A Simple Model for the Band Structure and D.C. Conductivity of an Infinite C═O···H—N Chain Perpendicular to the Protein Backbone

ATTILA BENDE,[1,2] FERENC BOGÁR,[2,3] JÁNOS LADIK[2]

[1]*Molecular 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*

[2]*Chair for Theoretical Chemistry and Laboratory of the National Foundation for Cancer Research, Friedrich-Alexander-University-Erlangen-Nürnberg, Egerlandstr. 3, 91058 Erlangen, Germany*

[3]*Supramolecular and Nanostructured Materials Research Group of the Hungarian Academy of Sciences, University of Szeged, Dóm tér 8, 6720 Szeged, Hungary*

Received 6 February 2008; accepted 10 June 2008
Published online 10 September 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/qua.21846

ABSTRACT: The[1] Hartree–Fock crystal orbital (CO) method in its linear combination of atomic orbitals form was applied to determine the band structure of histone proteins taking 0.041*e* charge transfer per nucleotide base from the PO_4^- groups of poly(guanilic acid) to the arginine, and lysine side chains in histones (see text). Assuming that there are infinite COs, perpendicular to the main chain, formed by the amide groups of one segment of the protein chain bound together by H-bonds with the C═O groups of another segment of the chain, we have calculated the band structure. From this, we have determined the mobility using the deformation potential approximation. Multiplying this with the mobile electron concentration due to the charge transfer between the PO_4^- groups of DNA and the positive side chains in histones, we have obtained for the direct current (D.C.) electron conductivity $\sigma_{fib} = 1.07 \times 10^{-9}\ \Omega^{-1}$ cm for a single fiber and after division by the cross-section of 9.10×10^{-16} cm^2, $\sigma_{spec} = 1.18 \times 10^6\ \Omega^{-1}\ cm^{-1}$ for the specific conductivity. Int J Quantum Chem 109: 612–617, 2009

Key words: HF band structure of a C═O···H—N chain; charge transfer between DNA and histones; D.C. conductivity of the C═O···H—N crystal orbitals; electric currents in DNA and proteins; probable mechanism of the initiation of cancer

Correspondence to: J. Ladik; e-mail: janos.ladik@chemie.uni-erlangen.de

Introduction

In two articles [1, 2], the crystal structure of the nucleosome core particle (NCP)[1] consisting of nucleohistone proteins and the structure of DNA around it [2] were reported. They were determined with the help of X-ray diffraction at 2.8 and 1.9 Å resolution, respectively.

NCP consists of eight histone molecules, which are in the form of α-helices and loops [1]. Around the NCP, a 147 base pairs long DNA-B superhelix is wrapped (for further details, see [1, 2]).

The interactions between DNA and the histone molecules is first of all caused by the charge transfer (CT) between the negative PO_4^- groups or of the O atoms of thymine and guanine bases and the positively charged side chains (lysine, arginine) of the protein molecules. These interactions are significant only, if the negative sites of the DNA double helix face the positive side chains of the histones. It was found that in a nucleosome this happens 120 times [3], that is, at every 1.3th base pair.

The authors of articles [1, 2] assume that if due to external perturbations (binding of molecules, radiation) the relative conformation of DNA and proteins may change in such a way that the CT cannot take place. In this case, the DNA superhelix unwraps from the nucleohistones. This happens, because the hole conduction in DNA and the electronic one in the proteins due to the CT gets interrupted and therefore the interaction between them disappears or weakens. If this happens (most probably simultaneously in a number of neighboring nucleosomes), the genetic information in DNA becomes freely readable. In reality, in the case of chemical carcinogens this takes place simultaneously at many places of a chromatin (which is built up from the nucleosomes). This procedure can lead to the occurrence of a number of proteins in the wrong time and location (oncoproteins). The occurrence of oncoproteins through a number of biochemical mechanisms can lead then to the disturbation of the self-regulation of the cell [4]. This may be one of the main mechanisms of cancer initiation.

To understand the details of these processes, we have started large-scale quantum mechanical calculations on the CT, on the electronic structure, and on the conduction properties both of DNA and proteins.

In a recent article [5], we have calculated the CT between the PO_4^- group and between the lysine and arginine molecule, respectively, at the HF+MP2 level, using a triple ξ basis with polarization functions on the valence shells of the atoms. We have found 0.26*e* and 0.21*e*, respectively, in the absence of water. The same calculation was repeated in the presence of 3 H_2O molecules and 1 K^+ ion. In this case, the CT was 0.10*e* for the first case and 0.095*e* in the second one [6]. Taking the average of the CT in the latter two cases (0.10*e*), the facts that in the 147 base pairs long superhelix there are 120 H-bonds between it and the histone molecules, we have to multiply this number by 120/147 = 0.82 (0.82 × 0.10 = 0.082). Finally, we have to divide this number by 2, if one considers that we calculate the number of electrons going from the PO_4^- groups of DNA single helix to the histones. In this way, one obtains finally 0.041*e* CT per nucleotide base.

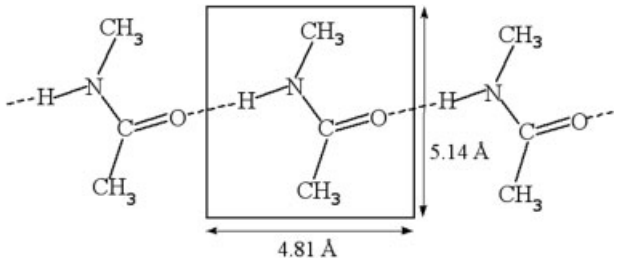


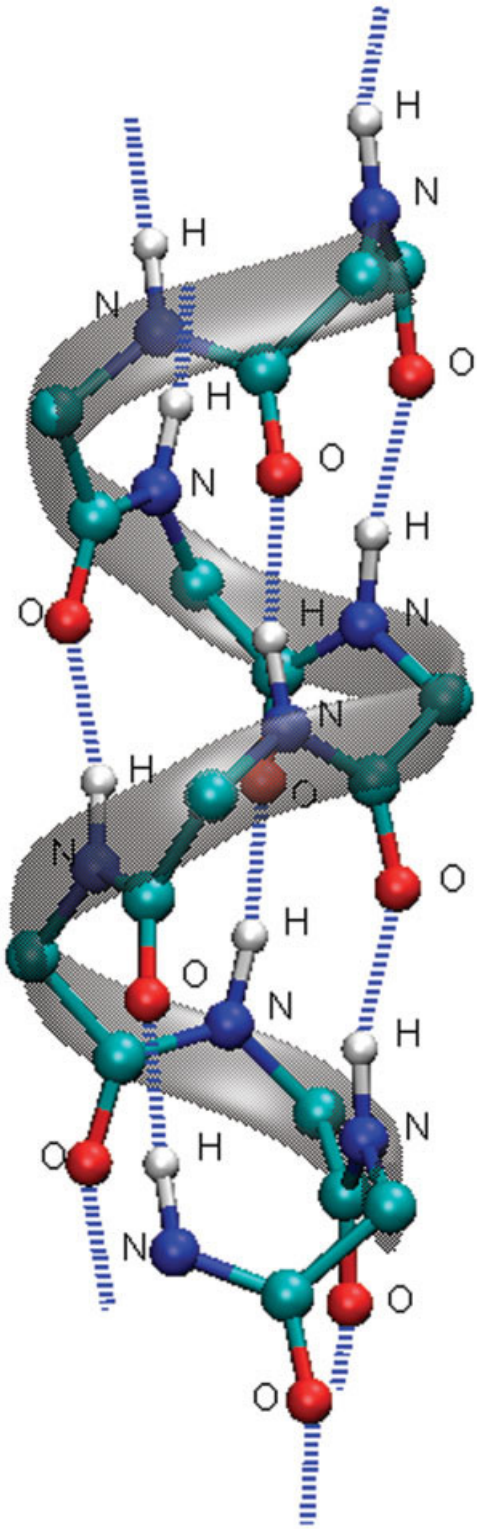
FIGURE 1. The infinite crystal orbitals through the H-bonds in the N—H···O═C units. These orbitals are perpendicular to the main protein chain in an α-helix. They are formed between the C═O group of the *n*th peptide unit and the HN group of the $n+4$th peptide unit in an α-helix (see text).

In a recent article, we have also calculated the ab initio HF band structures of the four homopolynucleotides in the presence of water and Na^+ ions [7, 8].

In this article, we report the band structure calculation of the C═O···H—N system (its elementary cell is shown in Fig. 1).

Here, we have depicted three unit cells of the infinite crystal orbitals (COs) that are perpendicular to the main chain of proteins. Such COs are formed between the C═O and H—N groups of neighboring turns of a protein main chain in an α-helix (Fig. 2) or between the units of an antiparallel β pleated sheet as it was assumed first on the basis of experimental biophysical evidence [9]. Afterward, the α-helix was discovered by Pauling through model building. In an α-helix, there are 3.6 residues per turn with a pitch of 5.4 Å [the helix built up from L-α-amino acid residues is right-handed (with torsion angels of $\Phi = -57°$ and $\Psi = +47°$)]. This arrangements leads to H-bonds between the C═O

group of the nth residue and the H—N group of the (n+4)th residue. These hydrogen bonds are strong (with an N···O distance of 2.8 Å) and have nearly the direction of the helix axis (see also Fig. 2) (for the description of the α-helix, see for instance [10]). The first tight binding band structure calculations for proteins were performed on the basis of this model [11–13]. It should be emphasized that by obtaining the D.C. conductivity through the H-bonds perpendicular to the protein main chain in an α-helix is only the first step. Most probably, the charge transport along the main chain (mostly by hopping) is more important to understand the movement of charges in a protein. To take into account both this and the electron transport in the ···O═C—N—H···O═C—N—H··· system, one has to treat proteins as 2D system in which in one-dimension it is highly aperiodic and the other one periodic.



Methods

For the band structure calculations, the standard ab initio Hartree–Fock (HF) crystal orbital method [13–15] has been used in the general case of combined symmetry operation [16]. (The symmetry group of a combined symmetry operation is isomorphic with the group belonging to simple translation. For more details, see [17]).

Having the band structure, the mobility of the breathing motion of the ···O═C—N—H···O═C—N—H··· system was calculated in complete analogy to the mobility calculation in the case of the breathing motion of a G-C base pair [18]. Multiplying this with the elementary charge and the concentration of the mobile electrons in the conduction band due to the CT from DNA, the D.C. conductivity in proteins has been computed (for details, see [18]).

For the ab initio HF CO method with not only a simple translation but with a combined symmetry operation for a periodic chain, the references were given in the Introduction. The basis set applied was Clementi's double ξ one [19]. The number of k-points in the half Brillouin zone was 12. The geometry was created with the help of Molden 4.6 program [20] and after that it was optimized with the help of the Gaussian 03 program [21].

FIGURE 2. Infinite periodic crystal orbitals on the mantle of a cylinder in the case of an α-helix (schematic). They are perpendicular to the protein main chain and nearly parallel to the main axis of the helix. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

For the determination of the D.C. conductivity the usual

$$\sigma_e = e\mu_e n_e \tag{1}$$

equation was used, where e is the elementary charge, μ_e the mobility of the electrons, and n_e their concentration.

The mobility μ_e was calculated from the expression

$$\mu_e = \sqrt{\frac{2}{\pi}} \frac{\tilde{G}\hbar^2 e}{\varepsilon_{c,l}^2 (m_e^*)^{3/2} (k_B T)^{\frac{1}{2}}} \tag{2}$$

derived before for quasi 1D systems [22]. Here $\varepsilon_{c,l}$ is the value of the deformation potential [23] (which approximates the electron-phonon interaction and is applicable if the band in question is broad enough; see below) at the lower edge of the conduction band, taking for the shifts of the unit cell length +0.1 Å and −0.1 Å, respectively,

$$\varepsilon_{c,l} = \frac{\delta\varepsilon_{c,l}}{\frac{\Delta l}{l}} = \frac{0.016}{\frac{0.1}{4.812}} = 0.75\, eV. \tag{3}$$

In the numerator, we have the average energy shift of the lower edge of the conduction band and in the denominator the length shift divided by the length of the unit cell 4.812 Å; see Fig. 1. For $\delta\varepsilon_{c,l}$ one obtains |0.016| eV for both Δl = +0.1 Å and Δl = −0.1 Å, respectively. We have taken so small values for Δl, because the deformation potential approximation works well only if the scattering of the electrons happens only near the band edge [23]. This is also the reason of the small (±0.1 Å) elementary cell length changes by the calculation of the total energy changes to obtain the value of the mechanical stress.

The effective mass at the lower edge of the conduction band was determined by fitting a fourth-order polynomial to the dispersion curve. From the $\varepsilon(k)$ function then we have obtained the effective mass in the usual way as 11.04 m_e. For T, we have substituted into Eq. (1) 300 K.

The elastic constant $\tilde{G}$ has been calculated in several steps:

1. Changing the length of the cell by 0.1 Å in both directions, we have computed the total energy belonging to these cell lengths and have taken the algebraic mean value of the total energy differences, ($|\Delta E_+|$ 7.58 × 10^{-3} eV, $|\Delta E_-|$ 9.38 × 10^{-3} eV. Taking the algebraic mean value of these two numbers and converting them to erg-s, one obtains ΔE 1.5 × 10^{-14} erg). Dividing this by the square of relative length changes $(0.1/4.812)^2$, one obtains for the mechanical stress $D = 3.13 \times 10^{-8}$ erg [24].
2. Dividing this by the volume of the unit cell $V = 4.812 \times 10^{-8} \times 5.14 \times 10^{-8} \times 1.77 \times 10^{-8}$ cm (length × breath × thickness) = 4.38×10^{-23} cm^3 (see Fig. 1), one obtains the theoretical elastic constant $G = D/V = 3.13 \times 10^{-8}/4.38 \times 10^{-23} = 7.15 \times 10^{14}$ dyn cm^{-2}. (We had to divide D by the volume of a unit cell to obtain the number of unit cells per cm^3 if we wish to calculate the specific conductivity).
3. Approximating the interactions of the electrons with the movement of the lattice in a fiber-like 1D system, we have finally multiply G by the cross section $q = 5.14 \times 10^{-8} \times 1.77 \times 10^{-8} = 9.1 \times 10^{-16}$ cm^2, $\tilde{G} = Gq$ = 0.65 dyn. This value has to be substituted into Eq. (2).

To obtain the concentration of mobile electrons one has to calculate it on the basis of equation [22]

$$n_e = \frac{L}{2\pi} \exp\left(-\frac{(\varepsilon_F - \varepsilon_{c,l})}{k_B T} \right) \frac{\sqrt{\pi m^* k_B T}}{\hbar\sqrt{2}}. \tag{4}$$

Here the Fermi energy ε_F is taken as its distance from the band edge. To determine it, we have to find out first the k-value (k_F) belonging to the q = 0.041e charge if one goes from the conduction band lower limit upwards. Taking into account that in the half Brillouin-zone one elementary charge belongs to a level in the band (because of the two spins), one finds the k-value which corresponds to the 0.041 charge. This is done using the expression of $k_F = (0.041/2)^2 \times \pi/a$ if one assumes a quadratic dependence of k on q. Afterwards one looks at the dispersion curve of the band to find out which energy belongs to this k-value (see Fig. 3).

The energy determined in this way gives the value of the Fermi level, ε_F = 3.674 eV. The corresponding Boltzmann factor will be $\exp(-(\varepsilon_F - \varepsilon_{c,l})/k_B T) = \exp(-(3.674 - 3.673)/0.025) = \exp(-0.04) = 0.96$.

Results and Discussion

From Table I, one can see that the valence (highest filled) and conduction (lowest unfilled) band

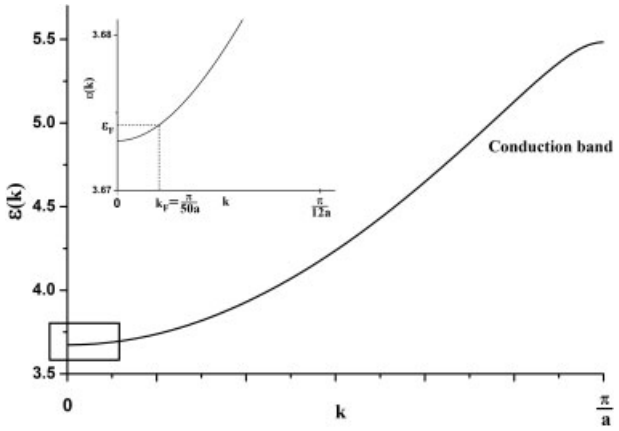


FIGURE 3. The dispersion curve of the conduction band belonging to the infinite ···O═C—N—H···O═C—N—H··· crystal orbitals in an α-helix.

widths (1.32 and 1.81 eV, respectively) are broad enough (one order of magnitude larger than the thermal energy $k_B \times 300$ K = 0.025 eV) for the applicability of the deformation potential approximation. The fundamental gap (13.84 eV) is by about 4.0 eV larger than in poly(guanilic acid) in the absence of water (9.99 eV) [7]. This HF gap is—as usual—far too large. In a previous calculation [25], we have found for a cytosine (C) stack a gap with a double ξ + polarization functions on the non-H atoms on the HF+MP2 level 8.17 eV. On the other hand, the difference of the experimental ionization potential and electron affinity of a single C molecule is only 7.0 eV [26]. This discrepancy is due first of all to the application of a medium level basis (with a better basis tailored for a stacked system we have found for the gap of C stack 6.60 eV [25]).

Because the HF approximation gives usually also too broad bands, one can expect that the deformation potential of the conduction band are in reality smaller. Because it occurs on the second power in the denominator of the mobility expression [see Eq. (2)] this means that the mobility and conductivity values given in Eqs. (5) and (6) are lower bounds of the real values.

From Table I, one can also see that the lower limit of the conduction band (3.67 eV) has the same value as in the case of poly(guanilic acid) in the presence of water (both at $k = 0$)[7] but the upper limit of the valence band (−10.2 eV) lies essentially deeper than at poly(guanilic acid) (−6.81 eV) [7]. Therefore, the fundamental gap is also by more than 3.0 eV larger, than in the former case.

Substituting into Eq. (2) $\varepsilon_{c,1} = 0.75$ eV, $\tilde{G} = 0.65$ dyn, and $m^* = 11.04m_e$ as well the values of the natural constants one obtains for the mobility value $\mu_e = 3.13 \times 10^3$ cm^2 V^{-1}s^{-1}. This we can compare with the mobility value of the electrons in a G-C base pair in the case of their interaction with the acoustic phonons of the breathing motion of the base pair [18], 2.55×10^2 cm^2 V^{-1}s^{-1}. The reason of this difference of one order of magnitude is that in the base pair case there are three H-bonds per unit cell while in our case only one.

Putting into Eq. (4) the value of the Boltzmann factor at $T = 300$ K, for $L = 1$ (we want to calculate specific conductivity) for m^* again $11.04m_e$ and the values of the natural constants one obtains for $n_e = 2{,}14 \times 10^6$ cm^{-1}. Substituting finally the elementary charge in Coulomb-s into Eq. (1) together with the found $\mu_{c,1}$ and $n_{c,1}$ values, we obtain for the D.C. conductivity of a single fiber:

$$\sigma_{\text{fib}} = \frac{4.8 \times 10^{-10}}{3 \times 10^9} \times 3.13 \times 10^3 \times 2.14 \times 10^6 = 1.07 \times 10^{-9} \Omega^{-1}\text{cm} \quad (5)$$

Dividing this by the cross section (the number of fiber surfaces in the unit cell perpendicular to the direction of the H-bonds) $q = 5.14 \cdot 1.77 \times 10^{-16} = 9.1 \times 10^{-16}$ cm^2, we obtain for the specific conductivity of

$$\sigma_{\text{spec}} = \frac{1.07 \times 10^{-9}}{9.1 \times 10^{-16}} = 1.18 \times 10^6 \Omega^{-1}\text{cm}^{-1} \quad (6)$$

This D.C. conductivity value seems to be somewhat large. The large mobility value [3.5 orders of magnitude larger than in the case of poly(guanilic acid)] and the same order of magnitude (10^6) for the free charge carrier concentration as in the polynucleotide explains the two orders of magnitude larger D.C. conductivity in our case for one filament than

TABLE I

The valence and conduction band edges with the gap of the ···O═C—N—H···O═C—N—H··· crystal orbitals perpendicular to the main chains of proteins (see also Figs. 1 and 2) in eV.

	Valence band	Conduction band[a]
Upper limit	−10.17 ($k = 0$)	5.48 ($k = \pi/a$)
Lower limit	−10.49 ($k = \pi/a$)	3.67 ($k = 0$)
Width	0.32	1.81

[a]Gap: 13.87 eV.

the σ-value of poly(guanilic acid), $\sigma = 2.51 \times 10^{-11}$ Ω^{-1} cm for one filament and 1.60×10^{3} Ω^{-1} cm^{-1} for 1 cm^3 poly(guanilic acid) [27].

Conclusions

The charge transport model with infinite COs perpendicular to the protein main chain seems to work in the case of mobile electrons obtained by CT from the PO_4^- groups to the positively charged side chains of histones.

One should emphasize that the results obtained are the outcome of a model calculation. In reality, proteins are highly disordered chains (simultaneous substitutional and spatial disorder). Therefore, for the calculation of the conductivity along the aperiodic main chain of proteins the application of a sophisticated hopping theory has to be worked out (which is still missing). One should mention, however, that for the frequency-dependent A.C. conductivity of different native proteins a variable range hopping theory has been worked out which uses a rather complicated quantum theoretical—statistical mechanical formalism [28]. If one extrapolates the $\sigma(\nu)$—ν curves (where ν is the frequency) to $\nu \rightarrow 0$, one obtains for the conductivity of a single chain the value of the order of 10^{-11} Ω^{-1} cm [28, 29].

A really serious protein conductivity calculation has to be two-dimensional and has to take into account both the hopping conductivity along the main chain and the Bloch-type conduction through the COs perpendicular to the main chain (considering in the latter also case the perturbative effects of the different side chains). A step-wise development of such a theory is one of our aims in our theoretical investigations of biopolymers.

ACKNOWLEDGMENT

We express our gratitude to Professor F. Beleznay for the very helpful discussions. We are further indebted to the High Performance Computing Center of the University of Szeged (Hungary) for providing free computer time.

References

1. Luger, K.; Mäder, W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J. Nature 1997, 389, 251.
2. Richmond, T. J.; Davey, C. A. Nature 2003, 423, 145.
3. Schalih, T.; Duda, S.; Sargent, F.; Richmond, T. J. Nature 2005, 436, 138.
4. Ladik, J.; Förner, W. The Beginnings of Cancer in the Cell; An Interdisciplinary Approach; Springer: Berlin, 1994; p 88.
5. Bende, A.; Bogár, F.; Ladik, J. Chem Phys Lett 2007, 437, 117.
6. Bende, A.; Bogár, F.; Ladik, J. Chem Phys Lett (in press).
7. Ladik, J.; Bende, A.; Bogár, F. J Chem Phys 2007, 127, 055102.
8. Ladik, J.; Bende, A.; Bogár, F. J Chem Phys 2008, 128, 105101.
9. Voet, D.; Voet, J. Biochemistry; Wiley: New York, 1995; p 146.
10. Evans, M.; Gergely, J. Biochim Biophys Acta 1949, 3, 188.
11. Ladik, J. Acta Phys Hung 1963, 15, 287.
12. Ladik, J. Nature 1964, 202, 1208.
13. Löwdin, P.-O. Adv Phys 1956, 5, 1.
14. Del Re, G.; Ladik, J.; Biczó, G. Phys Rev 1967, 155, 997.
15. André, J. M.; Gouverneur, L.; Leroy, G. Int J Quantum Chem 1967, 1, 427; 451.
16. Blumen, A.; Merkel, C. Phys Stat Solid B 1977, 83, 425.
17. Ladik, J. Phys Rep 1999, 313, 171.
18. Beleznay, F.; Szekeres, Zs.; Bogar, F.; Ladik, J. Chem Phys Lett 2006, 424, 399.
19. Gianolo, L.; Clementi, E. Gazz Chem Ital 1980, 110, 179.
20. Schaftenaar, G.; Noordik, J. H. J Comput-Aided Mol Des 2000, 14, 123.
21. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision B. 02. Gaussian, Inc.: Wallingford, CT, 2004.
22. Beleznay, F.; Bogár, F.; Ladik, J. J Chem Phys 2003, 119, 5690.
23. Schockley, W. Electrons and Holes in Semiconductors; Van Nostrand Press: New York, 1950.
24. El Haj Hassan, F.; Akbarzadeh, H. Compos Mater Sci 2006, 38, 362.
25. Bogár, F.; Ladik, J. J Chem Phys 1998, 237, 273.
26. Wiley, J. R.; Robinson, J. H.; Ehdale, S.; Chen, E. C. M.; Chen, E. S. D.; Wenworth, W. E. F. Biochem Biophys Res Commun 1991, 180, 841 (and reference therein).
27. Bende, A.; Bogár, F.; Beleznay, F.; Ladik, J. Phys Rev B (in press).
28. Ye, Y.-J.; Ladik, J. Phys Rev B 1993, 48, 5120.
29. Ye, Y.-J.; Ladik, J. Phys Rev B 1995, 51, 13091.