Hole mobilities of periodic models of DNA double helices in the nucleosomes at different temperatures

Attila Bende^{a,c}, Ferenc Bogár^{b,c}, János Ladik^{c,*}

^aMolecular and Biomolecular Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Str. Donath 65-103, C.P. 700, Cluj Napoca RO-400293, Romania ^bSupramolecular and Nanostructured Materials Research Group of the Hungarian Academy of Sciences, University of Szeged, Dóm tér 8., 6720 Szeged, Hungary

^c Chair for Theoretical Chemistry and Laboratory of the National Foundation for Cancer Research. Friedrich-Alexander-University-Erlangen-Nüzernberg, Egerlandstr. 3, 91058 Erlangen, Germany

Abstract

Using the Hartree-Fock crystal orbital method band structures of $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ were calculated (\tilde{G} , etc. means a nucleotide) including water molecules and Na⁺ ions. Due to the close packing of DNA in the ribosomes the motion of the double helix and the water molecules around it are strongly restricted, therefore the band picture can be used. The mobilities were calculated from the highest filled bands. The hole mobilities increase with decreasing temperatures. They are of the same order of magnitude as those of $poly(\hat{A})$ and $\operatorname{poly}(\tilde{T})$. For $\operatorname{poly}(\tilde{G})$ the result is ~5 times larger than in the $\operatorname{poly}(\tilde{G} - \tilde{C})$ case.

1. Introduction

We have performed large scale calculations on DNA because according to many experiments carcinogens binding to DNA or radiation hits do not act only locally but by different mechanisms they can have also long range effects (See Ref. 1). These long range effects can be strong if there is a hole conductivity in a DNA double helix.

Recently, using the *ab initio* HF crystal orbital theory we have calculated the band structures of the poly $(\tilde{G} - \tilde{C})$ and poly $(\tilde{A} - \tilde{T})$ periodic double helices. Here, \tilde{G} stands for the nucleotide G - S - P (S is the deoxyribose and P the phosphate group) [2].

Since in aqueous solution because of the structural distortions of the DNA helix and especially of the fluctuation of the water molecules around the DNA double helix, the band model is not adequate. Therefore many authors have applied different forms of hopping theories [3, 4, 5].

^{*}Corresponding author:

Tel.: +49 9131 996600; fax: +49 9131 85 27736. E-mail: Janos.Ladik@chemie.uni-erlangen.de

Email addresses: bende@itim-cj.ro (Attila Bende), bogar@sol.cc.u-szeged.hu (Ferenc Bogár), Janos.Ladik@chemie.uni-erlangen.de (János Ladik)

On the other hand according to high resolution X-ray investigations of DNA in the nucleosome is very closely packed [6, 7]. This strongly diminishes the possibilities both of structural distortions of the DNA molecule and the fluctuations of the water molecules around it. In our study we have applied the *ab initio* Hartree-Fock (HF) crystal orbital method [8, 9, 10] taking into account also the 36° rotation of each base pair perpendicular to the main axis of the helix [12] to obtain the band structures for different periodic DNA models. In the calculations the LCAO approximation was used [11]. One should observe that the HOMO levels of the double stranded DNA lie by about 5 eV higher than those of the single chains [2] (in both cases with the sugar and phosphate groups and in the presence of water). This is caused by the fact that there is a charge transfer of about 0.2e from deoxyribose to the nucleotide base to which it is bound and these excess charges repel each other in the double stranded cases [2].

On the basis of the obtained band structure we have computed the mobilities of the holes (at different temperatures) belonging to the contraction-dilatation motion of the double helix (we did not take into account the mobilities belonging to the torsional motion of the helix, because in our earlier studies we have found that the effect of carcinogens or radiation hits on DNA happens not only locally, but they can have also long-range effects (See [1]). For this reason the contraction-dilatation movement of the double helix is the only important one. Therefore we have computed only the mobility along the main axes of DNA.

2. Methods

The band structure of $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ systems were obtained using the *ab initio* HF crystal orbital method in its LCAO form [8, 9, 10, 11] by applying the helical symmetry operation [12]. For the HF calculations Clementi's double ξ basis set was used [13]. The unit cell was built taking the G-C and A-T nucleotides from the experimental geometrical structure of double-stranded DNA B reported by Olson *et al* [14]. The negatively charged PO₄⁻ groups of the nucleotides were neutralized by two Na⁺ ions, which positions were optimized using the GAUSSIAN 03 program package [15] using the HF method together with the same Clementi's basis set.

To determine the water structure around the nucleotides, first, a triple stacks of $\tilde{G} - \tilde{C}$ and $\tilde{A} - \tilde{T}$ in the DNA B geometry were generated. Around these systems a huge number of water molecules were randomly placed using the PACKMOL program [16] and after that, their geometries were optimized using molecular mechanics [Amber force field from the GAUSSIAN 03 ([15])], while the nucleotide sequences were kept frozen. As next step the two outer (highest and lowest) base pairs as well as their surrounding water molecules were eliminated. Finally, we kept only 15 water molecules around the remaining nucleotides in both of the $\tilde{G} - \tilde{C}$ and $\tilde{A} - \tilde{T}$ cases. Accordingly, the unit cell is built by a single $\tilde{G} - \tilde{C}$ and $\tilde{A} - \tilde{T}$ nucleotide pairs $+ 2 \text{ Na}^+$ ions + 15 water molecules (For details see Ref.-s [17, 18]). The number of k-points in the half Brillouin zone was 12 in both cases and the number of contracted basis functions per unit cell was 766 in the poly($\tilde{G} - \tilde{C}$) case and 754 for poly($\tilde{A} - \tilde{T}$). To reach self-consistency (7-8 decimals in the total energy per unit cell) we needed about 20 iteration.

It is well-known that the HF method gives too large band gaps and wrong energy positions for the conduction bands. In our previous work [18] we have shown that the electron correlation can decrease the too large HF gaps with about 2 eV as well as it moves down the conduction bands. Since, we are interested in the hole conductivity which occurs at the upper region of the valence band, the electron correlations would not change significantly the HF picture. One should point out that the DFT method [19] in its first form gives a too small gap, which could be increased during the further development of this method.

To the calculation of the mobilities of $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ systems we have started from their *ab initio* band structures. The valence band widths are ~0.7 eV in the $poly(\tilde{G} - \tilde{C})$ case and ~0.3 eV in the $poly(\tilde{A} - \tilde{T})$ case (see Table Ia and Ib) which are substantially broader that the thermal energy at body temperature ($k_BT = 8.617 \cdot 10^{-5} \cdot$ $310 = 2.67 \cdot 10^{-2}$ eV). Therefore the deformation potential approximation for the mobility calculations can be applied. One should point out that to perform a simultaneous translation and rotation one can proceed in the same way as in the case of simple translation because in both cases their symmetry groups have the same multiplication table (isomorphous groups). Therefore for the helix operation we have to put the nuclei in the right positions and rotate the basis functions (unless they are *s* functions) with the relevant rotation angle [12]. For more details see [2].

For the calculation of the mobilities the deformation potential method was used [20]. In the case of 1D systems (as a single DNA double helix) the original expression for a 3D systems had to be modified. This has been done in a previous paper [21]. The result for 1D system in the case of hole mobilities is

$$\mu_h = \sqrt{\frac{2}{\pi}} \frac{c_\perp \hbar^2 e}{\varepsilon_{1h}^2 m_h^{*3/2} \left(k_B T\right)^{1/2}} \tag{1}$$

Here c_{\perp} is the elastic constant for the contraction and dilatation of the chain. ε_{1h} is the deformation potential at the G or A-type upper limits of the valence bands of the poly $(\tilde{G} - \tilde{C})$ and poly $(\tilde{A} - \tilde{T})$ systems, respectively. The deformation potential for holes is defined as:

$$\varepsilon_{1h} = \frac{\delta \varepsilon_{v,u,l.}}{\frac{\Delta l}{l_0}}; \quad \left(l_0 = 3.32 \text{ \AA} \quad \Delta l = \pm 0.2 \text{ \AA}\right)$$
(2)

while m_h^* is the effective mass calculated from the dispersion curve of the corresponding

valence band

$$\frac{1}{m*} = \frac{1}{\hbar^2} \cdot \left. \frac{d^2 \varepsilon}{dk^2} \right|_{k_{v,u,k}}.$$
(3)

Finally, the elastic constant c_{\perp} was calculated from the contraction-dilatation displacement $(\Delta l = \pm 0.02 \text{ Å})$ of the two stacks using

$$c_{\perp} = l_0 \left. \frac{\partial^2 E}{\partial l^2} \right|_{l_0} \qquad l_0 = 3.32 \,\text{\AA} \tag{4}$$

expression, where l is the length of the DNA double helix, E its total energy per unit cell and l_0 the stacking distance in equilibrium.

One should mention that we have taken as the equilibrium stacking distance 3.32 Å instead of 3.36 Å (the Watson-Crick value), due to the very tight packing of the DNA molecule in the nucleosomes [6, 7].

The condition of the applicability of the deformation potential approximation is that the width of the bands for which it is applied should be at least four times larger than the thermal energy $(2.67 \cdot 10^{-2} \text{ eV})$. On the other hand the valence band widths for $\text{poly}(\tilde{G} - \tilde{C}) + \text{H}_2\text{O} + \text{Na}^+$ are 0.69 eV and 0.21 eV for $\text{poly}(\tilde{A} - \tilde{T}) + \text{H}_2\text{O} + \text{Na}^+$ systems [2], respectively. This means that the deformation potential method can be safely used because the thermal energy is at least by one order of magnitude smaller than the widths of the valence bands for which we intend to calculate the mobilities.

3. Results

In the Table I we present the conduction and valence bands of the $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ systems in the presence of water and Na⁺ ions at three different stacking distances. See our previous paper [2].

Table I.

Though the known d.c. conductivity measurements on DNA provided d.c. hole conductivities, it may happen in the future that due to the action of electron donors also d.c. electronic conduction takes place through the base stacks of DNA. This is the reason that we have included in Table I also the description of the first empty band of $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$, respectively, which are dominated by base stacks.

In Table II we give the effective masses, the elastic constants and the deformation potentials which occur in Eq. (1).

Table II.

Finally, in Table III the mobilities are presented at different temperatures

Table III.

4. Discussion

For the HF band structures of $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ the same methods were used as in Ref. [2]. In our previous calculation [2] the gap is 8.95 eV for the first system and 9.30 eV in the second one, respectively, while in the present one they are 9.02 eV and 10.19 eV, respectively, (see Table I)). This is caused by the different stacking distances used: in Ref. [2] we have taken 3.36 Å for it, while in the present paper considering the close packing in DNA in the nucleosomes we have used for its equilibrium value 3.32 Å in the deformation potential calculation.

The small differences in the widths of the valence bands and the base-dominant conduction bands as well in their upper and lower limits are most probably due to the same reason (compare Tables I and II of Ref. [2] with Table I of the present paper). Further, if one compares the upper and lower limits of the valence band of the poly($\tilde{A} - \tilde{T}$) in Ref. [2] and the values obtained in the present calculation we observe a significant discrepancy between them. A subsequent analysis of these results showed us that there are errors in Ref. [2] caused by a steric hindrance due to a H₂O molecule which becomes too close to the deoxyribose binding to adenine during the helix operation (translation-rotation). Another source of this discrepancy (but with much less influence on the valence band shift) could be the number of the water molecules around the nucleotide base in the unit cell as well as the slightly different stacking distance between the bases. In the present calculation we have considered 15 water molecules and 3.32 Å as the stacking distance, while in Ref. [2] we used 18 water molecules and 3.36 Å for the stacking distance.

The calculation of the band structures of $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ at smaller (3.30 Å) and larger (3.34 Å) stacking distances were necessary for the calculation of the deformation potential. The upper and lower limits as well as the band widths of valence and base-dominant conduction bands for $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ systems are presented in Table Ia and Ib. Considering the three different stacking distances (3.30 Å, 3.32 Å and 3.34 Å) the band widths of both the valence and conduction bands either remain unchanged or become slightly smaller at increasing stacking distances. At the same time, the upper and lower limits of the valence bands and of the conduction bands in both systems slightly decrease with increasing stacking distances. In each case the bands show monotonic behavior in the first Brillouin zone (without any singularity) their edges fit with the limits of the first Brillouin zone (k=0 or k=12).

Performing a detailed analysis of the eigenvectors we found that the crystal orbitals belonging to the so-called base-type lowest unfilled bands (which we call conduction bands), contain only about 60% contributions from those AO-s which belong to C or T, respectively. The remaining 40% has sugar-phosphate backbone or Na⁺ character. This analysis shows also that in the case of conduction bands there are significant overlaps between the base- and non-base-type AO-s and these conduction bands are spread both on base- and non-base-type molecular fragments. Finally, it should be mentioned that between the valence band and the base-dominant conduction bands there are (as in the previous calculation [2]) in the case of poly($\tilde{G} - \tilde{C}$) 16 bands which have no contributions from the AO-s belonging to the C bases and 13 bands for poly($\tilde{A} - \tilde{T}$) where the AO-s do not belong to T.

Turning to Table II the effective masses for both $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ were calculated from the dispersion curves of the valence bands. For this we have used the polynomial interpolation technique taking a 7th order polynomial function (three k-points both at the left and the right sides of the edge of the Brillouin zone) and taking the expression of the effective mass given by eq. (3). Since hole conduction is the more probable in DNA, we have calculated the quantities necessary for the computation of the deformation potentials of the $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ systems, respectively, only for their valence bands which are purely G- or A-type. We obtain an effective hole mass of $m_h^* = -1.68m_e$ for the $poly(\tilde{G} - \tilde{C})$ and $m_h^* = -3.69m_e$ for the $poly(\tilde{A} - \tilde{T})$ system, respectively.

The deformation potentials were determined from the valence band upper limits, according to eq. (2). The elastic constants were calculated from eq. (4) assuming a parabolic dependence of the total electronic energy along the z axis (the long axis of the DNA double helix). In the poly($\tilde{G} - \tilde{C}$) case we have obtained for the deformation potential a value of 9.36 eV and for the elastic constant $7.38 \cdot 10^{-2}$ dyn. The corresponding values for the poly($\tilde{A} - \tilde{T}$) system are 5.97 eV and $2.82 \cdot 10^{-2}$ dyn, respectively.

Finally, in Table III we show the hole mobilities (in $\text{cm}^2/\text{V} \cdot \text{sec units}$) computed from eq. (1) at three different temperatures: 310 K (the human body temperature), 300 K (room temperature) and 273 (water freezing temperature) using the previously computed values for the effective masses, deformation potentials and elastic constants. The hole mobilities increase with decreasing temperatures as expected for both investigated systems. One should point out that the obtained values refer only to the contraction-dilatation motion of DNA helix.

The hole mobilities obtained in the present calculation are of the same order of magnitude as those obtained for the previous single stranded DNA calculations [22] for poly(\tilde{A}) and poly(\tilde{T}). Only in the case of poly(\tilde{G}) was obtained in the previous work [22] a result by a factor of 5 larger than the one found in the present work for poly($\tilde{G} - \tilde{C}$). The reason of this discrepancy can lie in the difference of the band structures of the single stranded and double stranded DNA helices. To clear up this problem we are going to perform further calculations. On the other hand, our mobility value obtained for the poly($\tilde{G} - \tilde{C}$) system is in a good agreement with that obtained by Grozema *et al* [23] using a tight-binding model Hamiltonian, where the motion of holes is coupled to structural fluctuations of the poly(G - C). In the "frozen" twist motion case they obtained for the mobility a value of $22 \text{ cm}^2/\text{V} \cdot \text{sec}$, which has the same order of magnitude as our mobility result based on the band model.

The first dc conductivity measurements on native DNA were performed back in 1962 by Eley and Spivey [24]. From the temperature dependence of the specific conductivity, they have found for the activation energy of conductivities 4.8 eV. This value lies not very far from our HF+MP2 gap value calculated by the extended basis set for a C stack (6.6 eV). In an subsequent paper Porath *et al* [25] have performed a two point dc measurement of a 10.4 nm (30 base pairs) long poly($\tilde{G} - \tilde{C}$) double helix. In their theoretical analysis, they conclude that at least in such a comparatively shorter distance (using a number of experimental techniques) the charge transfer is mediated by energy bands. In recent papers [26, 27] of the Barton group they have found charge transport in DNA of 34 nm (over 100 base pairs) lengths. They point out that such a long distance CT cannot be explained by hopping or superexchange mechanisms. They suppose a coherent CT mechanism if the double helix is intact (not a simple distortion or mismatch). Coherent mechanism means in other words the existence of energy bands.

Acknowledgments

We should like to express our gratitude to Professor Ferenc Beleznay for the very useful discussions about the deformation potential approximation method. The numerical work has been carried out using the facilities at the High Performance Computing Group of the University of Szeged and the DataCenter of INCDTIM Cluj-Napoca. We gratefully acknowledge their support.

References

- J. Ladik and W. Förner, The Beginnings of Cancer in the Cell: An Interdisciplinary Approach (Springer, Berlin, 1994), Chapter 5, p. 104. – This book is out of print, but Springer has notified me (J. Ladik), that they put on internet and print again some selected books which are out of print. This book was chosen also for this purpose.
- [2] A. Bende, F. Bogár, J. Ladik, Solid State Comm. 151 (2010) 301.
- [3] J.C. Genereux, J.K. Barton, Chem. Rev, 110 (2010) 1642.
- [4] M. Bixon, J. Jortner, J. Am. Chem. Soc., 123 (2001) 12556

- [5] A.V. Voityuk, J. Jortner, M. Bixon, N. Rösch, J. Chem. Phys. 114 (2001) 5614.
- [6] K. Luger, W. Mäder, R.K. Richmond, D.F. Sargent, T.J. Richmond, Nature 389 (1997) 251
- [7] T.J. Richmond, C.A. Davey, Nature 423 (2003) 145
- [8] P.-O. Löwdin, Adv. Phys. 5 (1956) 1
- [9] G. Del Re, J. Ladik, G. Biczó, Phys. Rev. 155 (1967) 997
- [10] J.-M. André, L. Gouverneur, G. Leroy, Int. J. Quantum Chem. 11 (1967) 427; 11 (1967) 451.
- [11] J. Ladik, Phys. Rep. 313 (1999) 171.
- [12] A. Blumen, C. Merkel, Phys. Status Solidi 383 (1977) 425.
- [13] L. Gianolo, E. Clementi, Gazz. Chem. Ital. 110 (1980) 79.
- [14] W.K. Olson, M. Bansal, S.K. Burley, R.E. Dickerson, M. Gerstein, S.C. Harvey, U. Heinemann, X-J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C.-S. Tung, E. Westhof, C. Wolberger, H.M. Berman, J. Mol. Biol. 313 (2001) 229.
- [15] M.J. Frisch et al., GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford CT, 2004.
- [16] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, J. Comput. Chem. 30(13) (2009) 2157.
- [17] J. Ladik, A. Bende, F. Bogár, J. Chem. Phys. 127 (2007) 055102.
- [18] J. Ladik, A. Bende, F. Bogár, J. Chem. Phys. 128 (2008)105101.
- [19] W. Kohn, L. J. Sham (1965) Physical Review 140(4A) (1965) A1133.
- [20] W. Shockley, Electrons and Holes in Semiconductors, D. Van Nostrand Press, New York, 1950.
- [21] F. Beleznay, F. Bogár, J. Ladik, J. Chem. Phys. 119 (2003) 5690.
- [22] A. Bende, F. Bogár, F. Beleznay, J. Ladik, Phys. Rev. E 78 (2008) 061923.
- [23] F. C. Grozema, L. D. A. Siebbeles, Y. A. Berlin and M. A. Ratner, ChemPhysChem, 6 (2002) 536.
- [24] D. D. Eley, D. I. Spivey, Trans. Faraday Soc. 58 (1962) 411.
- [25] D. Porath, A. Bezradin, and S. de Vries, Nature 403 (2000) 635.
- [26] J. D. Slinker, N. B. Muren, S. E. Renfrew, J. K. Barton, Nature Chem. 3 (2011) 230.
- [27] N. B. Muren, E. D. Olmon, J. K. Barton Phys. Chem. Chem. Phys. 14 (2012) 13754.

Stacking	3.30 Å	3.32 Å	3.34 Å	
distance				
Cond. Band (C-type) - 16 impurity bands				
u.l.	$6.81 \ (k{=}12)$	$6.79 \ (k{=}12)$	$6.77 \ (k{=}12)$	
l.l.	$6.51 \ (k{=}0)$	$6.49 \ (k{=}0)$	$6.47 \ (k{=}0)$	
w.	0.30	0.30	0.30	
Valence Bands (G-type)				
u.l.	-2.41 (k=12)	-2.53 (k=12)	-2.58 (k=12)	
l.l.	-3.16 (k=0)	$-3.27 \ (k{=}0)$	$-3.29 \ (k=0)$	
w.	0.75	0.74	0.71	
Gap	8.92	9.02	9.05	

Table Ia. The band structures of $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{G} - \tilde{C}) + Na^+ + 15$ water molecules at different stacking distances.

Table Ib. The band structures of $poly(\tilde{A} - \tilde{T})$ and $poly(\tilde{A} - \tilde{T}) + Na^+ + 15$ water molecules at different stacking distances.

Stacking	3.30 Å	3.32 Å	3.34 Å	
distance				
Cond. Band (T-type) - 13 impurity bands				
u.l.	$6.35 \ (k{=}12)$	$6.32 \ (k{=}12)$	$6.29~(k{=}12)$	
l.l.	$5.90 \ (k{=}0)$	$5.89 \ (k{=}0)$	$5.88 \ (k{=}0)$	
w.	0.45	0.43	0.41	
Valence Bands (A-type)				
u.l.	-4.27 (k=12)	-4.30 (k=12)	-4.34 (k=12)	
l.l.	$-4.57 \ (k{=}0)$	$-4.60 \ (k=0)$	$-4.63 \ (k=0)$	
w.	0.30	0.30	0.29	
Gap	10.17	10.19	10.22	

Table II. The deformation potentials for holes (in eV-s), the elastic constants (in dyns) and the effective masses in m_e -s for poly($\tilde{G} - \tilde{C}$) and poly($\tilde{A} - \tilde{T}$) in the presence of water and Na⁺ ions.

	$\mathrm{poly}(\tilde{\mathrm{G}}-\tilde{\mathrm{C}})$	$\mathrm{poly}(\tilde{A}-\tilde{T})$
ε_{1h}	9.36	5.97
c_{\perp}	$7.38 \cdot 10^{-2}$	$2.82 \cdot 10^{-2}$
m_h^*	-1.68	-3.69

	$\mathrm{poly}(\tilde{\mathrm{G}}-\tilde{\mathrm{C}})$	$\mathrm{poly}(\tilde{A}-\tilde{T})$
μ_{310}	37.52	10.87
μ_{300}	38.14	11.05
μ_{273}	39.98	11.58

Table III. The hole mobilities of $poly(\tilde{G} - \tilde{C})$ and $poly(\tilde{A} - \tilde{T})$ in the presence of water molecules and a Na⁺ ion at different temperatures (in cm²/V · sec units).