 40. [d] Ref 41. [e] Combined with rotational constants of parent isotopomer. [f] Ref 42. [g] Mean value of C2−N1 and C2−N3. [h] Mean value of C4−N3 and C6−N1.

without scaling; the average percent differences excluding the out-of-plane bending vibration ν_3 are 2.3−2.8%, and the greatest deviations of 5−7% have the highest modes of 35−39 corresponding to C−H or N−H stretching. Although the MP2 values of ν_3 are closer to the experimental one of 242 cm^{-1} (IR (Ar))[17] (240 cm^{-1} in the gas phase[19]) than that from the B3LYP approximation which obviously poorly modeled this mode, there is still no reliable information about the potential energy function of the out-of-plane NH_2 vibration. Finally, the MP2/cc-pVTZ force constants were scaled by means of the SHRINK-07 program[28,35] so that the experimental frequencies ν_3, ν_{10}, ν_{21}, ν_{25}, $\nu_{27}-\nu_{29}$, ν_{31}, and $\nu_{33}-\nu_{39}$[17] (see Table 2) were reproduced exactly. After scaling, the average deviation between calculated and experimental frequencies decreased from 2.8 to 1.4%.

According to the results of all quantum-chemical calculations (see Table 1), the amino group is nonplanar (see Table 1). The absence of imaginary frequencies proved that the determined geometries correspond to the equilibrium conformation. To the contrary, the planar molecular configuration with one imaginary frequency of the out-of-plane vibration of the amino group (see Table 2) corresponds to a transition state. However, the deviation from planarity varies very strongly depending on the method of calculations: the C(N)−C6−N10−H angles were found to be up to 9 and 18° at the B3LYP and MP2 levels (cc-pVTZ basis set), respectively. It is remarkable that these angles decreased to 3 and 15°, respectively, when diffusion functions (aug-cc-pVTZ basis set) were included. Moreover, the barrier to inversion of the amino group is very small (only 0.1 kcal/mol (35 cm^{-1}) in the MP2/aug-cc-pVTZ approximation and even less at the B3LYP level); i.e., it is essentially lower than the frequency of the out-of-plane NH_2 vibration (see mode 3 in Table 2). Thus, the amino group in adenine is quasi-planar. Since the potential energy function of this vibration could not be determined reliably at either B3LYP or MP2 approximations, and the GED intensities are not sensitive to positions of the hydrogen atoms, the amino group was considered as effectively planar in the structural analysis of GED data.

According to the results of all quantum-chemical calculations (see Table 1), the molecular skeleton of the 9*H*-adenine tautomer is essentially planar in the equilibrium conformation: the C6−N10 bond deviates from the ring plane by only 0.5−2.6°. It should be noted that the transition-state molecular configuration (with a planar amino group) was found to be exactly planar (∠(N10−C6−N1−C2) = 180° (MP2/aug-cc-pVTZ)). Consequently, the heavy-atom skeleton was assumed to be planar in the GED analysis.

As can be seen from Table 1, although bond length variations depending on the level of the applied theory (B3LYP and MP2) reach 0.017 Å (see C8−N7 bond length values calculated with the cc-pVTZ basis set), the deviations of bond length differences in all MP2 calculations do not exceed a few thousandths of an angstrom unit. Therefore, it can be expected that these differences were obtained with high accuracy and could be successfully used as supporting data in the experimental structural analysis.

As for thymine,[8] the total vibrational corrections $\Delta(r_e - r_a)_{ij}$ to thermal-average bond lengths of adenine were estimated to be 0.002−0.012 Å for the bonds between heavy atoms and ∼0.015 Å for C−H and N−H distances, except for N10−H14 and N10−H15 ones. Since vibrational corrections exceed the experimental uncertainties of a few thousandths of angstroms, they have to be precisely calculated and taken into account in the molecular geometry determination by GED. It should be noted that $\Delta(r_e - r_a)_{ij}$ values (see Table 3) calculated from the B3LYP/cc-pVTZ and MP2/cc-pVTZ force constants using the same approximation (II) are very different in some cases, for example, for the C6−N1, C6−C5, and C6−N10 distances (B3LYP/MP2: −0.0111/−0.0035, −0.0132/−0.0003, and 0.0263/−0.0017 Å, respectively). The B3LYP force field proves to be not appropriate for calculations of the corrections because some of them were obtained to be unrealistically large and positive (see $\Delta(r_e - r_a)$ of 0.0263 Å for C6−N10). As can be seen from Table 3, both mathematical approximations (II and III) yield practically the same results. Maximal change of vibrational correction of 0.006 Å due to force field scaling was obtained for the C6−N10 bond, whereas for the other bonds the scaling effects were only a few thousandths of angstroms, i.e., less than the GED uncertainties of internuclear distances. Large corrections to N−H distances of the amino group (∼0.045 Å) point to large-amplitude motions.

Very low fitting factor R_f of 3.2% confirms that the applied anharmonic (cubic) theoretical model is appropriate for the description of the molecular dynamics of adenine in the GED analysis although the C−H and N−H stretching and C−N−H and N−C−H bending modes should have also large quartic terms.[38] At such high level of agreement between experimental data and GED theoretical model, the extension of this model by quartic terms seems to be not necessary and not possibly because of small sensitivity of experimental intensities to hydrogen positions. In each case, the corresponding anharmonic potential functions should be first reliably obtained by the ab initio method that could be a challenge for future investigations.

The molecular structure of adenine is determined experimentally for the first time. The obtained equilibrium (r_e) and thermal-average (r_g) parameters are presented in Table 5. Very good agreement between the experimental and theoretical (at the MP2 level with the cc-pVTZ, aug-cc-pVTZ, and cc-pVQZ basis sets) geometries (see Tables 5 and 1, respectively) is evidence of the high quality of both the GED experiment and ab initio calculations. The MW rotational constants (B_0) corrected for anharmonic effects agree with B_e values calculated from the GED and MP2 geometries (see Table 4) within uncertainties of the GED and MP2 data, respectively.

The experimentally determined C−C bond lengths of adenine (r_e(C4−C5) = 1.401(3) Å and r_e(C5−C6) = 1.409(3) Å) correspond to a typical aromatic bond length of 1.395(10) Å (see statistical distribution curve of experimental C−C bond lengths of organic gas-phase compounds implemented in the MOGADOC database,[39] Figure S4 of the Supporting Information). The C4−C5 bond length is very close to the C−C distance in pyrimidine (r_s(C−C) = 1.393(2) Å,[40] r_α(C−C) = 1.393(3) Å [41]) being remarkably longer than the C−C double bond in 1*H*-imidazole (r_s(C−C) = 1.3638 Å [42]) due to the condensation effect. The determined C−N bond lengths of adenine are in good agreement with those for both pyrimidine[40] and 1*H*-imidazole[42] (see Table 6). Being in agreement with the data from ref 40, our results confirmed that the two C−N bonds were erroneously separated from each other while the mean value of 1.336 Å was exactly determined in ref 41 (see Table 6).

In comparison with the distribution curve for the C−C bond lengths (Figure S4), the distribution curve for the C−N distances has no strongly pronounced position for aromatic bond lengths which range over a wide interval between 1.33 and 1.37 Å (see Figure S5 of the Supporting Information). This interval includes all the determined experimental r_e(C−N) distances of adenine (r_e(C−N)$_{mean}$ = 1.335 Å in the six-membered ring and r_e(C−N)$_{mean}$ = 1.362 Å in the five-membered ring).

Thus, the determined values of C−C and C−N bond lengths point to the aromatic nature of the adenine molecule. Additionally, the GIAO-MP2/cc-pVTZ calculated aromaticity indexes, NICS (abbreviation from nucleus-independent chemical shift; for definition see ref 43 and references therein), of −7.6 and −12.7 ppm for six-membered and five-membered rings, respectively, confirm the strong aromaticity of this molecule, which is of the same order as in benzene with NICS index of −8.2 ppm (MP2/cc-pVTZ) and stronger than in thymine with NICS index of −1.35 ppm (MP2/cc-pVTZ).

Acknowledgment. The authors thank the Dr. Barbara Mez-Starck Foundation (Germany) for financial support, Dr. M. Wunderlin (University of Ulm) for recording of MS (CI) spectrum, Dr. M. Popik (Moscow State University) for a gas chromatography−mass-spectrometric analysis, and Dr. J. Vogt (University of Ulm) for helpful discussions.

Supporting Information Available: Mass spectrum of adenine obtained by chemical ionization method (Figure S1), experimental intensity curves with final backgrounds (Table S1, Figures S2 and S3), distributions of experimental C−C and C−N bond lengths (Å) of organic gas-phase compounds implemented in the MOGADOC database (Figures S4 and S5, respectively), non-hydrogen interatomic distances (r_a) and vibrational amplitudes (u) at 484 K obtained from GED experiment and calculated using MP2/cc-pVTZ force field (Table S2). This information is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.
(2) Topal, M. D.; Fresco, J. R. *Nature* **1976**, *263*, 285.
(3) Kabeláč, M.; Hobza, P. *Phys. Chem. Chem. Phys.* **2007**, *9*−903.
(4) Ferenczy, G.; Harsányi, L.; Rozsonday, B; Hargittai, I. *J. Mol. Struct.* **1986**, *140*, 71.
(5) Vaguero, V.; Sanz, M. E.; López, J. C.; Alonso, J. L. *J. Phys. Chem. A* **2007**, *111*, 3443.
(6) Gahlmann, A.; Park, S. T.; Zewail, A. H. *J. Am. Chem. Soc.* **2009**, *131*, 2806.
(7) Lŏpez, J. C.; Peña, M. I.; Sanz, M. E.; Alonso, J. L. *J. Chem. Phys.* **2007**, *126*, 191103.
(8) Vogt, N.; Khaikin, L. S.; Grikina, O. E.; Rykov, A. N.; Vogt, J. *J. Phys. Chem. A* **2008**, *112*, 7662.
(9) Broo, A. *J. Phys. Chem. A* **1998**, *102*, 526.
(10) Menucci, B.; Toniolo, A.; Tomasi, J. *J. Phys. Chem. A* **2001**, *105*, 4749.
(11) Fonseca Guerra, C.; Bickelhaupt, F. M.; Saha, S.; Wang, F. *J. Phys. Chem A* **2006**, *110*, 4012.
(12) Hanus, M.; Kabeláč, M.; Rejnek, J.; Ryjáček, F.; Hobza, P. *J. Phys. Chem. B* **2004**, *108*, 2087.
(13) Roca-Sanjuán, D.; Merhán, M.; Serrano-Andrés, L.; Rubio, M. *J. Chem. Phys.* **2008**, *129*, 095104.
(14) Salter, L. S.; Chaban, G. M. *J. Phys. Chem. A* **2002**, *106*−4251.
(15) Ha, T. K.; Keller, M. J.; Günde, R.; Gunthard, H. H. *J. Mol. Struct. (THEOCHEM)* **1996**, *364*, 161.
(16) Choi, M. Y.; Dong, F.; Han, S. W.; Miller, R. E. *J. Phys. Chem. A* **2008**, *112*, 7185.
(17) Nowak, M. J.; Lapinski, L.; Kwiatkowski, J. S.; Leszczyński, J. *J. Phys. Chem.* **1996**, *100*, 3527.
(18) Nowak, M. J.; Rostkowska, H.; Lapinski, L.; Kwiatkowski, J. S.; Leszczyński, J. *J. Phys. Chem.* **1994**, *98*, 2813.
(19) Colarusso, P.; Zhang, K.; Guo, B.; Bernath, P. F. *Chem. Phys. Lett.* **1997**, *269*, 39.
(20) Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. *Chem. Phys. Lett.* **1989**, *156*, 61.
(21) Stepanian, S. G.; Sheina, G. G.; Radchenko, E. D.; Blagoi, Yu. P. *J. Mol. Struct.* **1985**, *131*, 333.
(22) Plützer, Chr.; Kleinermanns, K. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4877.
(23) Samijlenko, S. P.; Krechkivska, O. M.; Kosach, D. A.; Hovorun, D. M. *J. Mol. Struct.* **2004**, *708*, 97.
(24) *Structure of Free Polyatomic Molecules*, Landolt-Börnstein, New Series, Vol. II/28A-D; Kuchitsu, K., Vogt, N., Tanimoto, M. Eds.; Springer: Berlin, 2006 and 2007; pp 169 (A), 192 (B), 187 (C), and 229 (D).
(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R. ; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B. ; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O. ; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; 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.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision E.01; Gaussian: Pittsburgh, PA, 2003.
(26) Shibata, S.; Jijima, K.; Tani, R.; Nakamura, J. *Rep. Fac. Sci. Shizuoka Univ.* **1974**, *9*, 33.
(27) Vishnevskiy, Yu. V. *J. Mol. Struct.* **2007**, *833*, 30.
(28) Sipachev, V. A. In *Advances in Molecular Structure Research*; Hargittai, I. , Ed.; JAI: Greenwich, U.K., 1999; Vol. 5, p 263.
(29) Sipachev, V. A. *Struct. Chem.* **2000**, *11*, 167.
(30) Sipachev, V. A. *J. Mol. Struct.* **2001**, *567*−*568*, 67.
(31) Bartell, L. S. *J. Chem. Phys.* **1963**, *38*, 1827.
(32) Sipachev, V. A. *J. Mol. Struct.* **2004**, *693*, 235.
(33) Iwasaki, M.; Hedberg, K. *J. Chem. Phys.* **1962**, *36*, 2961.
(34) Arnol'd, V. I. In: *Advances in Science and Techology, Modern Problems of Mathematics*, Arnol'd, V. I. , Ed.; VINITI: Moscow, 1985.
(35) Sipachev, V. A. *J. Mol. Struct. (THEOCHEM)* **1985**, *121*, 143.
(36) Dorofeeva, O. V.; Vogt, N.; Vogt, J.; Popik, M. V.; Rykov, A. N.; Vilkov, L. V. *J. Phys. Chem. A* **2007**, *111*, 6434.

(37) Kuchitsu, K.; Nakata, M. In *Stereochemical Applications of Gas-Phase Electron Diffraction*; Hargittai, I. , Hargittai, M. , Eds.; VCH: New York, 1988; Part A, p 227.

(38) Lappi, S. E.; Collier, W.; Franzen, S. *J. Phys. Chem. A* **2002**, *106*, 11446.

(39) Vogt, J.; Vogt, N. *Struct. Chem.* **2003**, *14*, 137.

(40) Kisiel, Z.; Pszczółkowski; Lŏpez, J. C.; Alonso, J. L.; Maris, A.; Caminati, W. *J. Mol. Spectrosc.* **1999**, *195*, 332.

(41) Cradock, S.; Liescheski, P. B.; Rankin, D. W. H.; Robertson, H. E. *J. Am. Chem. Soc.* **1988**, *110*, 2758.

(42) Christen, D.; Griffiths, J. H.; Sheridan, J. *Z. Naturforsch.* **1982**, *37a*, 1378.

(43) Krygovski, T. M.; Cyranski, M. K. *Chem. Rev.* **2001**, *101*, 1385.

JP905755U