Electronic transport and thermopower in DNA sequences

ARTICLE *in* MODERN PHYSICS LETTERS B · APRIL 2004

Impact Factor: 0.69

2 AUTHORS:

Stephan Roche
Catalan Institute of Nanoscience and Nano…
203 PUBLICATIONS **4,087** CITATIONS
SEE PROFILE

Enrique Macia
Complutense University of Madrid
90 PUBLICATIONS **1,523** CITATIONS
SEE PROFILE

Electronic transport and thermopower in DNA sequences

Stephan Roche [1] and Enrique Maciá [2]

[1] *Commissariat à l'Énergie Atomique, DSM/DRFMC/SPSMS, 17 avenue des Martyrs, 38054 Grenoble, France*

[2] *Departamento de Fisíca Materiales, Facultad de Fisicas, Universidad Complutense de Madrid, E-28040 Madrid, Spain*

(July 20, 2004)

Abstract

A detailed study of charge transport properties of synthetic and genomic DNA sequences is reported. Genomic sequences of the Chromosome 22, λ-bacteriophage, and D1s80 genes of Human and Pygmy chimpanzee are considered in this work, and compared with both periodic and quasiperiodic (Fibonacci) sequences of nucleotides. Charge transfer efficiency is compared for all these different sequences, and large variations in charge transfert efficiency, stemming from sequence-dependent effects, are reported. In addition, basic characteristics of tunneling currents, including contact effects, are described. Finally, the thermoelectric power of nucleobases connected in between metallic contacts at differet temperatures is presented.

PACS numbers: 87.14.Gg, 72.20.Ee, 72.80.Le

I. INTRODUCTION

Desoxyribo-Nucleic-Acid (DNA) [1] is the biomolecular cornerstone that encodes the fundamental nature of living species, and its description is one of the main fascinating and challenging quest of biology. Once disclosed, the next step is to analyze the relation between the genetic code and its biological functionalism, an important concern related in particular with gene regulation and cell division [2]. A natural question is thus to better characterize the nature of DNA sequences, that is given by the local distribution of four nucleotides, guanine G, adenine A, cytosine C, thymine T, arranged by pairs through weak hydrogen-bonds, and weakly interacting along the DNA helix. Given that the four nucleotides have different ionization energies, the local chemical reactivity is sequence and scale dependent. For instance accumulation of G nucleotides (as triplex) is known to produce high spots for charge accumulation and consequently oxidation preferentially takes place at such sites. Short and long range correlations between base pairs further provide valuable informations to distinguish between almost random distributions, and more complex sequences, whose long range correlations might be also associated with some biological properties (folding, introns vs exons featuring, ...) [3–5]. Scale invariant properties in complex genomic sequences with thousands of nucleotide base pairs have been investigated with mathematical focus, and in particular wavelet analysis has revealed complex fingerprints [4,5]. Such studies clearly show that correlations are present at different scales, and that they are strongly sequence dependent. The statistical significance of the regular features in nucleotide sequences should be estimated with respect to the corresponding characteristics for random DNA sequences with the same nucleotide composition, since there are excluded- volume effects (each position in a sequence is occupied by only one nucleotide), which results in some correlations even for seemingly random sequences.

On the other hand, amongst the various physical, chemical or biological phenomena that might be inferred from sequence correlations, charge transfer properties deserve particular attention. Indeed, the nature of DNA-mediated charge migration has been related to the

understanding of damage recognition process, protein binding, or with the task of engineering biological processes (e.g. designing nanoscale sensing of genomic mutations) [6,2], opening new challenges for emerging *nanobiotechnologies*. In addition to photochemical experiments, mesoscopic transport measurements are extensively used to address the question of intrinsic conducting nature of DNA. However, in addition to its own structural complexity (that might also depend on its close chemical surrounding), measuring charge transport in a DNA chain is strongly biased by the invasive role of contacts, interaction with some inorganic substrate, as well as temperature and atmosphere experimental conditions. In consequence, given this fantastic complexity and wealth of parameters included in charge transfer experiments, undivided caution has to be taken when interpreting results of a particular approach.

In the following we will first present the basic formalism that is used to compute transmission coefficients of DNA sequences connected to two semi-infinite external metallic leads. Some basic properties of current-voltage characteristics in the coherent regime will be also addressed. Then we will investigate the effect of long range correlations on charge transfer in several type of DNA, from artificially aperiodic to long range correlated genomic sequence. We will investigate and compare charge transport in these several chains, mainly on coherent charge migration that would dominate at very low temperature, and assuming rigid DNA structure and non-invasive contacts. Our analysis of charge transfer will be driven by transmission coefficients, computed assuming a certain resonant configuration of some metallic-like contacts or electronic states to the DNA under study. In the present work, we disregard the possible role of phonon-assisted hopping conduction that would deserve a full study in itself, but to our opinion, would result much more insensitive to correlations because of its incoherent nature. Finally a section will be devoted to the thermopower of DNA sequences, as a property derived from conductance scaling.

II. CHARGE TRANSFER IN PERIODIC AND APERIODIC GENOMIC DNA SEQUENCES

Early on, the π-stacking in double strand DNA has been anticipated to favor long range charge migration along the double helix axis [8]. Notwithstanding, temperature dependent dynamical motions of the weakly interacting base pairs, together with sequence dependent energetic inhomogeneities, or ambient conditions are clear factors that may jeopardize such expectation.

Experimentally, photoexcitation experiments suggest that charge excitations can be transferred from tethered donor to acceptor metallointercalators, via the guanine highest occupied molecular orbitals (G-HOMO) of the DNA bridge [2,9]. On the other hand, mesoscopic transport measurements on single artificial or genomic DNA sequences (mainly bacteriophage λ) contacted in between metallic electrodes have also been the subject of intense and controversial debate, and measurements of its conductivity have ranged from zero (ideal insulator) to infinity (superconductor) [12,13]. While accurate determination of absolute values of conductivity is important, characteristic sequence dependences of charge transport could provide clues to mechanisms and biological functions of transport. In that respect, electronic delocalization is expected to depend on the periodicity or degree of randomness of sequences, as well as temperature effects, that in a first approximation will introduce even more random in the sequence through the disorientation of base pairs one respect to the other. The possible role of long range correlations on electronic delocalization has been recently discussed [10,11]. To gather deeper insight of effect of correlations in electronic transfer, we investigate the electronic transmission properties of various correlated and uncorrelated DNA sequences contacted with metallic leads. The electrode energetics is adjusted to mimic the resonance with G-HOMO, so that we restrict our study to coherent charge transfer, disregarding the phenomenon of charge injection. In fact, he current measured through a DNA macromolecule connecting two metallic electrodes results from the injection of carriers onto the stack of bases together with the intrinsic conduction along

the DNA sequence. At low voltage, the main contribution to the resistance comes from the metal-DNA junction potential mismatch, whereas at sufficiently higher bias, once the electrode Fermi level and the G-HOMO states of the molecule are lined up, the resistance of the junction is controlled by the DNA-mediated transfer rather than the injection process efficiency. Accordingly, the observed gap in a I(V) measurement is somehow related to the energy difference between the metallic work function and the energy of the G-HOMO [15].

Our Hamiltonian is an effective tight-binding model that describes the site energies of an hole located at nucleotide n [15–17]

$$\mathcal{H} = \sum_n -t_{\rm DNA}\cos(\theta_{n,n+1})(c_n^\dagger c_{n+1} + h.c.) + \varepsilon_n c_n^\dagger c_n$$

where $c_n^\dagger(c_n)$ gives the creation (annihilation) operator of an hole at site n. The variables ε_n describe the energy of the hole at site n, and will be given by the ionization potentials of respective bases, taken as $\varepsilon_A = 8.24eV$, $\varepsilon_T = 9.14eV$, $\varepsilon_C = 8.87eV$, $\varepsilon_G = 7.75eV$ [17], while $t_{\rm DNA}$ stands as the hopping integral between adjacent nucleotides. Such coupling describes the $\pi - \pi$-stacking interaction between base pairs. Ab-initio calculations find that between two interacting nucleotides, $t_{\rm DNA}$ are in order of $\sim 0.1 - 0.4eV$ [17], and it has been shown [15] that effective tight-binding Hamiltonian for the Poly(dG)-Poly(dC) with $t_{\rm DNA} = 0.4eV$ reproduce ab-initio results [17] and experiments [13]. When connecting the DNA chains to external leads, the case $t_{\rm DNA} \sim t_{\rm m}$ ($t_{\rm m}$ the hopping term for the metallic contacts) will allow a better characterization of intrinsic conduction properties (such as resistivity). Finally $\theta_{n,n+1}$ is the relative twist angle deviated from its equilibrium value between sites n and n+1, because of temperature [14]. Each $\theta_{n,n+1}$ is an independent random variable that follows a Gaussian distribution with average $\langle\theta_{n,n+1}\rangle = 0$, whereas the variance is taken according to the equipartition law, i.e. $\langle\theta_{n,n+1}^2\rangle = k_B T/I\Omega^2$, where I is the reduced moment of inertia for the relative rotation of the two adjacent bases and Ω is the oscillator frequency of the mode ($I\Omega^2 = 250K$). Following [14], we take $t_{\rm DNA}\cos(\theta_{n,n+1}) \sim t_{\rm DNA}(1 - \theta_{n,n+1}^2/2)$.

In the following, our DNA sequences are assumed to be connected to two semi-infinite electrodes described by another tight-binding Hamiltonian with site energies $\varepsilon_m = \varepsilon_G =$

$7.75eV$, whereas the hopping integral is taken as $t_m = 1eV$. This mono-molecular electronic device ideally should be ohmic and low in resistance,to ensure that the properties investigated are those of the molecule and not of the molecule - electrodes contacts. Assuming semi-infinite electrodes is enough to warrant that no charge carrier exiting the molecule will reenter it with the same phase (i.e. in the electrodes charges completely lose their phase coherence) [18].Besides we neglect the finite cross section of the electrodes (only one channel for charge transfer ε_m at the Fermi level). Our choice for modelling the external leads allow us to scan the transmission profile of a given DNA sequence over a spectrum with width $[\varepsilon_m - 2t_m, \varepsilon_m + 2t_m]$. An important issue is then related to the absolute value of $t_{\rm DNA}/t_{\rm m}$. Indeed, if $t_{\rm DNA}/t_{\rm m} \sim 1$ the potential barrier at contact interface is small, allowing for a better assessment of intrinsic scattering properties throughout DNA, whereas for $t_{\rm DNA}/t_{\rm m} \ll 1$, strong backscattering at interface might dominate the behavior of transmission coefficient and screen its intrinsic features (such as the typical resistivity). This motivates our choice to consider $t_{\rm DNA}$ within the range $[0.4, 1]eV$.

Now to compute the transmission coefficient, the transfer matrix formalism is used [19]. The time independent Schrödinger equation is projected into the localized basis and solved on the N-site sequence of concern, by properly taking the boundary conditions (to simulate the connection with leads). Sites between $[-\infty, 0] \cup [N+1, +\infty]$ belong to the leads, whereas sites $i = 1, N$ are associated to the DNA sequence under study. One easily gets

$$\begin{pmatrix} \psi_{n+2} \\ \psi_{n+1} \end{pmatrix} = T_n \begin{pmatrix} \psi_{n+1} \\ \psi_n \end{pmatrix} = T_n T_{n-1} .... T_1 \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix} = \mathcal{P}_n \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix} \tag{1}$$

with ψ_n the component of wavefunction for energy E at site n and

$$T_n = \begin{pmatrix} \dfrac{E - \varepsilon_n}{t_{n+1}} & -\dfrac{t_n}{t_{n+1}} \\ 1 & 0 \end{pmatrix} \qquad \text{and } \mathcal{P}(n) = \prod_{i=n}^{1} T_i$$

Now the incoming and outgoing scattering waves within the electrode metals expand as

$$\begin{aligned} \psi_n &= \mathcal{A} e^{{\rm i} k_L n a} + \mathcal{B} e^{-{\rm i} k_L n a} \qquad n \leq 0 \\ &= \mathcal{C} e^{{\rm i} k_R n a} + \mathcal{D} e^{-{\rm i} k_R n a} \qquad n \geq N + 2 \end{aligned}$$

taking a DNA sequence of N+1 bases.

$$\begin{pmatrix} \psi_{N+2} \\ \psi_{N+1} \end{pmatrix} = \begin{pmatrix} \frac{E-\varepsilon_{N+1}}{t_{dnamet}} & -\frac{t_{dna}}{t_{dnamet}} \\ 1 & 0 \end{pmatrix} \prod_{\alpha=N}^{2} T_\alpha \begin{pmatrix} \frac{E-\varepsilon_1}{t_{dna}} & -\frac{t_{dnamet}}{t_{dna}} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix} \tag{2}$$

with

$$\begin{aligned} \psi_{N+2} &= \mathcal{C}e^{\mathrm{i}k_R(N+2)a} + \mathcal{D}e^{-\mathrm{i}k_R(N+2)a} \\ \psi_{N+1} &= \mathcal{C}e^{\mathrm{i}k_R(N+1)a} + \mathcal{D}e^{-\mathrm{i}k_R(N+1)a} \\ \psi_1 &= \mathcal{A}e^{\mathrm{i}k_L a} + \mathcal{B}e^{-\mathrm{i}k_L a} \\ \psi_0 &= \mathcal{A} + \mathcal{B} \end{aligned}$$

while the bias voltage dependent electronic structure of the contacts is given by $E_L(k) = E_{met} + V_{bias} - 2t_{met}\cos(k_L a)$ and $E_R(k) = E_{met} - 2t_{met}\cos(k_R a)$, so that $k_L = Arcos(\frac{E-E_{met}-V}{2t_{met}})$ and $k_R = Arcos(\frac{E-E_{met}}{2t_{met}})$. Finally, the full transmission matrix equation can be recast to

$$\mathcal{T}_N \begin{pmatrix} \mathcal{B} \\ \mathcal{A} \end{pmatrix} = \begin{pmatrix} \mathcal{D} \\ \mathcal{C} \end{pmatrix} \tag{3}$$

with $\mathcal{T}_N = \Theta^{-1}\mathcal{S}^{-1}(k_R)\mathcal{P}_N\mathcal{S}(k_L)$, defining

$$\mathcal{S}(k_{L,R}) = \begin{pmatrix} e^{-\mathrm{i}k_{L,R}a} & e^{\mathrm{i}k_{L,R}a} \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \Theta = \begin{pmatrix} e^{-\mathrm{i}k_R(N+1)a} & 0 \\ 0 & e^{\mathrm{i}k_R(N+1)a} \end{pmatrix} \tag{4}$$

The transmission coefficient T(E) between metallic electrodes is then directly related to $|\mathcal{C}|/|\mathcal{A}|^2$ (having assumed $\mathcal{D} = 0$). The Landauer conductance [19] follows as

$$\begin{aligned} G_L &= \frac{2e^2}{h}\frac{|\mathcal{C}|}{|\mathcal{A}|}^2\frac{\sin(k_R a)}{\sin(k_L a)} \\ &= \frac{2e^2}{h}\frac{1}{\mathcal{T}_N(1,1)}\frac{\sin(k_R a)}{\sin(k_L a)} \end{aligned}$$

The transmission study between metallic electrodes allows to evaluate the energy and bias dependent transmission coefficient $T(E,V)$, that gives the fraction of tunneling electrons transmitted through the DNA. In the case of weak transmission, $T(E)$ can be further related

to the the Landauer resistance $h/2e^2(1-T(E))/T(E)$ ($h/2e^2$ the quantum resistance). Finally ,to compare transmission properties of different chains, we also study the behavior of Lyapunov coefficient $\gamma_N(E)$ (Ly) that is directly related to the transmission coefficient as [20]

$$\gamma_N(E) = \frac{1}{2N}\ln(T_N(E))$$

for a sequence with N sites. Ly has been widely used in the past as a useful and powerful probe to sort out the main features of complex localization patterns. For instance for systems with uncorrelated disorder, Ly gives a direct access to the energy-dependent localization length ($\xi(E) = 1/(\lim_{N\to\infty}\gamma_N(E))$ for E belonging to the electronic spectrum). In presence of scale invariance properties (such as in the case of quasiperiodic systems), the underlying structure of Ly reflects the self similar spectrum [21]

The electronic transmission coefficients is reflecting the strength of backscattering due to a given energetic profile. As our metallic leads are defined according to G-HOMO, at zero bias voltage hole transport will mainly experience a sequence dependent contribution of backscattering owing to the distribution and number of C, T, A potential barriers, upscaling with the sequence length. This will allow to discriminate between the relative contribution of sequence profiles to the resulting charge transfer properties.

We will hereafter first focus on the zero-bias case, and compare the case of sequences with artificial quasiperiodic chains made from two bases G and C arranged following the Fibonacci sequence, or with genomic sequences of ch22, λ-bacteriophage, as well as D1s80 gene of pygmy chimpanzee and human. Correlations in ch22 have been recently addressed [5]. As shown hereafter, correlation (artificial or genomic) will yield a more complex behavior of backscattering, with the possible resurgence of states with high transmissivities for larger chains, absent in sequences with uncorrelated disorder. Similarly, we confirm that D1s80 gene display very efficient charge transfer [23], given its huge concentration of GG-pairs.

STEPHAN: HE REESCRITO ESTE PARRAFO.

Quite interestingly, although it is supposed that this gene encodes similar functionalism

in both Human and Chimpanzee genomes, the corresponding charge transfer properties turn out to be quite different, as deduced from their related transmission fingerprints. In particular, charge transfer results very much efficient in chimpanzee's sequence, and it is enhanced by several order of magnitudes as compared to the λ-bacteriophage transmission coefficient corresponding to chains longer than 80nm.

III. PERIODIC AND QUASIPERIODIC CHAINS

To gain further insight into the nature of so complex phenomena like charge transport in DNA molecules for which exact or even approximate analytical results are not available, the study of quasiperiodic systems exhibiting characteristic self-similarity properties may be of considerable help. One typical aspect of any quasiperiodic system, possessing aperiodic long range order, is the presence of some repeating features which show up as some sort of building blocks of these structures. As a archetypal example, let us consider the Fibonacci polyGC sequence, which is constructed starting from a G base as a seed and following the inflation rule G$\rightarrow$ GC and C$\rightarrow$ G. This gives successively, G,GC,GCG,GCGGC,GCGGCGCG,GCGGCGCGGCGGC, $\cdots$, and so forth. In the thermodynamic limit the ratio of (majority) G basis over (minority) C basis will approach the the golden mean value $\tau = (1+\sqrt{5})/2 \sim 1.618$. A symmetrical inflation rule (C$\rightarrow$CG and G$\rightarrow$C) can be also used to generate a Fibonacci chain where the roles of majority and minority basis are exchanged. Since the basic building block in the Fibonacci inflation rule is the dimer GC, we briefly review some basic properties of the periodic polyGC chain for the sake of illustration. Its electronic structure is given the dispersion relation

$$4t^2 \cos^2 q = E^2 - (\varepsilon_{\rm C} + \varepsilon_{\rm G})E + \varepsilon_{\rm C}\varepsilon_{\rm G}, \tag{5}$$

so that the energy spectrum of a GC chain is composed of two wide bands separated by a gap of width $\Delta_{\rm GC} = \varepsilon_{\rm C} - \varepsilon_{\rm G} = 1.12$ eV, as it is illustrated in the left-top panel of Fig.1. The transmission coefficient for a GC chain embedded between guanine leads is given by

$$T_N(E) = \left[1 + \frac{\Delta_{\rm GC}^2}{4t^2 - (E-\varepsilon_{\rm G})^2} U^2_{\frac{N}{2}-1}(v)\right]^{-1}, \tag{6}$$

where $U^2_{m-1}(v)$ is a Chebyshev polynomial of the second kind, and $v \equiv (E-\varepsilon_{\rm C})(E-\varepsilon_{\rm G})/2t^2 - 1$. From the knowledge of the transmission coefficient we can obtain the corresponding Lyapunov coefficient from Eq.(**??**). In Fig.2 we show the energy dependence of $\gamma_N(E)$ for periodic polyGC chains of different lengths. As the system size grows an increasing number of extended states with $T(E_i) = 1$ and $\lim_{N\to\infty} \gamma_N(E_i) = 0$, progressively appear at the resonance energies given by the condition $U^2_{\frac{N}{2}-1}(v) = 0$.

A similar behavior is obtained for the next approximant in the series, namely the periodic polyGCG chain. In that case the energy spectrum is composed by three bands, as can be readily checked from the dispersion relation

$$2t^3 \cos 3q = (E-\varepsilon_{\rm G})^2(E-\varepsilon_{\rm C}) - t^2(3E - 2\varepsilon_{\rm G} - \varepsilon_{\rm C}). \tag{7}$$

As we proceed by considering higher order approximants to the Fibonacci DNA, new bands and gaps progressively appear in the energy spectrum, showing a hierarchical nested structure. Quasiperiodic correlations have been widely studied, and their ability to trigger power-law localization of eigenstates in the thermodynamic limit ($n \to \infty$) or power-law growth of Landauer resistance in finite samples has been shown by scaling analysis [24]. It has been shown that the global structure of the asymptotic electronic spectrum can be obtained in practice by considering very short approximants to infinite quasiperiodic chains [21]. Then, we will focus on the next approximant which correspond to the periodic poly(GCGGC) chain containing five nucleotides in its unit cell. Its dispersion relation is given by $2t^5 \cos(5q) =$

$$(E-\varepsilon_{\rm G})^3 (E-\varepsilon_{\rm C})^2 - t^2 (E-\varepsilon_{\rm G})(E-\varepsilon_{\rm C})(5E - 4\varepsilon_{\rm G} - \varepsilon_{\rm C}) + t^4 (5E - 3\varepsilon_{\rm G} - 2\varepsilon_{\rm C}), \tag{8}$$

so that the energy spectrum of a GCGGC chain is composed of three bands, whose centers are located at the energies $E_2 = 6.9155$ eV, $E_3 = 8.1432$ eV, and $E_4 = 9.5272$ eV, plus two narrower bands located at the edges of the spectrum at $E_1 = 6.1908$ eV and $E_5 = 10.213$

eV. This electronic structure is illustrated in the right-top panel of Fig.1, for the four bands included in our considered spectral window $[5.75, 9.75eV]$. In the bottom panels of Fig.1 we compare the energy dependence of the transmission coefficient numerically obtained for Fibonacci approximants with an increasing number of bases. It can be clearly appreciated that, as the system grows larger, several peaks with high transmission values remain at certain energy values. In addition, some degree of clustering around these resonant energies can be appreciated. [21,22] These resonant energies are robust enough to persist against backscattering effects due to the presence of C bases interspersed in the chain. Thus, one may be tempted of thinking that these states should exhibit good transport properties even in the thermodynamic limit. To further substantiate such a possibility we have analytically evaluated the transmission coefficient corresponding to the poly(GCGGC) chain embedded between guanine loads, which is given by

$$T_N(E) = \left[1 + q(x,y)U^2_{\frac{N}{5}-1}(w)\right]^{-1}, \tag{9}$$

where $x = (E - \varepsilon_{\rm C})/2t$, $y = (E - \varepsilon_G)/2t, w = 16x^2y^3 - 16xy^2 - 4yx^2 + 3y + 2x$, and $q(x,y) \equiv A^2/(1-y^2) + B^2 - 1$, with $A \equiv -24xy^3 - 16x^2y^2 + 6xy + 2x^2 + 32x^2y^4 + 4y^4 + y^2$ and $B \equiv 32x^2y^3 - 8x^2y - 24xy^2 + 4y^3 + 3y + 2x$. According to Eq.(9) the roots of the Chebyshev polynomial label a full transmission peak series given by $\cos(5k\pi/N) = w$,with $k = 0, ..., N$. Then, as the Fibonacci chain length is increased, less and less states will present good transmissivity, due to the progressive fragmentation of the spectrum. Nonetheless, a significant number of resonant states, satisfying the full transmission condition given by Eq.(9), will persist in the thermodynamic limit. Thus, we conclude that states belonging to the broader central bands around $E_2 \simeq 6.9$ eV and $E_3 \simeq 8.1$ eV are very robust to the progressive fragmentation of the energy spectrum. This point is conveniently illustrated in the bottom frames shown in Fig.1. The Lyapunov coefficient of the polyGC Fibonacci approximant (shown in Fig.2) display typical scale invariance properties, showing clear hints of self-similarity in the characteristic arrangement of dips and bumps. Indeed, the series of main elliptic bumps found in the chain with 60bp are reproduced at different scales in

the 240bp sequences, that also present superimposed features owing to the partitioning of spectrum. In agreement with the results obtained from the study of the transmission curve, chains as long as 160nm still exhibit nearly resonant transmission around the two main resonance energies.

IV. THE CH22, λ-BACTERIOPHAGE AND D1S80 SEQUENCES

The DNA sequence of the first completely sequenced human chromosome 22 contains about 33.4×10^6 nucleotides [26]. Statistical analysis have unveiled the presence of long range power law correlations which is inferred from scale invariance properties [5]. However, at variance to a quasiperiodic sequence, no construction rule allows to generate to whole chain, so that these correlations are of an intrinsically different nature. It is an important question to identify the limits of coherent charge transport in such a complex DNA sequence, that may be related with biological features. A recent study has anticipated on the role of long correlations in the ch22 sequence [10]. Given the huge amount of nucleotides, the physically relevant task would be rather to determine to which extent charge transfer takes place through the G-HOMO, in comparison with uncorrelated chains. Herebelow, charge transfer in finite sequences of chromosome 22 is first considered. The results previously obtained for quasiperiodic chains will help us in deepening our discussion about the relation between long range correlations and charge transport in aperiodic systems. Indeed, scale invariance characteristic of quasiperiodic sequences illustrate the way correlations among different bases at several scales give rise to a similar scaling of transmission properties.

We thus start focusing on ch22-based sequences extracted from the sequenced part entitled NT_{011520} containing 182.606 bp and retrieved from the Biotechnology Information (NCBI). Our first 20nm sequence is constructed by starting from site 1.500 of the full NT_{011520} sequence and then extracting the first 60 first bp, namely *gtgaaaccccatctctactaaaaatccaaaaaaaattagccgggtgtggtggcaggcgcct* (in doing so, we have fol-

lowed the choice made in a prior study to select the corresponding starting site [10]). Next sequences are constructed by adding the following next 60 bp of the sequence. Fig.3 presents the energy-dependent transmission coefficients and Lyapunov coefficients of chains with lengths between 20 nm and 90 nm. Lyapunov coefficients are computed for two ch22-based finite sequences (Fig.3-bottom right). Compared to quasiperiodic case (Fig.2), it is striking that self-similarity seems absent from the spectrum, so that the scaling in chromosome 22 relies on totally different kind of long range correlations. Self-similarity as present in quasiperiodic chains have been demonstrated to induce extended states at finite number of energies. The whole spectrum presents scale invariant features associated with the progressive partitioning of the spectrum with increasing length. In conclusion genomic do manifest long range correlations which however can not be assumed self-similar in the usual sense.

Next, in Fig.4, the temperature dependent transmission coefficients are shown for the the chain (cttcgggagg ctgaggcgga tgaatcacga ggtcaggagt tcaagaccag cctggccaac) extracted from the ch22 (starting site position = 150.000). In our approach, temperature yields missorientations of adjacent bases that result in temperature dependent base-base hoppings. At low temperature (see the case T=60K), the transmission spectrum present higher number of transmitting states, due to a breaking of level degeneracy. At higher temperature, the number of transmitting states decreases but interestingly there persist many states with high transmission coefficient at temperature as high as $\sim 160K$ (similar results are obtained for λ-based sequences [16]).

Let us now consider the case of the λ-phage DNA [25], whose transport properties have been been widely investigated experimentally [12]. The complete λ-phage DNA contains 48502 base pairs with a total length of about $16\mu m$. In Fig.5, we show for $t_{\rm DNA}/t_{\rm m} = 1$, the behaviors of $T(E)$ for short sequences of the chain but with increasing number of base pairs (bp), corresponding to systems from 20nm to 80nm long. The starting sequence with 60bp is λ_1-gggcggcgacctcgcgggttttcgctatttatgaaaattttccggtttaaggcgtttccg while larger sequence are constructed from λ_1 adding successively the next 60bp of the complete sequence. The transmission spectrum is critically scale dependent. By upscaling the sequence length, trans-

mission degrades and for length above $\sim$ 100nm, almost all states are strongly backscattered by the energetic profile. Different parts of the λ-phage sequence have also been investigated, suggesting that the exact details of a given transmission pattern critically depend on the exact structure of the sequence (see [16]).

We now briefly outline the effect of chemical mismatches that are known to produce strong damping of charge transfer. Herebelow mismatches are introduced in λ_1-gggcggcgacc tcgcgggttttcgc tatttatgaaaat tttccggtttaaggcgtttccg, by modifying all guanine to thymine nucleotide (Fig.6-top panel) in between ggg triplets present in the chain. Similar changes are done with $T \to G$ (middle panel) and $G \to A$ (bottom panel). As shown on the Figure, the transmission profile is also sensitive to the presence of chemical mismatches and the effect is merely predictive. If we focus on to the position of an effective Fermi level at HOMO-G energy (i.e. around 7.75 eV), one observes that whether replacement of guanine by thymine yields a vanishing transmission coefficient, the reverse situation leads to a strong transparency of the states at such energy. Different part of spectrum are differently affected depending on involved chemical mismatches. All this point towards a strong sensitivity of charge transport versus sequence details and correlations, in agreement with experiments [28].

V. THE CASE OF HUMAN AND PYGMY CHIMPANZEE D1S80 DNA

As shown before, genomic sequences have different long range correlations which result in more or less general trends in charge transfer. The case of Human and Pygmy Chimpanzee D1s80 DNA is particularly interesting since, despite their lack of periodicity, such sequences contain a large number of GG-pairs that have critical consequences on charge transfer. But more striking is the particularly strong efficiency of delocalization in the case of Pygmy Chimpanzee D1s80 DNA when compared to its Human counterpart.

Fig.7 and Fig.8 show the transmission coefficient spectra for increasing number of bp from 60bp up to 960bp in the case of Pygmy Chimpanzee D1s80. Under upscaling, the

charge transfer properties are shown to be particularly stable, and much more efficient when compared to λ or Ch22. But more strikingly is the significant difference with the Human D1s80 DNA, in which transmission is more quickly damped with increasing sequence length. This is an interesting observation for similar genes in however different species.

VI. DISCUSSION ON INTRINSIC RESISTIVITY AND TRANSFER RATES

Relating transmission coefficient results with measured resistivity is not an easy task. In principle, within the Landauer formalism, the intrinsic energy-dependent resistivity of DNA molecules would write

$$\rho(E) \sim \frac{h}{2e^2}\frac{1-T(E)}{T(E)}\frac{S_{DNA}}{L_{DNA}}$$

with L_{DNA} the length of the chain, $S_{DNA} = 3.10^{-18}m^2$ the typical average cross section and $h/2e^2 = 12k\Omega$ the quantum resistance. Such formula should not depend on external leads and has a range of applicability limited to the regimes where $T(E) \ll 1$. In the case of resonant (or ballistic) transmission, $\rho(E)$ obviously tends to zero, but in that regime, the resistance if fixed by the contacts plus molecule system, and will be limited to $\frac{h}{2e^2}\frac{1}{T(E)}$, while the resistivity has no more physical sense. One must underscore however that even in the regime $T(E) \ll 1$, such estimate of intrinsic resistivity is problematic. Indeed, as shown in most of the studied chains, as DNA length increases, transmission coefficient are scale-dependent and quickly tend to zero for typical length scale in the order of $100nm$, if we include temperature effect (following our approach). This means that deducing a "intrinsic resistivity" for a given length of the system is unsuitable. One may apply such argument to the experimental measurements that do generally assume either diffusive or localized regimes. Notice that for periodic Poly(dG)-Poly(dC), two well defined bands of Bloch-like states can be defined, and that resistivity can be extrapolated from the temperature dependent reduced transmission coefficient. For temperature as high as $240K$, one finds that the most conducting states roughly correspond to $T(E) \sim 1.510^{-2}$ for the 20nm long

chain. From this, one deduces $\rho \simeq 0.012\Omega$.cm [16], which turns out to be in good agreement with the experimental measurement [13].

Transfer rates are also interesting quantities since they can be directly extracted from photoluminescence experiments, and should be less affected by "chemical contacts" between donor/acceptor and the DNA bridge. Theoretically, much work has been done in generalizing a Fermi golden rule type formula (Marcus-Hush) [29] that directly computes a transfer rate between donor and acceptor, and treat inelastic effects on a semiclassical ground. Recently, a relation between the electronic transfer rate k_{DA} [**?**] and the quantum resistance has been derived by Nitzan [30]In this framework, the conductance writes $\sim e^2/\pi\hbar \times (10^{-13} k_{DA}(ps^{-1}))$ and allow some direct estimation of k_{DA}, once the quantum resistance has been computed. For a Poly(dG)-Poly(dC) with transmission coefficient $T(E) \sim 1.510^{-2}$ ($L_{DNA} = 20$nm), the transfer was found to be in order of $k_{DA} \sim 0.2ps^{-1}$ (for $t_{\rm DNA} = 1eV$). By taking $t_{\rm DNA} = 0.4eV$ closer to ab-initio predictions, $k_{DA} \sim 0.032ps^{-1}$ which is in the order of the transfer rate estimated with photochemical experiments, on smaller DNA sequences but in water environment [16,31]. The relation between molecular conduction and electron transfer rate was also derived for the case where both processes proceed via incoherent sequential hopping between nearest-neighbor bridge sites [32].

VII. TUNNELING CURRENTS

The transmission coefficient is an useful mathematical tool in order to describe the transport efficiency in quantum systems. Nonetheless, $T_N(E)$ is usually difficult to be directly assessed experimentally. Access to transmission properties can be performed by measuring current-voltage characteristics. However applying a voltage bias in between leads contacting the DNA has also some influence on the scattering properties inside the molecule, and direct information on intrinsic effects of sequences on transmission should thus be considered with care. In experiments on molecular wires the electric current I through the molecule is measured as a function of the voltage, V. It should rigorously be calculated by slightly

generalizing the Landauer formula which relates the electric current to the transmission coefficient as

$$I(V) = \frac{2e}{h} \int T_N(E,\delta)\ [f_L(E,T,V) - f_R(E,T,V)]\ dE, \tag{10}$$

where we assume the charges propagate from left to right. In this expression the transmission coefficient depends on the system's length, N, the charge carrier energy, E, and the lead coupling factor, δ. If δ were a function of the applied voltage then $T = T_N(V)$ as well. However, one usually assumes the transmission coefficient is evaluated a zero bias. This low bias approach is a very good description for short molecules connected to metallic leads by means of auxiliary molecules, and we will assume its validity in our study as a first approximation. Nevertheless, in the case of large DNA molecules the possible modification of the macromolecule electronic structure in the vicinity of the leads should be considered for moderated biases. Explicit evaluation of the Fermi distribution functions assuming a linear voltage drop (i.e. $\mu_{L,R} = \mu \pm eV/2$) in Eq.(10) leads to

$$I(V) = \left(\frac{2e}{h}\right) \sinh\left(\frac{eV}{2k_BT}\right) \int T_N(E,\delta) \left\{\cosh\left(\frac{E-\mu}{k_BT}\right) + \cosh\left(\frac{eV}{2k_BT}\right)\right\}^{-1}\ dE, \tag{11}$$

where k_B is the Boltzmann constant. Generally speaking, the I(V) curve can then be seen as the product of the hyperbolic function $\sinh(eV/2k_BT)$ modulating the integral factor expression.

STEPHAN: PARRAFO REESCRITO

In order to analyze the role of contacts in the charge transport, we shall consider the current density through our system as a function of the coupling strength, τ, between the DNA and the leads. The bias voltage dependence of the contact energies is given by $E_L(k) = \varepsilon_m + V + 2t_m \cos k$ and $E_R(k) = \varepsilon_m + 2t_m \cos k$, at the left ($L$) and right ($R$) lead, respectively. In the coherent regime, the electrical current $I(V)$ through the DNA molecule writes

$$I(V) = \frac{2e}{h} \int_{E_F}^{E_F+eV} T_N(E,V) dE$$

To extract the main features of tunneling currents in DNA chains, let us first compare the behavior of a PolyG chain with that corresponding to PolyAT and PolyGC sequences

under the resonance condition given by the coupling parameters choice $t = \tau = t_m = 1$ (Fig.9). In this case, if the potential barrier between the metallic contacts and the DNA is set to zero, a staircase increase of $I(V)$ is found (Fig.9-inset), in agreement with prior calculations [33]. In contrast, as soon as a potential barrier between the DNA and the metals is introduced, great changes are observed in the I(V) curves, which now exhibit typical *negative differential resistance*. Such phenomenon is a tunneling-related effect that has been observed in silicon-based heterostructures [35], as well as in small molecules [36]. The comparison of a pure G-based with PolyGC and PolyAT periodic chains is striking. The current density through a $N = 10$ PolyAT is several orders of magnitude lower that the one corresponding to PolyGC or PolyG of the same length, hence illustrating the influence of the considered DNA sequence in the charge transport. The voltage threshold at a given current scale, is also very sensitive to the presence of guanine, and can differ by several Volts. Such results are easily understood as the result of both deeper and larger tunneling barriers, related to the presence of basis with higher ionization potentials.

VIII. CONDUCTANCE AND THERMOELECTRIC POWER OF DNA BASED MOLECULAR JUNCTIONS

A. Physical Motivations

Due to the extreme sensitivity of thermopower to finer details in the electronic structure, the study of thermoelectric voltage over a molecular system attached to two metallic leads can provide new insights in nanoelectronics. On the one hand, one can gain valuable information regarding the location of the Fermi energy relative to the molecular levels. On the other hand, one can shift the Fermi level position in order to optimize the thermoelectric performance of a given molecular arrangement. The thermoelectric voltage over guanine molecules adsorbed on a graphite substrate was measured by Poler and co-workers using a STM tip [37]. The obtained Seebeck coefficient value ($S \simeq +18$ μV K^{-1} at room tem-

perature) indicates a p-type conduction. A similar figure has been recently derived from a theoretical study considering a phenyl-dithiol molecule chemisorbed on a gold surface [38]. These results naturally rise the question regarding the possible use of suitable organic molecules in designing novel thermoelectric devices. To this end, the recourse to nucleic acid nucleobases should take vantage of the vast knowledge gained from bioengineering research during the last decade.

In this Section we present a theoretical study of the Seebeck coefficient, S, and thermoelectric power factor ($S^2\sigma$, where σ is the electrical conductivity) of guanine (G), cytosine (C), adenine (A), and thymine (T) nucleobases connected in between metallic contacts at different temperatures. The role played by the chemical bonding at the interfaces is analysed in detail, since resonance effects may play a significant role in DNA based molecular systems [39].

B. Model and analytical expressions

Broadly speaking, one expects the binding to the metallic lead would affect the electronic structure of the molecule, so that we should consider the states belonging to the coupled molecular-metallic system rather than those of the molecular subsystem alone [40,41]. As a first approximation we shall consider the *weak coupling* case, so that the lead-molecule-lead junction can be properly described in terms of three non-interacting subsystems, [42,43] given by the single-band tight-binding Hamiltonian

$$\mathcal{H} = \varepsilon_0^{\upsilon} c_0^{\dagger} c_0 - \tau \left(c_0^{\dagger} c_1 + c_{-1}^{\dagger} c_0 + h.c. \right) + \sum_{k,l=\pm 1}^{\pm\infty} \left(\varepsilon_m c_{k,l}^{\dagger} c_{k,l} - t_m c_{k,l}^{\dagger} c_{k+1,l+1} + h.c. \right) \qquad (12)$$

where c_j ($c_j^{\dagger}$) is the creation (annihilation) operator for a charge at jth site in the chain, ε_0^{υ}, with $\upsilon = \{\mathrm{G,A,C,T}\}$, are the HOMO energies of the N-methylated G, A, C, and T bases, respectively, and τ measures the coupling strength between the leads and the nucleobase molecules. The leads are modeled as semi-infinite one-dimensional chains of atoms with one orbital per site, where ε_m is the on-site energy and t_m ($> \tau$) is the hopping term.

By considering nearest-neighbors interactions, within the transfer matrix framework, the Shrödinger equation can be expressed in the matrix form

$$M_{\upsilon} = \begin{pmatrix} 4J_{\upsilon}\cos k & -(2J_{\upsilon}+1) \\ 2J_{\upsilon}+1 & -(J_{\upsilon}+1)\sec k \end{pmatrix} \tag{13}$$

where $J_{\upsilon} \equiv 2\lambda x_{\upsilon}\cos k - 1$, $2\cos k = (E-\varepsilon_m)/t_m$, $\lambda \equiv t_m/\tau$, and $x_{\upsilon} = (E-\varepsilon_0^{\upsilon})/2\tau$. It is readily checked that $\det[M_{\upsilon}(E)] = 1$, so that the lead-base-lead transfer matrix belongs to the SL(2,R) group. The transmission coefficient is then given by the relationship $\mathcal{T}_{\upsilon}(E) = 4\sin^2 k/\mathcal{D}(E)$, where

$$\mathcal{D}(E) = [M_{12} - M_{21} + (M_{11} - M_{22})\cos k]^2 + (M_{11} + M_{22})^2 \sin^2 k. \tag{14}$$

Plugging the M_{ij} matrix elements given by Eq.(13) into Eq.(14), we obtain the analytical expression

$$\mathcal{T}_{\upsilon}(E) = \left[1 + \gamma^2 \frac{(E-\xi_{\upsilon})^2}{(E-E_-)(E_+-E)}\right]^{-1}, \tag{15}$$

where $E_{\pm} = \varepsilon_m \pm 2t_m$, define the allowed spectral window as determined by the metallic leads bandwidth, $\gamma \equiv \lambda^2 - 1 = (t_m^2 - \tau^2)/\tau^2$ measures the coupling strength, and

$$\xi_{\upsilon} \equiv \frac{\varepsilon_0^{\upsilon} t_m^2 - \tau^2 \varepsilon_m}{t_m^2 - \tau^2}, \tag{16}$$

is a base-dependent resonance energy. According to Eq.(15), the transmission amplitude is modulated by the metal-nucleobase coupling strength through the factor γ. The particular case $\gamma = 0$ ($\mathcal{T}_{\upsilon}(E) = 1$) corresponds to metallic conduction over the molecule. In that case, a very small thermoelectric voltage is expected from general principles. Consequently, we will consider the general case $\gamma \neq 0$. In this case Eq.(16) defines a resonance energy satisfying the full transmission property $\mathcal{T}_{\upsilon}(\xi_{\upsilon}) = 1$.

>From the knowledge of the transmission coefficient given by Eq.(15), the conductance through the lead-molecule-lead is determined using the Landauer formula [44]

$$G_{\upsilon} = \frac{2e^2}{h} \frac{\mathcal{T}_{\upsilon}(\mu)}{1 - \mathcal{T}_{\upsilon}(\mu)}, \tag{17}$$

where $2e^2/h \simeq 1/12906\ \Omega^{-1}$, and μ is the Fermi energy. We can also obtain the Seebeck coefficient by means of the expression [38]

$$S_v(T) = -|e|L_0 \left(\frac{\partial \ln \mathcal{T}_v(E)}{\partial E} \right)_\mu T, \tag{18}$$

where e is the electron charge, $L_0 = \pi^2 k_B^2/3e^2 = 2.44 \times 10^{-8}\ \mathrm{V^2K^{-2}}$ is the Lorenz number, and T is the temperature. Making use of Eq.(15) into Eq.(18) we obtain

$$S_v(T) = 2|e|L_0 T \left(1 + \frac{b a_v}{cd}\right) \frac{a_v \gamma^2}{cd + a_v^2 \gamma^2}, \tag{19}$$

where $a_v \equiv \mu - \xi_v$, $b \equiv \mu - \varepsilon_m$, $c \equiv E_+ - \mu$, and $d = \mu - E_-$.

C. Main results

According to Eqs.(15) and 19), in the cases $\gamma = 0$ or $\mu = \xi_v$ (i.e., $a_v = 0$) we have a perfect transmission and vanishing thermopower, as expected from general physical considerations about transport coefficients. Conversely, the Seebeck coefficient asymptotically diverges as μ approaches the band edges (i.e., $c = 0$ or $d = 0$). Therefore, very large thermopower values can be eventually reached by properly shifting the Fermi level through the electronic structure of the system.

In order to obtain quantitative results we evaluate Eqs.(17) and (19) at room temperature taking $\varepsilon_A = 8.24$ eV, $\varepsilon_T = 9.14$ eV, $\varepsilon_C = 8.87$ eV, and $\varepsilon_G = 7.75$ eV [45,46]. For the molecule-metal electronic coupling we consider values within the range $\tau = 0.1 - 0.5$ eV, corresponding to the $\mu - E_{\ HOMO}$ shift recently reported for oligothiophene derivatives adsorbed on gold [47]. The spectral window is given by the graphite π bandwidth $[-6.8, 0]$ eV, corresponding to the tight-binding parameters $t_m = 1.7$ eV, and $\varepsilon_m = -3.4$ eV [48]. With this parameter choice, the resonance energies given by Eq.(16) are located outside the allowed spectral window.

The resulting thermopower curves are very similar for the four bases considered. Thus, we get $\Delta S/S_G$ less than 1% for the different nucleobases over the entire spectral window, the

general trend being that the thermopower increases as the on-site base energy increases. We have also checked that the thermopower curves are rather insensitive to the adopted coupling strength value. Thus, we get $\Delta S/S_G$ less than 1% within the considered range, the general trend being that thermopower increases as τ increases. In the main frame of Fig.10 we show the guanine thermopower curve for $\tau = 0.5$ eV. As we can see, significantly large values of thermopower can be reached when the Fermi level is located close to the band edges. In particular, when the Fermi level is located at $E = -0.34$ we obtain the value $S_G \simeq +18$ μV K^{-1}, experimentally reported by Poler and co-workers. [37] In the inset of Fig.10, we compare the conductance of different nucleobases as a function of the Fermi level position. The four curves are very similar in shape, exhibiting a well defined maximum at the energy $E_\upsilon = \varepsilon_m + 4(t_m^2 - \tau^2)/(\varepsilon_0^\upsilon - \varepsilon_m)$, where $G_\upsilon(\mu)$ peak in the interval $G = 4 - 6 \times 10^{-8}$ Ω^{-1}, in reasonable agreement with usual experimental outcomes in metal-molecule-metal junctions. [49]

Making use of Eqs.(17) and (19) we can determine the magnitude S^2G, closely related to the thermoelectric power factor. In Fig.11 we compare the obtained graphite-G-graphite curve for two different values of the coupling strength τ. Far from the resonance energy E_ν, the overall shape of the power factor is mainly determined by the energy dependence of the Seebeck coefficient shown in Fig.10, while the power factor rapidly decreases close to the conductance peak. Noteworthy, at variance with the behavior observed for the Seebeck coefficient, the power factor is extremely sensitive to minor variations in the molecule-lead electronic coupling. In fact, S^2G is enhanced by three orders of magnitude by changing the coupling parameter from $\tau = 0.1$ eV to $\tau = 0.5$ eV, due to the conductance modulation. This result indicates that we can efficiently optimize the $S^2\sigma$ output by properly selecting the chemical group connecting the molecule to the leads

IX. CONCLUSION

STEPHAN: ALGUNAS MODIFICACIONES

In summary, we have reported on a extensive study of transport properties in DNA sequences of both biological and technological interest, and show that the transmission spectra are very sensitive to the nature and range of correlations. Huge differences are obtained for different genome samples, and sometimes albeit at the origin of similar biological functionality. If charge transport characteristics are anticipated to have some direct impact on mutagenesis, our first results thus opens new directions of works in relation with biological concerns.

We warmly thank E. C. Cox for sharing useful materials, and D. Bicout and E. Kats are acknowledged for interesting discussions. This work was partly supported by UCM through Project No. PR3/04-12450

REFERENCES

[1] J.D.H. Watson and F.H.C. Crick, Nature **171**, 737 (1953).

[2] C. Treadway, M.G. Hill and J.K. Barton, Chem. Phys. **281**, 409 (2002).

[3] C.-K. Peng, S.V. Buldyrev S. V., A.L. Goldberger, S. Havlin, F. Sciotino, M. Simons and H. E. Stanley, Nature **356**, 168 (1992). W. Li and K. Kaneko, Europhys. Lett. **17**, 655 (1992).

[4] R.F. Voss, Phys. Rev. Lett. **68**,3805 (1992). A. Arneodo, E. Bacry, P.V. Graves, J.F. Mury, Phys. Rev. Lett. **74**,3293 (1995). B. Audit et al., Phys. Rev. Lett. **86**, 2471 (2001).

[5] D. Holste, I. Grosse, H. Herzel, Phys. Rev. E **64**, 041917 (2001).

[6] N.C. Seeman, Acc. Chem. Res. **30**, 253 (1999). E. Braun, Y. Eichen, U. Sivan, and G. Ben-Yoseph, Nature **391**, 775 (1998). A. Pike, B. Horrocks, B. Connolly and A. Houlton, Aust. J. Chem. **55**, 191 (2002).

[7] D.B. Tippin, B. Ramakrishnan, M. Sundaralingam, J. Mol. Biol. **270**, 247 (1997). E. Braun, Y. Eichen, U. Silvan, G. Ben-Yoseph, Nature (London) **391**, 775 (1998).

[8] D.D. Eley and D.I. Spivey, Trans. Faraday Soc. **58**, 411 (1962).

[9] F. D. Lewis et al., Nature (London) **406**, 51 (2000).

[10] P. Carpena, P. Bernaola-Galvan, P. Ch. Ivanov, H.E. Stanley, Nature **418**, 955 (2002); ibidem Nature **421**, 764 (2003).

[11] S. Roche, D. Bicout, E. Maciá, E. I. Kats, Phys. Rev. Lett. **91**, 228101 (2003); ibidem Phys. Rev. Lett. **92**, 109901 (2004). M. Unge and S. Stafstrom, Nanolett. **3**, 1417 (2003) K. Hirayama, Int. J. Mod. Phys. B (to be published).

[12] H.W. Fink and C. Schönenberger, Nature (London) **398** 407 (1999). A.Y. Kasumov et al., Science **291**, 280 (2001). P. de Pablo et al., Phys. Rev. Lett. **85**, 4992 (2000). Y. Zhang, R. H. Austin, J. Kraeft, E. Cox, N. P. Ong, Phys. Rev. Lett. **89**, 198102 (2002).

A.J. Storm, J.V. Noort, S. de Vries, C. Dekker, Appl. Phys. Lett. **79**, 3881 (2002). B. Hartzell, B. McCord, D. Asare, H. Chen, J. J. Heremans, and V. Soghomonian, Appl. Phys. Lett. **82**, 4800 (2003). T. Heim, D. Deresmes, D. Vuillaume, *unpublished*

[13] D. Porath et al., Nature (London) **403**, 635 (2000). K.-H. Yoo et al., Phys. Rev. Lett. **87**, 198102 (2001). G. Cuniberti, L. Craco, D. Porath, C. Dekker, Phys. Rev. B **65**, 241314 (2002). R.G. Endres, D.L. Cox, R.R.P. Singh, Rev. Mod. Phys. **76**, 195 (2004).

[14] Z.G. Yu and X. Song, Phys. Rev. Lett. **86**, 6018 (2001)

[15] Y.A. Berlin, A.L. Burin and M.A. Ratner, Superlattices and Microstructures **28**, 241 (2000).

[16] S. Roche, Phys. Rev. Lett. **91**, 108101 (2003)

[17] H. Sugiyama and I. Saito, J. Am. Chem. Soc. **118**, 7063 (1996). H. Zhang et al., J. Chem. Phys. **117**, 4578 (2002). R.G. Endres, D.L. Cox, R.P. Singh, http://archiv.org/abs/cond-mat/0201404. Y. Yan and H. Zhang., J. Theor. Comput. Chem **1**, 225 (2002). F. Gervasio, P. Carloni, M. Parrinello, Phys. Rev. Lett. **89**, 108102 (2002). F. Gervasio, P. Carloni, M. Parrinello, *unpublished* R. Di Felice, A. Calzolari, E. Molinari, A. Garbesi, Phys. Rev. B **65**, 045104 (2002) Ch. Adessi, M.P. Anantram, Appl. Phys. Lett. **82**, 2353 (2003). Ch. Adessi, S. Walch, M.P. Anantram, Phys. Rev. B **67**, 081405 (2003) K. Iguchi, Intern. Journ. Mod. Phys. XXXXXX K. Iguchi, Intern. Journ. Mod. Phys. XXXXXX

[18] B.K. Nikolic, Phys. Rev. B, **64**, 165303 (2001).

[19] A. Stone, J. Joannopoulous, D. Chadi, Phys. Rev. B **24**, 5583 (1981). E. Maciá and F. Domínguez-Adame, *Electrons, Phonons and Excitons in low-dimensional aperiodic systems* (Colección Línea 300, Editorial Complutense, Madrid, 2000).

[20] J. L. Pichard, J. Phys.C: Solid State Phys. **19**, 1519 (1986). P. D. Kirkman and J. B. Pendry, J. Phys.C: Solid State Phys. **17**, 4327 (1986).

[21] M. Goda and H. Kubo, J. Phys. Soc. Jpn **58**, 2109 (1989). E. Maciá, F. Dominguez-Adame, A. Sanchez, Phys. Rev. B **49**, 9503 (1994). Bruno-Alfonso et al. Phys. Rev. B **59**, 2057 (1999). S. Roche and K. Moulopoulos, Phys. Rev. B **61**, 6048 (2000).

[22] J. Kollar and A. Süto, Phys. Lett. A **117**, 203 (1986). M. Baake, D. Joseph and P. Kramer, Phys. Lett. A **168**, 199 (1992).

[23] B. Hartzell et al. *to be published*

[24] B. Sutherland and M. Kohmoto, Phys. Rev. B **36**, 5877 (1987).

[25] F. Sanger et al., J. Mol. Biol. **162**, 729 (1982).

[26] Downloadable from the Sanger Center *http://www.sanger.ac.uk/HGP/Chr22/*

[27] TO BE COMPLETED

[28] S. R. Rajski, B.A. Jackson, J.K. Barton, Mutation Research **447**, 49 (2000). A. Okada, S. Yokojima, N. Kurita, Y. Sengoku, S. Tanaka, J. Mol. Struct. (Theochem) **630**, 283 (2003).

[29] R.A. Marcus, J. Chem. Phys. **24**, 966 (1956). R.A. Marcus and N. Sutin, Biochim. Biophys. Acta **811**, 265 (1985)

[30] A. Nitzan, J. Phys. Chem. A **105**, 2677 (2001).

[31] F.D. Lewis et al., J. Am. Chem. Soc. **125**, 4850 (2003).

[32] A. Nitzan, Israel Journal of Chemistry **42**, 163 (2002).

[33] V. Mugica, M. Kemp, A. Roitberg, M.A. Ratner, J. Chem. Phys. **104**, 7296 (1996).

[34] G.C. Liang, A.W. Ghosh, M. Paulsson, S. Datta, Phys. Rev. B **69**, 115302 (2004).

[35] R. Lake and S. Datta, Phys. Rev. B **45** , 6670 (1992). C. Strahberger, P. Vogl, Phys. Rev. B 62, 7289 (2000).

[36] N.P. Guisinger et al., Nano Lett. **4**, 55 (2003). S. Datta, S. Hong, R. Reifenberger, J.I. Henderson, C.P. Kubiak

[37] J. C. Poler, R. M. Zimmermann, and E. C. Cox, Langmuir **11**, 2689 (1995).

[38] M. Paulsson and S. Datta, Phys. Rev. B **67** 241403(R) (2003).

[39] E. Maciá, F. Trioçon, and S. Roche (unpublished).

[40] E. G. Emberly and G. Kirczenow, in *Molecular Electronics: Science and Technology,* edited by A. Aviram and M. Ratner, Ann. N. Y. Acad. Sci. **852** (New York Academy of Sciences, New York, 1998), p.54.

[41] E. G. Emberly and G. Kirczenow, Phys. Rev. B **58**, 10911 (1998).

[42] E. G. Emberly and G. Kirczenow, J. Phys. Condens. Matter **11**, 6911 (1999).

[43] T. Kostyrko, J. Phys. Condens. Matter **14**, 4393 (2002).

[44] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985).

[45] Y. J. Yan and H. Zhang, J. Theor. and Computational Chem.**1**, 225 (2002).

[46] H. Sugiyama and I. Saito, J. Am. Chem. Soc. **118**, 7063 (1996).

[47] L. Patrone, S. Palacin, J. Charlier, F. Armand, J. P. Bourgoin, H. Tang, and S. Gauthier, Phys. Rev. Lett. **91**, 096802 (2003).

[48] J. C. Charlier, X. Gonze, and J. P. Michenaud, Phys. Rev. B **43**, 4579 (1991).

[49] C. Kergueris, J. P. Bourgoin, S. Palacin, D. Esteve, C. Urbina, M. Magoga, and C. Joachim, Phys. Rev. B **59**, 12505 (1999).

[50] S. Reich, J. Maultzsch, C. Thomsen, and P. Ordejón, Phys. rev. B **66**, 035412 (2002).

FIGURES

FIG. 1. Energy dependent transmission coefficient for a finite length PolyGC approximant of quasiperiodic chain with increasing length. The spectrum is defined within $[\varepsilon_m - 2t_m = 5.75, \varepsilon_m + 2t_m = 9.75]$.

FIG. 2. Energy dependent Lyapunov coefficient for finite length PolyGC and PolyGCGGC (inset) approximants of quasiperiodic Fibonacci chain.

FIG. 3. Transmission coefficients for chromosome 22 based sequences with increasing numer of base pairs. *Right*-bottom : Energy dependent Lyapunov coefficient for two chromosome 22 base sequences

FIG. 4. Temperature dependent transmission coefficient for a 20nm chain extracted from Ch22 (see text).

FIG. 5. Energy dependent transmission coefficients for λ-bacteriophage-based sequences with increasing number of bp. *Right*-bottom : $T(E)$ at finite temperature $T = 20K$.

FIG. 6. Transmission coefficients for lambda bacteriophage based sequence with and without mismatches

FIG. 7. Transmission coefficients for D1s80-based sequences exctraced from the Pygmy chimpanzee genetic code.

FIG. 8. Transmission coefficients for D1s80-based sequences exctraced from the Human genetic code.

FIG. 9. Main frame: tunneling currents for several $N = 10$ sequences (PolyG, PolyAT and polyGC) with parameters $t_m = 1.0$ eV, $t = 1$ eV, and $\tau = 1$ eV and $\varepsilon_m = 5.36eV$. Inset: Tunneling current through a $N = 60$ PolyG sequence with $t_m = 1.0$ eV, $t = 1$ eV, and $\tau = 1$ eV and $\varepsilon_m = \varepsilon_G$.

FIG. 10. Room temperature dependence of the Seebeck coefficient (main frame) and the Landauer conductance (inset) as a function of the Fermi level energy for G (solid lines), A (dashed line), C (dotted line) and T (dot-dashed line) bases with $\tau = 0.5$ eV, $t_m = 1.7$ eV, and $\varepsilon_m = -3.4$ eV.

FIG. 11. Room temperature dependence of the thermoelectric power factor as a function of the Fermi level energy in a graphite-G- graphite junction for two different values of the molecule-lead coupling strength.