stitution in the perpendicular allyl cation XXIII provides a slight stabilization (−1.9 kcal mol^{-1}, eq 12), but 1-methyl substitution produces a much larger effect in going from III to XXIV (−20.4 kcal mol^{-1}, eq 13). Stereomutation of XXII should proceed through the 1-methylcyclopropyl cation (XXIV) since methyl substitution favors path B over path A by 18.5 kcal mol^{-1}.

Electron releasing substituents, R'', which stabilize carbonium ions to a greater extent than methyl should favor path B even more. In the extreme such substituents might even render the 1-substituted cyclopropyl cations more stable than their 2-substituted allyl counterparts. From known thermochemical data[59] and theoretical stabilization energies of substituted methyl cations,[60] it would appear that methoxy, hydroxy, and amino groups should be such substituents. Abundant experimental evidence is available already. Many cyclopropane substitutions are known involving 1-RO- and 1-R_2N-cyclopropyl cation intermediates; these proceed *without* ring opening.[61] The stable 1-dimethylaminocyclopropyl cation has been observed directly.[62] In addition, reactions involving cyclopropyl cations stabilized by 1-aryl,[63] 1-cyclopropyl,[64] 1-alkenyl,[65] and 1-thiophenoxy[55b] groups are known which proceed with only partial ring opening. However, despite attempts,[66] no cases of closure of 2-substituted allyl cations to 1-stabilized cyclopropyl cations have been discovered yet.[67] The problems appear to be practical rather than thermodynamic.

Acknowledgment. We thank Professor J. Hauser (Duquesne University) for helpful discussions. This work was supported in Princeton by National Science Foundation Grant GP 29078X, the Petroleum Research Fund, administered by the American Chemical Society, and Hoffmann-LaRoche, Inc., Nutley, N. J. Work at Carnegie–Mellon was supported by the National Science Foundation under Grant GP25617.

(59) R. H. Martin, F. W. Lampe, and R. W. Taft, *J. Amer. Chem. Soc.*, **88**, 1353 (1966).
(60) J. A. Pople, submitted for publication.
(61) W. J. M. van Tilborg, S. E. Schaafsma, H. Steinberg, and Th. J. deBoer, *Recl. Trav. Chim. Pays-Bas*, **86**, 417 (1967); J. Szmuskovica, D. J. Duchamp, E. Cerda, and C. G. Chidester, *Tetrahedron Lett.*, 1309 (1969); H. H. Wasserman and M. S. Baird, *ibid.*, 1729 (1970), 3721 (1971); W. J. M. van Tilborg, G. Dooyewaard, H. Steinberg, and Th. J. deBoer, *ibid.*, 1677 (1972); a case which may involve a 1-fluorocyclopropyl cation is also known: P. Weyerstahl, G. Blume, and C. Miller, *ibid.*, 3869 (1971).
(62) E. Jongejan, W. J. M. van Tilborg, Ch. H. V. Dussen, H. Steinberg, and Th. J. deBoer, *Tetrahedron Lett.*, 2359 (1972).
(63) D. B. Ledlie and E. A. Nelson, *Tetrahedron Lett.*, 1175 (1969); D. T. Clark and F. Smale, *Chem. Commun.*, 868 (1969); W. J. M. van Tilborg, J. R. van der Vecht, H. Steinberg, and Th. J. deBoer, *Tetrahedron Lett.*, 1681 (1972).
(64) J. A. Landgrebe and L. W. Becker, *J. Amer. Chem. Soc.*, **89**, 2502 (1967); **90**, 395 (1968); B. A. Howell and J. G. Jewett, *ibid.*, **93**, 798 (1971).
(65) M. L. Poutsma and P. A. Ibaria, *Tetrahedron Lett.*, 4967 (1970).
(66) H. M. R. Hoffmann, private communication, Sept 1971.
(67) For an apparent exception, see D. Cantacuzène and M. Tordeux, *Tetrahedron Lett.*, 4807 (1971); J. C. Blazjewski, D. Cantacuzène, and C. Wakselman, *Tetrahedron*, in press.

Molecular Orbital (CNDO/2 and MINDO) Calculations on Protonated Deoxyribonucleic Acid Bases. The Effects of Base Protonation on Intermolecular Interactions

Frank Jordan* and H. Dirk Sostman

Contribution from the Department of Chemistry, Rutgers University, Newark, New Jersey 07102. *Received December 19, 1972*

Abstract: The electronic reorganization accompanying monoprotonation of DNA bases was examined employing all-valence-electron SCF molecular orbital methods. σ as well as π reorganization upon protonation is evident in all four bases. Stacking and hydrogen-bonding interactions were calculated including monopole–monopole, monopole–induced dipole, and dispersion terms between the various bases in their neutral and monoprotonated states for all possible combinations. The intermolecular interactions are invariably more favorable for half-protonated pairs (*i.e.*, one base protonated) than for neutral pairs. Stacking interactions are always unfavorable in doubly protonated pairs (*i.e.*, both bases protonated).

As part of a continuous program in this laboratory to examine the effect of the state of ionization of nucleic acid components on their electronic structures and intermolecular interactions, recently all-valence-electron CNDO/2 and MINDO SCF calculations were reported on some adenine tautomers and their protonated analogs.[1]

This contribution completes our studies at the level of approximation of base interaction only. Electronic structures of all four protonated bases, adenine (A), guanine (G), thymine (T), and cytosine (C), are described and an attempt is made to predict their interbase interactions both in the vertical (stacking) and horizontal (in-plane) hydrogen-bonding mode.[2]

The theoretical approaches employed are the same as those previously reported, the CNDO/2[3a] and

(1) F. Jordan and H. D. Sostman, *J. Amer. Chem. Soc.*, **94**, 7898 (1972).
(2) Some other abbreviations used are poly T, poly A, poly G, poly C, and poly U for the homopolymers and ApA, ApG, ApC, etc., for the dinucleoside monophosphates. For example, ApA would have a O–3′ and a O–5′ bound adenosine attached to the phosphate.
(3) (a) J. A. Pople and G. A. Segal, *J. Chem. Phys.*, **44**, 3289 (1966); Quantum Chemistry Program Exchange No. 91; (b) N. C. Baird and M. J. S. Dewar, *ibid.*, **50**, 1262 (1969); Quantum Chemistry Program Exchange No. 137.

Table I. Energy, Dipole Moments, and Ionization Potential of Neutral and Protonated DNA Bases

Molecule or ion	Method[a]	Total energy[b] or ΔH_f, kcal/mol	Total[c] dipole moment, D	π[d] dipole moment, D	IP,[e] eV	Geometry employed[f]
Adenine	C	−60,964.1	2.99	1.25	10.3	P, H;[g] I, C[h]
Adenine	M	−101.13	2.30	1.78	9.0	P, H;[i] I, C[h]
Adenine N-1 H^+	C	−61,171.1		3.58	16.2	P, C;[h] I, C[h]
Adenine N-1 H^+	M	−14.87		2.66	13.9	P, C;[h] I, C[h]
Guanine	C	−72,651.2	7.25	3.93	9.74	P, O;[i] I, O[i]
	M	−83.11	7.09		9.22	
Guanine N-9 CH_3	C	−78,064.8	7.00		9.57	P, O;[j] I, O + M[d]
Guanine N-7 H^+	C	−72,764.9		4.70	15.78	P, ST;[k] I, ST[k]
	M				14.33	
Guanine N-9 CH_3 N-7 H^+	C	−78,225.3			15.69	P, ST;[k] I, ST[k]
Cytosine	C	−55,374.4	7.71	2.65	10.83	BM[l]
	M	−221.00	8.14		10.1	BM[l]
Cytosine N-1 CH_3	C	−58,872.6	7.39		10.6	BM + M[j]
Cytosine N-3 H^+	C	−53,920.3		3.74	18.1	BT[m]
	M	+10.98			15.4	BT[m]
Cytosine N-1 CH_3, N-3 H^+	C	−59,399.0			17.0	BT[m]
Thymine	C	−62,739.07	4.23	2.00	11.79	H[g]
	M	−220.57			10.51	
Thymine N-1 CH_3	C	−68,137.92	4.18		11.45	H[g]
Thymine O-4 H^+	C	−63,078.3		5.15	17.68	ST[n]
	M	−87.27			15.34	ST[n]
Thymine O-4 H^+, N-1 CH_3	C	−68,364.21			16.99	ST[n] + M[j]

[a] C = CNDO/2; M = MINDO. [b] According to CNDO/2 including nuclear repulsions, according to MINDO heat of formation. [c] Including net charges and hybrid moments as in ref 3a. [d] Based on π charges only assuming two π electrons on an imidazole N–H type nitrogen and one π electron on a pyridine like nitrogen. [e] Given by the negative of the energy of the highest occupied molecular orbital according to Koopmans' theorem. [f] P = pyrimidine geometry; I = imidazole geometry. [g] H. K. Hoogsteen, *Acta Crystallogr.*, **16**, 907 (1963). [h] C. W. Cochran, *ibid.*, **4**, 81 (1951). [i] O: O. J. O'Brien, *ibid.*, **23**, 92 (1967). [j] Methyl constructed according to Hoogsteen's parameters in *g*. [k] ST: H. M. Sobell and K. I. Tomita, *Acta Crystallogr.*, **17**, 126 (1964). [l] BM: D. L. Barker and R. E. Marsh, *ibid.*, **17**, 1581 (1964). [m] BT: C. F. Bryan and K. I. Tomita, *ibid.*, **15**, 1174 (1962). [n] ST: H. M. Sobell and K. I. Tomita, *ibid.*, **17**, 122 (1964).

MINDO[3b] methods with their original parametrizations. The very simple approximation employed in calculating the intermolecular interactions was described earlier.[1] Briefly, it uses CNDO/2 or MINDO charge densities, isotropic bond polarizabilities, and MINDO ionization potentials (the much more reasonable value of the two) to calculate the interaction energy as a sum of monopole–monopole, monopole–induced dipole, and dispersion terms.

Since the calculations are performed for ionic species, the dipole–dipole approximation cannot be used, the dipole moment of a charged species being origin dependent.

Partially due to economy (since a large number of calculations had to be performed), we only report on results using CNDO/2 charge densities. This we can justify since our previous report indicated no major differences in interaction energies based on CNDO/2 or MINDO charges and because others have shown that CNDO/2 is more reliable in its dipole moment predictions than is MINDO.[4]

The report is logically broken down into two parts; the effect of protonation on the DNA base electronic structure, and the effect of protonation on interbase interactions. Several experimental contributions called our attention to this topic and certainly the well known pK_a of the conjugate acids[5] of the DNA bases assures that a significant, if small, number of cytosine and adenine components would be protonated under physiological conditions.

(4) G. Klopman, "Topics in Current Chemistry," Vol. 15, No. 4, Springer-Verlag, West Berlin, 1970.

(5) R. M. Izatt, J. J. Christensen, and J. H. Rytting, *Chem. Rev.*, 71, 439 (1971).

The Electronic Effects of Base Protonation

One can note the effects of base protonation on a number of molecular orbital parameters. Among these are the electronic reorganization (in the σ and π distributions and in the π moment which unlike the total molecular moment remains origin independent on protonation) and ionization potential changes.

Table I presents all the gross features of the CNDO/2 and MINDO calculations employing X-ray crystallographic geometries. The total energies are not considered to be of great significance in this study but are quoted for comparison with other studies on these molecules.[6]

The ionization potentials (given by the negative of the highest occupied molecular orbital energy level according to Koopmans' theorem) are consistently increased upon protonation of the base by as much as 50% or more in the four cases studied. Such increases in ionization potentials were also found by the *ab initio* calculations on H_2O and H_3O^+.[7] The ionization potentials enter the intermolecular energy calculations, their relative magnitudes being of great importance in the dispersion calculations.

Charge Reorganization upon Protonation. CNDO/2 and MINDO calculations were performed for each system for several reasons. MINDO provides more realistic ionization potentials than CNDO/2 but usually less realistic electron density pictures.[4]

The questions of interest in comparing the neutral and protonated bases are: how do the σ and π frame-

(6) D. B. Boyd, *J. Amer. Chem. Soc.*, **94**, 64 (1972).

(7) P. A. Kollman and L. C. Allen, *J. Amer. Chem. Soc.*, **92**, 6104 (1970).

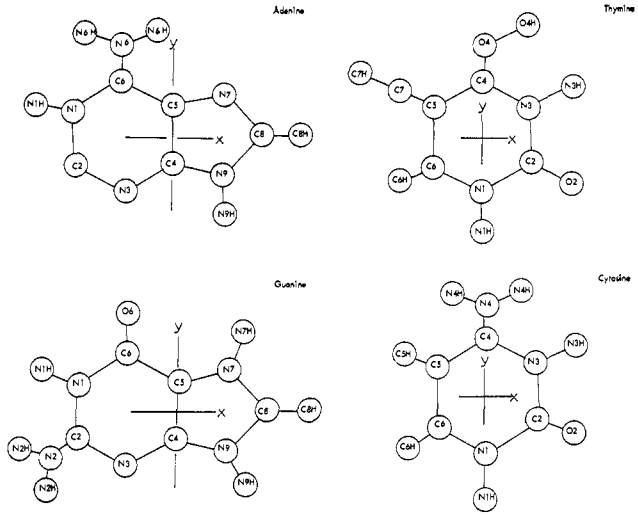


Figure 1. Numbering system and unrotated geometries of bases.

works react to protonation, and how does the attacking proton adjust to its new environment. While many calculations have been published on the neutral bases,[6] no comparison with protonated ones is available; therefore, such a comparison is presented below. Figure 1 describes the numbering system employed.

Adenine. The π-electron density on the nitrogen being protonated decreases substantially while its total electron density decreases (CNDO/2) or changes very little (MINDO) (Tables II and III). Protonation decreases the total electron density at all atoms to varying degrees, except for the slight increase predicted at N-9 according to CNDO/2. On the whole, all atoms undergo changes in charge upon protonation, the changes being more pronounced according to CNDO/2 than according to MINDO. These substantial changes in SCF base charges are at variance with some findings on nucleotides[8] and nucleotide zwitterions[9] (N-1 protonated, monophosphate monodeprotonated) obtained by extended Hückel calculations. The latter confined the effects of protonation to the pyrimidine ring; however, these calculations did not allow for charge iteration so that these results are not surprising.

As Table I indicates the π moment increases upon protonation independent of the method employed.

Guanine. The π-electron density is increased substantially at N-7 (site of protonation), N-1, and N-3 upon protonation according to both methods (Tables IV and V). The total electron density is increased at N-1 and at the incoming proton again upon protonation. All other atoms distribute the resulting loss in electron density. The π moment increases upon protonation but shows a much smaller fractional increase than does the π moment of adenine (Table I).

Table II. π and Net Atomic Charges in Adenine (A) and N-1 H^+ Adenine (AH^+), CNDO/2[a]

Atom	π charges[b] A	π charges[b] AH^+	π charges[b] Transfer	Total atomic charges A	Total atomic charges AH^+	Total atomic charges Transfer[c]
N-1	−0.276	−0.611	−0.335	−0.296	−0.138	+0.158
C-2	+0.109	+0.094	−0.014	+0.209	+0.236	+0.027
N-3	−0.205	−0.137	+0.067	−0.226	−0.153	+0.072
C-4	+0.044	+0.061	+0.016	+0.218	+0.221	+0.003
C-5	−0.158	−0.136	+0.022	−0.067	−0.017	+0.049
C-6	+0.154	+0.221	+0.066	+0.256	+0.344	+0.087
N-7	−0.232	−0.528	−0.296	−0.261	−0.139	+0.122
C-8	+0.010	+0.359	+0.348	+0.167	+0.174	+0.007
N-9	+0.397	+0.430	+0.033	+0.084	−0.119	−0.034
N-6	+0.155	+0.247	+0.091	−0.248	−0.207	+0.041
N-1 H					+0.171	−0.828
C-2 H				−0.030	+0.047	+0.078
C-8 H				+0.009	+0.046	+0.036
N-9 H				+0.109	+0.178	+0.068
N-6 H[d]				+0.122	+0.178	+0.055

[a] Geometries used are the same as those in Table I. [b] π charges derived from the orbital electron population of the atomic orbital perpendicular to the molecular planes. [c] Transfer of electron density on protonation; a negative value indicates that the atom has become electron richer on protonation; a positive value shows electron density loss on protonation. [d] The average of two amino hydrogen values.

Table III. π and Net Atomic Charges in Adenine (A) and N-1 H^+ Adenine (AH^+), MINDO[a]

Atom	π charges A	π charges AH^+	π charges Transfer	Total atomic charges A	Total atomic charges AH^+	Total atomic charges Transfer
N-1	−0.563	−0.727	−0.164	−0.703	−0.742	−0.039
C-2	+0.405	+0.413	+0.008	+0.690	+0.710	+0.020
N-3	−0.509	−0.470	+0.038	−0.638	−0.501	+0.137
C-4	+0.246	+0.272	+0.025	+0.570	+0.577	+0.006
C-5	−0.255	−0.249	+0.006	−0.101	−0.078	+0.023
C-6	+0.397	+0.387	−0.009	+0.708	+0.741	+0.032
N-7	−0.515	−0.529	−0.013	−0.490	−0.421	+0.069
C-8	+0.322	+0.358	+0.035	+0.597	+0.641	+0.044
N-9	+0.316	+0.334	+0.017	−0.605	−0.590	+0.014
N-6	+0.156	+0.211	+0.055	−0.708	−0.703	+0.004
N-1 H					+0.319	−0.680
C-2 H				−0.031	+0.032	+0.063
C-8 H				−0.045	+0.029	+0.074
N-9 H				+0.221	+0.313	+0.091
N-6 H[b]				+0.268	+0.337	+0.069

[a] See footnotes *a–c* in Table II. [b] The average of two amino hydrogen values.

Table IV. π and Net Atomic Charges in Guanine (G) and N-7 H^+ Guanine (GH^+), CNDO/2[a]

Atom	π charges G	π charges GH^+	π charges Transfer	Total atomic charges G	Total atomic charges GH^+	Total atomic charges Transfer
N-1	+0.287	+0.262	−0.025	−0.221	−0.249	−0.028
C-2	+0.200	+0.242	+0.042	+0.382	+0.422	+0.040
N-3	−0.392	−0.432	−0.040	−0.328	−0.310	+0.018
C-4	+0.083	+0.101	+0.017	+0.218	+0.266	+0.048
C-5	−0.240	−0.253	−0.013	−0.111	−0.092	+0.018
C-6	+0.188	+0.156	−0.032	+0.355	+0.381	+0.026
N-7	−0.134	−0.457	−0.323	−0.160	−0.001	+0.159
C-8	−0.045	+0.125	+0.170	+0.140	+0.241	+0.101
N-9	+0.345	+0.410	+0.065	−0.142	−0.117	+0.024
N-2	+0.171	+0.207	+0.035	−0.251	−0.241	+0.009
O-6	−0.464	−0.362	+0.101	−0.385	−0.300	+0.085
N-1 H				+0.119	+0.182	+0.063
N-7 H					+0.193	−0.806
C-8 H				−0.009	+0.078	+0.087
N-9 H				+0.116	+0.195	+0.079
N-2 H[b]				+0.139	+0.175	+0.036

[a] See footnotes *a–c* in Table II. [b] The average of two amino hydrogen values.

(8) D. B. Boyd and W. N. Lipscomb, *J. Theor. Biol.*, **25**, 403 (1969).
(9) F. Jordan, unpublished results.

Table V. π and Net Atomic Charges in Guanine (G) and N-7 H^+ Guanine (GH^+), MINDO[a]

Atom	π charges			Total atomic charges		
	G	GH^+	Transfer	G	GH^+	Transfer
N-1	+0.230	+0.204	−0.026	−0.783	−0.838	−0.054
C-2	+0.462	+0.482	+0.019	+0.919	+0.928	+0.009
N-3	−0.647	−0.676	−0.029	−0.658	−0.576	+0.082
C-4	+0.262	+0.241	−0.020	+0.577	+0.591	+0.013
C-5	−0.329	−0.329	−0.000	−0.148	−0.128	+0.020
C-6	+0.451	+0.425	−0.026	+0.882	+0.922	+0.039
N-7	−0.447	−0.598	−0.151	−0.399	−0.456	−0.057
C-8	+0.289	+0.382	+0.092	+0.605	+0.660	+0.055
N-9	+0.271	+0.290	+0.019	−0.640	−0.630	+0.010
N-2	+0.148	+0.184	+0.036	−0.696	−0.687	+0.008
O-6	−0.693	−0.606	+0.086	−0.653	−0.574	+0.079
N-1 H				+0.307	+0.373	+0.065
N-7 H					+0.332	−0.667
C-8 H				−0.061	+0.034	+0.096
N-9 H				+0.247	+0.364	+0.116
N-2 H[b]				+0.251	+0.341	+0.090

[a] See footnotes *a–c* in Table II. [b] The average of two amino hydrogen values.

Cytosine. The only point of general agreement between CNDO/2 and MINDO in this case concerns the decreased π charge on the nitrogen (N-3) being protonated. Almost all net atomic charges (except attacking proton) become more positive upon protonation, N-3 and O-8 bearing most of the change (Tables VI and VII). A rather large increase in π moment ac-

Table VI. π and Net Atomic Charges in Cytosine (C) and N-3 H^+ Cytosine (CH^+), CNDO/2[a]

Atom	π charges			Total atomic charges		
	C	CH^+	Transfer	C	CH^+	Transfer
N-1	+0.321	+0.346	+0.025	−0.188	−0.135	+0.053
C-2	+0.200	+0.166	−0.034	+0.422	+0.472	+0.050
N-3	−0.366	−0.645	−0.279	−0.345	−0.143	+0.202
C-4	+0.217	+0.287	+0.070	+0.326	+0.414	+0.088
C-5	−0.201	−0.217	−0.016	−0.184	−0.144	+0.039
C-6	+0.136	+0.235	+0.098	+0.188	+0.276	+0.088
N-4	+0.203	+0.240	+0.037	+0.244	−0.187	+0.057
O-8	−0.512	−0.413	+0.098	−0.414	−0.272	+0.141
N-1 H				+0.134	+0.205	+0.070
N-3 H					+0.192	−0.807
C-5 H				−0.001	−0.089	−0.087
C-6 H				+0.043	+0.045	+0.001
N-4 H[b]				+0.131	+0.182	+0.051

[a] See footnotes *a–c* in Table II. [b] The average of two amino hydrogen values.

companies protonation. Of the four neutral bases cytosine appears to have the smallest fractional π-moment contribution to the total moment.

Thymine. Tables VIII and IX provide the charge densities on neutral and O-4 protonated thymine. Interestingly, the π-electron density substantially increases on the oxygen to which the proton is being attached, while its net atomic charge changes much less in the opposite direction. CNDO/2 indicates electron density increases at some carbon atoms upon protonation. Again a substantial π-moment increase is noted upon protonation.

As we showed in our previous work,[1] the dipole moments and ionization potentials appear to be less sensitive to geometric changes than are the total energies

Table VII. π and Net Atomic Charges in Cytosine (C) and N-3 H Cytosine (CH^+), MINDO[a]

Atom	π charges			Total atomic charges		
	C	CH^+	Transfer	C	CH^+	Transfer
N-1	+0.236	+0.250	+0.013	−0.752	−0.705	+0.046
C-2	+0.498	+0.469	−0.029	+1.068	+1.079	+0.011
N-3	−0.579	−0.731	−0.151	−0.720	−0.731	−0.011
C-4	+0.425	+0.437	+0.012	+0.756	+0.785	+0.029
C-5	−0.299	+0.372	+0.671	−0.314	−0.270	+0.043
C-6	+0.271	−0.325	−0.597	+0.505	+0.533	+0.028
N-4	+0.180	+0.207	+0.026	−0.713	−0.655	+0.058
O-8	−0.734	−0.680	+0.054	−0.667	−0.562	+0.104
N-1 H				+0.295	+0.387	+0.092
N-3 H					+0.353	−0.646
C-5 H				+0.055	+0.134	+0.079
C-6 H				−0.083	+0.011	+0.094
N-4 H[b]				+0.285	+0.319	+0.034

[a] See footnotes *a–c* in Table II. [b] The average of two amino hydrogen values.

Table VIII. π and Net Atomic Charges in Thymine (I) and O-4 H^+ Thymine (TH^+), CNDO/2[a]

Atom	π charges			Total atomic charges		
	T	TH^+	Transfer	T	TH^+	Transfer
N-1	+0.249	+0.393	+0.144	−0.207	−0.130	+0.076
C-2	+0.196	+0.164	−0.032	+0.447	+0.445	−0.002
N-3	+0.260	+0.322	+0.062	−0.259	−0.212	+0.046
C-4	+0.187	+0.269	+0.081	+0.346	+0.429	+0.083
C-5	−0.108	−0.162	−0.053	−0.095	−0.102	−0.006
C-6	+0.062	+0.233	+0.171	+0.137	+0.247	+0.109
O-2	−0.451	−0.403	+0.047	−0.360	−0.287	+0.073
O-4	−0.396	−0.818	−0.421	−0.343	−0.188	+0.154
C-5′				−0.016	−0.014	+0.002
N-1 H				+0.150	+0.197	+0.046
N-3 H				+0.006	+0.191	+0.185
C-6 H				+0.141	+0.064	−0.076
C-5′ H[b]				+0.017	+0.042	+0.025
O-4 H					+0.231	−0.768

[a] See footnotes *a–c* in Table II. [b] The average value for the three methyl hydrogens.

Table IX. π and Net Atomic Charges in Thymine (T) and O-4 H^+ Thymine(TH^+), MINDO[a]

Atom	π charges			Total atomic charges		
	T	TH^+	Transfer	T	TH^+	Transfer
N-1	+0.195	+0.284	+0.088	−0.731	−0.695	+0.036
C-2	+0.492	+0.482	−0.009	+1.115	+1.090	−0.025
N-3	+0.227	+0.250	+0.023	−0.810	−0.795	+0.015
C-4	+0.477	+0.465	−0.011	+0.892	+0.897	+0.005
C-5	−0.246	−0.312	−0.065	−0.325	−0.353	−0.028
C-6	+0.234	+0.376	+0.142	+0.473	+0.579	+0.106
O-2	−0.704	−0.681	+0.022	−0.670	−0.543	+0.127
O-4	−0.675	−0.866	−0.190	−0.670	−0.665	+0.005
C-5′				+0.363	+0.360	−0.003
N-1 H				+0.332	+0.359	+0.027
N-3 H				−0.067	+0.346	+0.413
C-6 H				+0.293	+0.016	−0.277
C-5′ H[b]				−0.064	−0.017	+0.046
O-4 H					+0.455	−0.544

[a] See footnotes *a–c* in Table II. [b] The average value for the three methyl hydrogens.

of the systems. While not much experimental data are available on the dipole moments of the bases, the two methods MINDO and CNDO/2 usually are in qualitative agreement with each other, the latter giving closer agreement with experiment.[1] Perhaps what is remark-

able is the relative agreement between the moments given by MINDO and CNDO/2 in spite of the grossly different charge densities predicted by the two methods. We have also shown[10] that with a reasonable parametrization even the uniterated extended Hückel theory can provide dipole moments in good agreement with those provided by the present SCF methods. Notwithstanding the correct dipole moment predictions, one would expect SCF charges to be more correct.

Although protonation takes place in the plane of the σ system, a substantial perturbation of the π system is evident. π reorganization is more extensive in adenine than in guanine, the imidazole ring of both being strongly affected even though only guanine is protonated on this ring (N-7). π reorganization is somewhat more extensive in thymine than in cytosine; *e.g.*, in thymine the N-1 atom is more susceptible to O-4 protonation than is the N-1 atom in cytosine to N-3 protonation. The σ system appears to be less active in charge delocalization than is the π system. The charge of the heteroatom being protonated absorbs a large fraction of the plus charge introduced by the incoming proton.

Another interesting observation is that the change in net atomic charge upon protonation does not follow the accepted electronegativity relationships of the particular atoms.

Doty, *et al.*,[11] showed that some changes occur in the uv spectra of the bases upon protonation. Presumably, such changes would be related to π reorganization. ^{13}C nmr studies should prove helpful in checking the validity of our findings.

Intermolecular Interactions

Theoretical Considerations. To our knowledge this work represents the initial attempt to account for the effects of protonation on DNA interactions.

Because of the size of the molecules and the approximations involved in gathering the various input parameters needed for an intermolecular interaction calculation (such as net atomic charges, ionization potentials, bond polarizabilities), it was felt that a relatively approximate solution is in order. Many groups have calculated the base–base interactions between neutral species[12] starting with the classic work of De Voe and Tinoco.[12a] We are here attempting to incorporate the effects of protonation on such interactions.

A considerable amount of theoretical work has been performed on intermolecular interactions and the progress is summarized by Hirschfelder, Curtiss, and Bird[13] and Margenau and Kestner.[14]

As in our previous contribution, we again adopt an approximation for intermolecular forces easily applicable to such large systems for a variety of geometrical variations.

We assume point charges at the atoms as given in Tables II, IV, VI, and VIII (CNDO/2) to calculate the monopole–monopole interaction energy ($\rho\rho$) between groups of such charges. The monopole–induced dipole (polarization) energy ($\rho\alpha$) is calculated using isotropic bond polarizabilities.[15] Finally, the dispersion energy ($\alpha\alpha$) is calculated with the use of molecular (or ionic) ionization potentials and isotropic bond polarizabilities. The appropriate equations are given in ref 1 and are a further simplified variation of a recent calculation by Rein, *et al.*[16]

We have employed MINDO ionization potentials since these appear to be closer (see Table I) to the experimentally reported values where such are available (calculated in electron volts by MINDO for A, 9.0; C, 10.1; T, 10.5; experimental[17] for A, 8.90; C, 8.90; T, 9.43). Rein, *et al.*,[18] in their most recent contribution showed that in the interaction of two pyridine molecules quadrupole terms may be dominant at interaction distances of less than 10 Å. No such calculations for DNA bases are available. According to electrostatic theory both dipole and quadrupole moments of ions vary with the origin assumed, and the dipole moment is zero at the center of charge.

In order to retain the simplest consistent approximation we only calculate monopole–monopole, monopole–induced dipole, and dispersion terms as before. These terms can be calculated with ease for any interacting system irrespective of the charge of the species.

More rigorous theoretical attempts have tried to decompose hydrogen-bonding interactions into electrostatic, exchange, charge-transfer, polarization, and dispersion terms.[19] Each term was shown to lead to significant contributions to the total energy. While we are accounting for most of these terms, the possibility of charge transfer at hydrogen-bonding or stacking distances is not taken into account in the present work. In order to see if there is significant charge transfer according to CNDO/2, one needs to compute eigenvalues and eigenfunctions of the entire system. This we will attempt in the future.

Stacking Interactions. In order to be able to compare the stacking interactions between adjacent bases as a function of the state of protonation, all calculations were performed at 3.30-Å interplanar separation. This value represents an average of the interplanar separations in similar molecules compiled by Bugg, *et al.*[20] Calculations were performed for a large variety of rotational and translational variations at the 3.30-Å interplanar separation.

The zero rotation places all molecules in the X–Y plane and centers the C-4–C-5 bond of purines and the N-1–C-4 bond of pyrimidines along the Y axis as indicated in Figure 1. Molecule (or ion) 1 was held stationary and molecule (or ion) 2 was rotated clockwise in

(10) F. Jordan, *Biopolymers*, **12**, 243 (1973).
(11) D. Voet, W. B. Gratzer, R. A. Cox, and P. Doty, *Biopolymers*, **1**, 193 (1963).
(12) (a) H. De Voe and I. Tinoco, Jr., *J. Mol. Biol.*, **4**, 500 (1962); (b) H. A. Nash and D. F. Bradley, *Biopolymers*, **3**, 261 (1965); (c) A. Pullman and B. Pullman, *Advan. Quantum Chem.*, **4**, 267 (1968); (d) B. Pullman, "Molecular Associations in Biology," B. Pullman, Ed., Academic Press, New York, N. Y., 1968; (e) R. Rein, P. Claverie, and M. Pollak, *Int. J. Quantum. Chem.*, **2**, 129 (1968).
(13) J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, "Molecular Theory of Gases and Liquids," Wiley, New York, N. Y., 1964, Chapters 12 and 13.
(14) H. Margenau and N. R. Kestner, "Theory of Intermolecular Forces," Pergamon Press, Oxford, 1969.
(15) R. J. W. Le Fèvre, *Advan. Phys. Org. Chem.*, **3**, 1 (1965).
(16) M. S. Rendell, J. P. Harlos, and R. Rein, *Biopolymers*, **10**, 2083 (1971).
(17) C. Lifschitz, E. D. Bergmann, and B. Pullman, *Tetrahedron Lett.*, 4583 (1967); employing mass spectrometry.
(18) R. Rein, J. R. Rabinowitz, and T. J. Swissler, *J. Theor. Biol.*, **34**, 215 (1972).
(19) P. A. Kollman and L. C. Allen, *Chem. Rev.*, **72**, 283 (1972).
(20) (a) C. E. Bugg, J. M. Thomas, M. Sundaralingam, and S. T. Rao, *Biopolymers*, **10**, 175 (1971); (b) C. E. Bugg, "The Purines Theory and Experiment, Jerusalem Symposia on Quantum Chemistry and Biochemistry IV," The Israel Academy of Sciences and Humanities, Jerusalem, 1972.

Table X. Optimal Stacking Interaction Energies and Corresponding Geometries[a]

Stationary	Rotated, translated species	Geometrical[b] definition of variable species RA, °	TRA, °	RTD, Å	Interaction energies, kcal/mol $\rho\rho$[c]	$\rho\alpha$ (1)[d]	$\rho\alpha$ (2)[e]	$\alpha\alpha$[f]	E_{total}[g]
				Homogeneous Pairs					
A	A	90	270	0.5	−0.41	−0.09	−0.08	−7.85	−8.43
		180	270	0.5	−0.30	−0.11	−0.11	−8.03	−8.54
AH^+	A	45	225	0.5	−1.97	−0.10	−4.92	−9.97	−16.96
		270	0	1.0	−1.94	−0.08	−4.75	−9.79	−16.56
AH^+	AH^+	180	90	5.0	44.07	−1.20	−1.20	−2.17	39.50
G	G	135	270	1.5	−1.48	−0.34	−0.37	−7.51	−9.69
GH^+	G	90	270	1.0	−5.35	−0.34	−3.90	−9.26	−18.85
GH^+	GH^+	180	0	5.0	43.28	−1.31	−1.31	−1.90	38.77
T	T	180	0	0.5	−4.49	−0.57	−0.57	−7.33	−12.96
TH^+	T	270	135	1.5	−5.17	−0.45	−4.85	−6.70	−17.17
TH^+	TH^+	180	135	5.0	43.46	−1.04	−1.04	−1.45	39.93
C	C	135	45	0.5	−1.97	−0.33	−0.34	−6.31	−8.95
CH^+	C	270	180	1.0	−3.91	−0.40	−5.32	−7.12	−16.75
CH^+	CH^+	180	135	5.0	46.71	−1.15	−1.15	−1.38	43.02
				Heterogeneous Pairs					
A	G	135	225	0.5	−1.61	−0.23	−0.09	−7.58	−9.51
AH^+	G	315	135	0.5	−4.59	−0.36	−3.99	−9.03	−17.97
A	GH^+	180	315	1.0	−2.46	−4.88	−0.10	−10.10	−17.54
AH^+	GH^+	270	90	5.0	44.08	−1.17	−1.28	−1.96	39.68
A	T	90	225	1.0	−1.63	−0.72	−0.08	−7.42	−9.84
AH^+	T	180	45	1.5	−4.52	−0.50	−3.14	−7.54	−15.70
A	TH^+	0	270	0.5	−2.31	−6.72	−0.08	−9.24	−18.35
AH^+	TH^+	135	90	5.0	44.44	−1.26	−1.03	−1.78	40.38
A	C	45	0	0	−1.65	−0.37	−0.07	−6.95	−8.99
AH^+	C	180	90	1.0	−4.93	−0.50	−3.14	−7.75	−16.32
A	CH^+	225	315	1.0	−2.93	−6.67	−0.06	−8.89	−18.54
AH^+	CH^+	135	90	5.0	45.86	−1.44	−1.01	−1.74	41.67
G	C	135	315	1.0	−2.92	−0.37	−0.27	−6.71	−10.26
GH^+	C	315	225	1.5	−4.27	−0.47	−3.28	−8.20	−16.22
G	CH^+	315	0	1.0	−5.47	−5.81	−0.33	−7.67	−19.27
GH^+	CH^+	45	0	5.0	45.18	−1.49	−0.97	−1.57	41.15
G	T	135	315	0.5	−3.87	−0.62	−0.24	−7.35	−12.08
GH^+	T	135	0	1.5	−4.65	−0.55	−3.98	−7.98	−17.16
G	TH^+	135	0	1.0	−4.46	−5.99	−0.28	−8.09	−18.83
GH^+	TH^+	45	0	5.0	43.54	−1.33	−0.99	−1.60	39.63
T	C	135	0	0.5	−4.04	−0.35	−0.54	−6.84	−11.77
TH^+	C	225	135	1.5	−5.20	−0.45	−4.63	−6.36	−16.64
T	CH^-	45	270	1.5	−4.53	−4.81	−0.45	−7.02	−16.81
TH^+	CH^+	180	135	5.0	45.08	−1.02	−1.18	−1.42	41.46

[a] All interplanar distances set at 3.3 Å; starting geometrical arrangement as in Figure 1. [b] Geometrical variations on variable geometry species defined according to clockwise rotation angle (RA) followed by clockwise translational angle (TRA) and translational distance (TRD). [c] Monopole–monopole. [d] Monopole–induced dipole (induced by species 2 in 1). [e] Monopole–induced dipole (induced by species 1 in 2). [f] Dispersion forces calculated from the species' own ionization potential. [g] Total interaction energy.

45° increments over 360°. For each rotational angle (RA) translations were performed in 45° increments over 360° clockwise (TRA), and for each of these angles 0.5-Å incremental variation up to a total of 5-Å translational distance (TRD) was allowed. Thus, the three geometrical variations RA, TRA, and TRD exactly define the position of the rotated group with respect to the stationary group.

Table X quotes some of our results. Due to space limitations in most cases only the optimal interaction energies and the corresponding geometries are quoted. In most cases there were several regions of nearly optimal energies, however.[21]

Our main interest is to obtain relative values for the interaction energies as a function of base protonation. One needs to consider the pK_a's for the conjugate acids of the bases.[5] These are 4.15 for N-1 protonation of adenine, 3.30 for N-7 protonation of guanine, 4.45 for N-3 protonation of cytosine, and unknown (very low) for O-4 protonation of thymine.[5] While the values vary subtly in going from base to nucleoside to nucleoside monophosphate, the relative orders remain the same. Thus, as the pH is lowered the relative ease of protonation follows the order cytosine > adenine > guanine > thymine. For completeness all possible pairing permutations were considered and are presented in Table X. We also calculated the stacking energies for doubly protonated pairs although as in our previous study[1] the results again indicate substantial destabilization of such pairs compared with the monoprotonated or neutral pairs.

A natural separation of both stacking and hydrogen-bonding interactions is in terms of self-stacking (homogeneous pairs) or mixed-stacking (heterogeneous pairs) interactions.

Strictly speaking, such calculations apply only to the gaseous state. The effect of solvent is very difficult to predict. However, it has been experimentally established that stacking interactions are much stronger in an

(21) A list of optimal stacking energies and corresponding geometries as well as the Fortran program performing the calculations are available from the authors upon request.

aqueous medium[22a] and hydrogen bonding is observable both in nonaqueous[22b] and aqueous[22c] solutions of the bases.

We have not considered the geometrical restrictions on the stacked bases present in polynucleotides but allowed complete freedom of the two stacked bases in searching out their optimum interaction energies. The values referred to in Table X represent such optimal values.

Very significantly, our optimum interaction energies are in the range reported for stacking interactions by various authors on dinucleotides[23] in aqueous solution and not very different from the theoretical values reported in ref 12c and 12d to which they may be compared (subject to lack of knowledge of exact geometries in these references). Our own interest is in the effect of protonation on such interactions, and in this area there are relatively little experimental data reported. However, for sake of comparison with experimental results, the interactions of neutral species are also quoted, since to our knowledge the CNDO/2 charges have not to date been used for this purpose.

Homogeneous Pairs. Among neutral pairs the order of relative stability appears to be T–T > G–G > C–C > A–A, and all optimal geometries require rotation of the rotatable group to some extent. While the absolute magnitudes for A–A, G–G, and C–C are similar, for T–T they appear to be rather high. According to ref 23d UpU (dinucleoside monophosphate) does stack.

Some further experimental evidence indicates that methylation has significant effects on interactions. 5-Methyl poly C is more ordered than poly C.[24] Poly T is helical; poly U is a random coil under the same conditions.[25] Poly G, poly C, poly U, and poly A are all capable of stacking under neutral pH conditions[26] in agreement with the relatively large association energies here reported for all four homodimers. $\Delta H_{\text{stacking}}$ values reported[23] for the various neutral homodimers are in the range of 4–10 kcal/mol depending on the experimental determination employed. We could not find any experimental evidence for the self-stacking processes as a function of pH (especially near the pK).

Protonation of one partner changes the order somewhat with decreasing order of stability, GH^+–G > TH^+–T > CH^+–C > AH^+–A. All half-protonated pairs have increased stabilities over their neutral counterparts, the enhanced stability being due to a combination of more favorable monopole–monopole and monopole–induced dipole terms in the half-protonated pairs.

The increased induction energies of the four pairs upon protonation are relatively similar, whereas the per cent charge–charge interaction is much larger in purines (AH^+–A > GH^+–G) than in pyrimidines (CH^+–C > TH^+–T).

(22) (a) S. Hanlon, *Biochem. Biophys. Res. Commun.*, **23**, 861 (1966); (b) Y. Kyogoku, R. C. Lord, and A. Rich, *J. Amer. Chem. Soc.*, **89**, 496 (1967); (c) M. Raszka and N. O. Kaplan, *Proc. Nat. Acad. Sci. U. S.*, **69**, 2025 (1972).

(23) (a) J. Applequist and V. Damle, *J. Amer. Chem. Soc.*, **88**, 3895 (1966); (b) M. M. Warshaw and I. Tinoco, Jr., *J. Mol. Biol.*, **20**, 29 (1966); (c) J. Brahms, J. C. Maurizot, and A. M. Michelson, *ibid.*, **25**, 481 (1967); (d) R. C. Davis and I. Tinoco, Jr., *Biopolymers*, **6**, 223 (1968).

(24) W. Szer, M. Swierkowski, and D. Shugar, *Acta Biochim. Pol.*, **10**, 87 (1963).

(25) W. Szer and D. Shugar, *J. Mol. Biol.*, 17, 174 (1966).

(26) A. M. Michelson, J. Massoulie, and W. Guschlbauer, *Progr. Nucl. Acid Res. Mol. Biol.*, **6**, 83 (1967).

Poly C and poly A both form double helices under acidic conditions[26] (latter was discussed in ref 1); however, the ability to form the double helix is usually attributed to hydrogen bonding (*vide infra*). We are now proposing that the stacking interactions also become more favorable upon protonation of one partner.

Double protonation of the pairs leads to substantial electrostatic repulsions in all situations studied both in the hydrogen bonding and in the stacking modes. However, with the often invoked salt bridge between a base proton and a neighboring phosphate charge it is not too difficult to rationalize the existence of such fully protonated forms on the basis of these calculations. To overcome perhaps, *e.g.*, 40 kcal/mol or more destabilization, the average distance between the phosphate -1 charge and base $+1$ charge must be less than 331/40 or about 8 Å. Depending on the charge concentration in the phosphate and base moieties, such salt bridge stabilization could be of paramount importance for these structures. In the absence of the double protonation the singly protonated pair could be further stabilized by the salt bridge. The problem of double protonation will not be further pursued since the experimental results are not very clear on this point either. Warshaw and Tinoco[23b] indicated that only a few dinucleoside monophosphates were stacked at pH 1 (GpC, CpG, UpG, GpU). In our view perhaps these represent half-protonated species.

Heterogeneous Pairs. Stacking with Adenine. All neutral pairs containing adenine are more stable than A–A. On half-protonation all pairs become more stable than their neutral counterparts. Protonation of pyrimidine partners leads to greater stabilization than protonation of purine partners in a purine–pyrimidine stack. In an A–G stack AH^+–G is more stable than A–GH^+

Guanine-Containing Pairs. The neutral G–G or G–A interaction is less stable than the G–C or G–T interaction, though only the G–T interaction appears to be substantially different from the other three possibilities. Among half-protonated pairs purine–pyrimidine interactions are much more favorable upon pyrimidine protonation (especially for CH^+–G pair *vs.* C–GH^+ pair). This latter pair may be particularly important at ordinary pH conditions considering the pK of cytosine protonation. It may be the one predominating at pH 1[23b,c] if monoprotonation (on C) perhaps suppresses the pK of G protonation in G–CH^+, so that at this pH one is observing the half-protonated pair. Recently GpU was reported to stack only at low pH;[27] we also predict stabilization of the stack by G protonation.

Thymine–Cytosine Pairs. This pair exhibits the smallest per cent stabilization of the half-protonated pair over the neutral pair of all the combinations studied. The stabilization of the stack due to protonation of either partner leads to remarkably similar values. Obviously, on thermodynamic grounds the CH^+ is much more available than the TH^+.

Hydrogen-Bonding Interactions. The results presented in Table XI represent values for the geometric arrangements drawn in structures I–XX. Most theoretically feasible arrangements have been accounted for with the assumption of the ribose (or deoxyribose)

(27) W. Guschlbauer, I. Fric, and A. Holy, *Eur. J. Biochem.*, **31**, 1 (1972).

Table XI. Hydrogen-Bonding Interactions between DNA Bases as a Function of the State of Protonation

Number[a]	Structure[b]	H-bond distances,[c] Å 1	2	3	Interaction energies, kcal/mol $\rho\rho$[d]	$\rho\alpha(1)$[e]	$\rho\alpha(2)$[f]	$\alpha\alpha$[g]	E_{total}[h]
				Homogeneous Pairs					
I-1	A–A	2.81	2.81		−0.93	−0.16	−0.16	−3.36	−4.60
		2.90	2.90		−0.80	−0.13	−0.13	−2.82	−3.88
I-2	AH^+–A	2.87	2.89		−1.73	−0.14	−2.43	−4.14	−8.45
		2.95	2.96		−1.64	−0.12	−2.22	−3.50	−7.47
II	A–A	2.93	2.93		−1.41	−0.13	−0.13	−2.43	−4.11
III-1	A–A	2.88	2.87		−0.69	−0.15	−0.16	−3.21	−4.21
		2.96	2.97		−0.63	−0.13	−0.13	−2.72	−3.60
III-2	AH^+–A	2.90	2.92		−4.18	−0.14	−2.24	−3.50	−10.05
IV-1	G–G	2.86	2.86		−7.39	−0.63	−0.63	−1.87	−10.53
IV-2	GH^+–G	2.84	2.79		−8.08	−0.71	−1.27	−2.92	−12.98
V	C–C	2.92	2.92		−3.84	−0.33	−0.33	−2.81	−7.31
VI	CH^+–C	2.90	2.84*	2.85	−16.17	−0.90	−4.51	−4.40	−25.97[i]
VII-1	T–T	2.98	2.98		2.66	−0.28	−0.28	−1.38	0.73
VII-2	TH^+–T	2.87	2.97		−9.27	−0.33	−1.31	−1.86	−12.77
VIII-1	$T–T_{rev}$	2.87	2.90		3.12	−0.34	−0.38	−1.68	0.72
VIII-2	$T–TH^+_{rev}$	2.87	2.80		−11.28	−1.60	−0.46	−2.27	−15.61
				Heterogeneous Pairs					
IX	A–T	2.92	2.85		0.04	−0.40	−0.14	−2.47	−2.98
X-1	$A–T_{rev}$	2.93	2.89		0.27	−0.33	−0.14	−2.45	−2.66
X-2	$A–TH^+_{rev}$	2.94	2.89		−6.57	−4.27	−0.17	−3.81	−14.82
XI	$A–C_{rev}$	2.92	2.90		−2.28	−0.36	−0.13	−2.63	−5.41
XII	$A–CH^+_{rev}$	2.85*	2.86		−7.11	−4.94	−0.21	−4.45	−16.71
XIII-1	$A–G_{rev}$	2.92	2.84		−3.36	−0.76	−0.17	−2.95	−7.23
XIII-2	$A–GH^+_{rev}$	2.89	2.84		−4.85	−1.91	−0.16	−3.34	−10.26
XIV	$A–GH^+_{rev}$	2.86	2.94*		−2.68	−3.04	−0.12	−3.27	−9.10
XV-1	G–T	2.86	2.86		−0.15	−0.36	−0.60	−2.16	−3.27
XV-2	GH^+–T	2.83	2.82		−8.45	−0.35	−1.35	−3.38	−13.54
XV-3	$G–TH^+$	2.88	2.89		−10.98	−1.85	−0.60	−2.33	−15.76
XVI-1	$G–T_{rev}$	2.83	2.83		−0.53	−0.46	−0.62	−2.05	−3.64
XVI-2	$GH^+–T_{rev}$	2.87	2.87		−8.88	−0.40	−1.01	−2.46	−12.75
XVII-1	G–C	2.92	2.87	2.86	−8.00	−0.85	−0.78	−3.12	−12.75
XVII-2	GH^+–C	2.86	2.88	2.87	−11.19	−0.77	−2.09	−3.94	−18.00
XVIII-1	T–C	2.84	2.79		2.35	−0.55	−0.41	−3.00	−1.61
XVIII-2	TH^+–C	2.85	2.82	2.85*	−15.29	−0.76	−4.33	−4.27	−24.65[i]
XIX	CH^+–T	2.87*	2.85		−11.45	−0.37	−2.22	−2.60	−16.64
XX	$CH^+–T_{rev}$	2.83*	2.85		−13.27	−0.49	−2.45	−2.76	−18.96

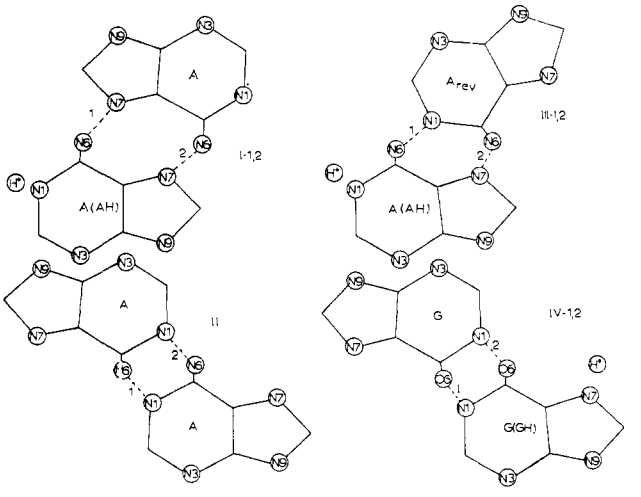
[a] Numbering refers to the structures. [b] Protonation always on N-1 in AH^+, N-7 in GH^+, N-3 in CH^+, and O-4 in TH^+; "normal" structures as drawn in Figure 1; reversed (rev) structures are the mirror image structures of those in Figure 1. [c] Corresponding to the numbers in structures I–XX, starred hydrogen bond with attached proton. [d] Monopole–monopole. [e] Monopole–induced dipole, induced in the first partner. [f] Monopole–induced dipole induced in the second partner. [g] Dispersion energy based on calculated ionization potential of each partner. [h] Energy sum. [i] Structures with three hydrogen bonds.

being attached to N-9 of the purines and N-1 of the pyrimidines (see Figure 1 for numbering) and the site of protonation assumed at N-1 in A, N-7 in G, N-3 in C, and O-4 in T.

A geometric search was performed for a particular coplanar hydrogen-bonded pair which would allow for the most symmetrical and, in most cases, most linear hydrogen bond. Hydrogen-bonding distances for the A–T and G–C pairs are available and are approximately adhered to throughout.[28]

The energy for any other separation along these essentially linear hydrogen bonds can be easily calculated for each contribution recalling that $\rho\rho$ varies with d^{-1}, $\rho\alpha$ with d^{-4}, and $\alpha\alpha$ with d^{-6} where the significance of the distances, d, associated with each term is discussed in ref 1.

Homogeneous Pairs. Adenines. The configurations corresponding to structures I–III were calculated. Scheme I-2 corresponds to the Hoogsteen scheme found in poly A with hydrogen bonds[29] at the N-6 and N-7 atoms. In this arrangement N-1 can be protonated



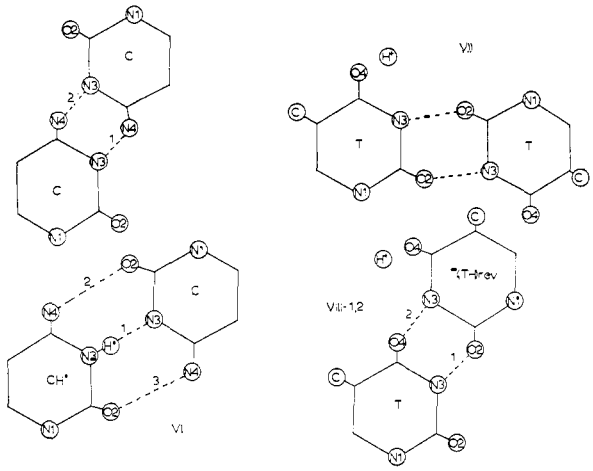
and this protonation is shown to lead to stabilization arising from both the $\rho\rho$ and $\rho\alpha$ terms. At variance with the Pullman and Pullman theoretical results[12] the three neutral pairs show remarkably similar stabilities. Protonation in arrangement III appears to lead to more stability than the protonation in the Hoogsteen scheme. (I-2). Only arrangements I and

(28) S. Arnott, S. D. Dover, and A. J. Wonacott, *Acta Crystallogr., Sect. B*, **25**, 2192 (1969).
(29) A. Rich, D. R. Davies, F. H. C. Crick, and J. D. Watson, *J. Mol. Biol.*, **3**, 71 (1961).

III have precedents interestingly,[12c] and these we predict to be slightly more stable than II. Some of these schemes supposedly help stabilize the double helical structures of monoprotonated and slightly more acidified polyriboadenylic acids.[30]

Guanine. G–G hydrogen bonding *via* N-1 and O-6 leads to an electrostatically ($\rho\rho$) much more stabilized system than any A–A pairs. Protonation of one partner on N-7 gives very small extra stabilization in this case (IV-1 and IV-2 and Table IV). This structure presumably occurs in poly G.[31]

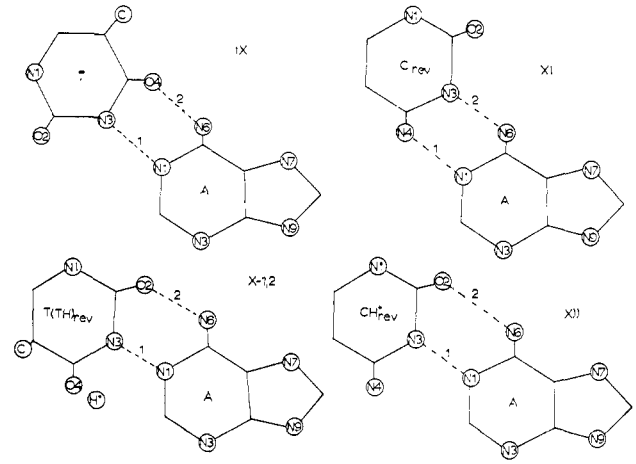
Cytosine. While the C–C hydrogen bond is much stronger than the A–A hydrogen bond due to the $\rho\rho$ term, it is somewhat weaker than the G–G hydrogen bond as suggested by ir experiments.[32] Structure VI



is the scheme found in 1-methylcytosine.[33] The half-protonated pair (cannot be accomplished without movement of one base with respect to the other) has capabilities of forming three hydrogen bonds (VI), one of which utilizes the proton acquired during protonation (on N-3). The CH^+–C pair is shown to have an extraordinary stabilization (also suggested by Pullman[12c] and by several pieces of experimental evidence[34]). It is due in great part to the presence of three hydrogen bonds. However, it appears that protonation of cytosine is especially favorable for hydrogen-bond stabilization, since a comparison with the GH^+–C structure (also with three hydrogen bonds) indicates that for essentially similar hydrogen-bond lengths the GH^+–C system is very much less stabilized than the C–CH^+ system due to both $\rho\rho$ and $\rho\alpha$ type interactions.

Thymine. There is a serious disagreement between our result and those of Pullman and Pullman[12c] on the neutral pair. We predict no stabilization at all for this couple. Upon half-protonation both structures VII and VIII become substantially stabilized, much less, however, than is CH^+–C. Most recent evidence in H_2O indicates H-bonding self-association only for C–C and G–G pairs[22c] as suggested here.

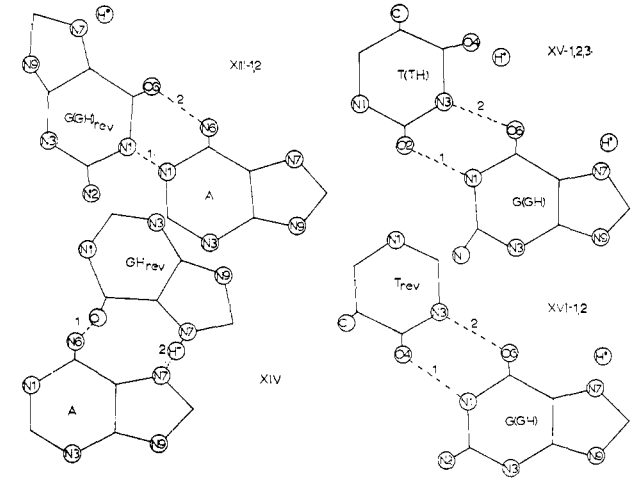
Heterogeneous Pairs. Adenine Pairs. Of greatest interest is A–T, structure IX, the Watson–Crick pairing scheme. To our great surprise this pair is not very greatly stabilized. In fact, it is less stable than the A–A pair, the A–G pair, or the A–C pair but more stable than the T–T pair. It appears very likely that thymine is a relatively poor hydrogen-bonding partner (though it is an excellent stacking partner). In the Watson–Crick scheme (IX) A cannot be protonated at N-1 and



T cannot be protonated at O-4. The scheme X-1,2 can accommodate T protonation leading to substantial stabilization again. This neutral scheme (X) is found in the X-ray structures of 9-ethyladenine with 1-methyl-5-iodouracil.[35]

Cytosine and adenine can form a hydrogen-bonding scheme somewhat reminiscent of the Watson–Crick A–T pair (XI) but with the pyrimidine ring turned around. Upon protonation of the N-3 atom of cytosine this ring can slide up to form a very stable CH^+–A complex (XII).

Adenine and guanine can form a symmetrical hydrogen-bonding scheme (XIII). Two half-protonated



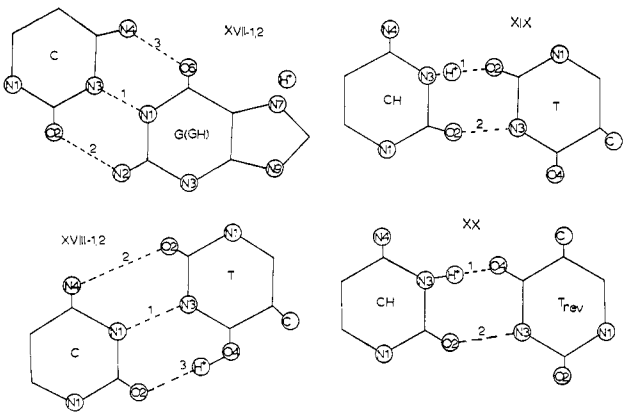
pairs may be found between A and GH^+ (XIII-2 and XIV), their relative energies fairly similar to each other. The neutral A–G pair appears to be the most stable one adenine is capable of forming.

Guanine Pairs. Of greatest interest is the G–C pair suggested by Watson and Crick (scheme XVII). This is indicated to be the strongest neutral one of all those calculated as also indicated by experiments.[32,36] The pairing allows N-7 protonation of G and this protona-

(30) A. J. Adler, L. Grossman, and G. D. Fasman, *Biochemistry*, **8**, 3846 (1969).
(31) F. Pochon and A. M. Michelson, *Proc. Nat. Acad. Sci. U. S.*, **53**, 1425 (1965).
(32) Y. Kyogoku, R. C. Lord, and A. Rich, *Science*, **154**, 518 (1966).
(33) F. S. Mathews and A. Rich, *Nature (London)*, **201**, 179 (1964).
(34) (a) E. O. Akinrimisi, C. Sander, and P. O. P. Ts'o, *Biochemistry*, **2**, 340 (1963); (b) T. M. Garestier and C. Hélène, *ibid.*, **9**, 2865 (1970).
(35) T. D. Sakore, S. S. Tavale, and H. M. Sobell, *J. Mol. Biol.*, **43**, 361 (1969).
(36) L. Katz and S. Penman, *J. Mol. Biol.*, **15**, 220 (1966).

tion leads to further stabilization. The great stabilization of the G–C pair is undoubtedly due to the existence of three hydrogen bonds in scheme XVII. The calculations indicate that the C=O···NC hydrogen bond is somewhat stronger than are the CN···NC ones.

Guanine is also suggested as a partner to thymine (schemes XV–XVII) with the G–T hydrogen bond being nearly the same energy as the A–T one. Protonation of either G (N-7) or T (O-4) under a variety of schemes indicates substantial further stabilization of the hydrogen bonds with the G–TH$^+$ pairs preferred over the GH$^+$–T pairs theoretically, while thermodynamically



(*i.e.*, pK) it appears to be much easier to protonate G than T.

Cytosine–Thymine. Possessing very similar stability of the neutral pair to T–T, protonation of C–T on either T (leads to three hydrogen bonds and exceptional stability) or C (leads to two hydrogen bonds only but very great stability, *e.g.*, schemes XIX and XX) again is very favorable thermodynamically. One would, of course, expect the C to be much more easily protonated than T.

Realistically, one needs to consider the ease of protonation of each partner (*i.e.*, pK_a's of the conjugate acid) to decide if any of the proposed schemes are feasible or not. Intrinsically, the calculations apply to the gaseous state only. Undoubtedly the solvent would have substantial effects on the pK_a's in the di-, tri-, and polynucleotides. To our knowledge no exact values of first and second protonation pK's are known for any of the cases of interest, undoubtedly due to the experimental difficulties encountered in determining sites and equilibrium constants of protonation in these systems. For example, it is certainly not unreasonable to expect that the pK for second protonation will be profoundly influenced by the first base protonation in the pair.

The more recent aqueous work indicates hydrogen bonding in GMP–UMP, AMP–CMP, and CMP–UMP pairs along with weaker bonds in GMP–CMP and AMP–UMP.[22c]

Critical Evaluation of Results

As all other calculations before ours, the present work also suffers from some obvious theoretical approximations.[1]

It is worth emphasizing that we have chosen a method which is uniformly applicable to both neutral and protonated interacting species. While results on neutral pair interactions abound,[12] ours is the first attempt to compare neutral pair interactions with half- and fully protonated pair ones. Many of our results on neutral pairs qualitatively follow those of Pullman and Pullman,[12c] but one still should not forget that even different physical methods predict quite different interaction enthalpies for the neutral pairs.[37]

The present approach includes only the monopole–monopole, monopole–induced dipole, and induced dipole–induced dipole terms in the intermolecular potential. The results are reported for fixed intermolecular separations, such separations being taken from available X-ray data. We do not expect that the calculations, as a function of intermolecular separation, would lead to a minimum energy geometry, since no repulsion (for example, of a Lennard-Jones or other nonbonded type) term is included. Our aim is, as a first approximation, to compare the interaction energies for very similar stacking and hydrogen-bonding distances varying only the possible arrangement of interacting species. Our interaction energies for neutral pairs are somewhat lower than those quoted from other theory, probably because of our underestimate of net atomic charges.[37a] Our calculations give interaction energies closer to experimental values.

Disagreement between our results and those of Pullman and Pullman[12c] are due to several factors (aside from our lack of knowledge of the precise geometric input employed by the other group).

Our monopole–monopole term is calculated with CNDO/2 net atomic charges. The reasonable dipole moments predicted by CNDO/2 are, however, due to the inclusion of the hybrid moment contribution[3,4] (monoatomic overlap charges). Without this second term the dipole moment is too small and the electron density map is not invariant to rotation of coordinate axes.[38]

Previous studies[12c] employed $\sigma + \pi$ net atomic charges which directly yielded reasonable dipole moments in a point charge approximation. Thus our net atomic charges (hence $\rho\rho$ and $\rho\alpha$ terms) are bound to be much smaller than those quoted by Pullman and Pullman.[12c] The variation in $\rho\rho$ terms is most evident in our much smaller hydrogen-bonding interaction energies.

Using isotropic bond polarizabilities and the ionization potentials previously quoted, we obtain larger dispersion terms ($\alpha\alpha$) than those referenced[12c] which employed smaller ionization potentials.

We feel that, for the present, the CNDO/2 net atomic charges represent a consistent approach to the monopole–monopole and monopole–induced dipole terms when comparing neutral to protonated substrate interactions. While the $\sigma + \pi$ approach, mimicking correct dipole moments in neutral species, is useful, its application to charged species is of doubtful value since in the latter one has no obvious experimental quantity to mimic. We are not at all certain that an iterated extended Hückel charge density calculation is any more meaningful than the one here employed, assuming such gives reasonable dipole moments for neutral species. Improved intermolecular potential calculations with *ab*

(37) (a) N. K. Kochetkov and E. I. Budovskii, "Organic Chemistry of Nucleic Acids," Plenum Press, London, 1971, Chapter 4; (b) "Molecular Associations in Biology," B. Pullman, Ed., Academic Press, New York, N. Y., 1968.

(38) H. L. Hase, H. Meyer, and A. Schweig, *Theor. Chim. Acta*, **28**, 99 (1972).

initio molecular orbital parameters are contemplated for the future.

Throughout the calculations the effects of solvation have been neglected since we have no simple method of accounting for such effects within the molecular orbital framework. That solvation can have a dramatic effect on base–base interactions has been shown by various groups. Most recently, based on solubility studies, it was shown that transfer of the DNA bases from organic solvents to water is accompanied by both positive ΔH and ΔS values.[39] It was suggested that the water–DNA base interactions are at least partially due to structure breaking on water by the bases. This further emphasizes the fact that base–base interactions are made up of several contributions of which we have only calculated three.

Summary and Conclusions

Eventual recalculation, using *ab initio* net atomic populations (not yet available) and more sophisticated intermolecular potentials, will undoubtedly change the quantitative, and perhaps even qualitative, predictions of our study. For the moment we suggest that our results warrant some of the following experiments.

The effects of protonation on electron distribution could be checked by ^{13}C nmr spectroscopy since it is now known that CNDO/2 atomic charges correlate with chemical shifts qualitatively. With the recent development of ^{15}N nmr measurements in natural abundance the sites of protonation can hopefully be pinned down.

While the sophisticated studies on dimers, trimers, and polymers have been of great value, one would like to know the pK_a's for various base-protonation processes in polymer (at least dimer or trimer) environments. Only with an exact knowledge of pK's can one hope to decide whether a certain process occurs between two protonated bases or between a monoprotonated one and a neutral one. Or, whether Tinoco, *et al.*'s, results near pH 1 refer to a doubly protonated pair or a monoprotonated one, or perhaps a mono base protonation and a primary phosphate protonation process.

The determination of the pK's is very difficult but perhaps feasible on microquantities using ORD, CD, or absorption spectra (the latter show fairly small changes on protonation).

One can finally suggest that, with the discovery of significant H-bonding in H_2O,[22c] it should be feasible to check if our most consistent prediction, that half-protonated pairs are more stable than neutral ones in both hydrogen-bonding and stacking modes, is correct. Such studies, however, must be done near the pK_a's of the species.

Our findings on doubly protonated pairs are surprising since apparently there is ample precedent for these in the solid state.[20b] We can, at least, be assured that in these molecules phosphate–base electrostatic interactions must be important.

We acknowledge the tentative nature of our results which must be considered as a first step toward the elucidation of the effects of base ionization on base–base interactions. Our results suggest that the selectivity found in hydrogen-bonding schemes of neutral bases is decreased upon protonation. Protonation of cytosine, for example, could introduce stabilization of schemes not considered significant ordinarily.

That protonation affects not only base–base interactions but perhaps even nucleoside and nucleotide conformational preferences has recently been suggested.[40]

Acknowledgment. Computer time was generously provided by the Rutgers University Center for Computer and Information Service. We thank Professor C. E. Bugg for sending us an offprint of his contribution quoted in the manuscript.

(39) R. L. Scruggs, E. K. Achter, and P. D. Ross, *Biopolymers*, 11, 1961 (1972).

(40) T. D. Son, W. Guschlbauer, and M. Guéron, *J. Amer. Chem. Soc.*, **94**, 7903 (1972).