April 15, 1992

UV Photoelectron and Theoretical Characterization of 2'-Deoxyguanosine-5'-phosphate Valence Electronic Properties: Changes in Structure Associated with the B to Z-DNA Conformational Transition

Min Yu, Ho Soon Kim and Pierre R. LeBreton

Department of Chemistry, The University of Illinois at Chicago Chicago, Illinois 60680

Received January 21, 1992

Summary: He(I) UV photoelectron spectroscopy and *ab initio* SCF molecular orbital calculations with the 4-31G basis set have been employed to characterize the valence electronic structures of 2'-deoxyguanosine-5'-phosphate (5'-dGMP'). In 5'-dGMP', the electron distributions of the upper occupied orbitals are localized and similar to those appearing in 1,9-dimethylguanine (1), 3-hydroxytetrahydrofuran (2) and CH₃HPO₄ (3). Theoretical ionization potentials (IP's) of 5'-dGMP' (4) have been obtained by applying Koopmans' Theorem to the 4-31G SCF results. The IP's of seven orbitals in the base and sugar groups in 4, predicted from the 4-31G SCF calculations, have been individually corrected by comparison to results from 4-31G SCF calculations on neutral 5'-dGMP, and to HeI photoelectron spectra of the model compounds, 1 and 2. The IP's of six of the highest occupied orbitals of the phosphate group in 4 and in the model anion 3, predicted from 4-31G SCF calculations, have been corrected by comparing 4-31G SCF results for PO₂: to theoretical IP's obtained from second-order Moller-Plesset perturbation calculations on PO₂: For 4 in the conformation occurring in B-DNA, the first IP's associated with the phosphate, base, and sugar groups occur at 5.1, 5.6 and 6.6 eV, respectively. A comparison of the valence electronic structures of 4 in geometries associated with the B and Z-DNA conformations indicates that in B-DNA the base and sugar orbitals have lower IP's than in Z-DNA, while the phosphate orbitals have higher IP's.

Gas phase UV photoelectron spectroscopy, used in conjunction with molecular orbital calculations, has provided detailed, experimentally based pictures of valence orbital structures of the smallest DNA nucleotide, 2'-deoxycytidine-5'-phosphate (5'-dCMP'), which contains 160 electrons (1). Here we report results in which we have employed this approach to describe the valence electronic structure of the largest DNA nucleotide, 2'-deoxyguanosine-5'-phosphate (5'-dGMP), which contains 180 electrons.

Photoelectron spectra (PES) of five methyl substituted guanines have been previously assigned (2,3). Results from past investigations of methyl substituted guanines, and of other nucleotide components, demonstrate that there is often good agreement between PES and the energetic ordering and spacing of valence orbitals predicted by SCF calculations (1,3,4). However, only molecules of limited size can be examined by directly comparing PES with results from theoretical calculations. It is not currently possible to obtain highly resolved, gas phase PES of nucleotides. An approach (1) which was previously applied

to 5'-dCMP' and which is employed here for 5'-dGMP' (4) is less direct. With this method, nucleotide ionization potentials (IP's) are obtained from results of 4-31G SCF calculations by a correction procedure in which base, sugar and phosphate orbitals in the anionic nucleotides are correlated with orbitals in neutral nucleotides and in small model molecules and anions. For the models, IP's can be measured or accurately calculated. The correlation relies on the observation that SCF calculations with a split valence basis set describe upper occupied orbitals in nucleotides which are localized on either the base, sugar or phosphate groups. For each of the upper occupied base and sugar orbitals in 4, the IP predicted by 4-31G SCF calculations was corrected by using a method (1) which compares the IP of the correlating orbital in neutral 5'-dGMP, obtained from 4-31G SCF calculations, to the experimental IP of the correlating orbital in 1,9-dimethylguanine (1) or 3-hydroxytetrahydrofuran (2). For phosphate orbitals in 4 and CH₃HPO₄ (3), IP's predicted by 4-31G SCF calculations were corrected by using a scaling equation (1) derived by comparing the values of the first five IP's of PO₂, obtained from 4-31G calculations, with values obtained by using second-order Møller-Plesset perturbation theory (MP2) (5). The MP2 calculations were performed on PO₂ because this anion is small enough to permit post-SCF calculations of multiple IP's using available computational resources, and because, like 3 and the phosphate group in 4, PO₂ contains phosphorous and oxygen.

Prior to our combined use of PES and computational results to describe the valence structure of 5'dCMP' (1), valence and core electron photoionization spectra for several single base containing fragments
of DNA were predicted from results of *ab initio* SCF calculations using a single zeta basis set (11). However,
in the earlier study, no comparison of experimental IP's and theoretical results was carried out.

MATERIALS AND METHODS

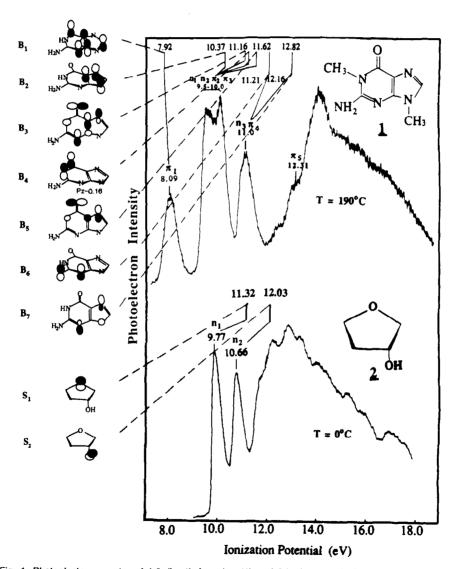
Experimental Procedures. He(I) photoelectron spectra of 1,9-dimethylguanine (1) and 3-hydroxytetrahydrofuran (2) were measured with a Perkin-Elmer PS-18 spectrometer. The spectrometer probe temperatures used in the experiments are given in Figure 1. The spectra were calibrated by using the $^2P_{3/2}$ and $^2P_{1/2}$ bands of argon and xenon. A sample of $\underline{1}$ was synthesized as described previously (3). A sample of $\underline{2}$ was obtained from Aldrich Chemical Co. For $\underline{1}$ and $\underline{2}$, spectra measured from a single sample over a period of three hours were identical, indicating that no decomposition occurred. Molecular Orbital Calculations And Geometries. For 5'-dGMP' (4) and for the model compounds and anions, IP's have been obtained from results of 4-31G SCF (12) calculations through the application of Koopmans' Theorem (13). Computations were carried out on Cray 2, and IBM 3090/300J/VF and 600J/VF computers. The GAUSSIAN 88 program (14) was used for all calculations. Molecular orbital diagrams were drawn, as described previously (1), from 4-31G SCF molecular orbital coefficients.

The geometry of 1 employed in the calculations was obtained from crystallographic data (3). The geometry of 2 was obtained by combining electron diffraction data for tetrahydrofuran (15) with standard OH bond lengths and bond angles (16). The C-O bond length in the hydroxyl group was 1.43 Å; the CCO bond angle was 108.5°. For the 4-31G SCF calculations, geometries of PO₂, PO₃, H₂PO₄, and CH₃HPO₄ (3) were the same as those used in the earlier investigation of 5'-dCMP (1). The geometry of the phosphate group of neutral 5'-dGMP was based on that for trimethylphosphate (1,17). The O-H bond lengths and HOP bond angles are 1.06 Å and 107.6°. Heavy atom bond angles and bond lengths in 5'-dGMP'(4) and in the base and sugar groups of 5'-dGMP were taken from crystallographic data for deoxycytidylyl-(3'-5')-deoxyguanosine (18). Bond lengths and bond angles involving the hydrogen atoms

in $\underline{1}$ and $\underline{2}$, and in the base and sugar groups of 5'-dGMP and 5'-dGMP ($\underline{4}$) were: C-H = 1.11, O-H = 0.95, N-H =1.01 Å, HCH = 109.5°, HOC = 105.0°, HNH=120° and HNC=120° (16). Torsional angles used to define the geometries of 5'-dGMP and $\underline{4}$ in the B-DNA conformation were taken from optimized crystallographic parameters for B-DNA (19). The geometry of $\underline{4}$ in the Z-DNA conformation was taken from crystallographic data for a d(CpGpCpGpCpG) hexamer (20).

RESULTS

Photoelectron Spectra of 1,9-Dimethylguanine (1) and 3-Hydroxytetrahydrofuran (2): Figure 1 contains the photoelectron spectrum of 1,9-dimethylguanine (1) along with previous assignments (3) and IP's obtained from 4-31G SCF calculations. Theoretical IP's are given above the experimental vertical IP's. The figure



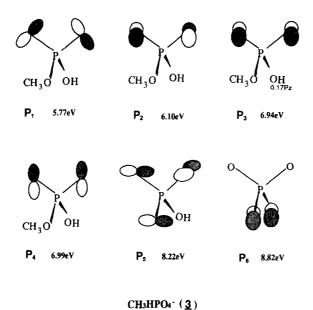
<u>Fig. 1.</u> Photoelectron spectra of 1,9-dimethylguanine (1) and 3-hydroxyletrahydrofuran (2). Theoretical ionization potentials and orbital diagrams obtained from 4-31G SCF calculations. For $\underline{1}$, orbital diagrams illustrate major atomic orbital contributions to the seven highest occupied molecular orbitals (B₁ to B₇). In π and lone-pair orbital diagrams the viewing angle is different. For $\underline{2}$, diagrams are given for the two highest occupied oxygen atom lone-pair orbitals, S_1 and S_2 .

also shows orbital diagrams. A comparison of the theoretical and experimental results indicates that the first IP of $\underline{1}$, which arises from a π orbital, is accurately calculated to within 0.17 eV, and that the theoretical IP's of the second, third and fourth highest occupied π orbitals are larger than the experimental values by 0.4-0.9, 1.2-1.7 and 1.2 eV, respectively. For the three highest occupied lone-pair orbitals in $\underline{1}$, the theoretical IP's are 1.2-1.7, 1.6-2.1 and 1.8 eV higher than the experimental values.

Figure 1 also shows the spectrum of 3-hydroxytetrahydrofuran (2) together with theoretical IP's, and orbital diagrams for two bands which arise from O atom lone-pair orbitals. As in the case of 1, 4-31G SCF calculations on 2 yield IP's for lone-pair orbitals which are too large. For 2, the theoretical IP's are 1.37 and 1.43 eV larger than the experimental IP's.

<u>Ionization Potentials of Model Anions Predicted by SCF and MP2 Calculations</u>: Calculations of anion IP's were carried out by two different methods: (Method I) SCF calculations on the anions at the 4-31G level using Koopmans' Theorem (13); and (Method II) MP2 calculations on both anions and radicals at the 6-31+G level, using optimized geometries for anions and the same geometries for the corresponding radicals (1,5,13). In the MP2 calculations, the core electrons were frozen and the nonsinglet states were described using the unrestricted Hartree-Fock formalism (21).

Figure 2 shows orbital diagrams and IP's for the six highest occupied orbitals in CH₃HPO₄ (3) obtained from 4-31G SCF calculations. In order to evaluate IP's of closed shell anions containing phosphorus and



<u>Fig. 2.</u> Orbital diagrams and ionization potentials of the seven highest occupied orbitals in CH_3HPO_4 (3) predicted by 4-31G SCF calculations. For the geometry employed in the calculation torsional angles are the same as those associated with phosphate group in 5'-dGMP'(4) in the conformation occurring in B-DNA.

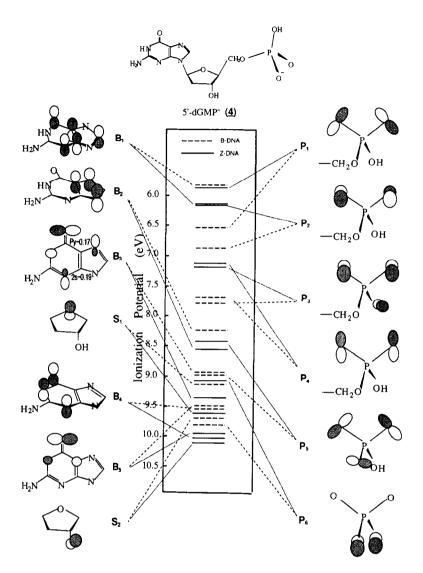
oxygen, the IP's of PO₂, PO₃ and H₂PO₄ obtained from 4-31G SCF calculations were compared to IP's obtained from the MP2 calculations. The IP's of PO₂ and PO₃, obtained from the calculations, were also compared to the experimental values (6-10) of 3.3 ± 0.2 eV and 4.9 ± 1.3 eV, respectively. For PO₂, the IP's obtained using either the SCF method (3.20 eV) or the MP2 method (3.34 eV) are similar to one another, and both agree well with experiment. On the other hand, the two methods yield values for the IP's of PO₃ and H₂PO₄, which differ significantly. For PO₃ and H₂PO₄, the first IP's obtained from the 4-31G SCF calculations are 6.60 and 5.31 eV. The IP's obtained from the MP2 calculations are 4.90 and 4.89 eV. In general, the MP2 method is more reliable. For example, the IP of PO₃ obtained from the MP2 calculation agrees better with the experimental value than the IP obtained from the 4-31G SCF calculation.

Theoretical Ionization Potentials of 5'-dGMP'(4) Predicted by 4-31G SCF Calculations: Figure 3 shows IP's and orbital diagrams for the thirteen highest occupied orbitals in 5'-dGMP' (4) in the B and Z-DNA conformations as predicted by 4-31G SCF calculations. The 4-31G SCF results indicate that the total charges on the base and sugar groups in neutral and anionic 5'-dGMP are similar. In B-DNA, the difference between the charges on the base in the neutral and the anion is only 0.021 e. For the sugar group, the difference is 0.101 e. In addition to similarities between charges on the base and sugar groups in neutral and anionic 5'-dGMP, the electron distributions in the upper occupied base (B₁ to B₅) and sugar orbitals (S₁ and S₂) in 5'-dGMP' (4) are similar to corresponding orbitals in neutral 5'-dGMP and in the model compounds 1,9-dimethylguanine (1) and 3-hydroxytetrahydrofuran (2). For 1, 2 and 4, these similarities are demonstrated in Figures 1 and 3.

A examination of the orbital diagrams in Figures 2 and 3 indicates that in $CH_3HPO_4^-$ (3) and 5'-dGMP' (4) the phosphate orbitals (P_1 to P_6) are made up primarily from O atom 2p orbitals. Figures 2 and 3 also indicate that corresponding phosphate orbitals in 3 and 4, like corresponding base and sugar orbitals in 1, 2 and 4 have nearly identical orbital diagrams.

DISCUSSION

Valence Orbital ionization Potentials of Base and Sugar Orbitals in 5'-dGMP'(4): Ionization potentials of the B_1 to B_5 , and S_1 and S_2 orbitals in $\underline{4}$ predicted by the 4-31G SCF calculations have been corrected by employing eqs 1 and 2, which are analogous to equations previously derived for 5'-dCMP' (1). Equations 1 and 2 are based on the finding that the difference between the true IP of one of the upper occupied base or sugar orbitals in anionic nucleotides and the IP obtained from 4-31G SCF calculations is approximately equal to the difference between the true IP and the calculated IP of the corresponding orbital in the neutral nucleotide (1). Equations 1 and 2 also rely on the observation that the true IP of one of the upper occupied



<u>Fig. 3.</u> Valence orbital IP's, obtained from 4-31G SCF calculations, for 5'-dGMP'(4) in the conformations occurring in B and Z-DNA. Orbital diagrams are drawn for <u>4</u> in the conformation occurring in B-DNA. Orbitals located primarily in the base, sugar, and phosphate groups are denoted B, S and P, respectively.

base or sugar orbitals in a neutral nucleotide is nearly equal to that of the corresponding orbital in a model compound (1).

$$IP_{cor,\underline{4}}(i) = IP_{calc,\underline{4}}(i) - \Delta IP$$
 (1)

$$\Delta IP = IP_{calc,5'-dGMP}(i) - IP_{exp}(i)$$
 (2)

In eq 1, IP_{cor.4} (i) is the corrected IP for the i'th base or sugar orbital in 5'-dGMP⁻ (4), and IP_{calc.4}(i) is the IP obtained from 4-31G SCF calculations on <u>4</u>. In eq 2, IP_{calc.5'-dGMP} (i) is the IP of the i'th orbital obtained from 4-31G SCF calculations on neutral 5'-dGMP; IP_{exp} (i) is the experimental IP of the i'th orbital in either 1,9-dimethylguanine (1) or 3-hydroxytetrahydrofuran (2). Values of IP_{exp} (i) were obtained from the PES in

Orbital	Orbital Type	B-DNA		Z-DNA	
		IP _{calc} b	IP _{cor} c	IP _{clac}	IP _{cor}
		Bas	se Orbitals	/n:	
В,	π	5.73	5.6	6.17	6.2
B ₂	π	8.24	6.8-7.3	8.56	7.4-7.9
B_3	n	8.87	6.9-7.4	9.48	7.6-8.1
B ₄	π	9.53	7.3-7.8	10.07	7.8-8.3
B ₅	n	9.50	7.2-7.7	9.95	7.7-8.2
·		Sug	gar Orbitals		
S,	n	9.14	6.6	9.61	7.1
S ₁ S ₂	n	9.68	7.6	10.11	8.4
-		Pho	sphate Orbitals		
P ₁	n	6.54	5.1	5.88	4.8
P,	n	6.88	5.3	6.19	4.9
P,	n	7.79	5.8	7.20	5.5
P ₂ P ₃ P ₄	n	7.71	5.8	7.13	5.4
P ₅	n	8.98	6.5	8.42	6.4
P ₆	n	9.84	7.0	9.09	6.5

Table I. Theoretical and Estimated Ionization Potentials of 5'-dGMP-a

Figure 1. Table I lists corrected IP's of the B_1 to B_5 , and S_1 and S_2 orbitals in $\underline{4}$, obtained via eqs 1 and 2. <u>Valence Orbital Ionization Potentials of Phosphate Orbitals in 5'-dGMP'(4)</u>: Ionization potentials predicted by the 4-31G SCF calculations for the P_1 to P_6 orbitals in $\underline{4}$ have been corrected by noting that the true IP's for the P_1 to P_6 orbitals in CH_3HPO_4 (3) and in $\underline{4}$ are nearly equal. Corrected IP's for the P_1 to P_6 orbitals in $\underline{3}$ are given by equation 3, which was derived by comparing values of the first five IP's in PO_2 , predicted by MP2 calculations, with values predicted by 4-31G SCF calculations (1). Equation 3 is the result of a linear regression analysis of these two sets of values (22,23).

$$IP_{cor,3}(j) = 0.55 IP_{calc,3}(j) + 1.54$$
 (3)

In eq 3, IP_{cor,3} (j) is the corrected IP for the j'th phosphate orbital (j=1-6) in $\underline{3}$ and IP_{calc,3} (j) is the IP obtained from the 4-31G SCF calculation. A test of the accuracy of the procedure used for correcting the phosphate anion IP's is provided by considering the corrected IP (5.05 eV) obtained by applying eq 3 to 4-31G SCF results on PO₃. Here the corrected IP differs from the experimental IP by only 0.2 eV (7-10). Corrected values for the IP's of orbitals P₁ to P₆ in $\underline{4}$ are given in Table I.

A Comparison 5'-dGMP'(4) Ionization Potentials in B and Z-DNA: Figure 3 and Table I compare IP's for the thirteen highest occupied orbitals in $\underline{4}$ occurring in B and Z-DNA. The results indicate that the B₁ to B₅, and S₁ and S₂ orbitals in B-DNA have smaller IP's than corresponding orbitals in Z-DNA. The sum of the corrected IP's of the three highest occupied π orbitals (B₁, B₂ and B₄) of 5'-dGMP' in B-DNA is 1.7 eV smaller than the sum of the IP's of the corresponding orbitals in Z-DNA. As demonstrated by previous

^aAll ionization potentials in eV.

bTheoretical ionization potentials obtained from 4-31G SCF calculations.

^cCorrected ionization potentials. See text.

investigations of association constants for the stacking of nucleosides, and for the formation of nucleosideriboflavin complexes, base stacking interactions generally increase as base IP's decrease (24). Our results indicating that π IP's of 5'-dGMP' are lower in B than in Z-DNA, suggest that base stacking interactions are stronger in B than in Z-DNA. While the results of this comparison of 5'-dGMP' in B and Z-DNA may be altered when the influence of adjacent nucleotides, water molecules and counterions is considered, the results are consistent with the finding that the distance between stacked bases in B-DNA is smaller than in Z-DNA. The rise per base pair in B and Z-DNA is 3.3 and 3.8 Å, respectively (25).

ACKNOWLEDGMENTS: Support of this work by The American Cancer Society (Grant #CN-37), the Petroleum Research Fund (Grant #21314-AC), The National Cancer Institute of the National Institutes of Health (Grant #CA41432) and Cray Research, Inc. is gratefully acknowledged. Computer access time has been provided by the Computer Center of The University of Illinois at Chicago, the Cornell National Supercomputer Facility, and the National Center for Supercomputing Applications, at The University of Illinois at Urbana-Champaign. The authors would like to thank Professor Nelson Leonard (The University of Illinois at Urbana-Champaign) for supplying a sample of 1,9-dimethylguanine, and Ms. Sharon Fetzer (The University of Illinois at Chicago) for measuring the photoelectron spectrum of 3-hydoxytetrahydrofuran.

REFERENCES

- Tasaki, K.; Yang, X.; Urano, S.; Fetzer, S.; LeBreton, P.R. (1990) J. Am. Chem. Soc. 112, 538-548.
- Lin, J.; Yu, C.; Peng, S.; Akiyama, I.; Li, K.; Lee, L. -K.; LeBreton, P. R. (1980) J. Phys. Chem. <u>84</u>, 1006-1012.
- LeBreton , P.R.; Yang, X.; Urano, S.; Fetzer, S.; Yu, M.; Leonard, N.J.; Kumar, S. (1990) J. Am. Chem. Soc. <u>112</u>, 2138-2147.
- 4. For a review see: Urano, S.; Yang, X.; LeBreton, P.R. (1989) J. Mol. Struct. 214, 315-328.
- 5. Møller, C.; Plesset, M. S. (1934) Phys. Rev. 46, 618-622.
- 6. Wu, R. L. C.; Tiernan, T.O. (1982) Bull. Am. Phys. Soc. 27, 109 (Session JA-6).
- Henchman, M.; Vigianno, A. A.; Paulson, J. F.; Freedman, A.; Wormhoudt, J.C. (1985) J. Am. Chem. Soc. <u>107</u>, 1453-1455.
- 8. Unkel, W.; Freedman, A. (1983) AIAA 21, 1648-1651.
- Freedman, A.; Warmhoudt, J. C.; Kolb, C. E. (1982) in Metal Bonding and Interactions in High Temparature Systems, ACS Symposium Series 179; Gole, J. L.; Stwalley, W. C. Eds.; American Chemical Society: Washington P. 609.
- 10. Lohr, L. L.; Boehm, R. C. (1987) J. Phys. Chem. 91, 3203-3207.
- 11. Vercauteren, D. P.; Clementi, E. (1983) Int. J. Quantum. Chem.: Quantum. Biol. Symp. 10, 11-30.
- 12. W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. (1986) Ab Initio Molecular Orbital Theory; John Wiley & Sons: New York p. 86 and references cited therein.
- 13. Koopmans, T. (1933) Physica 1, 104-113.
- 14. The Gaussian 88 program was obtained from J. A. Pople of Carnegie-Mellon University.
- 15. Geise, H. J.; Adams, W. J.; Bartell, L. S. (1969) Tetrahedron 25, 3045-3052.
- Tables of Interatomic Distances and Configurations in Molecules and Ions; The Chemical Society, Burlington House: London (1965).
- LeBreton, P. R.; Fetzer, S.; Tasaki, K.; Yang, X.; Yu, M.; Slutskaya, Z.; Urano, S. (1988) J. Biomol. Str. Dyn. 6, 199-222.
- Cruse, B.T.; Egert, E.; Kennard, O.; Sala, G.B.; Salisbury, S.A.; Visamitra, M.A.(1983) Biochemistry 22, 1833-1839.
- 19. Arnott, S.; Hukins, D. W. L. (1972) Biochem. Biophys. Res. Commun. 47, 1504-1509.
- Wang, A., H.-J.; Quigley, J.G.; Kolpak, F.J.; van der Marel, G.; van Boom, J.H.; Rich, A. (1981) Science <u>211</u>, 171-176.
- 21. Pople, J. A.; Nesbet, R. K. (1954) Chem. Phys. 22, 571-572.
- Lambert, J. B.; Xue, L.; Bosch, R. J.; Taba, K. M.; Marko, D. E.; Urano, S.; LeBreton, P. R. (1986) J. Am. Chem. Soc. 108, 7575-7588.
- 23. Santiago, C.; Houk, K. N.; DeCicco, G. J.; Scott, L. T. (1978) J. Am. Chem. Soc. 100, 692-696.
- 24. Yu, C.; O'Donnell, T.J.; LeBreton, P.R. (1981) J. Phys. Chem. 85, 3851-3855.
- Dickerson, R.E.; Drew, H.R.; Conner, B.N.; Wing, R.M.; Fratini, A.F.; Kopka, M.L. (1981) Science <u>216</u>, 171-176.