Hopping Conductivity in Nucleotide Base Stacks*

J. LADIK,[1] Y.-J. YE[2,†]

[1]*Chair for Theoretical Chemistry and Laboratory of the National Foundation for Cancer Research, Friedrich-Alexander University Erlangen-Nürnberg, Egerlandstr. 3, D-91058 Erlangen, Germany*
[2]*Department of Protein Engineering, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China*

Received 19 June 2001; revised 3 July 2001; accepted 23 July 2001

DOI 10.1002/qua.1117

ABSTRACT: The frequency-dependent complex hopping conductivities of single nucleotide base stacks and base pair stacks were calculated on the following basis: (1) The determination of the density of states of these disordered systems with the help of the matrix block negative counting method. (2) Using the inverse iteration technique, the Anderson localization of the one-electron functions belonging to the physically interesting levels were computed. (3) With the help of the latter quantities the hopping frequencies corresponding to the relative motion of the base pairs with respect to each other (acoustic phonons) were determined. (4) Finally, using a random-walk theory the frequency-dependent complex conductivities $[\sigma(\omega)]$ were computed. The $|\sigma(\omega)| - \omega$ curves show a saturation at $\omega \approx 10^{10}$ s^{-1}. The $|\sigma(\omega)|$'s have the values in the base pairs case (taking into account the role of the sugar–phosphate backbone, basis set, and correlation effects) are a few times 1 Ω^{-1} cm^{-1} for $\omega > 10^{10}$ s^{-1}. This room temperature result is in excellent agreement with the energy loss (in a resonant cavity) experiment on lambda phage deoxyribonucleic acid (DNA) of Gruner and co-workers. © 2002 Wiley Periodicals, Inc. Int J Quantum Chem 90: 838–847, 2002

Key words: conduction in DNA-frequency; dependent conductivity in base (pairs) stacks; conductivity in aperiodic base (pairs) stacks

*In memoriam of Professor Per-Olov Löwdin.

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

†Present address: Baker Laboratory of Chemistry and Chemical Biology, Chemistry Department, Cornell University, Ithaca, NY 14853-1301 USA.

© 2002 Wiley Periodicals, Inc.

Introduction

In previous works [1, 2] the density of states (DOS) and hopping conductivities, respectively, of single- and double-stranded nucleotide base (pair) stacks were calculated. The published hopping conductivity results refer to the valence-band regions (hopping of holes [2]) using two different aperiodic sequences. In the first case a 100- or 200-base-long stack of the C end of a human oncogene [3] has been calculated (seq.1), while in the second case a random sequence of 100 bases has been constructed in the proportion A:C:G:T = 1:1:1:1 avoiding repeated bases (seq.2).

While the DOS calculations were extended also to the conduction bands regions [1], here we present only hopping conductivity curves for this part of the level spectrum.

Several significant developments have occurred in the literature since the appearance of previous studies [1, 2]. Tran and co-workers [4] have published energy loss experiments at microwave frequencies as a function of temperature of liophylisized lambda phage DNA in a buffer (for a more detailed discussion of these experiments see below). On the theoretical side Ye and Jiang [5] have published hopping conductivity calculations of a 16-nucleotide-long segment of an operon [6]. These latter calculations have taken into account explicitly besides the base stacks also the sugar and phosphate parts of the nucleotides.

Hole conduction in base stacks cannot have a long range because the guanine molecules act as traps for the holes; to understand the conduction in deoxyribonucleic acid (DNA) one has to concentrate on the electronic charge transport in the conduction band region. Since according to previous [7] and most recent [5] calculations there is $\sim 0.2e$ excess negative charge on the nucleotide bases, these charges can serve as charge carriers in the conduction band regions.

Further we intend to compare the results on hopping conductivities of the full polynucleotide calculations with the previous base (pair) stack calculations and discuss the possible reasons of the still larger experimental hopping conductivity values as compared to our theoretical results.

Methods

In the calculations the geometry of the base stacks was taken from a data set determined by X-ray diffraction on a single crystal of DNA [8]. The rotation angle between the nearest neighboring bases (base pairs) is as usual 36° and the stacking distance is 3.36 Å. In this way we have constructed the geometry of the whole stack in which the main features of a DNA molecule could be maintained, though the sugar–phosphate backbones were not taken into account explicitly. In this connection it should be mentioned that according to earlier band structure calculations of periodic homopolynucleotides (sugar–phosphate and always the same base repeated), the resulting bands can always be classified as sugar–phosphate bands and base stack bands [9, 10]. The valence (highest filled) band and the conduction (lowest unfilled) band of these systems originate in all the three calculated cases [with cytosine (C), thymine (T), and adenine (A) as nucleotide base] from the highest filled or lowest unfilled levels of the single bases, respectively. The highest filled or lowest unfilled sugar–phosphate bands were always below or above the valence and conduction bands, respectively. Inspecting in more detail these homopolynucleotide band structures one could notice that they are in all three cases in a very good approximation a superposition of the band structures of the sugar–phosphate chain and of the periodic base stacks. This fact is due to the effect of the mutual screening of the charges on the subunits of a nucleotide ($\sim -1.2e$ on the phosphate group [11], $\approx +0.4e$ on the sugar units, and $\approx -0.2e$ on the bases [7, 9, 10] if one takes into account also the $+1.0$ charge on the counterions). These alternating charges on the sugar and phosphate units *do not* change the charge distributions significantly on the base pairs (otherwise the band structures of the homopolynucleotides could not be practically identical with those of the superposition of the sugar–phosphate chain and of the base stacks).

An early Hartree–Fock band structure calculation [12] on a periodic C stack has shown that taking into account the effect of the water molecules (in the form of five water clusters in positions determined

by a Monte Carlo calculation [13] in each plane of the C molecules), the water environment has hardly influenced the band structure of the stack.

On the basis of the above described results one can conclude that the presence of the sugar–phosphate chains in DNA and the water environment together with the counterions influences only in a very small amount the band structure of the periodic base (pairs) stacks and therefore most probably also the level distributions of an aperiodic stack. In other words the electronic structure of the free base or base pair stacks (both periodic or aperiodic) provides a very good approximation of the electronic structures of the same systems in DNA (especially if one takes into account in the aperiodic stack at least 50, but possibly 100, units). This does not mean, however, that the presence of the sugar–phosphate groups do not influence the hopping conductivity (see below).

In the calculations the overlapping dimer approximation was used as was done in the calculations of proteins [14] (the molecular systems were partitioned into overlapping dimers). The 16 dimers were calculated (in a single DNA strand one has 16 dimers along the strand, because in a $\overset{X}{Y}\uparrow$ dimer X = C, T, G, or A and Y is also C, T, G, or A) treating the dimers as supermolecules. The Fock (overlap) matrix of the whole system was constructed by putting all dimers in the same local coordinate system defined by the first dimer. To build up the helical structure the nuclei have to be rotated and translated in the direction of the z axis (assuming that the helical axis coincides with the z axis) by appropriate multiples of 36° and 3.36 Å, respectively. In addition the basis functions have to be rotated, too. That is, all matrix elements in which p_x and/or p_y functions occur have to be transformed. It should be mentioned that the overlapping dimer approximation gives practically the same total DOS for the Fock matrix constructed in this way than if one performs the negative factor counting (NFC) calculation on a Fock matrix constructed directly as it was checked on systems with smaller unit cells [15].

The generalized eigenvalue equations of the Fox matrix of the whole chain constructed in the above-described way can be solved by the NFC method [16–18]. The program ENFC [19] (extended matrix block NFC method), which can take into account also cross links, was used to obtain the DOS and the energy levels and orbitals of the whole system. In the calculations 5377 and 5275 basis functions (using a Clementi's minimal basis [20]), respectively, were applied for the two different sequences in the single-stranded and 10,539 basis functions in the case of the double-stranded chain with 100 units. The number of basis functions was 10,693 for a single-stranded and 21,088 for a double-stranded chain in the case of 200 units, respectively.

To investigate the effect of a better basis set we have performed also a calculation with Clementi's double ζ basis [21] for the single-strand Sseq.1 of 100 units and for a double helix with the same basis, number of base pairs and sequence.

The Anderson localization of the different orbitals was investigated with the help of the inverse iteration method [22]. We have found the orbitals to be localized on one or two bases. These localized wave functions were substituted then into Eq. (1).

The calculation of the primary hopping frequencies (which are the basic quantities in the random-walk theory for the calculation of the hopping conductivity) was performed using the generalization of the expression (given in the book of Mott and Davis [23]) for many orbitals per site (see, e.g., [24]).

$$h_{n,i\to n',j} = \nu_{\text{phonon}}\left(\sum_{r\in n}\sum_{s\in n'} C_{i,r}C_{j,s}\langle\chi_r^n|\chi_s^{n'}\rangle\right)^2 = \begin{cases}1, & \text{if } \varepsilon_j \le \varepsilon_i,\\ \exp[-\Delta E_{i,j}/k_BT], & \text{if } \varepsilon_j > \varepsilon_i.\end{cases} \quad (1)$$

Here ν_{phonon} is the acoustic phonon frequency for which we have taken the typical value of 10^{12} s^{-1} [23]; $\Delta E_{i,j} = \varepsilon_i - \varepsilon_j$ is the energy difference of the one-electron levels $\varepsilon_{n',j}$ (the jth level of site n') and $\varepsilon_{n,i}$ (the ith level of site n), k_B is the Boltzmann constant and T the absolute temperature. The level differences $\Delta E_{i,j}$ were determined by studying the grid points of the NFC procedure in the different regions of the DOS curves.

One should point out that Eq. (1) is a rather crude approximation for the primary jump rates. It would be more correct to express them with the help of electron–phonon matrix elements as

$$h_{n,i\to n',j} = \langle\Psi_i^{(n)\text{el}}|\hat{H}_{\text{el-ph}}|\phi_{\text{ph}}^{\text{ac}(n)}\rangle \times\left(\sum_{r\in n}\sum_{s\in n'} C_{i,r}^{(n)}C_{j,s}^{(n')}\langle\chi_r^{(n)}|\chi_s^{(n')}\rangle\right)^2,$$
$$\langle\phi_{\text{ph}}^{\text{ac}(n')}|\hat{H}_{\text{el-ph}}^{(n')}|\Psi_j^{(n')\text{el}}\rangle = \begin{cases}1, & \text{if } \varepsilon_j \le \varepsilon_i,\\ \exp[-\Delta E_{i,j}/k_BT], & \text{if } \varepsilon_j > \varepsilon_i.\end{cases} \quad (2)$$

Here $\hat{H}_{\text{el-ph}}$ is the operator of the electron–acoustic phonon interaction (in the case of the base stacks the acoustic phonons describe the vibrations of the stacked molecules with respect to each other), $\Psi^{\text{el}}_{n,i}$ and $\Psi^{\text{el}}_{n',j}$, respectively, describe the one-electron wave functions at site n belonging to the energy level i and at site n' of the energy level j, respectively.

A separate detailed calculation is needed to determine the vibrational wave functions $\varphi^{\text{ac}(n)}_{\text{ph}}$ of the normal modes belonging to the acoustic vibrations in a base stack and the matrix elements $\langle \Psi^{\text{el}}_{n,i} | \hat{H}_{\text{el-ph}} \varphi^{\text{ac}}_{\text{ph},t} \rangle$, $\langle \varphi^{\text{ac}}_{\text{ph},t} \hat{H}_{\text{el-ph}} | \Psi^{\text{el}}_{n',j} \rangle$. We plan to do this in the future, but it could not have been done within the framework of this study.

Further it should be mentioned that besides the dominant acoustic phonons (the movement of the whole base pairs with respect to each other) also optical phonons (vibrations in the planes of the base pairs or out of plane vibrations of single groups) certainly have some contributions to the primary jump rates (hopping frequencies). Though these vibrations have larger frequencies (due to the smaller masses of the vibrating units) and their number can be also quite large, through the much smaller probabilities of their contributions, they most probably will not have an order of magnitude effect. In the future a detailed analysis should be performed both for the effects of using Eq. (2) instead of Eq. (1) and for those of the optical phonons. We can say, however, already now that due especially to the second effect, Eq. (1) gives only lower bounds of the primary jump rates.

Using Einstein's relation we can write

$$\sigma(\omega) = \frac{n_v e^2}{k_B T} D(\omega). \tag{3}$$

Here n_v is the number density of free charge carriers (e.g., in the case of a single-stranded chain with 100 bases $n_v = 100/32121$ Å^3 [8]). The number in the denominator is the volume of the single chain with 100 bases in the case of the oncogene sequence; further $D(\omega)$ is the diffusion constant of the electrons. For the temperature $T = 300$ K was taken.

A random-walk theory [25–29] was applied for the hopping conductivities (for the detailed description of this formalism, which was again generalized for the case of many orbitals per unit cells, see references [24] and [30]).

It should be pointed out that there are more refined methods for the calculation of hopping conductivities as given by Lax and co-workers [25–28] (see, e.g., [29]). These theories are rather complicated and are usually valid only in such special cases that do not occur at ab initio calculations of complicated biopolymers (see also [31]).

Results

Since the density of states of different aperiodic nucleotide base stacks was published together with a detailed discussion, we refer here to this study [1].

In the calculations of the *hopping conductivities* of the different systems, we have taken into account also the three-dimensional structures of the different chains [8]. As an approximation only those hoppings were included, which happen between first- and second-neighbor units. The volume of the double-stranded (base pair) segment (with 100 units) was calculated to be 64,242 Å^3. (The same volume for the single-stranded case is as mentioned before 32,121 Å^3.) Assuming that all the levels originating from the highest occupied molecular orbitals (HOMOs) and lowest unoccupied molecular orbitals (LUMOs), respectively, of the bases may take part in the hopping conductivity (of course not simultaneously), the 100 highest filled and 100 lowest unoccupied orbitals were taken into account in the calculations of the conductivities of the two single-stranded stacks. The number of charge carriers was assumed to be also 100. This means that we have calculated both electronic and hole conduction. The hole conduction is most probably due to the charge transfer between the positive arginine groups and the negative PO_4 groups if a nucleohistone binds to DNA. Since in this charge transfer obviously only one electron per arginine $-PO_4^-$ participates, the number of holes will be 100 and not 200 (though on the first 100 HOMO levels 200 electrons are present with opposite spins). In the same way one can see that if one wishes to take into account 200 levels originating from the HOMOs of the single bases of 100 base pairs, one should perform the calculation for 200 holes.

To see the effect of n doping on DNA we have extended the calculations also to electronic hopping conduction using the DOS of the unfilled bands region. In the double-stranded case, we assume also that the first 200 levels take part in the conductivity. The number of charge carriers is now also 200. For comparison with the results of the single stranded systems, we have also calculated the hopping conductivity of the double-stranded system with 100 highest filled orbitals. The results are presented in

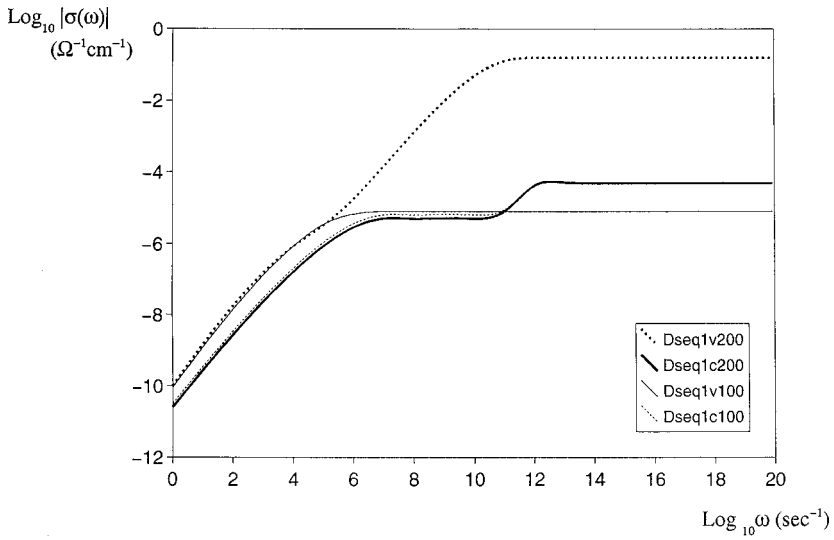


FIGURE 1. Logarithm of the absolute values of the complex hopping conductivities ($|\sigma(\omega)|$) as a function of the frequency. Double strand in the valence-band region taking into account 200 levels (Dseq1v200) (first curve), double strand in the conduction-band region with 200 levels (Dseq1c200) (second curve), (Dseqv100) (third curve), and (Dseq1c100) (fourth curve).

Figures 1 and 2 (in Fig. 1 $|\sigma(\omega)|$ and in Fig. 2 $\mathrm{Im}[\sigma(\omega)]$).

In Figure 3 results are shown for the valence-band region of a single-stranded stack with 200 units and 200 and 100 levels, respectively, using the sequence of the human oncogene.

Finally in Figure 4 $\log|\sigma(\omega)|$ as function of ω is given for a double-stranded stack with sequence 1 for 300 levels in the valence-band region (Dseq1v300) using Clementi's double-ζ basis [21] (solid line). The broken line refers to 200 levels in the same system (Dseq1v200).

Discussion

From the hopping conductivity versus frequency ($|\sigma(\omega)| - \omega$) curves one sees that the base pair stack

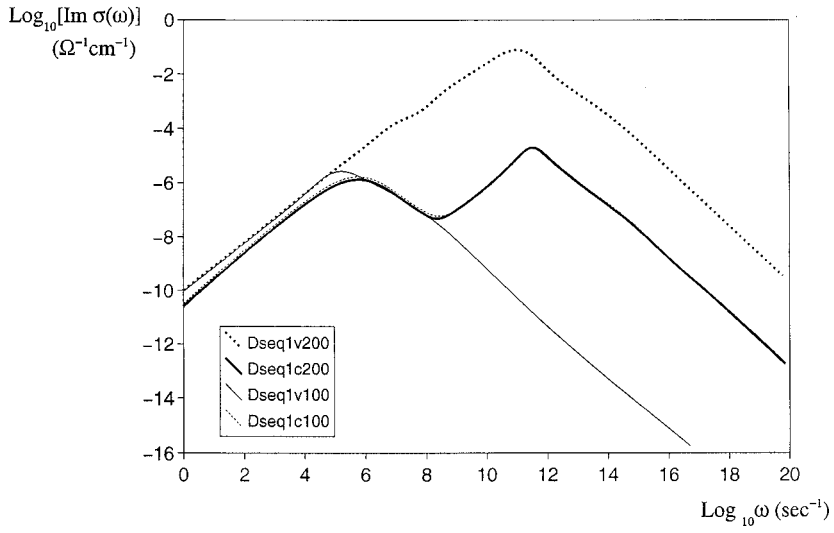


FIGURE 2. Log[Im $\sigma(\omega)$] in the cases given in Figure 1.

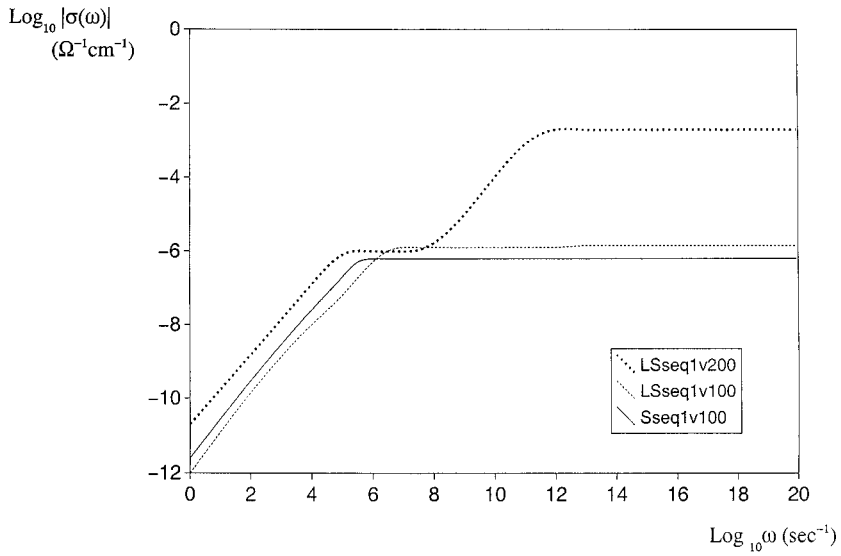


FIGURE 3. Log$|\sigma(\omega)|$ for the valence-band region of single-stranded DANN with sequence 1 of 200 units taking into account 200 levels (LSseq1v200) (first curve) and 100 levels (LSseq1v100) (second curve). For comparison the results for a single chain with 100 units and 100 levels in the valence-band region (Sseq1v100) are also shown (third curve).

with 200 levels taken into account in the valence-band region has the largest value ($|\sigma(\omega)| \approx 5 \times 10^{-1}\ \Omega^{-1}\ \mathrm{cm}^{-1}$) at the frequency of $\omega = 5 \times 10^{10}\ \mathrm{s}^{-1}$ (using sequence 1). In the conduction-band region the saturation value is at a somewhat larger frequency of only $\approx 5 \times 10^{-5}\ \Omega^{-1}\ \mathrm{cm}^{-1}$.

One can ask the question what is the biological relevance of hopping conductivity at the high-frequency range, $\omega \geq 10^{10}\ \mathrm{s}^{-1}$? To answer this it should be pointed out that the time scale of the elementary steps of many biochemical reactions is in the picosecond range [32]. Therefore $\sigma(\omega)$ in the high-frequency range between 10^{10} and $10^{12}\ \mathrm{s}^{-1}$ is of biological interest. The reason that we have extended the computations until the physically uninteresting frequency range of $10^{20}\ \mathrm{s}^{-1}$ was only to

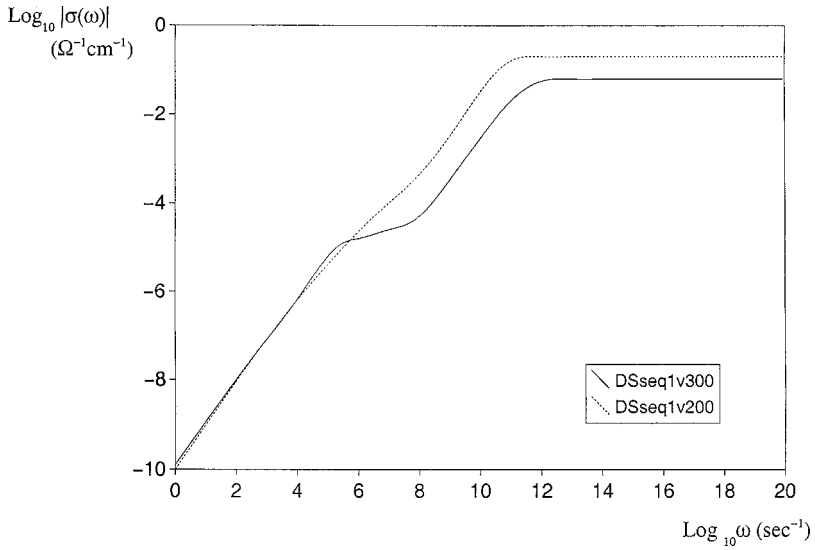


FIGURE 4. $|\sigma(\omega)|$ for 300 levels of 100 base pairs (seq.1) in the valence-band region using the double-ζ basis (DSeq1v300) (solid line). The broken line gives the result for 200 levels for the same system and basis set (DSseq1v200).

test whether Im[$\sigma(\omega)$] and Re[$\sigma(\omega)$] behave in a way as the theory predicts it (they do).

Further one can observe from the DOS curves [1] that around the edges of the gap the DOS contain much more levels at the valence-band region than in the conduction-band one. This explains that due to the larger level density and therefore smaller $\Delta E_{i,j}$ values the Boltzmann factors in Eq. (2) become much larger. Therefore $|\sigma(\omega)|$ is also much larger (by the same number of units, same sequence and number of levels considered) in the valence-band than in the conduction-band regions. On the other hand, if the number of levels taken into account is only 100, in both cases (the valence- and conduction-band regions, respectively) $|\sigma(\omega)|$ has only values between 10^{-5} and 5×10^{-5} Ω^{-1} cm^{-1} at $\omega = 5 \times 10^{10}$ s^{-1}. (The larger value belongs to the conduction-band region in this case.) This can be explained again by the effect of Boltzmann factors.

One should mention that $|\sigma(\omega)|$ has values of about 10^{-5} Ω^{-1} cm^{-1} at $\omega = 5 \times 10^{10}$ s^{-1} in the single stacks with 100 units (using both sequences) and 100 levels. Namely, in this case having 100 levels originating from the HOMOs of the single free bases, the Boltzmann factors become again smaller due to the larger $\Delta E_{i,j}$ level spacings (to save place we do not show these curves here).

One should point out that the log $|\sigma(\omega)|$ curves at the measured frequency [33] ($\omega \approx 5 \times 10^7$ s^{-1}) have similar values (between 5×10^{-6} and 10^{-5} Ω^{-1} cm^{-1}) for the base pair stack in the conduction-band region (200 levels) and for the single base stack both in the conduction- and valence-band regions (100 levels) than the theoretical ones. On the other hand the base pair stack in the valence-band regions (200 levels taken into account again) has a much larger theoretical $|\sigma(\omega)|$ value of $\approx 10^{-4}$ Ω^{-1} cm^{-1} at the mentioned frequency. At lower frequencies ($\sim 5 \times 10^3$ s^{-1}) our $|\sigma(\omega)|$ values are substantially larger (10^{-7}–5×10^{-7} Ω^{-1} cm^{-1}) than the experimentally found value of 5×10^{-10} Ω^{-1} cm^{-1}. One should mention, however, that the experimental conductivity values, especially at lower frequencies, are not very reliable. Namely, because of the difficulties in their purification and characterization the samples are not very well defined. Serious biomaterial science developments would be needed to obtain such samples of biopolymers, which would be well applicable for solid-state physical measurements (providing reliable experimental data).

One should mention further that for the base pair stack with 100 units we have performed also $|\sigma(\omega)|$ calculations using 300 and 400 levels for their valence-band regions. Since the filled levels over 200 do not originate from the HOMOs of the single bases but from the next filled π levels, the hopping conductivities obtained are more than an order of magnitude smaller than in the case of 200 levels. One can write down the matrix containing the hopping frequencies [which is necessary for the calculation of $D(\omega)$] as

$$\underline{\underline{H}} = \begin{pmatrix} \underline{\underline{L}}_1 & \underline{\underline{S}} \\ \underline{\underline{S}} & \underline{\underline{L}}_2 \end{pmatrix}. \tag{4}$$

Thus the matrix $\underline{\underline{H}}$ constructed from the hopping conductivities in the case of 400 levels can be partitioned into 200 × 200 blocks of $\underline{\underline{L}}_i$(arge) ($i = 1, 2$) and $\underline{\underline{S}}$(mall) parts. The submatrices $\underline{\underline{L}}_i$ contain the hopping frequencies originating from the HOMO and HOMO-1 levels of the free bases, and therefore they are large. Namely, one knows from ab initio Hartree–Fock correlation-corrected band structure calculations of periodic nucleotide base stacks that the resulting π bands do not overlap [34]. Therefore the corresponding Bolzmann factors are large due to the small level spacings. On the other hand in the cases of hoppings between levels that can be deduced from different π levels of the free bases, the level spacings $\Delta E_{i,j}$ are larger and the corresponding Boltzmann factors occurring in expression (2) are smaller. This causes that most elements of the submatrix $\underline{\underline{S}}$ will be small.

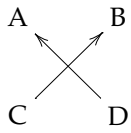
Next, one should turn to the case of 200 units of a single base stack shown in Figure 3. Taking into account 200 levels in the valence-band region of a single base stack $|\sigma(\omega)|$ is about 10^{-2} Ω^{-1} cm^{-1} again at $\omega = 5 \times 10^{10}$ s^{-1} (after having a saddle region between $\omega = 10^5$ and 10^9 s^{-1} with a $|\sigma(\omega)|$ value of 5×10^{-7} Ω^{-1} cm^{-1}). If we take in the 200 units sequence only 100 levels into account, $|\sigma(\omega)|$ decreases also to the above given value of 5×10^{-7} Ω^{-1} cm^{-1}. On the other hand for a single stack with 100 units and 100 levels (again in the valence-band region), we have obtained previously a value of $\approx 10^{-7}$ Ω^{-1} cm^{-1} (see above).

Finally we have calculated $|\sigma(\omega)|$ for a single stack with sequence 1 also with Clementi's double-ζ basis [21] taking into account 100 levels in the valence-band region. In this case in the base dimer calculations we have corrected the energy levels for correlation using the inverse Dyson equation in its diagonal approximation with an second-order Møller–Plesset (MP2) self-energy (for details see [1] and [3] and references therein). In the case of the double-ζ basis the saturation value of $|\sigma(\omega)|$ is about

an order of magnitude larger ($\approx 6 \times 10^{-6}\ \Omega^{-1}\ \mathrm{cm}^{-1}$) and occurs at $\omega = 10^{12}\ \mathrm{s}^{-1}$. Finally, the double-$\zeta$ curve with correlation corrections has a saturation value of $|\sigma(\omega)| \approx 2 \times 10^{-3}\ \Omega^{-1}\ \mathrm{cm}^{-1}$. Thus the better basis increases $|\sigma(\omega)|$ by one order of magnitude and the correlation effects by additional 2.5 orders of magnitude. These results can be again easily interpreted if we take into account that both the better basis and the correlation effects [38] decrease the level spacings $\Delta E_{i,j}$. In this way they increase through the Boltzmann factors the hopping frequencies (for the corresponding $|\sigma(\omega)| - \omega$ curves see [1]).

In Figure 4 $|\sigma(\omega)|$ is shown for a stack of 100 base pairs with sequence 1 taking into account 300 or 200 levels in the valence-band region, respectively, using the double-ζ basis. One can see that in the case of 300 levels $|\sigma(\omega)|$ is about $7 \times 10^{-2}\ \Omega^{-1}\ \mathrm{cm}^{-1}$, while taking into account only 200 levels (which all can be deduced from the HOMOs of the free bases) $|\sigma(\omega)|$ increases to about $2 \times 10^{-1}\ \Omega^{-1}\ \mathrm{cm}^{-1}$. This result can be explained in the same way as was done before for a base pair stack minimal basis calculation with 300 and 400 levels, respectively [see the text above and below Eq. (4)].

There is, however, an interesting phenomenon, namely, that the (Dseq1v200) calculation with a minimal basis gives about twice as large $|\sigma(\omega)|$ values than the double-ζ one (compare the corresponding curves in Figs. 1 and 4). This can be probably explained in the case of base pair stacks if we take into account also diagonal hoppings, which have small $\Delta E_{i,j}$ values (hopping in



from D to A or from C to B). In such cases the magnitude of the hopping frequencies in the 200-level case depends first of all on the overlap integrals and not on the Boltzmann factors [see Eq. (5)]. These overlap integrals have also an $\exp(-kR_{n,n'})$ dependence. This means that if the exponent of a Gaussian belonging to a minimal basis of say base A is α_{A}, this exponent falls usually between the orbital exponents of the corresponding Gaussian double-ζ basis, $\alpha_{\mathrm{A}_1} < \alpha_{\mathrm{A}} < \alpha_{\mathrm{A}_2}$. Therefore if we add the overlap integrals between $n = \mathrm{A}$ and $n' = \mathrm{D}$ of both members of the double-ζ basis belonging to the same type of basis function (say 2s) centered on the same type of atom (say N) we obtain the sum

$$\begin{aligned}
&\langle \chi_{2s'}^{(\mathrm{A,N})} | \chi_{2s'}^{(\mathrm{D,N})} \rangle + \langle \chi_{2s''}^{(\mathrm{A,N})} | \chi_{2s'}^{(\mathrm{D,N})} \rangle \\
&\quad + \langle \chi_{2s'}^{(\mathrm{A,N})} | \chi_{2s''}^{(\mathrm{D,N})} \rangle + \langle \chi_{2s''}^{(\mathrm{A,N})} | \chi_{2s''}^{(\mathrm{D,N})} \rangle \\
&\approx e^{-[\alpha_{2s'}^{(\mathrm{A,N})} + \alpha_{2s'}^{(\mathrm{D,N})}]_{R_{\mathrm{N(A)N(D)}}}} + e^{-[\alpha_{2s'}^{(\mathrm{A,N})} + \alpha_{2s'}^{(\mathrm{D,N})}]_{R_{\mathrm{N(A)N(D)}}}} \\
&\quad + e^{-[\alpha_{2s'}^{(\mathrm{A,N})} + \alpha_{2s'}^{(\mathrm{D,N})}]_{R_{\mathrm{N(A)N(D)}}}} + e^{-[\alpha_{2s'}^{(\mathrm{A,N})} + \alpha_{2s'}^{(\mathrm{D,N})}]_{R_{\mathrm{N(A)N(D)}}}}.
\end{aligned} \tag{5}$$

Here $R_{\mathrm{N(A)N(D)}}$ is the distance between the two N atoms on bases A and D, respectively [for the sake of simplicity we have left out from (5) the linear coefficients occurring in (1)]. Since at "diagonal" hoppings $R_{\mathrm{N(A)N(D)}}$ is rather large and because of the large orbital exponents $\alpha_{2s'}^{\mathrm{N(A\ or\ D)}}$, Eq. (5) may be substantially smaller than the corresponding minimal basis overlap integral. This may overcompensate the effect of the smaller $\Delta E_{i,j}$ level spacings in the Boltzmann factors caused by double-ζ basis. Of course, more detailed calculations and theoretical analysis would be needed to prove this assumption.

It should be mentioned that in the recent calculation of the hopping conductivity of an aperiodic sequence of 16 nucleotides, it was found that taking into account also the sugar–phosphate units $|\sigma(\omega)|$ increases in both of the valence- and conduction-band regions by one order of magnitude at comparable conditions (basis set, ω, T) [5]. Though the sugar–phosphate bands are below and above the valence and conduction bands of the homopolynucleotides, this result can be easily understood in the following way if one looks at Eq. (1). If the hopping occurs from the lower lying sugar–phosphate valence band (or bunch of levels) to the valence band of the stack, this requires energy. On the other hand a hopping from the stack to a sugar–phosphate unit does not need energy and therefore the corresponding much larger hopping frequency will be dominant. The opposite happens in the conduction-band regions: The hopping from the sugar–phosphate levels to the levels belonging to the stack requires no energy, and therefore this much larger hopping frequency will be the dominant one.

Finally, it should be pointed out that according to recent energy loss experiments in a resonance cavity of lyophilized phage DNA resulted in $|\sigma(\omega)| \approx 2\ \Omega^{-1}\ \mathrm{cm}^{-1}$ at $\omega = 10^{10}\ \mathrm{s}^{-1}$ at 310 K [4]. Above $\omega = 10^{10}\ \mathrm{s}^{-1}$ σ is frequency independent [4]. The measurements were performed at different temperatures. The results show that σ strongly increases with T. The authors highlight that the charge

transport at room and higher temperatures is definitely electronic and a tunneling mechanism (which would be T-independent) can be ruled out. They concluded that the most probable mechanism of charge transport in a stack with a random sequence is hopping [4].

Conclusions

If one assumes electronic conduction of a base pair stack with 200 levels, using a minimal basis, one obtains for $|\sigma(\omega)| \approx 10^{-4}\ \Omega^{-1}\ \mathrm{cm}^{-1}$ (see Fig. 1). On the other hand if one takes into account also the sugar–phosphate chain (as we have seen above) this increases σ by 1 order of magnitude [5]. Further using a double-ζ basis and taking into account also correlation effects, this would increase σ at high ω by a further ~3.5 orders of magnitude [2]. Finally, the hopping frequencies [and with it the $\sigma(\omega)$ values] are somewhat underestimated because only a single (acoustic) phonon frequency was taken into account (the contribution of the optical phonons was completely neglected). On the basis of all these considerations we can estimate $|\sigma(\omega)|$ at $\omega \approx 10^{10}\ \mathrm{s}^{-1}$ a few times $1\ \Omega^{-1}\ \mathrm{cm}^{-1}$ in excellent agreement with the room-temperature results of Gruner et al. [4].

In a nucleohistone the negative phosphate groups of DNA are in close contact with the positive guanidium groups of the arginine residues of the histone molecules. This rather probably leads to a charge transfer from DNA to the protein molecule resulting in a hole conduction along the periodic sugar–phosphate chain.

Finally it should be mentioned that recently Barton and coworkers [37, 38] have found a hole conduction across the DNA base pairs for a 37-Å distance (though the conduction may persist for substantially larger distances of ~200 Å [39]) if they bind or intercalate to the DNA stacks electron acceptors. This conduction can be interrupted by kinks or strand breaks [38]. Our previous results on hopping conductivity in proteins have shown that this is influenced first of all by the conformations of the proteins and not by their sequences [31, 40]. Obviously there is a similar situation in aperiodic DNA.

Of course, the problem of charge transport across a nucleotide base stack is a complicated one. Some authors think that above 150 K the main mechanism is hopping, which could be interrupted only by scavengers like 5-bromo-6-hydroxy-5,6-dihydrothymine (TOHBr), 5-bromocytosine (CBr), etc. On the other hand according to our findings in the case of proteins [31] conformational changes can also decrease very strongly the hopping distance.

On the basis of time and frequency resolved fluorescence spectrum measurements of electron acceptors, other authors [41] have come to the conclusion that for smaller distances coherent charge transport (tunneling) and for larger distances incoherent hopping transport is the dominant mechanism. These authors have taken into account also solvent (water) dynamics.

There are experiments that show a distance dependence of the electron transfer between an acridine derivative (which is intercalated into DNA instead of a base) and guanine [42]. There are also photoinduced charge transport investigations of DNA using different techniques (see, e.g., [43]).

In our opinion even in the electronic ground state of DNA besides hopping, Bloch-type conduction, multichannel tunneling, and simultaneously other mechanisms, like sequential tunneling and hopping, charge transport along the sugar–phosphate chain influenced by the structure of surrounding water (which would be mostly ionic or protonic possibly coupled to the motion of the electrons along the double strand) can play a role. In the very necessary farther experimental and theoretical studies, to reach a better understanding of the dominant factors determining the charge transport in DAN, all these mechanism have to be taken into account.

ACKNOWLEDGMENT

One of us (J.L.) should like to express his deep gratitude to the late Professor Per-Olov Löwdin for the many very inspiring discussions about the electronic structure of DNA and by introducing him to the International Quantum Chemistry Community by his numerous invitations to Uppsala and to the Sanibel Symposia.

References

1. Ye, Y.-J.; Chen, R.-S.; Martinez, A.; Otto, P.; Ladik, J. Physica B 2000, 279, 246.
2. Ye, Y.-J.; Chen, R.-S.; Martinez, A.; Otto, P.; Ladik, J. Solid State Comm 1999, 112, 139.
3. Bell, G. I.; Picklet, R.; Writter, J. Nucleic Acid Res 1980, 8, 4091.
4. Tran, I. P.; Alavi, B.; Gruner, G. Phys Rev Lett 2000, 85, 1564.
5. Ye, Y.-J.; Jiang, Y. Int J Quantum Chem 2000, 78, 112.
6. Lawson, C. L.; Carey, J. Nature 1993, 366, 178.

7. Ladik, J.; Suhai, S. Int J Quantum Chem OBS 1980, 7, 181; Clementi, E.; Corongiu, G. Int J Quantum Chem OBS 1982, 9, 213.
8. Dickerson, R. A.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. I. Science 1982, 216, 475.
9. Ladik, J.; Suhai, S. Phys Lett 1980, 77A, 25.
10. Otto, P.; Clementi, E.; Ladik, J. J Chem Phys 1980, 8, 454; Clementi, E.; Corongiu, G. Int J Quantum Chem OBS 1982, 9, 213.
11. Ye, Y.-J., unpublished results.
12. Otto, P.; Ladik, J.; Corongiu, G.; Suhai, S.; Förner, W. J Chem Phys 1982, 77, 5026.
13. Corongiu, G.; Clementi, E. Biopolymers 1981, 20, 551.
14. Ye, Y.-J.; Ladik, J. J Math Chem 1993, 14, 141.
15. Gazdy, B.; Seel, M.; Ladik, J. Chem Phys 1984, 86, 41.
16. Dean, P.; Martin, J. I. Proc Roy Soc A 1960, 259, 409; Dean, P. Rev Mod Phys 1972, 44, 127.
17. Seel, M. Chem Phys 1979, 43, 103.
18. Day, R. S.; Martino, F. Chem Phys Lett 1981, 84, 86.
19. Ye, Y.-J. J Math Chem 1993, 14, 121.
20. Gianolo, L.; Pavoni, R.; Clementi, E. Gazz Chim Ital 1978, 108, 181.
21. Gianolo, L.; Clementi, E. Gazz Chim Ital 1980, 110, 179.
22. Wilkinson, J. H. The Algebraic Eigenvalue Problem; Clarendon: Oxford, 1965.
23. Mott, N. F.; Davis, E. A. Electronic Processes in Non-Crystalline Materials; Clarendon: Oxford, 1971; p. 215.
24. Ye, Y.-J.; Ladik, J. Int J Quantum Chem 1994, 52, 491.
25. Odagaki, T.; Lax, M. Phys Rev B 1982, 26, 6480.
26. Scher, H.; Lax, M. Phys Rev B 1973, 7, 4491.
27. Odagaki, T.; Lax, M. Phys Rev B 1981, 24, 5284.
28. Odagaki, T.; Lax, M.; Day, R. S. Phys Rev B 1984, 30, 6911.
29. Haus, J. W.; Kehr, K. W. Phys Rep 1987, 150, 263.
30. Ye, Y.-J.; Ladik, J. Phys Rev B 1993, 48, 5120.
31. Ye, Y.-J.; Ladik, J. Phys Rev B 1995, 51, 1309.
32. Netzel, T. L.; Nafini, K.; Headrick, J.; Eaton, B. E. J Phys Chem 1995, 51, 17948.
33. O'Konski, C. T. Rev Mod Phys 1963, 35, 5721.
34. Suhai, S. Int J Quantum Chem 1984, 11, 223; Bogar, F.; Ladik, J. Chem Phys 1998, 237, 273.
35. Ladik, J. J. Phys Rep 1999, 6, Polymer of Solids and Quantum Mechanical Treatment.
36. Ladik, J. J. Quantum Theory of Polymers as Solids; Plenum: New York, 1988; Chapter V.
37. Hall, D. B.; Holmkin, R. E.; Barton, J. K. Nature 1996, 382, 731.
38. Hall, D. B.; Barton, J. K. J Am Chem Soc 1997, 119, 5045; Arkin, M. R.; Stamp, E. D. A.; Pulver, S. C.; Barton, J. K. Chem and Biol 1997, 4, 369.
39. Barton, J. K. Lecture at the 25th Anniversary Conference of the National Foundation of Cancer Research, Washington, 1998.
40. Jiang, Y.; Ye, Y.-J.; Chen, R.-S. Biophys Chem 1996, 59, 95.
41. Okada, A.; Chernyak, V.; Mukamel, S. J Phys Chem A 1998, 102, 1241.
42. Fukui, K.; Iwane, K.; Shimidzu, T.; Tanaka, K. Tetrahedron Lett 1996, 37, 4893.
43. Shimomura, M.; Karthaus, O.; Ijiro, K. Synth Metals 1996, 81, 351.