
**EFFECTS OF HYDRATION AND STACKING INTERACTIONS
ON THE ELECTRONIC STRUCTURE OF DNA MODELS**

Christopher-Maria LIEGENER, Rung Shen CHEN*, Peter OTTO and Janos LADIK

*Chair for Theoretical Chemistry and Laboratory of the National Foundation for Cancer Research,
Friedrich-Alexander-University Erlangen-Nürnberg, D-8520 Erlangen, F.R.G.*

Received August 10th, 1987

Accepted August 24th, 1987

Dedicated to Dr R. Zahradnik on the occasion of his 60th birthday.

The energy band structures of a cytosine, adenine, and guanine stack in the presence of water have been calculated by the *ab initio* crystal-orbital method. The surrounding water molecules have been simulated by arrays of point charges, using for their positions the results of previous Monte-Carlo calculations of the corresponding polynucleotides. Furthermore, the effects of internucleotide interactions have been studied on the basis of calculations on base dimer stacks in comparison to the corresponding dimers and monomers.

DNA like most biomolecules occurs *in vivo* in hydrated form. This means that the hydration shell has to be included in some way in a realistic study of the electronic structure of DNA or models of DNA.

Some of the most frequently studied models for DNA are nucleotide base stacks¹⁻¹². The effects of hydration on the electronic structure of base stacks have been studied in a preliminary calculation using the water cluster around a cytidine (cytosine + + sugar + phosphate) unit to simulate the hydration shell around a cytosine stack¹⁰. In the present work we have calculated the band structure of polycytosine, polyadenine, and polyguanine base stacks in the effective field produced by water molecules the positions of which were obtained from Monte-Carlo calculations on a periodic single helix of nucleotide base stacks in water environment¹¹.

The simplest extension of this polynucleotide model is the inclusion of interactions between different nucleotides. Two simple possibilities for more complex models exist: First, the interaction of nucleotide bases belonging to two different strands can be included. This would model the interstrand interaction of Watson-Crick type double helices. Calculations on such models have already been performed^{8,9}. Another extension is to consider stacks of pairs of nucleotide bases on top of each other (i.e. a kind of periodic copolymers of two nucleotide bases). This second possibility has

* Permanent address: Institute for Biophysics, Academia Sinica, Beijing, People's Republic of China.

been investigated by cluster calculations¹² and will be investigated now by polymeric crystal orbital calculations in the second part of the present work.

THEORETICAL

The *ab initio* crystal-orbital method^{13,14} was used for the calculation of the electronic band structures. This means one has to solve the matrix equation

$$\mathbf{F}(k) \cdot c_n(k) = \varepsilon_n(k) \mathbf{S}(k) \cdot c_n(k) \quad (1)$$

with

$$\mathbf{F}(k) = \sum_{q=-\infty}^{\infty} \exp(ikqa) \mathbf{F}(q), \quad (2)$$

$$\mathbf{S}(k) = \sum_{q=-\infty}^{\infty} \exp(ikqa) \mathbf{S}(q), \quad (3)$$

where $\varepsilon_n(k)$ gives the band structure, $\mathbf{F}(q)$ and $\mathbf{S}(q)$ are the Fock- and overlap matrices, defined by

$$[\mathbf{S}(q)]_{rs} = \langle \chi_r^0 | \chi_s^q \rangle, \quad (4)$$

$$\begin{aligned} [\mathbf{F}(q)]_{rs} = & \langle \chi_r^0 | -\Delta/2 - \sum_{q_1=-\infty}^{\infty} \sum_{n=1}^M (Z_n / |r - R_n^{q_1}|) + V_{\text{eff}} | \chi_s^q \rangle + \\ & + \sum_{q_1, q_2=-\infty}^{\infty} \sum_{u, v=1}^m P(q_1 - q_2)_{uv} [\langle \chi_r^0 \chi_u^{q_1} | \chi_s^q \chi_v^{q_2} \rangle - (1/2) \langle \chi_r^0 \chi_u^{q_1} | \chi_v^{q_2} \chi_s^q \rangle], \end{aligned} \quad (5)$$

where χ_r^q is the r -th atomic basis function placed in cell q , m is the number of basis functions, M is the number of atoms per cell and $P(q_1 - q_2)_{uv}$ are the elements of the density matrix for the polymer.

In the case of a hydrated polymer V_{eff} is the electrostatic potential of arrays of point charges simulating the surrounding water molecules¹⁵; in the case of a free polymer $V_{\text{eff}} = 0$. The positions of the water molecules have been taken from ref.¹¹ and each molecule has been represented by an array of 23 point charges¹⁵. Similar calculations have recently been performed on several polypeptides in the β -pleated sheet^{15,16} and α -helical¹⁷ conformations.

All band structure calculations have been performed in a strict second-neighbors' interaction approximation using Clementi's $7s3p/2s1p$ (for C, N and O) and $4s/1s$ (for H) minimal basis sets¹⁸.

RESULTS AND DISCUSSION

Nucleotide Base Stacks with Water Interaction

In Table I the results are listed of the band structure calculations on cytosine, guanine and adenine stacks in the free state and in the presence of water. We abbreviate the system names by $(C)_x$ for cytosine, $(G)_x$ for guanine and $(A)_x$ for adenine stack. For all systems upper and lower edges for valence and conduction bands, band widths, the band gap and the total energy per elementary cell are given. It is interesting to note that there is no uniform trend neither for the band shifts nor for the changes of band widths, gaps or total energy. This means that one can make predictions of such effects only on the basis of explicit calculations. In particular, the total energy (which refers only to the stack, not to the hydration shell) is in the case of $(A)_x$ destabilized, in the case of $(C)_x$ and $(G)_x$ stabilized, due to the presence of water.

The band widths and band gaps are not changed much by the effects of hydration (the changes for the widths are about 0.05 eV, the changes for the gaps about 0.2 to 0.4 eV). The band positions, on the other hand, are changed by up to 4 eV (in the case of $(G)_x$). This means that the bands are shifted more or less parallel. The limited changes of band gap and band widths imply that hydration has only limited influence on the conduction properties of DNA.

TABLE I

Band structure results for single base stacks in the presence of water (total energies in a.u.^a, all other energies in eV)

System	Band type ^b	Maximum	Minimum	Band width	Gap	Total energy per unit cell
$(C)_x$ free	V	-9.405	-9.423	0.018		
	C	2.212	1.802	0.410	11.207	-391.160147
$(C)_x$ hydrated	V	-7.535	-7.557	0.022		
	C	4.385	3.907	0.478	11.442	-387.031925
$(G)_x$ free	V	-8.775	-9.312	0.537		
	C	2.808	2.672	0.136	11.447	-537.460240
$(G)_x$ hydrated	V	-4.797	-5.314	0.517		
	C	6.410	6.293	0.117	11.090	-528.748393
$(A)_x$ free	V	-9.966	-10.106	0.140		
	C	2.413	2.309	0.103	12.275	-462.834059
$(A)_x$ hydrated	V	-12.556	-12.773	0.218		
	C	-0.352	-0.517	0.164	12.039	-470.432819

^a 1 a.u. = 2 625.5 kJ mol⁻¹, ^b V stands for valence, C for conduction band.

TABLE II
Results for molecules and clusters (total energy in a.u. all other energies in eV)

System	Total energy a.u.	Orbital no.	Orbital energy	Gap	Type
C	-391.166548	n + 2	3.895	11.582	π^*
		n + 1	1.636		π^*
		n	-9.946		$\pi(O; N)$
		n - 1	-11.902		π
		n - 2	-11.929		$\sigma(O; N)$
T	-449.847671	n + 2	3.778	12.569	π^*
		n + 1	1.585		π^*
		n	-10.924		π
		n - 1	-12.332		$\pi(O, N)$
		n - 2	-12.992		$\sigma(O, N)$
G	-537.470635	n + 2	3.389	11.916	π^*
		n + 1	2.932		π^*
		n	-9.484		π
		n - 1	-12.311		$\sigma(O, N)$
		n - 2	-12.402		π
A	-462.830384	n + 2	3.026	12.291	π^*
		n + 1	1.923		π^*
		n	-10.368		π
		n - 1	-12.035		π
		n - 2	-12.221		$\sigma(N)$
$\begin{array}{c} G \\ \\ C \end{array} \begin{array}{c} (C) \\ \\ (G) \end{array}$	-928.640423 (-928.637734) ^a	n + 2	2.633	11.018 (11.194)	$\pi^*(G) \begin{pmatrix} 3.887 & \pi^*(G) \\ 2.125 & \pi^*(C) \\ -9.068 & \pi(G) \\ -9.515 & \pi(C) \\ -11.419 & \pi(C) \end{pmatrix}^a$
		n + 1	1.646		$\pi^*(C)$
		n	-9.372		$\pi(G)$
		n - 1	-10.002		$\pi(C)$
		n - 2	-11.792		$\pi(C)$
$\begin{array}{c} G \\ \\ A \end{array} \begin{array}{c} (A) \\ \\ (G) \end{array}$	-1 000.300372 (-1 000.298902) ^a	n + 2	2.877	11.205 (11.581)	$\pi^*(G) \begin{pmatrix} 2.641 & \pi^*(G) \\ 2.285 & \pi^*(A) \\ -9.296 & \pi(G) \\ -10.002 & \pi(A) \\ -11.645 & \pi(A) \end{pmatrix}^a$
		n + 1	1.861		$\pi^*(A)$
		n	-9.344		$\pi(G)$
		n - 1	-10.348		$\pi(A)$
		n - 2	-12.067		$\pi(A)$
$\begin{array}{c} T \\ \\ G \end{array}$	-987.317653	n + 2	2.495	11.542	$\pi^*(G)$
		n + 1	2.096		$\pi^*(T)$
		n	-9.446		$\pi(G)$
		n - 1	-10.504		$\pi(T)$
		n - 2	-11.875		$\pi(T)$

^a The numbers in parenthesis refer to the dimer with the opposite arrangement.

TABLE III
Results for stacks (total energy in a.u., all other energies in eV)

System	Total energy a.u.	Band no.	Minimum	Maximum	Width	Gap	Type
(C) _x	-391.160147	n + 2	4.294	4.487	0.193	11.207	
		n + 1	1.802	2.212	0.410		
		n	-9.423	-9.405	0.018		
		n - 1	-11.394	-11.323	0.071		
		n - 2	-11.666	-11.391	0.276		
(T) _x	-449.848888	n + 2	3.841	4.252	0.412	11.994	
		n + 1	1.408	1.885	0.477		
		n	-11.047	-10.580	0.461		
		n - 1	-12.222	-12.029	0.193		
		n - 2	-12.763	-12.709	0.054		
(G) _x	-537.460240	n + 2	3.772	3.898	0.126	11.447	
		n + 1	2.672	2.808	0.136		
		n	-9.312	-8.775	0.537		
		n - 1	-11.859	-11.850	0.009		
		n - 2	-11.886	-11.869	0.017		
(A) _x	-462.834059	n + 2	3.234	3.636	0.402	12.275	
		n + 1	2.309	2.413	0.103		
		n	-10.106	-9.966	0.140		
		n - 1	-11.630	-11.568	0.092		
		n - 2	-11.853	-11.735	0.118		
$\begin{pmatrix} G \\ \\ C \end{pmatrix}_x$	-928.762925	n + 2	3.143	3.252	0.109	11.209	$\pi^*(G)$
		n + 1	1.961	2.024	0.063		$\pi^*(C)$
		n	-9.264	-9.248	0.016		$\pi(G)$
		n - 1	-9.723	-9.720	0.003		$\pi(C)$
		n - 2	-11.158	-11.155	0.003		$\pi(C)$
$\begin{pmatrix} G \\ \\ A \end{pmatrix}_x$	-1 000.461916	n + 2	2.782	2.810	0.028	11.670	$\pi^*(G)$
		n + 1	2.297	2.305	0.008		$\pi^*(A)$
		n	-9.417	-9.373	0.043		$\pi(G)$
		n - 1	-10.308	-10.258	0.050		$\pi(G)$
		n - 2	-11.564	-11.525	0.039		$\pi(A)$
$\begin{pmatrix} T \\ \\ G \end{pmatrix}_x$	-987.438756	n + 2	2.558	2.781	0.223	11.634	$\pi^*(G)$
		n + 1	2.114	2.224	0.110		$\pi^*(T)$
		n	-9.555	9.520	0.035		$\pi(G)$
		n - 1	-10.449	-10.440	0.009		$\pi(G)$
		n - 2	-11.885	-11.763	0.012		$\pi(T)$
$\begin{pmatrix} C \\ \\ T \end{pmatrix}_x$	-841.016065	n + 2	1.946	1.973	0.027	11.452	
		n + 1	1.693	1.797	0.104		
		n	-9.762	-9.759	0.002		
		n - 1	-10.658	-10.602	0.056		

TABLE III
(Continued)

$\begin{pmatrix} A \\ \\ T \end{pmatrix}_x$	-912.681505	n + 2	2.123	2.135	0.012	12.036
		n + 1	1.875	1.888	0.013	
		n	-10.171	-10.161	0.010	
		n - 1	-10.689	-10.675	0.014	
$\begin{pmatrix} A \\ \\ C \end{pmatrix}_x$	-853.995854	n + 2	2.541	2.551	0.010	11.505
		n + 1	1.996	2.043	0.047	
		n	-9.510	-9.509	0.001	
		n - 1	-9.847	-9.843	0.004	

Base-Pair Stack Calculations and Comparison with Clusters and Hydration

Band structure calculations on stacks of base dimers (consisting each of two different nucleotide bases) have been performed and compared with the results for the corresponding clusters describing the elementary cell. The dimers were guanine-cytosine (both possible sequences), guanine-adenine (both possible sequences), thymine-guanine, cytosine-thymine, adenine-thymine, and adenine-cytosine. Furthermore, the results on nucleotide base and base dimer stacks have been compared to the corresponding molecules and the base dimers. All the molecular (or cluster) results are listed in Table II, the stack results in Table III.

Similarly to the observations for the Watson-Crick double-strand polynucleotides^{8,9} the bands of the base dimer stacks are approximately obtainable by superposition of the bands of the constituting single base stacks. The assignment of the bands to the components is the same as can be obtained from the corresponding dimer calculations. Due to this superposition effect the band gaps have the same order of magnitude as for the single base stacks. On the other hand due to the alternating base sequence in the stack both the conduction and valence bands have in most cases about one order of magnitude smaller widths, than the corresponding single base stacks. This is in accordance with the results of previous π electron calculations on the same systems¹⁹. This means that, even if the generation of free charge carriers has a similar probability in the base dimer stack as for the simpler periodic DNA models, the mobility in the conduction bands becomes essentially smaller for these more complex periodic DNA models.

CONCLUSIONS

Different possible modifications of the simplest DNA models have been considered, and the effects on the electronic structure studied. Hydration shifts the bands, these shifts, however, are more or less parallel. Band widths and band gaps, therefore, do

not change drastically. This means that hydration should have only limited influence on the conduction properties of DNA. Copolymerization of two different nucleotides on the other hand, leads to essentially narrower valence and conduction band widths, while the band positions are approximately obtainable from the corresponding dimers and the band gaps are in the same order of magnitude as for the single base stacks. This means that though the generation of free charge carriers in these more complex DNA models is of similar probability as in the single base stacks, the charge carrier mobility is, however, strongly reduced.

Our thanks are due to Dr H. Statz and IBM Germany for the interest in this project and the financial support. We wish to express our gratitude to Dr P. Franchi, Dr P. Di Chio and Dr P. Sguazzero and the IECSEC Rome (IBM Italy) for their hospitality and financial support and putting in a generous way the computer facilities to our disposal. One of us (R. S. Chen) is very grateful to the Alexander-von-Humboldt Foundation for the award to a Fellowship that made his stay at the Chair for Theoretical Chemistry at the University Erlangen-Nürnberg possible. We also acknowledge the support of the "Fonds der Chemischen Industrie".

REFERENCES

1. Ladik J.: *Int. J. Quantum Chem.* 4, 307 (1971).
2. Ladik J.: *Adv. Quantum Chem.* 7, 377 (1973).
3. Ladik J.: *Int. J. Quantum Chem. QBS I*, 651 (1974).
4. Suhai S., Ladik J.: *Int. J. Quantum Chem.* 7, 547 (1973).
5. Tinoco L., jr: *J. Am. Chem. Soc.* 82, 4785 (1960); 83, 5047 (1969).
6. De Voe H., Tinoco L., jr: *J. Mol. Biol.* 4, 5181 (1962).
7. Rhodes W.: *J. Am. Chem. Soc.* 83, 3609 (1961).
8. Otto P., Clementi E., Ladik J.: *J. Chem. Phys.* 78, 4547 (1983).
9. Otto P.: *Int. J. Quantum Chem.* 30, 275 (1986).
10. Otto P., Ladik J., Corongiu G., Suhai S., Förner W.: *J. Chem. Phys.* 77, 5026 (1982).
11. Otto P.: *Int. J. Quantum Chem.* 28, 895 (1985).
12. Förner W., Otto P., Ladik J.: *Chem. Phys.* 86, 49 (1984).
13. Del Re G., Ladik J., Biczó G.: *Phys. Rev.* 155, 997 (1967).
14. André J.-M., Gouverneur L., Leroy G.: *Int. J. Quantum Chem.* 1, 427 (1967).
15. Chen R. S., Otto P., Ladik J.: *J. Chem. Phys.* 85, 5365 (1986).
16. Liegener C.-M., Otto P., Chen R. S., Ladik J.: *Theor. Chim. Acta*, in press.
17. Chen R. S., Liegener C.-M., Otto P., Ladik J.: *Biochim. Biophys. Hung.* 22, 205 (1987).
18. Gianolio L., Pavani R., Clementi E.: *Gazz. Chim. Ital.* 108, 181 (1978).
19. Rozsnyai B. F., Martino F., Ladik J.: *J. Chem. Phys.* 52, 5708 (1970).