Calculations and Characterization of the Electronic Spectra of DNA Bases Based on *ab Initio* MP2 Geometries of Different Tautomeric Forms

Anders Broo* and Anders Holmén

Department of Physical Chemistry, Chalmers University of Technology, 412 96 Göteborg, Sweden

Received: November 26, 1996; In Final Form: February 26, 1997[⊗]

The absorption spectra and transition moment directions of the four DNA bases and uracil were calculated, utilizing the semiempirical INDO/S-CI method and the *ab initio* CIS/6-31G* method. A linear scaling of the CIS/6-31G* transition energies was necessary to get good agreement with the observed energies. CASPT2 $\pi \rightarrow \pi^*$ transition energies and polarizations of adenine were also calculated. Fully MP2/6-31G*-optimized geometries were used in the calculations of the electronic spectra. The effect of possible tautomerism was investigated by calculation of the spectrum of several tautomers of guanine and cytosine. Specific and nonspecific solvent effects were considered in the calculations. Nonspecific solvent effects were modeled by a continuum model. Specific solvent interactions were modeled by inclusion of one water molecule in "supermolecule" calculations of the geometry and the spectrum of cytosine and guanine. Experimental transition energies were reproduced with an average error of less than 0.3 eV, and in most cases even better. Transition moment directions were calculated in good agreement with observations and other theoretical results in most cases. In some cases directions do not agree with observations. In these cases we have shown that the observed directions might be affected by the presence of more than one tautomeric form or by solvent or crystal effects. We propose a new interpretation of the $\pi \rightarrow \pi^*$ absorption spectrum of cytosine and guanine due to solvent effects and tautomerism. A simple model, based on symmetry and geometry arguments, was used to rationalize the computed transition moment directions.

1. Introduction

Quantum chemistry has been very successful in predictions of a broad range of ground state properties, especially regarding organic molecules. Excited state properties have been more problematic to reproduce. In principle only semiempirical methods have been able to reproduce absorption spectra of molecules with more than about 10 atoms. Recent developments of *ab initio* theory for excited states, where the dynamic electron correlation of both the ground state and the excited state is properly taken care of, have demonstrated the *ab initio* predictions of absorption spectra have a promising future.[1] Unfortunately, *ab initio* methods still have problems when the chromophore is large or has low symmetry. Furthermore, the calculation of the electronic transitions of medium-sized molecules using the rather accurate CASPT2 method is still very time consuming and very demanding on computer resources.[1] At present, we do not have access to computer resources that allow a systematic investigation of several related chromophores using the CASPT2 method. Instead, we rely on methods that are less costly like semiempirical or *ab initio* CIS/6-31G* methods.

Absorption spectra of the DNA bases have been calculated using many different semiempirical methods by others; see ref 2 for a good review. *Ab initio* predictions of the absorption spectra have also been published.[1b,3−5] In earlier works the agreement between theory and experiments has been satisfactory for some of the transitions, but often the disagreement between calculated and observed transition moment directions has been large. The reason for this disagreement is probably manifold but may roughly be divided into two kinds: (a) experimental difficulties in determining the moment directions from polarized spectroscopy and possible solvent effects on the measured spectra that have not been modeled correctly in the calculations and (b) shortcomings of the theory, mainly due to insufficient treatment of the electron correlation.

The free DNA bases can occur in more than one tautomeric form. The equilibrium between the different tautomeric forms depends on the character of the environment; for instance, base pairing and hydrogen bonding can exclude one or several possible tautomers. Thus, it is important to model both the polarity of the solvent as well as the specific solvent interactions to get a good understanding of the equilibrium. The existence of more than one tautomeric form of many of the DNA bases further complicates the interpretation of the absorption and emission spectra taken in solution and in stretched film. Tautomerism of the DNA bases has been investigated with computational methods by many groups.[6,7] Lately, computational models for solvent−solute interactions have been developed and used to study tautomerism of the DNA bases.[8,9] It has been shown that for uracil and thymine only one tautomer, the 2,4-keto tautomer, is thermally accessible at room temperature and in solution. For adenine the 9*H* and 7*H* tautomers are the only tautomers that are accessible at room temperature. For guanine and cytosine the situation remains somewhat unclear since several tautomers are calculated to be close in energy. For instance, Colominas et al.[8] found four tautomers of cytosine within 8 kcal/mol of total free energy in the gas phase, and the relative ordering of the stability was found to be method dependent. However, on the basis of reaction field calculations, they argued that the 2-keto-4-amino tautomer of cytosine and the 2-amino-6-keto-9*H*-guanine and 2-amino-6-keto-7*H*-guanine tautomers were the only tautomers accessible in water solution.[8] In this work we will investigate the importance of specific solvent interactions for relative stability of some of the possible tautomers of cytosine and guanine.

A consistent treatment of the absorption spectra of the DNA bases is presented, using geometries obtained at the MP2/6-31G* level of theory, in contrast to previous theoretical works

[⊗] Abstract published in *Advance ACS Abstracts,* April 15, 1997.

S1089-5639(96)03928-X CCC: $14.00

where idealized geometries have been used. The MP2/6-31G* geometries are very close to experimental geometries where reliable experimental data are available.[10] The use of accurate geometries proved to be important since the calculated spectral properties are sensitive to the choice of the ground state geometry. The semiempirical INDO/S method is used for an extensive study of the absorption spectra of several tautomeric forms of the DNA. We also report CIS/6-31G*//MP2/6-31G* calculated absorption spectra for the free DNA bases. The $\pi \rightarrow \pi^*$ transitions of adenine calculated with the CASPT2 method are also presented. A simple model is used in an attempt to rationalize the computed and observed transition moment directions. Environmental effects have been considered using a self-consistent reaction field to account for nonspecific solvent effects on the computed spectra. In some cases we have included specifically bonded water molecules to model specific solvent effects. Furthermore, crystal effects are estimated by perturbing the gas-phase molecule with an external electric field and by dimer interactions. In the latter type of calculations two identical chromophores are included in the calculation and several different relative orientations are considered.

In section 2 the computational details are described. In section 3 we first investigate the geometry dependency on the predicted spectroscopical properties of adenine. Then a simple model is outlined that rationalizes the computed transition polarizations. In subsection 3.4 the INDO/S spectra are reported and compared with observations. In subsection 3.5 the CIS/6-31G* spectra are presented together with the CASPT2 spectrum of adenine. In the last sections the conclusions about the theoretical treatment of the absorption spectra of the nucleotides are summarized.

2. Computational Methods

The geometry optimizations were performed using the Gaussian94 program package.[11] The basis set used in the MP2 geometry optimizations was the standard split-valence double-zeta basis set, augmented with a d-type polarization function on the non-hydrogen atoms (6-31G*).[12] The geometries of the enol forms of guanine and cytosine, four conformers of the cytosine−water complex, and three conformers of the guanine−water complex were fully optimized in this work. The remaining MP2/6-31G* geometries of the DNA bases have already been published.[10]

Semiempirical calculations of absorption spectra were done with the ZINDO program, using the intermediate neglecting of differential overlap model Hamiltonian with the spectroscopical parametrization (INDO/S).[13] The absorption spectra were calculated using a configuration interaction between singly excited configurations (CIS). The active space consisted of all the π, π^*, and nonbonding molecular orbitals. The importance of doubly excited configurations was tested for all the calculated spectra. Only minor changes with respect to transition energies and directions were found.

The CIS/6-31G* absorption spectra were calculated using a CIS where the active space consisted of all occupied valence MOs and all the virtual MOs. The Gaussian94 program was used for these calculations.

The MOLCAS program[14] was used for the CASPT2 calculation of the absorption spectrum of adenine. The basis set was the built-in so-called small ANO basis set contracted to [3s2p1d/2s] and augmented with a set of Rydberg orbitals contracted to [1s,1p,1d]. The Rydberg orbitals were centered at the center of charge of the adenine cation. This basis set is somewhat smaller than the basis set used in the CASPT2 calculations for the other DNA bases.[1b,3a,4] The geometry was a C_s symmetric

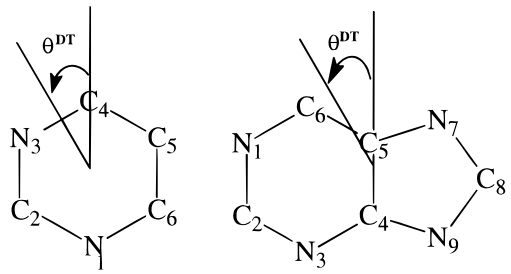

Figure 1. Convention suggested by Devoe and Tinoco that is used in this work for polarizations of the transition moments of the DNA bases.[16]

MP2/6-31G*-optimized geometry of 9*H*-adenine. The so-called intruder states were identified in a preliminary CASSCF calculation and deleted in the remaining calculations. The procedure is described in detail in ref 1b. The 12 π electrons were allowed to distribute in all possible ways in 12 active π orbitals in the CASSCF step. This active space generates 60 984 configuration state functions. The density matrix in the CASSCF step was averaged over the first 10 roots with π symmetry.

Nonspecific solvent effects on the predicted spectra were modeled by the self-consistent reaction field model (SCRF).[15] In the SCRF model, used in this work, the solvent is represented by a dielectric continuum separated from the solute by a spherical cavity. The cavity radius, a_0, was determined by the mass−density formula proposed by Karelson and Zerner.[15] The radii used are as follows: thymine $a_0 = 3.68$ Å, uracil $a_0 = 3.54$ Å, cytosine $a_0 = 3.53$ Å, cytosine + H_2O $a_0 = 3.71$ Å, guanine $a_0 = 3.91$ Å, guanine + H_2O $a_0 = 4.06$ Å, and adenine $a_0 = 3.77$ Å.

All the transition moment directions reported in this work are according to the convention first suggested by Devoe and Tinoco;[16] see Figure 1.

3. Results and Discussion

3.1. Geometry Dependence on the Predicted Absorption Spectrum of Adenine. In Table 1, the INDO/S-CIS-predicted $\pi \rightarrow \pi^*$ absorption transitions of adenine, using geometries obtained at different levels of theory, are collected. A comparison between most of the geometries and the observed geometry of 9-methyladenine[17] can be found in ref 18. In most cases, at least one bond length deviates more than 0.04 Å from the corresponding neutron diffraction-determined bond length, except in the MP2/6-31G* and DFT/6-31G** geometries. The rather large difference in bond lengths affects the conjugation and the electron structure of the molecule. The calculated $\pi \rightarrow \pi^*$ transition energies and oscillator strengths show a noTable geometry dependency, especially in the high-energy region. The energy spread of an individual transition could be as large as 2900 cm^{-1} (0.36 eV), due to the geometrical differences. The calculated transition moment directions and oscillator strengths are in agreement except for the MM+ geometry. However, the directions are not in accord with the observed directions. The average difference between predicted and observed $\pi \rightarrow \pi^*$ transitions is between 2700 and 1300 cm^{-1}. We further note that the oscillator strength of the third $\pi \rightarrow \pi^*$ transition is rather sensitive to the geometry. The computed polarization of that transition indicates that the transition density arises from the N_7-C_8 double bond and the pyrimidine ring system. The N_7-C_8 bond length is predicted to be much too short by the HF/6-31G* method and a little too long with the two semiempirical methods. The adjacent C_8-N_9 bond length is also predicted quite differently by these methods. We conclude that it is

TABLE 1: Summary of $\pi \rightarrow \pi^*$ Transition Energies and Polarizations of 9*H*-Adenine as Predicted by the INDO/S-CI Model Using Theoretically Predicted Geometries[a]

AM1[b]		PM3[c]		HF/6-31G* [d]		MM+[e]		B3LYP/6-31G** [f]		observed geometry[g]		MP2/6-31G* [d]		observed[i]	
E	Θ^{DT}	E	Θ^{DT}	E	Θ^{DT}	E	Θ^{DT}	E	Θ^{DT}	E	Θ^{DT}	E	Θ^{DT}	E	Θ^{DT}
33.3	32	34.6	40	36.1	32	35.5	63	35.5	34	34.7	40	35.2	38	36.4	83
(0.14)		(0.10)		(0.11)		(0.36)		(0.09)		(0.13)		(0.12)		(0.10)	
35.7	76	37.7	73	38.7	70	37.0	−42	37.7	70	37.0	69	37.9	75	37.7	25
(0.31)		(0.30)		(0.31)		(0.07)		(0.33)		(0.30)		(0.31)		(0.20)	
42.0	−11	43.8	−27	44.8	−22	42.6	38	44.0	−23	43.2	−22	43.9	−16	46.9	−45
(0.07)		(0.18)		(0.13)		(0.07)		(0.10)		(0.10)		(0.06)		(0.25)	
46.1	−25	46.9	−33	48.4	−30	49.0	−22	48.3	−44	47.5	−38	48.5	−59	49.0	15
(0.68)		(0.61)		(0.65)		(0.57)		(0.75)		(0.74)		(0.80)		(0.11)	
48.3	69	48.5	62	50.3	64	51.9	87	50.0	40	49.7	57	51.0	67	52.9	64
(0.53)		(0.53)		(0.46)		(0.15)		(0.34)		(0.45)		(0.47)		(0.33)	
2.7[j]		2.3[j]		1.3[j]		1.4[j]		1.5[j]		2.2[j]		1.4[j]			

[a] Energies are given in units of 1000 cm^{-1}. The numbers within parentheses are the oscillator strengths. [b] The AM1 semiempirical method was used to optimize the geometry. Mean deviation between calculated and experimental bond lengths: $b_{diff} = 0.028$ Å. [c] The PM3 semiempirical method was used to optimize the geometry. $b_{diff} = 0.025$ Å. [d] *Ab initio* Hartree−Fock-optimized geometry. $b_{diff} = 0.016$ Å. [e] Molecular mechanics-optimized geometry with the MM+ as implemented in the HyperChem 4.5 program. $b_{diff} = 0.050$ Å. [f] DFT/B3LYP-optimized geometry. $b_{diff} = 0.008$ Å. [g] The neutron diffraction geometry from ref 17. [h] *Ab initio* MP2-optimized geometry. $b_{diff} = 0.008$ Å. [i] Single-crystal absorption spectrum of adenosine, ref 41. [j] Average difference between calculated and observed $\pi \rightarrow \pi^*$ transitions.

important to use an accurate geometry to obtain consistent results that could be used for predictions. The first aim is to estimate the "error bars" for the transition moment directions and transition energies obtained in the following way: The geometry was optimized at the MP2/6-31G* level of theory, without any symmetry constraints; then the absorption spectrum was calculated using the INDO/S-CIS model or the *ab initio* CIS/6-31G* model. The active space in the INDO/S calculations consists of single excitations from all the occupied π orbitals and nonbonding orbitals into the unoccupied π^* orbitals, while the *ab initio* CIS/6-31G active space spans all possible single excitations except excitations from the innermost orbitals (frozen core).

3.2. Parent Chromophores for the DNA Bases. Hug and Tinoco[19] tried to analyze the absorption spectra of the DNA bases using benzene as the "parent chromophore". However, only very small similarities were found between the benzene spectrum and the spectra of the DNA bases. They suggested that pyrimidine is a better choice of "parent chromophore" for cytosine and thymine (uracil) and purine for adenine and guanine. A minimum requirement for an approximate method, such as the INDO/S method, is that the spectrum of "parent chromophore" must be well reproduced in order to get reliable results for the more complicated DNA bases.

Pyrimidine was one of the molecules that was used in the parametrization of the INDO/S method. Consequently, the electronic spectrum of pyrimidine is well reproduced by the INDO/S model. Due to the symmetry (C_{2v}) of pyrimidine, only two in-plane transition moment directions are possible. These are 30 and −60° in the Devoe and Tinoco notation.

A comparison between the predicted $\pi \rightarrow \pi^*$ transitions of 9*H*-purine (9H-P) and 7*H*-purine (7H-P) and the experimental linear dichroism study by Albinsson and Nordén[20] of 9-methylpurine and 7-methylpurine shows that the INDO/S method reproduces energies and transition moment directions rather well. The changes of the absorption profile in the low-energy region when the methyl group are moved is reproduced (see Table 2). The first five calculated transitions of the purine spectrum are dominated by excitations from the two highest occupied π orbitals into the three lowest unoccupied π^* orbitals. The third transition is predicted to be weak, and the transition moment direction is, as a consequence of the low intensity, very sensitive to changes in the geometry and environment. The lowest $n \rightarrow \pi^*$ transition is predicted to be located at a lower energy than the first $\pi \rightarrow \pi^*$ transition in accord with observations.

TABLE 2: INDO/S-Predicated $\pi \rightarrow \pi^*$ Transition Energies and Polarizations of 7*H* and 9*H*-Purine Compared to the Stretched Film Measurements of 7- and 9-Methylpurine[20] [a]

	INDO			experiment			polarization differences		
	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}
				7*H*-Purine					
I	34.9	0.182	56	37.2	6930	**55**/−25	1	6	5
II	41.4	0.030	−66	39.8	2110	−**62**/−88	4	26	22
III	44.9	0.080	−2	44.8	1200	−**38**/68	36	32	4
IV	48.3	0.962	−68	48.1	15600	−46/76	22	28	6
V	50.0	0.403	31						
				9*H*-Purine					
I	35.7	0.142	68	37.5	4230	**53**/−73	15	18	3
II	40.3	0.212	88	39.1	4460	−**63**/43	29	12	17
III	45.6	0.045	−47	44.4	950	−**53**/33	6	7	13
IV	50.1	0.686	−43	48.3	17600	−50/40	7	3	10
V	52.8	0.823	55						

[a] Bold numbers indicate the selected transition moment direction in the experimental work. The experimental polarizations are given using the Devoe−Tinoco notation[10] as opposed to the long axis/short axis coordinate system used in ref 20. Energies are given in units of 1000 cm^{-1}. The predicted polarizations by the simple fragment model are for 7*H*-purine 70°, 50°, −40°, and −34° and for 9*H*-purine 50°, 34°, −40°, and −80°. [b] The difference between the INDO/S and the observed polarizations. [c] The difference between the INDO/S and the simple model-predicted polarizations. [d] The difference between the observed and the simple model-predicted polarizations.

3.3. Are There Any Preferred Transition Directions for the DNA Bases? The transition moment directions in pyrimidine are dictated by the symmetry of the molecule. What will happen with the transition moment directions if the symmetry is reduced? We expect the directions to be affected, but is there any preferred direction if the molecule possesses near-C_{2v} symmetry? We have performed a series of INDO/S calculations on pyrimidine with slightly distorted geometries, formally C_1 symmetric geometries. The directions are rotated but not more than 3−5° when we introduce an asymmetry of 0.09 Å in one of the former symmetry-related C−N bonds. Thus, a rather distorted pyrimidine still has transition moment directions close to 30 and −60° that were dictated by the C_{2v} symmetry. On the other hand, if we consider a molecule like cyclohexene with one localized double bond, only one possible in-plane transition moment direction is possible due to the symmetry, namely, along the double bond. In a molecule like cyclohexadiene there are two formal double bonds. Is it possible to tell the transition moment directions beforehand in such a molecule? Cyclohexa-

diene can be in two different isomeric forms: the 1,3-cyclohexadiene and the 1,4-cyclohexadiene. In 1,4-cyclohexadiene, the two formal double bonds are parallel and the molecule can possess C_2 symmetry (boat form) or C_s symmetry (chair form). If we now divide the molecule into two localized π fragments, the symmetry determines the polarizations of the transitions to be along the fragments and along a line that connects the two fragments. In 1,3-cylcohexadiene the two formal double bonds form an angle of 60°, and the molecule possesses C_2 symmetry. By inspecting the geometry we divide the molecules into one π fragment, containing C_1, C_2, C_3, and C_4 since the C_1-C_2 and C_3-C_4 bond lengths are close to true double bonds and the C_2-C_3 bond is significantly shorter that a true single bond. If this simple model holds, we expect transitions to be polarized along the long axis and short axis of the fragment. If we count the transition moment relative to the C_1-C_2 axis, the calculated transition moments for 1,3-cylohexadiene are 30° and −60°, and for 1,4-cylohexadiene the angles are 0° and 90°, which were predicted by the simple fragment-partitioning model. An interesting test of the simple model is to substitute the C_1-H group with a nitrogen in, for example, 1,3-cyclohexadiene. If the model holds, the transition moment polarizations should remain at approximately 30° and −60°. The INDO/S transition moment directions for the two lowest $\pi \rightarrow \pi^*$ transitions are −60 and 24°; all other transitions have mixed character due to the nonplanarity of the molecule. We conclude that it seems possible to predict possible INDO/S transition moment directions using the simple fragment model. We will use the model to rationalize the calculated transition moment directions for the much more complicated DNA bases.

Purine contains a pyrimidine unit. Is it possible to see geometrical and spectroscopical similarities between the two molecules or does the additional ring system dramatically change the geometry and electron structure of the pyrimidine unit? The five-membered ring system introduces only small geometrical changes in the six-membered ring. The largest difference between corresponding bond lengths in the two molecules is only 0.02 Å, and the largest difference in bond angles is 3° at the MP2/6-31G* level of theory. Thus, the local symmetry of the six-membered ring of purine is rather close to C_{2v}. The five-membered ring has a C−N bond that is rather short. We partition purine into two π type fragments: the first fragment is the pyrimidine ring system, and the second fragment is the N_7-C_8 group (9H-P) and the N_9-C_8 group (7H-P). Assuming transitions to be polarized between and along the two fragments, we get the following preferred directions for 9H-P 34°, 30°, −60°, and −80°, and for 7H-P we get the following directions: 70°, 30°, −34°, and −60°. The 30° and −60° directions come from the near C_{2v} symmetry of the pyrimidine ring for both tautomers. If we further assume that the additional five-membered ring affects the pyrimidine-like transitions in such a way that the directions are rotated 20°, the preferred directions are for 9H-P: 50°, 34°, −40°, and −80° and for 7H-P 70°, 50°, −34°, and −40°. The polarizations obtained with this simple model are within ±30° of both the experimental and INDO/S polarizations; see Table 2. The largest differences $|\Delta\Theta^{DT}|$ are for the weak transitions, indicating a larger degree of uncertainty in both the theoretical and observed polarizations of these transitions.

3.4. Spectra of the Individual Bases. Thymine and Uracil. Uracil and thymine are the two less complicated RNA and DNA bases, respectively. The geometry of the ring system of both chromophores deviates rather much from a local C_{2v} symmetry. The carbonyl groups break the conjugation in the ring system, and a π fragment between C_5-C_6 might be identified. A second fragment is built from the two carbonyl groups together with N_3. The preferred transition moment directions between and along the fragments are then about 80°, 0°, and −30°. The predicted INDO/S spectrum is reported in Table 3. The simple model accounts very well for the calculated directions.

TABLE 3: INDO/S-Predicted $\pi \rightarrow \pi^*$ Transition Energies and Polarizations of Thymine and Uracil Compared with Observed Spectra[a]

	INDO			CASPT2 (ref 4)			experiment		
	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}
					Thymine				
	40.2	0.386	−5	39.4	0.17	15	37.6[b]		−**31**/51
I	(39.8)	(0.395)	(−4)				36.6[e]	0.19	−19/14
							37.7[c]		
II	46.9	0.231	85	47.4	0.17	−19	48.3[e]	0.28	
	(45.4)	(0.248)	(86)				46.5[c]		
III	52.8	0.261	−25	49.2	0.15	67	51.3[c]		
	(53.6)	(0.310)	(−26)						
IV	57.8	0.576	−29	57.5	0.85	−35			
	(58.4)	(0.580)	(−27)						
					Uracil				
I	41.1	0.363	−3	40.3	0.19	−7	38.5[b]		−**11**/7
	(40.4)	(0.366)	(−5)			−2[j]	37.1[d]	0.195	−**9**/16
							38.2[c]		
							41.0[h]		
II	48.3	0.207	82	46.9	0.08	−29[j]	47.0[d]	0.260	**59**/−53
	(47.0)	(0.214)	(83)			−24[j]	46.5[c]		
III	51.8	0.160	−17	52.1	0.29	23	51.6[b]	0.162	
	(52.6)	(0.198)	(−22)			62[j]	51.3[c]		
IV	56.1	0.319	−21	56.5	0.76	−42	56.5[c]		
	(56.5)	(0.113)	(−19)			−32[j]			
					1,3-Dimethyluracil				
I	39.4	0.311	−8				37.7[f]		−14[g]
							39.0[i]		
II	46.7	0.218	65				49.0[f]		**35**/−63[g]
							48.8[i]		
III	48.1	0.166	−38						
IV	52.4	0.421	−18				53.5[f]		
							53.5[i]		

[a] Bold numbers indicate the selected transition moment direction in the experimental work. The number within parentheses corresponds to SCRF calculations that account for solvent effects. Energies are given in units of 1000 cm^{-1}. The predicted polarizations by the simple fragment model are 0°, −30°, and −80°. [b] Stretched film LD measurements of thymine and uracil, ref 26. [c] CD spectrum of deoxythymidine-5′-phosphate and deoxyuridine-5′-phosphate, ref 23. [d] Single-crystal absorption spectrum of 1-methyluracil, ref 25. [e] Aqueous solution spectrum of 1-methylthymine, directions from crystal spectrum of 1-methylthymine, ref 39. [f] Aqueous solution spectrum of 1,3-dimethyluracil, ref 21. [g] Stretched film measurements of 1,3-dimethyluracil, ref 27. [h] Gas-phase absorption spectrum of uracil, ref 21. [i] Gas-phase absorption spectrum of 1,3-dimethyluracil, ref 21. [j] Transition directions obtained with the CASSI method; for details, see reference 4.

The absorption spectrum of uracil possesses a rather large red shift of the first strong absorption band upon solvation. The first strong band of the gas-phase absorption spectrum of uracil is located around 41 000 cm^{-1}.[21] The observed solution absorption spectrum of both thymine and uracil has a broad band with the band maximum around 38 000 cm^{-1}.[21,22] The remaining bands are not affected very much by the solvent. The CD spectra taken in solution suggest that the first band is due to two transitions: one at about 38 000 cm^{-1} and a second transition at about 42 000 cm^{-1} with unknown character.[23] The second observed band is composed of two $\pi \rightarrow \pi^*$ transitions, and a third band, located at about 56 000 cm^{-1}, has also been observed.[23] Thymine and uracil show both fluorescence with the initial state having $n \rightarrow \pi^*$ character and phosphorescence.[22] Furthermore, Fujii et al.[24] observed two low-energy $n \rightarrow \pi^*$ transitions in the fluorescence excitation spectrum of uracil taken in a supersonic jet. The gas-phase absorption spectrum of uracil also hints on an $n \rightarrow \pi^*$ transition as the lowest observed state.[22]

TABLE 4: INDO/S-Predicted $\pi \rightarrow \pi^*$ Transition Energies and Polarizations of Cytosine Using Geometries Optimized at the MP2/6-31G* Level of Theory Compared with Experimental Spectra and with a Spectrum Obtained with the CASPT2 Method[2b] [a]

	INDO			CASPT2 (ref 3a)			experiment		
cytosine	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}
I	36.0	0.170	46	35.1	0.061	61	37.1[c]	0.14[c]	**6**/54
	(36.1)[g]	(0.205)	(46)				37.3[d]		**25**/−46
							36.0[e]		35/−28
							37.0[f]		
II	44.9	0.156	−15	40.3	n → π^*		42.7[c]	0.03[c]	−35
	(45.8)	(0.165)	(−12)				41.7[d]		**6**/−27
							43.0[e]		15/−6
							41.2[f]		
III	47.2	0.007		43.2	0.108	−2	45.2[c]	0.13[c]	**76**/−17
	(49.5)	(0.177)	(−36)				46.1[f]		
IIV	49.5	1.045	−39	49.7	0.863	−40	50.0[c]	0.36[c]	86/−26
	(49.1)	(0.757)	(−42)				49.0[e]		50/−43
							50.0[f]		
V	53.8	0.090	32	54.3[b]	0.147[b]	15[b]	53.5[f]		
	(53.8)	(0.093)	(30)						

[a] Bold numbers indicate the selected transition moment direction in the experimental work. Energies are given in units of 1000 cm^{-1}. The predicted polarizations by the simple fragment model are 45°, −30°, −60°, and −85°. [b] This state has a low weight of the reference state after the perturbation calculation; thus the energy and properties have a rather large degree of uncertainty. [c] Single-crystal absorption spectrum of cytosine, ref 29. [d] Stretched film LD measurements, ref 26. [e] Stretched film measurements, ref 28. [f] CD spectrum of deoxycytidine-5′-phosphate, ref 23. [g] The number within parentheses corresponds to SCRF calculations that account for solvent effects.

The INDO/S model predicts the first $\pi \rightarrow \pi^*$ transition of thymine and uracil at too high energy. This is in contrast to all other molecules considered in this work where the first $\pi \rightarrow \pi^*$ transition is calculated at slightly too low energy. The INDO/S transition energies agrees well with the CASPT2 and MRCI energies[3] and the gas-phase spectrum.[22] The INDO/S calculations predict two n → π^* excited states in the vicinity of the first $\pi \rightarrow \pi^*$ excited state. In all theoretical calculations the first excited state has n → π^* character and is located between 33 000−36 000 cm^{-1}. Experimentally, a small red shift of the first absorption band is observed when the C_5 hydrogen (uracil) is changed to a methyl group (thymine).[22,23] That shift is reproduced with the INDO/S model. In all the calculations, the first $\pi \rightarrow \pi^*$ transition is predicted to have the transition density along the C_4-C_5 double bond, in accord with observations. Only a few measurements of the polarization of the second band have been reported. The reported values are only preliminary in most cases due to experimental difficulties in measuring polarizations in the high-energy region of the UV spectrum. Novros and Clark[25] reported an angle of 59° for the second band in the spectrum of 1-methyluracil, which seems to be consistent with the LD measurements made by Matsuoka and Nordén on uracil.[26] However, the latter authors did not assign an angle to the second band due to incompleteness of the experimental data in the high-energy region. Holmén et al.[27] found an angle of 35° for the second transition of 1,3-dimethyluracil.

Tautomerism seems to be of little importance for uracil and thymine since all the possible tautomers of uracil are predicted to be much less stable than the 2,4-keto tautomer used in this work.[6,9] The estimation of the solvent effects on the predicted spectra suggest that the first $\pi \rightarrow \pi^*$ transition is red shifted and the second $\pi \rightarrow \pi^*$ transition is blue shifted for both thymine and uracil when increasing the polarity of the solvent (gas phase to water).

We conclude that the INDO/S model reproduces the observed $\pi \rightarrow \pi^*$ spectra of the uracil and thymine chromophores. The INDO/S transition energies agree well with the CASPT2 energies, but the polarizations of the second and third transitions are not in agreement. Judging from the calculated dipole moments of the excited states, the INDO/S-calculation and the CASPT2 calculation do not give the same ordering of the excited states. Thus, if we compare the INDO/S predicted second $\pi \rightarrow \pi^*$ transition with the third $\pi \rightarrow \pi^*$ transition from the CASPT2 calculation and vice versa, both methods predict similar transition moment directions.

Cytosine. In cytosine the C_4 carbonyl group has been substituted with an amino group compared to uracil. As a consequence of the substitution, the N_3 hydrogen has been removed. At the MP2/6-31G* level of theory the C_4-C_5 bond is predicted to be somewhat longer than in uracil and thymine but still close to what we expect for a C−C double bond. The amino group is predicted to be rather nonplanar which causes the N_3-C_4 and C_4-C_5 bond lengths to be somewhat shorter than what is found in uracil and thymine. The local symmetry lowering of the ring system is not as marked as for uracil and thymine. In our simple model we divide cytosine into three fragments. The first fragment contains the carbonyl group, the second contains the N_3-C_4 group, and the last fragment contains N_1, C_6, and C_5. The preferred directions will then be 45°, −30°, −60°, −85°. The model rationalizes the computed directions reported in Table 4.

Circular dichroism measurements on deoxycytidine-5′-phosphate identifies five $\pi \rightarrow \pi^*$ transitions below 54 000 cm^{-1}.[23] Two different stretched film measurements are summarized in Table 4.[26,28] In the work by Matsuoka and Nordén two transitions were assigned and the corresponding transition moments were deduced.[26] In the work by Bott and Kurucsev[28] three bands were identified and the corresponding transition moments were assigned. In the crystal work by Zaloudek et al. four transitions were identified in the spectrum of cytosine with energies lower than 51 000 cm^{-1}.[29] However, they noted that there was a large degree of uncertainty in the assignment of the second transition. Furthermore, the oscillator strength reported for transition II is rather different in solution and in the solid phase. In solution, the oscillator strength of transition II is about 50% of the oscillator strength of transition I, while in the solid phase the oscillator strength of transition II is only 21% of the oscillator strength of transition I. Zaloudek et al.[29] also indicated the possiblity of two n → π^* transitions at 43 000 and at 45 000 cm^{-1}. The angle between the first two transitions has been estimated to be 40 ± 15° from fluorescence anisotropy measurements of 5-methylcytosine.[30] In the stretched film work by Matsuoka and Nordén the lower bound angle, 25°, was used

TABLE 5: INDO/S-Predicted $\pi \rightarrow \pi^*$ Transition Energies of Different Cytosines Compared with Observed Spectra[a]

	cytosine−H_2O 1			cytosine(enol)			cytosine(enol)−H_2O 4		
	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}
I	36.0	0.180	46	37.1	0.150	4	37.4	0.136	6
	(36.3)[b]	(0.198)	(46)						
II	45.4	0.111	−8	45.9	0.207	60	45.3	0.215	53
	(45.5)	(0.167)	(−10)						
III	46.7	0.003		46.8	0.004		46.2	0.003	
	(48.2)	(0.034)	(−47)						
IV	49.1	1.053	−41	51.3	0.800	−42	50.7	0.840	−43
	(48.9)	(0.916)	(−42)						
V	53.1	0.143	33	54.8	0.398	38	54.3	0.358	34
	(53.1)	(0.150)	(24)						
	1-methylcytosine			protonated 1-methylcytosine			expt protonated 1-methylcytosine		
I	36.6	0.145	47	33.1	0.322	11	35.6[c]	0.21	15/−8
II	44.9	0.284	−18	47.0	0.146	30	45.0[c]	0.24	57/−50
IV	49.7	0.902	−45	48.5	0.609	−53			
V	52.4	0.171	22	56.1	0.051	−35			

[a] Energies are given in units of 1000 cm^{-1}. [b] The number within parentheses corresponds to SCRF calculations that includes the solvent effects. [c] Single-crystal absorption spectrum of protonated 1-methylcytosine, reference 31.

to discriminate one of the two possible choices of the polarization. However, if the 40° angle is used in the analysis, the second transition direction is more likely to be −27° (the rejected choice) and not 6° as was selected. For the third transition (fourth band of Zaloudek et al.) it was not possible to discriminate between the two alternative directions. Furthermore, Callis concluded in his review paper from 1983 that the general features of the cytosine spectrum appear to be solvent dependent.[2]

A large number of theoretical investigations of the absorption spectrum of cytosine have been published.[2,3] With a few exceptions the calculations failed to reproduce the observed transition moment directions for cytosine. However, both crystal measurements and stretched film measurements of the transition moment directions usually give two alternative directions for each transition. A better agreement can be found if the rejected choice from the crystal measurements is compared with the theoretical works. The transition moment directions predicted with the INDO/S method are in best agreement with stretched film measurements; see Table 4. The INDO/S results are very similar to the CASPT2, MRCI, and RPA results.[3] All four methods predict the transition moment direction of the first $\pi \rightarrow \pi^*$ transition to be between 41 and 68°. The theoretically predicted angles are in reasonably good agreement with the stretched film measurements and the rejected alternative angle from the crystal measurements. The relative angle between the first and second transition moment is predicted in accord with the observed fluorescence anisotropy measurements on 5-methylcytosine. The assignment of the first band in the crystal spectrum is based on a comparison between the reflection spectrum of cytosine and 1-methylcytosine.[30] In later single-crystal investigations[30,31] and in stretched film measurements,[26,28] the transition moment direction for the first peak is chosen in accord with the earliest crystal study.[30] At about 50 000 cm^{-1}, a strong $\pi \rightarrow \pi^*$ transition is predicted by all the theoretical methods and the direction is found to be between −47° and −25°, in good agreement with observations. Thus, theory gives a rather consistent picture of the electronic spectrum of cytosine.

Gould et al.[7] showed, theoretically, that two tautomeric forms of cytosine are accessible at room temperature; these are the keto and enol forms with the enol form favored by 4.7 kJ/mol over the keto form at the QCISD(T)/6-31G*//MP2/6-31G* level of theory (energy/basis set//geometry/basis set), which is in accord with observations in inert gas matrix,[32] while Colominas et al.[8] reported a value of 3.35 kJ/mol at the MP4/6-311++G-(d,p)//MP2/6-31G* level of theory. In a semiempirical investigation, Katritzky and Karelsson found that the keto form is stabilized more by a polar solvent than the enol form.[9] Colominas et al. arrived at the same conclusion in a more recent *ab initio* work on solvent effects on the tautomeric equilibrium for cytosine and guanine.[8] Thus, if only nonspecific solvent effects are accounted for, only the keto form will be found in water solution. We will discuss the importance of specific solvent interactions on the enol−keto tautomerism in the next subsection. However, substitution of the N_1 hydrogen with a methyl group blocks this tautomerism. We have calculated the MP2/6-31G* geometry and the INDO/S spectrum of the enol tautomer; see Table 5. The transition energies are not very different between the keto and enol forms, but the directions for the first two $\pi \rightarrow \pi^*$ transitions are certainly affected by the keto−enol tautomerism. In Table 4, two sets of transitions are presented for the keto form of cytosine: the first set corresponds to the vacuum calculation, and the second set of transitions is calculated including solvent effects using the SCRF method. The influence of the solvent on the first two transitions is minor. The third transition has mostly $n \rightarrow \pi^*$ character. Because of the pronounced nonplanarity of the amino group, the third transition shows intensity in the ring plane due to a weak coupling to $\pi \rightarrow \pi^*$ transitions. This third transition is blue shifted in a polar solvent and gains much intensity due to larger mixing with a $\pi \rightarrow \pi^*$ transition. These two states are closer in energy in a polar solvent than in vacuum or a nonpolar solvent. On the basis of these two calculations, we suggest that the cytosine absorption spectrum is due to four true $\pi \rightarrow \pi^*$ transitions and one transition with mixed character. The mixing occurs because of the nonplanarity of the molecule. Furthermore, the third band position is very sensitive to the environment. Both the INDO/S and the CIS/6-31G* calculations predict three $n \rightarrow \pi^*$ transitions with energies less than 50 000 cm^{-1}. At the CASPT2 level, Fülscher et al.[3] calculated an $n \rightarrow \pi^*$ transition which they assigned to the out-of-plan polarized intensity at 43 000 cm^{-1} hinted in the work of Zaloudek et al.[29] They suggested that the second transition is due to the $n \rightarrow \pi^*$ transition. We suggest that an $n \rightarrow \pi^*$ transition mixes with a $\pi \rightarrow \pi^*$ state and becomes the third state. Both calculations and experiments seem to be in agreement with the fact that three transitions are necessary to describe the absorption spectrum in the 40 000−50 000 cm^{-1} region.

To investigate solvent effects on the geometry of cytosine, we optimized the geometry using the Onsager SCRF model with

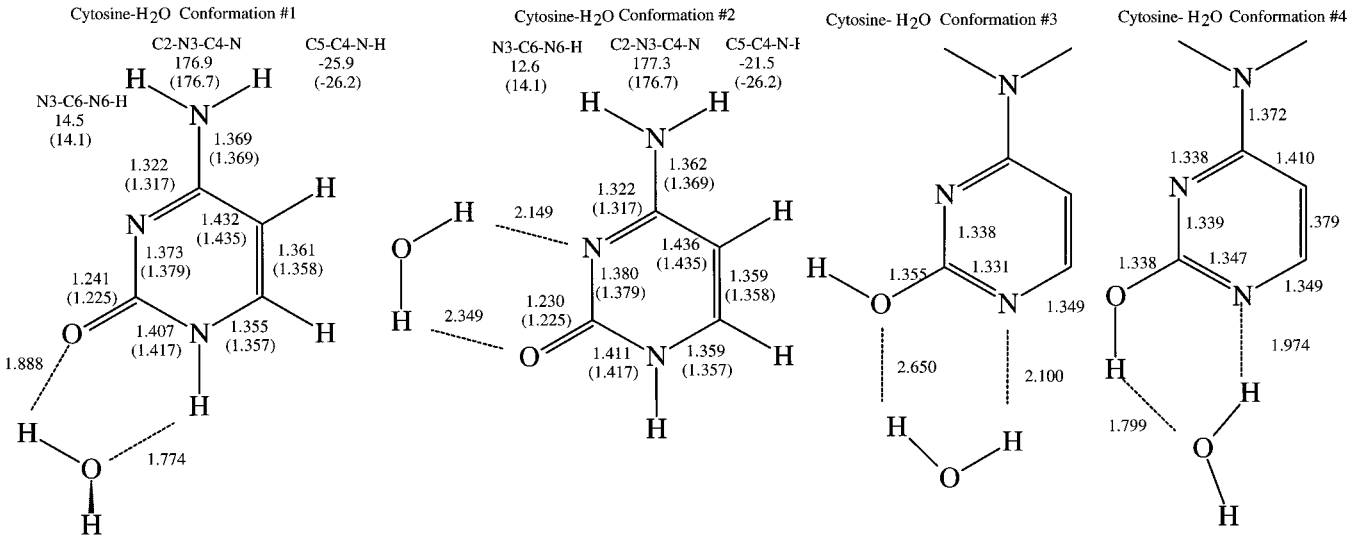

Figure 2. MP2/6-31G*-predicted geometries of the cytosine−H_2O complexes. Numbers within parentheses correspond to the gas-phase MP2/6-31G* geometry of cytosine.

the hybrid density functional model B3LYP with the 6-31G* basis set. Unfortunately, at present there are no analytical gradients implemented in the Gaussian94 program for the MP2/Onsager SCRF model making a geometry optimization extremely expensive. A direct comparison between the MP2 gas-phase geometry and the B3LYP solution geometry might be a little doubtful, but in a previous study on adenine and purine the two methods gave very similar gas-phase geometries.[10] The largest change is found in the amino group region. The solution geometry is much more planar than the gas-phase geometry; consequently, the C_4-N_8 bond length is shorter in solution. With the B3LYP-SCRF/6-31G* geometry the INDO/S calculation give $\pi \rightarrow \pi^*$ transition energies and polarizations that are very similar to the gas-phase geometry results. The largest difference is found for the $n \rightarrow \pi^*$ transitions that get much less in plane intensity.

Cytosine−Water Complex. To study the importance of specific solvent−solute interactions on the predicted absorption spectrum, we have optimized the geometry of four cytosine water conformers, two with the keto form and two with the enol form of cytosine. The fully MP2/6-31G*-optimized geometries of the cytosine−H_2O conformers are depicted in Figure 2. The conformers 1, 3, and 4 will be affected by substituting the N_1 hydrogen with a methyl group (1-methylcytosine). In conformer 1 two relative short hydrogen bonds are formed. The C−O bond and the N−H bond are elongated somewhat when compared to the gas-phase cytosine, while the N_1-C_2 bond is decreased. Otherwise, only minor changes are observed upon binding of one water molecule. For conformer 2, two hydrogen bonds are also formed. However, the hydrogen bonds are much longer than the hydrogen bonds in conformer 1. The amino group is predicted to be somewhat more planar in conformer 2 compared to free cytosine and conformer 1. Conformer 1 is found as the most stable cytosine(keto)−H_2O conformer, favored over conformer 2 with 22.4 kJ/mol. Furthermore, conformer 1 is also more stable than the two cytosine-(enol)−H_2O conformers but only favored with 4.47 kJ/mol over conformer 4. The relative energy difference between conformers 1 and 4 gives a relative population of 14% at 298 K assuming a Boltzmann distribution over the two states and neglecting differences in entropy. Evenmore, the dipole moment of the two conformers differs only by 0.25 D. This small dipole moment difference indicates that differences in ΔH_{solv} estimated from SCRF continuum models will be small. We conclude that it is likely that there are measurable amounts of the enol form of cytosine in water solution and stretched films. This will have a notable impact on the measured transition polarizations of cytosine. The absorption spectrum of the cytosine−H_2O conformers does not deviate much from the spectrum of cytosine; see Table 5. The directions are not changed by more than a few degrees by the specific solvent interactions. However, the weak third transition is affected by the water bonding. A somewhat smaller blue shift of the third peak is found in the reaction field calculations of the cytosine−H_2O complex. The intensity increase is not as dramatic for the complex as for the free keto form of cytosine.

1-Methylcytosine and Protonated 1-Methylcytosine. For 1-methylcytosine the enol−keto tautomerism is blocked by the methyl group. The effect due to the substitution on the absorption spectrum was investigated using the planar MP2/6-31G** geometry from the work by Stewart et al.[10] The predicted $\pi \rightarrow \pi^*$ transition energies of 1-methylcytosine are also shown in Table 5. Since the σ and π orbitals are not allowed to mix, due to the C_s-symmetry, the third transition of the nonplanar cytosine, which has some in-plane intensity, does not have any in-plane intensity for the 1-methylcytosine. Hence, only four $\pi \rightarrow \pi^*$ transitions are predicted. The energies and directions are about the same as predicted for the keto form of cytosine.

The crystal spectrum of protonated 1-methylcytosine has been reported by Clark.[31] We constructed the geometry of a protonated 1-methylcytosine by adding a proton to N_3, using the MP2/6-31G** geometry of 1-methylcytosine. The neighboring N_3-C_4 bond length was adjusted to allow for the change from a double to a single C−N bond. The relative change upon protonation seems to be consistent with the observed change in the absorption profile.[31] A considerable rotation of the transition moment directions is caused by protonation of the ring nitrogen (N_3).

In summary, the INDO/S method reproduces other theoretical results on the electronic spectrum of cytosine very well. An explanation of the apparent five $\pi \rightarrow \pi^*$ transitions is proposed. The theoretical transition moment directions for the first $\pi \rightarrow \pi^*$ transition do not agree well with the observed direction, but they agree rather well with the symmetry argument discussed in the beginning of this subsection.

Guanine. In guanine, the conjugation of the six-membered ring is broken by the carbonyl group, and we make the following

TABLE 6: INDO/S-Predicted $\pi \rightarrow \pi^*$ Transition Energies and Polarizations of Guanine Compared with Observed Spectra[a]

	INDO			CAST2 (ref 1b)			experiment		
	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}
					Guanine				
I	32.8	0.318	−58	36.2	0.20	−64	36.8[e]	0.15	**−24**/60
	(33.2)[b]	(0.326)	(−51)				35.7[d]		**4**/−61
							36.0[f]	0.16	**−4**/35
							36.4[g]		
II	37.8	0.317	52	42.7	0.09	52	40.6[e]	0.24	**88**/−56
	(37.2)	(0.362)	(59)				40.3[d]		**−88**/31
							39.4[f]	0.25	−75
							39.7[g]		
M	41.0	0.023					44.0[f]	0.01	
	(41.2)	(0.014)					44.4[g]	0.03	
III	49.5[c]	0.113	40	45.4	0.05	−90	49.9[e]	0.40	**86**/−50
	49.7[c]	0.282	48				49.0[f]	0.41	−71/−79
	(49.5)	(0.333)	(48)				50.0[g]		
IV	53.6	0.427	−70	55.1	0.26	0	53.9[e]	0.48	−8/44
	(52.0)	(0.250)	(88)				53.0[f]	0.50	−9/41
							53.5[g]		
	guanine(enol)			guanine−H_2O			guanine(enol)−H_2O		
I	33.9	0.15	47	33.0	0.33	−57	34.1	0.14	26
II	37.3	0.40	−83	37.4	0.35	54	37.4	0.42	80
M	43.3	0.02		41.2	0.02		43.4	0.01	
III	48.4	0.73	−31	49.3	0.21	46	48.6	0.74	−49
IV	51.6	0.67	66	52.6	0.30	59	51.7	0.67	50

[a] Bold numbers indicate the selected transition moment direction in the experimental work. Energies are given in units of 1000 cm^{-1}. The predicted polarizations by the simple fragment model are 65°, 35°, −10°, and −60°. [b] The number within parentheses corresponds to SCRF calculations that includes the solvent effects. [c] An n $\rightarrow \pi^*$ transition is mixing with the $\pi \rightarrow \pi^*$ transition and gives rise to two transitions with mixed character. The accidental degeneracy is easily removed by a small perturbation by a reaction field or a solvent molecule hydrogen bonded to the carbonyl oxygen. [d] Stretched film LD measurements, reference 26. [e] Single-crystal absorption spectrum of guanosine and aqueous solution spectrum, reference 33. [f] Single-crystal absorption spectrum of 9-ethylguanine and guanine·HCl, reference 38. [g] CD spectrum of deoxyguanosine-5′-phosphate, reference 23.

partitioning into π type fragments. The first fragment contains the carbonyl group, the second fragment contains C_2 and N_3, and the third fragment contains N_7 and C_8. The preferred directions will then be 65°, 35°, −10°, and −60°. The calculated $\pi \rightarrow \pi^*$ transition energies and polarizations of guanine are summarized in Table 6. Again, there is good agreement between the INDO/S transition moment directions and those obtained with the simple fragment model.

Both CD measurements[23] and single-crystal absorption measurements[33] on guanine have indicated that five $\pi \rightarrow \pi^*$ transitions are present in the region between 36 000 and 54 000 cm^{-1}. Clark reported also a weak n $\rightarrow \pi^*$ transition at about 33 300 cm^{-1} in 9-ethylguanine.[33] However, CD measurements on guanosine derivatives do not indicate any n $\rightarrow \pi^*$ transitions. In the crystal spectrum of 9-ethylguanine, a weak band at about 44 000 cm^{-1} was reported by Clark. The CD measurements do not contradict the assignment of a $\pi \rightarrow \pi^*$ transition in that wave length region.

The INDO/S calculation also give five $\pi \rightarrow \pi^*$ transitions in that energy region. The first two transitions are calculated too low in energy compared to the observed spectra. Clark[31,33] noted that crystal effects on the transition moment directions could be as large as 12°. Theiste, Callis, and Woody[34] included crystal effects in a INDO/S calculation and observed that the transition moment directions were changed rather dramatically by the crystal field. The effect of the crystal field was that the lowest unoccupied molecular orbital (LUMO) and LUMO + 1 changed order compared to the gas-phase ordering of the MOs. In the gas phase both the first and second transitions were dominated by one configuration. In the crystal-field calculation, the two configurations are mixed in a symmetric and an unsymmetric combination. Thus, the resulting transition moment direction is given by a weighted linear combination of the two gas-phase-calculated directions (−44 and 57). It is possible to reproduce this effect by applying a static electric field along the long axis of the molecule. It is thus very important that the energy separation between the first two transitions is well reproduced to get a good description of the transition moment directions. In our gas-phase spectra the separation is 5000 cm^{-1}, whereas the observed separation is only between 3400 and 3700 cm^{-1}. The SCRF calculation reduces this gap to 4000 cm^{-1}. Only a small change in the directions is induced by the reaction field, but the change is in the same direction as the rotation of the transition moments by a static electric field or a crystal field. Increasing the solvent influence by using an artificial small solute cavity in the SCRF calculation, we obtain an almost identical rotation of the transition moments as Theiste et al.[34] in the crystal-field simulation. We conclude that a rather large perturbation is necessary to obtain a configuration mixing that leads to a rotation of the transition moment directions. A somewhat more serious crystal effect is the dimer interaction between two neighboring chromophores. We have estimated this effect by calculating the INDO/S spectrum for several guanine dimers constructed from the MP2/6-31G* geometry. The relative orientation between the two chromophores was varied. Not surprisingly, the largest effect on the spectrum was obtained with the dimer where the two chromophores were stacked on top of each other in an antiparallel way. The first two excited states change character, and a new configuration interaction pattern is obtained. Thus, depending on the relative orientation, different transition moment directions are obtained. Unfortunately, this type of stacking is just what is observed in crystals of guanosine. We conclude that crystal effects are probably very large in guanine crystals and that the solid state reflection spectrum is not representative for the gas-phase spectrum.

The keto−enol tautomerism discussed in connection with cytosine is also possible for guanine. This tautomerism is not excluded by substitutions in any of the experimental works that are referred to in this work. In the gas phase the enol form of guanine has been calculated to be very close in energy to the keto form of guanine.[7] The keto form is favored by only 5.4 kJ/mol,[7] which gives a relative population of the enol form of 11% at 298 K from a Boltzmann distribution over the two states assuming that there is no contribution from differences in entropy. In inert gas matrices both forms are observed, with the enol form dominating in N_2 matrix.[35,36] For 9-methylguanine the enol form is observed to be even more dominant than in guanine.[36] However, the keto form is expected to be the most stable form in solution due to the larger ground state dipole moment compared with the ground state dipole moment of the enol form. Colominas et al.[8] concluded that the keto form was favored by about 25 kJ/mol over the enol form in water solution. Specific solvent interactions act in the same way for the keto−enol tautomerism of guanine as for cytosine. In analogy with the cytosine case we have optimized three conformers of the guanine−H_2O complex, two with the enol form and one with the keto form. The important bond lengths are reported in Figure 3. The keto form of the guanine−H_2O complex is found to be the most stable tautomer, favored by 12.9 kJ/mol over conformer 1 of the enol form and by 5.4 kJ/mol over conformer 2 of the enol form. The relative population of the two enol conformers is 0.5% and 10%, respectively. The dipole moment difference between the keto form and conformer 1 of the enol

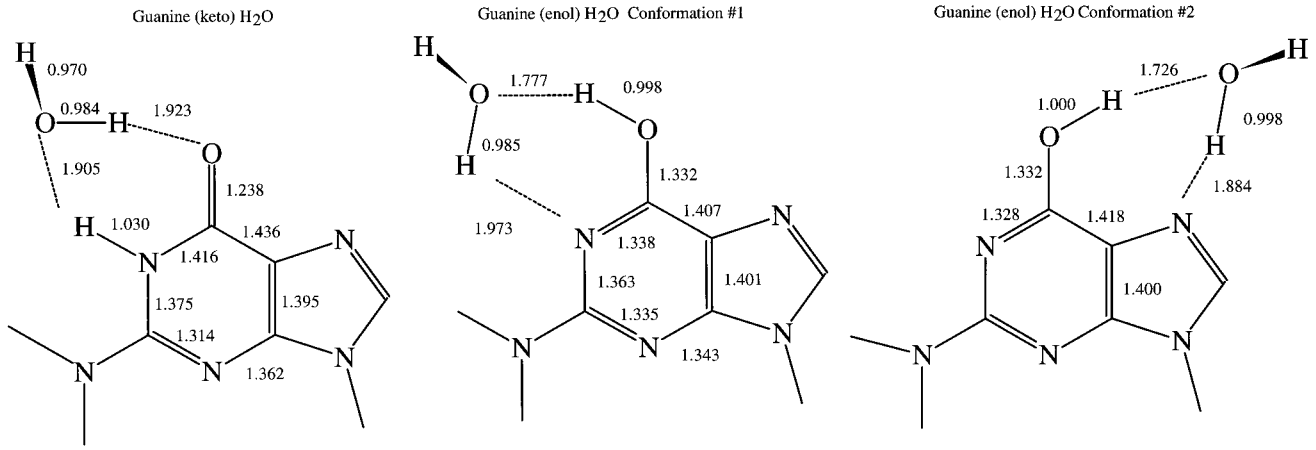

Figure 3. MP2/6-31G*-predicted geometries of the guanine−H_2O complexes.

form is about 1.5 D, with the largest dipole moment for the keto form. The dipole moment difference between the keto form and conformer 2 of the enol form is only 0.4 D, which indicates that the effect of the solvent continuum should be about the same for the two tautomers. Thus, we expect that small amounts of the enol form of guanine will be found in hydrogen-bonding polar solvents.

If more than one tautomer is present when an experimental spectrum is recorded, the resulting spectra will be a weighted average of the species. The 7*H* ↔ 9*H* tautomerism is also important which further complicates the analysis of the observed transition moment directions for guanine but not for 9-ethylguanine. Colominas et al.[8] showed that it is likely that both tautomers are equally stable in water solution (without specific solvent interactions). It is evident from the work by Matsuoka and Nordén[26] and by Clark[33] that the band form absorption spectrum is sensitive to substitution in the N_9 position. In water solution of guanine the 7*H* and 9*H* tautomers are both observed and calculated to be present in equal amounts.[37,8] Thus, we can only rely on experiments where this tautomerism is blocked.

To get a good agreement between the INDO/S-predicted polarizations for the keto form and the experimental selected directions, we have to rotate the transition moment directions by about 35°. In Table 6, the predicted $\pi \rightarrow \pi^*$ transitions of the enol form of guanine and the corresponding water complex are reported. The transition energies of the keto−enol tautomerism are not affected very much by the specific solvent interactions, but the polarizations of the transitions are certainly changed.

In all INDO/S calculations, the first $n \rightarrow \pi^*$ transition is predicted to be very close to the first $\pi \rightarrow \pi^*$ transition. The keto form is predicted to have a $n \rightarrow \pi^*$ transition at about 33 100 cm^{-1} ($f_{osc} = 0.02$), in good agreement with what Clark reported for the 9-ethylguanine crystal.[38]

Adenine. The six-membered ring of adenine is close to having local C_{2v} symmetry. The geometry of both the six-membered and five-membered rings deviates only very little from the geometry of purine. We make the same partitioning into fragments as for purine. Thus, we expect to get transition moment directions in adenine that are very similar to those of 9*H*-purine.

The CD spectrum of deoxyadenosine-5′-phosphate resolves seven transitions with energies below 54 000 cm^{-1}.[23] Five transitions are assigned to $\pi \rightarrow \pi^*$ transitions. Between 37 000 and 41 000 cm^{-1} Sprecher and Johnson[23] found two $\pi \rightarrow \pi^*$ transitions and a $n \rightarrow \pi^*$ transition. The weak absorption at 43 500 cm^{-1} was assigned to a $n \rightarrow \pi^*$ transition. The analysis of the observed absorption spectrum of adenine is rather complicated due to the two almost overlapping transitions in the first absorption band. These two states have been addressed as accidentally degenerate. In the first published work on the transition moment directions for adenine, Stewart and Davidson assumed only one transition in the first band.[39] Their analysis gave two alternative directions: −3° and 45°. The first angle was selected since that choice agreed best with the observed adenine:thymine absorption spectrum.[39] Later, Clark[40] reanalyzed the adenine transition moment directions by comparing the crystal spectrum of 9-methyladenine with the crystal spectrum of 6-(methylamino)purine. He reassigned the first band by arguing that two transitions are needed to explain the band shape. The complete spectrum was later assigned by Clark.[40,41] In almost all the crystal work, not only for adenine, the position of the band maximum is not found at the same energy when recording the spectra along the different crystal axes. The peak position difference is some times in the order of several hundred wave numbers.[34,40] Furthermore, to obtain an oscillator strength for each individual transition, some type of band fitting is necessary. In the case of adenine where two transitions are almost overlapping, the band fitting is rather complicated, and the assumed oscillator strengths have a large degree of uncertainty. The measurements by Matsuoka and Nordén of adenine oriented in stretched PVA film have been questioned, and a new investigation of the transition moment directions in adenine is under way.[42]

Part of the motivation for this work is to try to clarify the ambiguity of the assignment of the transition moment directions for adenine. As for the other DNA bases, several theoretical works on the transition moment directions for adenine have been published.[2,5] Most of the earlier theoretical works are in reasonably good agreement with the INDO/S spectrum for adenine reported here but in less good agreement with the experimental assignments of the transition moment directions; see Table 7. At around 43 500 cm^{-1} we calculate a $\pi \rightarrow \pi^*$ transition, but also some $n \rightarrow \pi^*$ transitions are found in that energy region. The disagreement between the transition moment directions determined by Clark[40] and the gas-phase INDO/S directions has been attributed by Callis and co-workers[43] to crystal effects leading to strongly perturbed crystal transitions. However, Clark has estimated that these effects should be relatively small and in contrast to the large perturbations needed to obtain an improvement.[40] The transition moment directions displayed in Table 7 are almost unchanged by external perturbation with a reaction field. Furthermore, specific hydrogen bonding to solvent water molecules does not alter the reported directions in a significant way. In similarity with the purine spectrum, the first two excited states are composed of two

TABLE 7: INDO/S-Predicted $\pi \rightarrow \pi^*$ Transition Energies and Polarizations of Adenine Compared with Observed Spectra[a]

	INDO			experiment		
adenine	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}
I	35.2	0.12	38	36.5[b]	0.09	67
				36.4[c]	0.10	83
				37.0[d]		
II	37.9	0.31	75	39.4[b]	0.18	35
				37.7[c]	0.20	25
				40.8[d]		
				43.5[d,e]		
III	43.9	0.06	−16	46.9[b]	0.25	−45
				46.9[c]	0.25	−45
				46.3[d]		
IV	48.5	0.80	−59	49.5[b]	0.11	15
				49.0[c]	0.11	15
				51.3[c]		
V	51.0	0.47	67	52.9	0.33	64

[a] Energies are given in units of 1000 cm^{-1}. The predicted polarizations by the simple fragment model are 50°, 34°, −40°, and −80°e [b] Single-crystal absorption spectrum of adenosine, reference 41. [c] Single-crystal absorption spectrum of 9-methyladenine and N^6-(methylamino)purine, reference 40. [d] CD spectrum of deoxyadenosine-5′-phosphate, reference 23. [e] Assigned as an n $\rightarrow \pi^*$ transition in reference 23.

configurations. In order to alter the computed transition moment directions, it is necessary to affect this mixing by an external perturbation. If a static electric field is applied along the long axis of the molecule, the two accidental degenerate states start to separate. At high field strengths the two states interact very little. In the noninteracting limit the transition with the largest oscillator strength is forming an angle of about 65° while the polarization of the other, now very weak, transition is very field dependent.

Petke et al.[5] reported MRCI and RPA calculations on the electronic spectrum of adenine (and guanine) which agree reasonably well with the present INDO/S spectrum. However, they used scaled transition energies to get qualitative agreement with observed transition energies. The scaling is necessary since the dynamic electron correlation is not sufficiently accounted for in the MRCI and RPA methods. Both methods predict a weak and a stronger $\pi \rightarrow \pi^*$ transition that contributes to the first observed band. The *ab initio* methods predict a rather small oscillator strength of the first transition.[5] Due to the small oscillator strength, the predicted polarization direction is uncertain. The direction of the stronger second transition is predicted to be 41° (MRCI) and 51° (RPA). Noteworthy is that the RPA method predicts a larger oscillator strength for the first $\pi \rightarrow \pi^*$ transition, in better accord with the INDO/S result, and the direction of the second peak is rotated toward the direction predicted by the INDO/S method. The MRCI results are in better accord with the original assignment of the first observed band.[39] The polarization of the peak found around 47 000 cm^{-1} is predicted to be between −59° and −68° by all theoretical methods, which is in agreement with observations (−45°).

***Ab Initio* Predicted Spectra of the DNA Bases.** Due to lack of sufficient computer resources, we were not able to perform *ab initio* calculations on transition energies for all the tautomers included in the semiempirical study. For validation of the INDO/S model, *ab initio* CIS calculations were performed for one tautomer of each DNA base using the MP2/6-31G* optimized geometries. It Table 8, the CIS spectra are compared with the INDO/S spectra and the observed crystal spectra. In an *ab initio* CIS treatment of the absorption spectrum of a molecule, no dynamic electron correlation is introduced in either of the states. Apparently, the CIS method overestimates the transition energies due to the different size of the dynamic electron correlation energy in the ground state and the excited states. If we assume that the difference in electron correlation between the ground state and the excited states is constant, a linear scaling of the calculated CIS/6-31G* energies can be applied. We have used a scaling factor of 0.72 on the computed transition energies to compare the CIS energies with the INDO/S and the observed transition energies. The scaled CIS energies and polarizations agree well with the INDO/S-predicted values and experimentally determined values, with the same exceptions as for the INDO/S predictions.

In Table 9, we report CASPT2-computed $\pi \rightarrow \pi^*$ transition energies and polarizations of adenine. In the 35 000−54 000 cm^{-1} energy region we calculate six transitions: one with Rydberg character and the other five with valence character. Between 35 000 and 40 000 cm^{-1} we found two $\pi \rightarrow \pi^*$ transitions in all calculations, in accord with observations. The polarization of the strongest transition is not calculated consistently by the different methods. It seems that the interaction between the two states is somewhat larger in the INDO/S calculation than in the *ab initio* calculations. As mentioned earlier, if the interaction between the two states is reduced by an external perturbation, the INDO/S direction for the strongest transition is rotated in such a way that the polarization is more like what we obtained from the *ab initio* calculations and in better accord with observations. With this in mind we look closer at the CASPT2 results and note that at the CASSCF step of the CASPT2 calculation the two transitions are much more separated than after the correlation energy correction. The transition moments are calculated using the CASSCF wave function and the CASPT2 energies. Thus, due to the large energy splitting of the two states at the CASSCF level, the character of the states does not correspond to an accidentally degenerate case but rather to a nondegenerate case. In other words, the computed transition moment directions of the two states do not correspond to the accidentally degenerate situation that the PT2 energies indicate. We conclude that states that are found to be close in energy at the PT2 level must be close in energy also at the CASSCF level, in order to calculate reliable polarizations.

With that knowledge we turn back to the guanine spectrum. The polarizations of the first two transitions are calculated to be about the same with all methods. The agreement between different methods is not that good for the two remaining transitions with large oscillator strength (transitions III and IV). The INDO/S and The CIS/6-31G* calculations suggest that at least three configurations are important to describe the two excited states. The CASPT2 method overestimates the excited state separation compared to the observed spectrum for guanine. Consequently, the computed transition moment directions have a larger degree of uncertainty compared to the first two transitions. We believe that the necessary restriction of the active space at the CASSCF level of the calculation introduces too large errors in the zeroth-order description of the wave function so that the final properties of the excited states are not well reproduced by the CASPT2 method in cases of near-degenerate excited states.

Conclusions

The $\pi \rightarrow \pi^*$ transition energies of the DNA bases obtained with the INDO/S model, using geometries obtained at the MP2/6-31G* level of theory, reproduce observed transitions with an average error in energy of 0.3 eV or less. The corresponding transition moment directions are in most cases reproduced within

TABLE 8: Comparision between Observed and INDO/S- and CIS/6-31G*-Calculated $\pi \rightarrow \pi^*$ Transition Energies and Polarizations of the DNA Bases[a]

	INDO			CIS			observed		
	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}	E	f_{osc}	Θ^{DT}
					Uracil				
I	41.1	0.36	−3	53.4 (38.5)	0.46	−6	37.1[b]	0.20	**−9**/16
II	48.3	0.21	82	68.9 (49.6)	0.13	44	47.0	0.26	**59**/−53
III	51.8	0.16	−17	71.8 (51.7)	0.40	−60	51.6	0.16	
IV	56.1	0.32	−21	77.7 (55.9)	0.25	−10	56.5		
					Cytosine				
I	36.0	0.17	46	49.5 (35.6)	0.17	30	37.1[c]	0.14	**6**/54
II	44.9	0.16	−15	62.2 (44.8)	0.32	−16	42.7	0.03	−35
III	47.2	0.007		69.9 (50.3)	6×10^{-4}	$n \rightarrow \pi^*$	45.2	0.13	**76**/−17
IV	49.5	1.045	−39	65.9 (47.5)	0.75	−45	50.0	0.36	86/−26
V	53.8	0.09	32	73.9 (53.2)	0.06	36	53.5		
					Guanine				
I	32.8	0.32	−58	49.4 (35.6)	0.29	−46	36.8[d]	0.15	**−24**/60
II	37.8	0.32	52	55.9 (40.3)	0.50	63	40.6	0.24	**88**/−56
M	41.0	0.02		64.6 (46.4)	0.07	50	44.0	0.01−0.03	
III	49.7	0.28	48	71.9 (51.7)	0.03	−33	49.9	0.40	**86**/−50
IV	53.6	0.43	−70	73.0 (52.5)	0.46	72	53.9	0.48	−8/44
					Adenine				
I	35.2	0.12	38	51.5 (37.1)	0.41	38	36.5[e]	0.09	67
II	37.9	0.31	75	52.0 (37.5)	0.03	−72	39.4	0.18	35
III	43.9	0.06	−16	64.1 (46.2)	0.28	43	46.9	0.25	−45
IV	48.5	0.80	−59	67.6 (48.7)	0.49	−80	49.5	0.11	15
V	51.0	0.47	67	69.0 (49.7)	0.59	8	52.9	0.33	64

[a] Bold numbers indicate the selected transition moment direction in the experimental work. Energies are given in units of 1000 cm^{-1}. [b] Single-crystal absorption spectrum of 1-methyluracil, reference 25. [c] Single-crystal absorption spectrum of cytosine, reference 29. [d] Single-crystal absorption spectrum of guanosine and aqueous solution spectrum, reference 33. [e] Single-crystal absorption spectrum of adenosine, reference 41.

TABLE 9: CASPT2-Computed $\pi \rightarrow \pi^*$ Transition Energies and Polarizations of Adenine Compared with Observed Transition Energies and Polarizations[a]

	E_{CAS}	E_{PT2}	f_{osc}	Θ^{DT}	ω^b	obsd[c]	f_{osc}	Θ^{DT}
I	52.2	38.1	0.236	43	0.67	36.4	0.09	67
II	42.4	38.2	0.001	−56	0.68	37.7	0.18	35
II	64.5	47.0	0.307	−38	0.60	46.9	0.25	−45
IV	61.0	47.6	0.167	−40	0.29	49.0	0.11	15
Rydberg	48.5	50.6	0.001	−89	0.70			
V	67.7	53.5	0.055	89	0.45	52.9	0.33	64

[a] Energies are given in units of 1000 cm^{-1}. [b] ω is the relative weight of the CAS reference in the perturbative expansion. A small number indicates that the final state interacts with so-called intruder states, and the final energy and transition polarization are somewhat uncertain. [c] Single-crystal absorption spectrum of adenosine, reference 41.

±30° of experimentally determined directions. In a few cases the agreement is not within these limits. In such cases we have shown that tautomerism and environmental effects can account for the discrepancy.

The $\pi \rightarrow \pi^*$ transition energies obtained at the CIS/6-31G* level of theory agree well with observed energies and the INDO/S energies if a linear scaling of 0.72 is applied to the computed transition energies. In most cases the polarizations are in good agreement with INDO/S polarizations. The CASPT2-computed transition energies of adenine agree well with the observed spectrum. The accidental degeneracy of the first two excited states of adenine is not reproduced at the CASSCF level of calculation, and since the CASSCF wave function is used to compute the oscillator strengths at the PT2 level, we do not expect that the reported transition polarization is reliable. Furthermore, we note that this problem might affect the computed polarizations reported by Roos et al.[1b,3a,4] for the other DNA bases.

Simple geometry arguments and substitution effects on the predicted transition moment directions using pyrimidine as a starting point have proven to be useful to rationalize most of the calculated and observed transition moment directions. This very simple model predicts transition moment directions with an uncertainty of less than ±30°.

For cytosine, we have shown that the spectrum between 35 000 and 51 000 cm^{-1} is due to three $\pi \rightarrow \pi^*$ transitions and one transition with mixed character. The position of the mixed transition (III) is very sensitive to the environment. The

B3LYP-SCRF/6-31G* calculation of the geometry of cytosine indicates that the pronounced nonplanarity of the amino group found in the gas-phase-optimized geometry is not necessarily expected to be found in solution.

The electronic spectra of thymine and uracil are calculated in good agreement with experiments. The relative shift upon methyl substitution is also reproduced. The INDO/S and CIS/6-31G* spectra agree well with earlier *ab initio* predictions. Tautomerism is not important for uracil and thymine. A rather large shift of the second band was shown to occur in solution compared to gas phase.

The observed $\pi \rightarrow \pi^*$ transition energies and polarizations of guanine are reproduced with less good accuracy than for the other DNA bases. The predicted transition moment directions are in agreement with observations if a 35° rotation is applied to the theoretical directions. However, it is likely that environmental effects and possible tautomerism are not sufficiently accounted for in the analysis of the observed data.

For adenine, the calculated transition moment directions do not agree very well with the polarizations determined in the solid phase. The *ab initio* calculated polarizations agree better with the observed polarizations than the INDO/S polarizations. Transitions I + II and III + IV are rather close in energy. This near-degenerate situation is rather hard to reproduce with all the computational methods used in this work. We believe that the exact mixing of individual components is rather sensitive to the surrounding. Thus, to reproduce the observed polarizations it is necessary to includ crystal effects or solvent effects in a high-level *ab initio* calculation.

From previous calculations and observations it is evident that the DNA bases are not entirely planar, leading to some in-plane intensity for the $n \rightarrow \pi^*$ transitions. The in-plane component of the transition moment is perhaps not large in most cases but might be of importance for weak transitions. The major effect of the nonplanarity of the molecules is that the transitions will have mixed character.

It has been shown by model calculations that the crystal effects could be rather large.[34,43] However, in order to reproduce these rather large effects by a simple static electric field, we have to apply very large electric fields. Specific solvent interactions do not affect the directions in a considerable way, but the tautomeric equilibrium is affected. The general impression is that the calculated results agree better with stretched film measurements than with crystal measurements.

For all the reported molecules in this work, usually one $n \rightarrow \pi^*$ transition is calculated to be close in energy to the reported $\pi \rightarrow \pi^*$ transition energies. Most of these $n \rightarrow \pi^*$ transitions have not been observed since the much stronger $\pi \rightarrow \pi^*$ transition overlaps with the $n \rightarrow \pi^*$ transitions.

Acknowledgment. This work was financed by grants from the Swedish Natural Science Research Council (NFR). The National Supercomputer Center (NSC) in Linköping, Sweden, and the Parallel Computer Center (PDC) in Stockholm, Sweden, are acknowledged for supply of computer time.

References and Notes

(1) (a) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218. (b) Roos, B. O.; Fülscher, M. P.; Malmqvist, P.-Å.; Merchán, M.; Serrano-Andrés, L. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dortrecht, The Netherlands, 1995.

(2) Callis, P. R. *Annu. Rev. Chem.* **1983**, *34*, 329; *Photochem. Photobiol.* **1986**, *44*, 315.

(3) (a) Fülscher, M. P.; Roos, B. O. *J. Am. Chem. Soc.* **1995**, *117*, 2089. (b) Petke, J. D.; Maggiora, G. M.; Christoffersen, R. E. *J. Phys. Chem.* **1992**, *96*, 6992. (c) Jensen, H. J. A.; Koch, H.; Jorgensen, P.; Olsen, J. *Chem. Phys*, **1988**, *119*, 297. (d) Matos, J. M.; Roos, B. O. *J. Am. Chem. Soc.* **1988**, *110*, 7664.

(4) Lorentzon, J.; Fülscher, M. P.; Roos, B. O. *J. Am. Chem. Soc.* **1995**, *117*, 9265.

(5) Petke, J. D.; Maggiora, G. M.; Christoffersen, R. E. *J. Am. Chem. Soc.* **1990**, *112*, 5452.

(6) Norinder, U. *THEOCHEM* **1987**, *151*, 259. Sabio, M.; Topiol, S.; Lumma, W. C., Jr. *J. Phys. Chem.* **1990**, *94*, 1366.

(7) Gould, I. R.; Hillier, I. H. *Chem. Phys. Lett.* **1989**, *161*, 185. Gould, I. R.; Burton, N. A.; Hall, R. J.; Hillier, I. H. *J. Mol. Struct. Theochem* **1995**, *331*, 147.

(8) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811.

(9) Katritzky, A. R.; Karelson, M. *J. Am. Chem. Soc.* **1991**, *113*, 1561.

(10) Broo, A.; Holmén, A. *Chem. Phys.* **1996**, *211*, 147. Stewart, E. L.; Foley, G. K.; Allinger, N. L.; Bowen, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 7282.

(11) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J.A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian94; Gaussian, Inc.: Pittsburgh, PA, 1995.

(12) The 6-31G* basis set: Hariharan, P. C.; Pople, J. A. *Theoret. Chim. Acta* **1973**, *28*, 21.

(13) ZINDO; M. C. Zerner, Quantum Theory Project, University of Florida, Gainesville, FL. Ridley, J. E.; Zerner, M. C. *Theoret. Chim. Acta* **1973**, *32*, 111; **1976**, *42*, 223.

(14) Andersson, K.; Fülscher, M. P.; Karlström, G.; Lindh, R.; Malmquist, P.-Å.; Olsen, J.; Roos, B. O.; Sadlej, A. J.; Blomberg, M. R. A.; Siegbahn, P. E. M.; Kellö, V.; Noga, J.; Urban, M.; Widmark, P.-O. MOLCAS version 3.

(15) Karelson, M. M.; Zerner, M. C. *Int. J. Quant. Chem. Symp*. **1986**, *20*, 521; *J. Phys. Chem*. **1992**, *96*, 8991. Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486.

(16) Devoe, H.; Tinoco, I. *J. Mol. Biol.* **1962**, *4*, 518.

(17) Mc Mullan, R. K.; Benci, P.; Craven, B. M. *Acta Crystallogr*. **1980**, *B36*, 1424.

(18) Holmén, A.; Broo, A. *Int. J. Quant. Chem.* **1995**, *QBS22*, 113.

(19) Hug, W.; Tinoco, I. *J. Am. Chem. Soc.* **1973**, *95*, 2803.

(20) Albinsson, B.; Nordén, B. *J. Am. Chem. Soc.* **1993**, *115*, 223.

(21) Clark, L. B.; Peschel, G. G.; Tinoco, I., Jr, *J. Phys. Chem.* **1965**, *69*, 3615.

(22) Becker, R. S.; Kogan, G. *Photochem. Photobiol.* **1980**, *31*, 5.

(23) Sprecher, C. A.; Johnson, W. C., Jr. *Biopolymers* **1977**, *16*, 2243.

(24) Fujii, M.; Tamura, T.; Mikami, N.; Ito, M. *Chem. Phys. Lett.* **1986**, *126*, 583.

(25) Novros, J. S.; Clark, L. B. *J. Phys. Chem.* **1986**, *90*, 5666.

(26) Matsuoka, Y.; Nordén, B. *J. Phys. Chem.* **1982**, *86*, 1378.

(27) Holmén, A.; Broo, A.; Albinsson, B. *J. Phys. Chem*. **1994**, *98*, 4998.

(28) Bott, C. C.; Kurucsev, T. In *Molecular Optical Dichroism and Chemical Applications of Polarized Spectroscopy*; Nordén, B., Ed.; Lund University Press: Lund, Sweden, 1977; p 81.

(29) Zaloudek, F.; Novros, J. S.; Clark, L. B. *J. Am. Chem. Soc.* **1985**, *107*, 7344.

(30) Callis, P. R.; Simpson, W. T. *J. Am. Chem. Soc.* **1970**, 92, 3593.

(31) Clark, L. B. *J. Am. Chem. Soc.* **1986**, *108*, 5109.

(32) Szczesniak, M.; Szczesniak, K.; Kwiatkowski, J. S.; Kubulat, K.; Person, W. B. *J. Am. Chem. Soc.* **1988**, *110*, 8319.

(33) Clark, L. B. *J. Am. Chem. Soc*. **1977**, *99*, 3934.

(34) Theiste, D.; Callis, P. R.; Woody, R. W. *J. Am. Chem. Soc.* **1991**, *113*, 3260.

(35) Sheina, G. G.; Stepanian, S. G.; Radchenko, E. D.; Blagoi, Y. P. *J. Mol. Struct.* **1987**, *158*, 275.

(36) Sowers, L. G.; Fazakerley, G. V.; Eritja, R.; Kaplan, B. E.; Goodman, M. F. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 5434.

(37) Pfleiderer, W. *Liebigs Ann. Chem.* **1961**, *647*, 167.

(38) Clark, L. B. *J. Am. Chem. Soc.* **1994**, *116*, 5265.

(39) Stewart, R. F.; Davidson, N. *J. Chem. Phys.* **1963**, *39*, 255.

(40) Clark, L. B. *J. Phys. Chem*. **1989**, *93*, 5345; **1990**, *94*, 2873.

(41) Clark, L. B. *J. Phys. Chem.* **1995**, *99*, 3366.

(42) Holmén, A.; Broo, A.; Albinsson, B.; Nordén, B. Manuscript not published.

(43) Sreerama, N.; Woody, R. W.; Callis, P. R. *J. Phys. Chem.* **1994**, *98*, 10397.