PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Optical Properties and Electronic Transitions of DNA Oligonucleotides as a Function of Composition and Stacking Sequence

Jacob B. Schimelman[1#], Daniel M. Dryden[2#], Lokendra Poudel[3], Katherine E. Krawiec[1], Yingfang Ma[2], Rudolf Podgornik[4,5,6], V. Adrian Parsegian[6], Linda K. Denoyer[7], Wai-Yim Ching[3], Roger H. French[2,9,10*], Nicole F. Steinmetz,[1,2,8,9,*]

1. Department of Biomedical Engineering, Case Western Reserve University, Schools of Medicine and Engineering, Cleveland, Ohio 44106, USA,

2. Department of Materials Science and Engineering, Case Western Reserve University, 10900 Euclid Ave., Cleveland, Ohio 44106,

3. Department of Physics and Astronomy, University of Missouri-Kansas City, Kansas City, MO, 64110, USA,

4. Department of Theoretical Physics, J. Stefan Institute, SI-1000 Ljubljana, Slovenia.

5. Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana, SI-1000 Ljubljana, Slovenia,

6. Department of Physics, University of Massachusetts, Amherst, Massachusetts 01003, USA.

7. Deconvolution and Entropy Consulting, 755 Snyder Hill, Ithaca, New York 14850, USA

8. Department of Radiology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA.

9. Department of Macromolecular Science and Engineering, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA.

10. Department of Physics, Case Western Reserve University, 10900 Euclid Ave., Cleveland, Ohio 44106

# both authors contributed equally
*corresponding authors: roger.french@case.edu and nicole.steinmetz@case.edu

Abstract

The role of base pair composition and stacking sequence in the optical properties and electronic transitions of DNA is of fundamental interest. We present and compare the optical properties of DNA oligonucleotides (AT)10, (AT)5(GC)5, and (AT-GC)5 using both *ab initio* methods and UV-Vis molar absorbance measurements. Our data indicate a strong dependence of both the position and intensity of UV absorbance features on oligonucleotide composition and stacking sequence. The partial densities of states for each oligonucleotide indicate that the valence band edge arises from a feature associated with the PO_4^{3-} complex anion, and the conduction band edge arises from anti-bonding states in DNA base pairs. The results show a strong correspondence between the *ab initio* and experimentally determined optical properties. These results highlight the benefit of full spectral analysis of DNA, as opposed to reductive methods that consider only the 260 nm absorbance (A260) or simple purity ratios, such as A_{260}/A_{230} or A_{260}/A_{280}, and suggest that the slope of the absorption edge onset may provide a useful metric for the degree of base pair stacking in DNA. These insights may prove useful for applications in biology, bioelectronics, and mesoscale self-assembly.

Introduction

DNA, the fundamental building block of life, has long been the focus of research where biology, chemistry and physics come together to explain both its structure and function.[1] Bridging the traditional sciences and engineering disciplines, DNA has become a versatile tool for materials science and applications. For example, DNA is used as a structural building block for the controlled self-assembly and spatial positioning of colloidal and metallic nanoparticles, yielding novel materials with unique applications in sensing, plasmonics, and photonics.[2–8] The study of optical and electronic properties of DNA has opened new paradigms in electron and charge transport.[9] While DNA has been considered as an example of a molecular wire, its capability to mediate charge transport via coherent tunneling and diffusive thermal hopping[10] depends crucially on the details of the experimental assay, sometimes yielding contradictory results.[11–13]

Understanding the optical and electronic properties of DNA on a fundamental level provides the opportunity to advance materials design and function.[14] The combination of experimental and computational research in particular opens up the door toward realization of the rational materials design framework, which is guided by theory and informed by experiment. As a first step toward this paradigm, we set out to study the *ab initio* and experimental optical properties of DNA in the context of sequence and stacking order. The primary goal of this work is to determine, identify, and map the spectral features, *viz.* the electronic structure and optical dispersion properties of DNA oligonucleotides in the ultraviolet and visible (UV-Vis) energy range. We consider the spectral imprint of the differences in the base pair composition and stacking (Figure 1) as well as the contributions from the sugar-phosphate backbone. The latter is important because our previous work implicated a role for phosphate complex anions in the electronic structure and spectral properties of organic and inorganic systems.[15,16] Furthermore, in prior investigations using co-polymers with which DNA shares certain structural similarities, we found that the energy

of UV absorption peaks is dependent not only on the average composition but more specifically on the exact polymer block sequence.[17,18] A better understanding of how the structural composition and stacking sequence of DNA affects its electronic structure and subsequent intermolecular interactions would enhance the available methodologies for the rational design of molecular self-assembly in nano- and mesoscale systems.[19–21]

Previous studies establish the optical and near-UV spectral properties as a reliable way to characterize materials.[22,23] Accurate spectroscopic data are available for single nucleotides, nucleosides and derivatives, synthetic polynucleotides (polyuridylic acid (poly-U), polyadenylic acid (poly-A), poly-AU), and various nucleic acids, such as RNA and native bacterial DNAs in aqueous solutions[22] as well as for wet and dry M-DNA.[24] Dry DNA spectroscopic data are available for polymerized oligonucleotides and mononucleotides[25,26] and dry DNA thin films.[27] Despite this seeming abundance of experimental data, there have been few detailed experimental investigations of the DNA or oligonucleotide base-pair stacking order effect on its UV spectral properties. Using a combined *ab initio*-experimental methodology, in this work we address the possibility and extent to which DNA absorption spectra may be deconvoluted into specific features pertaining to the sequence and stacking order of base pairs. Specifically, we correlate the *ab initio* partial densities of states (PDOS), obtained by quantum mechanical calculations using the orthogonalized linear combination of atomic orbitals (OLCAO) method[28] and partitioned according to structural subgroup, with the optical features obtained by UV-Vis decadic molar absorbance measurements. We consider three distinct duplex DNA oligonucleotides: (AT)10, (AT)5(GC)5, and (AT-GC)5 (Figure 1). Apart from absorption spectra of single DNA base molecules adenine, thymine, guanine, and cytosine, calculated from gradient-corrected density-functional theory,[29] to our best knowledge no calculations of *ab initio* optical properties on realistic DNA models have been attempted. By comparing these three duplexes, we seek to address the effects of composition and stacking sequence in order to elucidate the structure-function relationship of these important biomolecular systems.

Methods

The structure of the three DNA duplex structures is shown schematically in Figure 1, colored by structural group and by element.

Ab Initio Quantum Mechanical Calculations of DNA Oligonucleotides

We used the orthogonalized linear combination of atomic orbitals (OLCAO) variant of the density functional theory (DFT) method.[28] The model construction, relaxation, and electronic structure calculation are reported in detail elsewhere.[30] Briefly, the models were constructed starting with the Amber force field[31] and then fully relaxed to high accuracy and converged total energy using the Vienna *Ab initio* Simulation package (VASP).[32,33] The final structures are periodic in the axial (z-) direction, with 10 base pairs composed of 660, 650 and 655 atoms in (AT)10, (AT)5(GC)5 and (AT-GC)5 models, respectively, and have 2220 valence electrons each. The relaxed structures (Figure 1) are then used to calculate the electronic structure and optical properties using the OLCAO method. The PDOS for each structural group in (AT)10, along with their peak identifications, are given in Figure 2a. The PDOS for each structural group in (AT)5(GC)5 and (AT-GC)5, along with their peak identifications, are given in Figures 2b and 2c. The DOS results are finally broadened by a Gaussian of 0.1 eV, the standard implementation of the OLCAO method.[28] The ability to resolve the DOS into PDOS of each group is very important in the interpretation of the measured absorption spectra.

During the past decades, there have been tremendous developments toward more accurate calculation of optical excitations in crystals and molecules that go beyond the simple DFT calculations appropriate for ground state properties. These methods include many body perturbation method (MBPM)[34–37] or the GW method followed by the solution of Bethe-Salpeter equation (BSE), time-dependent density functional theory (TDDFT),[38–43] linear-scaling DFT theory,[44,45] configuration interaction-singles method (CIS)[46] and their various approximate versions that have already been implemented in some of the

operational packages.

As these advanced methods are related among themselves, they also share a commonality of drawbacks stemming from the fact that they all start with plane-wave based DFT methods. This results in undesirable consequences despite many concerted efforts: they apply mostly to small molecules or crystals (less than 100 atoms at most), and they apply only to a limited frequency range of optical absorption (mostly up to 10-15 eV).

These drawbacks effectively limit their role in the verification of well documented smaller systems instead of enabling exploratory studies, especially those of highly complex biomolecular systems where the structure, the environment, and the nature of many body effects on the excited state are much less clear.

In comparison, it has been well documented that OLCAO-DFT calculations can be effectively applied to some very large systems with results, in most cases, that are in good agreement with experiment. For example, OLCAO-DFT calculations correctly predict the location and the strong optical anisotropy in herapathite crystal $(C_{20}H_{24}N_2O_2H_2)_4{\cdot}C_2H_4O_2{\cdot}3SO_4{\cdot}2I_3{\cdot}6H_2O)$,[47] a historically important optical material discovered in 1852, whose structure was only resolved recently. It consists of 988 atoms of six different species in five distinct structural groups.[48] Another example is the calculation of core-level excitations in a model of super-cooled water with 340 water molecules (1020 atoms), revealing the network structures when compared with experimental XANES data.[49] A most recent example is the prediction of the optical dielectric function in the RGD molecule, with spectacular sharp absorption peaks at 240 nm and 204 nm arising from optical transitions tied to specific amino acids.[50] In light of these previous successes, we have full confidence that the application of the same OLCAO-DFT optical calculation to the large DNA models can yield meaningful and useful insights. The ability of the OLCAO-DFT method to do such massive calculations, covering a much larger frequency range of up to 30 eV and relying on relatively modest computational resources, can be traced to the use of the localized atomic basis expansion and efficiency in the evaluation of multicenter integrals

with Gaussian type orbitals (GTO).[28]

Using the OLCAO method, we have calculated the imaginary part of the complex dielectric response function $\varepsilon''(\hbar\omega)$ within the random phase approximation (RPA) of the one-electron theory according to the following equation:[28]

$$\varepsilon''(\hbar\omega) = \frac{e^2}{\pi m \omega^2} \sum_{n,l} \left| \langle \psi_n(\vec{r}) | -i\hbar\vec{\nabla} | \psi_l(\vec{r}) \rangle \right|^2 \delta(E_n - E_l - \hbar\omega) \qquad (1)$$

where $\left| \langle \psi_n(\vec{r}) | -i\hbar\vec{\nabla} | \psi_l(\vec{r}) \rangle \right|$ is the dipolar matrix element for electronic transitions from the occupied molecular orbital (valence band) state ψ_l to the unoccupied molecular orbital (conduction band) state ψ_n with $\hbar\omega$ being the transition energy. The $\varepsilon''(\hbar\omega)$ spectra are then scaled according to the ratio of the calculation supercell volume to the oligonucleotide solvent-excluded volume.[51]

The absorption coefficient α is a function of both ε'' and ε'. Since ε' may be recovered from ε'' via Kramers-Kronig transform,[52] α is then given by

$$\alpha = \frac{E\sqrt{2}}{\hbar c} \left[(\varepsilon'(E)^2 + \varepsilon''(E)^2)^{\frac{1}{2}} - \varepsilon'(E) \right]^{\frac{1}{2}} \qquad (2)$$

These absorption coefficients are equivalent to those used in the Beer-Lambert law, discussed below. Figure 3 shows the *ab initio* absorption coefficients; Table 1 shows the main features that were indexed in these spectra as well as the transition assignments from the PDOS.

Melting Curve of DNA Duplexes

To confirm the thermal stability of DNA duplexes at room temperature (at which molar absorbance is measured), we calculated the melting curves of the DNA duplexes using uMELT melting curve prediction software.[53] Since the stability of DNA duplexes depends on the length of the molecule, the thermal stability of both 10 base-pair oligonucleotides and 40 base-pair oligonucleotides are studied. The calculated melting

curves are illustrated in Figure 4.

Molar Absorbance of DNA

We obtained all DNA samples from IDT Technologies (Coralville, IA). (AT)10, (AT)40, (AT)5(GC)5, ((AT)5(GC)5)4, (AT-GC)5, and (AT-GC)20 were considered. Based on the thermal stability of the various complexes, only (AT)40, (AT)5(GC)5, and (AT-GC)20 are discussed in detail; these sequences form stable duplex structure at room temperature. Samples were stored at stock concentrations of 100 μM in sodium-phosphate buffer at pH 7.4. Concentration gradients were prepared in sodium-phosphate buffer at pH 7.4 to yield 50 μM, 40 μM, 25 μM, 20 μM, 12.8 μM, 6.4 μM, and 3.2 μM samples of each DNA oligonucleotide. The molarity range was determined in accordance with the absorbance limitations of the spectrophotometer used, which is most sensitive for absorbance values within the range 0.1 - 1.0. UV-Vis absorbance measurements were taken using the NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific, Waltham, MA) at room temperature. All measurements were performed in triplicate. The spectrometer was baselined between each measurement to reduce correlated errors. The molar absorption coefficient spectrum for each duplex DNA was obtained using the absorption spectra taken for seven concentrations of each duplex DNA, processed using GRAMS spectral software from Thermo Scientific (Waltham, MA) and Electronic Structure Tools from Deconvolution & Entropy Consulting (Ithaca, NY). Absorbance values between 0 and 0.8 were considered; data points, which deviated significantly from linearity due to detector saturation were automatically excluded. Absorbance spectra were fitted linearly with respect to concentration according to Beer's Law, $\alpha = c\ l$, using maximum likelihood spectral estimation.[54,55] Errors were assumed to be heteroscedastic, *i.e.* the error in the absorbance spectrum at each concentration c_i is dependent on the concentration. The molar absorption was fitted via weighted least-squares fitting, minimizing the function

$$F \equiv \sum_i \frac{A_i(E) - \alpha_i(E) d_i(E)}{w_i(E)} \quad (3)$$

where $A_i(E)$ is the measured absorbance spectrum at the concentration c_i, d_i is the product of the concentration c_i and path length of the spectrometer (1 mm in this case), α_i is the decadic molar absorbance from a single absorbance spectrum at concentration c_i, and the weight $w_i(E)$ is proportional to the error in each observation $A_i(E)$. The heteroscedastic standard error, displayed as error bars, is given by

$$SE_{\alpha}(E) = \sum_j \frac{\left(\frac{\sigma_j(E)}{w_j(E)^2} \left(\sum \frac{d_i}{w_i(E)^2} \right) - \frac{\sigma_j(E) d_j}{w_j(E)^2} \left(\sum \frac{1}{w_i(E)^2} \right) \right)^2}{\left(\sum \frac{d_i^2}{w_i(E)^2} \sum \frac{1}{w_i(E)^2} - \sum \frac{d_i}{w_i(E)^2} \sum \frac{d_i}{w_i(E)^2} \right)} \quad (4)$$

where $\sigma_j(E)$ is the difference at each energy E between each input spectrum $A_j(E)$ and the model constructed using the best-fit algorithm. The molar absorption coefficient spectra are shown in Figure 5.

Results

The PDOS for (AT)10 is shown in Figure 2a. Significant features are indexed with a capital letter indicating an occupied (valence band) state, and a lower-case letter indicating an unoccupied (conduction band) state. Similarly annotated PDOS for (AT)5(GC)5 and (AT-GC)5 are shown in Figures 2b and 2c, respectively. In all three DNA duplex structures the highest occupied molecular orbital (HOMO) is defined by P1, a feature associated with the phosphate group in the DNA backbone. In (AT)10 the lowest unoccupied molecular orbital (LUMO) is defined by a combination of a1 and t1, features arising from the adenine and thymine base pairs; in (AT)5(GC)5 the valence band edge is defined by features arising from cytosine and thymine, and in (AT-GC)5, from thymine and guanine. The Na+ counterion does not contribute to available states until the PDOS feature at 5.35 eV.

The *ab initio* optical properties of all three oligonucleotides, given as absorption coefficient α, are shown in Figure 3. Characteristic features are designated in three regions

I, II, and III for clarity of discussion. The features that correspond to the characteristic 230 nm, 260 nm, and 280 nm features seen experimentally in DNA UV-Vis absorbance are also marked. The absorbance at these three single energies, rather than whole spectra, are often used to estimate DNA concentration and sample purity.[1,56] While DFT methods are very effective for determining the relative energies of features within individual bands, they characteristically underestimate the band gap significantly.[28] Therefore characteristic features in the *ab initio* absorbance spectra are shifted downwards in energy by approximately 1 eV from their equivalent features in the experimental spectra. The calculated *ab initio* HOMO-LUMO gaps for (AT)10, (AT)5(GC)5 and (AT-GC)5 are shown in Table 1. Direct gap energies were determined for the experimental results by fitting the steepest portion of the fundamental absorption edge on plots of α^2E^2;[15] these experimentally determined gap energies and their respective fitting ranges are also shown in Table 1.

The sequence of gap values between the calculation and measurement is the same, or E_g ((AT)10) > E_g ((AT)5(GC)5) > E_g ((AT-GC)5), which is significant given the aforementioned limitation of the DFT calculation regarding the estimation of the band gap and the complexity of the invoked DNA models.

The *ab initio* transition energies of characteristic features and valence band-to-conduction band transition assignments are given in Table 2. For instance, the P2 $\rightarrow$ t1 transition is an electronic transition from the occupied P1 feature to the unoccupied t1 feature. Individual features are indexed for each spectrum, *e.g.*, feature $e_1^{(AT)10}$ is the first prominent transition feature in the *ab initio* absorbance spectrum for (AT)10.

The melting curves of DNA duplexes are illustrated in Figure 4. At room temperature, both 10 base-pair and 40 base-pair oligonucleotides (AT)5(GC)5 retain helicity; however, 10 base-pair (AT-GC)5 and (AT)10 duplex structures are thermally less stable, with significant decrease of duplex structure. Therefore thermally stable (AT-GC)20 and (AT)40 duplexes were used for the study of molar absorbance.

The experimentally determined molar absorption coefficients for (AT)40, (AT)5(GC)5 and (AT-GC)20 DNA duplexes are shown in Figure 5. Regions I, II, and II are marked according to ranges of features corresponding to those which are prominent in Figure 3.

Discussion

(AT)n

(AT)40, (AT)5(GC)5 and (AT-GC)20 were considered for the study of molar absorbance; duplex structures considered for the *ab initio* calculations are infinitely long structural models. In the following, we will refer to the sequences as (AT)n, (AT)n(GC)n and (AT-GC)n; the length of the duplex only affects the thermal stability not the optical features.

Of the three duplex DNA structures studied, (AT)n has the simplest composition, and correspondingly its optical properties are the simplest to correlate with its constituent structural groups. The first feature in Region I (Figure 3) consists of a shoulder, $e_1^{(AT)n}$, and a peak, $e_2^{(AT)n}$ (Table 2). This feature defines the fundamental absorption edge in (AT)n. Indexing according to the PDOS makes it clear that the shoulder arises from transitions from occupied states P2 and T1 to unoccupied states t1 and a2 (Figure 2a). This first transition feature is therefore dependent on the electronic structure not only of the DNA bases, but of the phosphates in the backbone as well. The absorbance maximum is a result of the P1 $\rightarrow$ a2, P1 $\rightarrow$ t2, and A1 $\rightarrow$ a1 transitions. These two features together define the "260 nm peak", the characteristic absorption feature seen at 4.79 eV in the experimental absorption spectrum (Figure 5). The shift of this *ab initio* feature downwards in energy by 1 eV with respect to its experimentally observed counterpart is expected due to the tendency of DFT calculations to underestimate the band gap.[28]

Region II of the *ab initio* spectrum is defined by a characteristic minimum $m_1^{(AT)n}$, which is in agreement with the relative lack of intensity in the PDOS at that energy range (Figure 2). There exists a corresponding trough at 5.17 eV, or 230 nm, in the experimental optical properties (Figure 5) that is well known in the literature,[1] and a similar one for the

260 nm peak, with the corresponding *ab initio* feature shifted downward by approximately 1 eV. While the *ab initio* spectrum shows a distinct feature, $e_3^{(AT)n}$, within region II, no such well-defined feature is apparent in the experimental spectrum. A great deal of broadening is seen in the experimental spectrum in comparison to the *ab initio* spectrum, resulting from solvent-solute interactions and vibrational fine structure.[57] Since there is still a great deal of absorption within this trough, and the ratio of intensities A_{260}/A_{230} is smaller in the experimental than in the *ab initio* spectrum, it is reasonable that the transitions that constitute this feature are still present but experimental broadening effects have obscured the local maximum, resulting instead in the broad shoulder seen on the experimental absorption feature around 5.41 eV.

In Region III (4.6 eV and above, Figure 3), seemingly well-defined absorption features are the result of an amalgam of constituent electronic transitions originating from a range of bands. The *ab initio* spectrum shows the distinct features $e_4^{(AT)n}$ and $e_5^{(AT)n}$, which may be attributed to S3 $\rightarrow$ a1, P5 $\rightarrow$ a1, P1 $\rightarrow$ t3, A2 $\rightarrow$ t1, A2 $\rightarrow$ a2, T1 $\rightarrow$ t3, and P2 $\rightarrow$ t3 transitions at or near the observed maximum (Figures 2 and 3). At higher energies, it becomes prohibitively unwieldy to ascribe absorption features to specific states, with all observed features arising from the continuum of dozens of transitions across a range of energies. These transitions result in the experimental broadened absorption feature seen at 5.4 eV and above (Figure 5), which exhibits broadening and secondary spectral features such as small shoulders and slope changes between 5.5 and 6 eV. Overall the spectrum of (AT)n shows a complexity of features in Region II and Region III as a result of a great many more possible electronic transitions.

(AT)n(GC)n

It is experimentally challenging to obtain duplex DNA structures consisting of more than five consecutive GC base pairs, as the consecutive guanines have a strong propensity to form stable g-quadruplex conformations instead of duplex DNA.[58] We reasoned that (AT)n(GC)n DNA duplex structures would provide a reasonable basis for comparison, as the

shorter length of consecutive guanine bases is expected to preserve the duplex DNA structure, and the resulting optical properties are expected to be a superposition of the (AT)n features and those resulting from the additional GC base pair segment. While the results are complex due the number of PDOS features and transitions that contribute to the features, certain observations can nevertheless be made: within region I the experimental "260 nm" feature (Figure 5) resolves into two well-defined maxima, $e_3^{(AT)n(GC)n}$ and $e_4^{(AT)n(GC)n}$ (Figure 3). While the intuitive temptation is to assign one peak to the GC segment and one to the AT, each proves to be a result of a variety of transitions, and neither specific base pair segment contributes solely to either feature. This overall broadening is replicated in the experimental results. Furthermore, new low-energy shoulders $e_1^{(AT)n(GC)n}$ and $e_2^{(AT)n(GC)n}$may be observed, which are not observed in the (AT)n spectra (Figure 5), and from the PDOS appear to be resultant solely from the GC contribution (Figure 2). While these features are not separate and distinguishable in the experimental spectrum, they likely contribute to the broadening and resulting reduction in the fundamental absorption edge energy seen when compared to (AT)n.

In Region II a similar trough corresponding to the "230 nm" minimum in the experimental spectrum is observed in (AT)n(GC)n as in (AT)n, with a slight frequency shift. The features in Region II result from a wide range of transitions, which, between the occupied and unoccupied states involved, span all structural groups except the Na^+ counter ion. Both the *ab initio* and experimental spectra consequently exhibit a complex absorption spectrum in Region II, with broadening effects attenuating the extremity of the 230 nm trough. A small shoulder at 5.41 eV is observed. The most notable difference between (AT)n(GC)n and (AT)n in Region II is the decrease in the A_{260}/A_{230} ratio and band gap energy for (AT)n(GC)n, compared to (AT)n, owing to the contributions of GC bases. DNA with GC pairs has been known to exhibit a lower band gap than DNA consisting solely of AT bases.[59]

Similar to (AT)n, the absorption features of (AT)n(GC)n in Region III are the result

of a continuum of many overlapping transitions. While the absolute magnitude of these absorption features shows increasing uncertainty with increasing energy, spectral features such as changes in slope in an otherwise smooth spectrum are less sensitive to experimental error, and well defined spectral features may be observed at energies up to 6 eV before the spectrometer loses sensitivity.

It is noticed that the experimental molar absorptivity of (AT)5(GC)5, consisting of 10 base-pairs, shows a magnitude comparable to the molar absorptivity of (AT)40, comprised of 40 base-pairs.

(AT-GC)n

The study of DNA duplex structures consisting of segments of consecutive like base pairs as seen in (AT)n and (AT)n(GC)n minimizes the effect of interactions between differing base pairs, and emphasizes the contributions of composition alone. Also of interest, however, is the effect of stacking sequence on the electronic structure and optical properties of DNA. By comparing an alternating base-pair segment of (AT-GC)n with the block oligonucleotide (AT)n(GC)n, which shares its net composition exactly, we hypothesize that effects of stacking sequence and the subsequent effects that arise from the interactions of AT with GC base pairs could be isolated.

The same small shoulders are seen in Region I in the *ab initio* spectrum (Figure 3) at the onset of absorption as were observed in (AT)n(GC)n, albeit shifted slightly higher in energy. The feature corresponding to the "260 nm" experimental absorbance feature is now defined by a shoulder $e_3^{(AT\text{-}GC)n}$ and peak $e_4^{(AT\text{-}GC)n}$. Many of the same experimental features are observed in both the (AT)n(GC)n and (AT-GC)n experimental spectra (Figure 5) in Regions II and III, including the same position of 230 nm trough, the small shoulder at 4.9 eV, and the previously noted shoulder at 5.41 eV.

With regard to the experimental spectrum, since the (AT-GC)20 duplex has 40 base-pairs (compared with (AT)5(GC)5 which has 10 base-pairs), the experimental molar absorbance of (AT-GC)20 has a substantially higher magnitude than that of (AT)5(GC)5.

Nevertheless, the major feature difference between the experimental absorption of (AT)5(GC)5 and (AT-GC)20 is the band gap energy, which could be caused by base-base interactions.

The overall shift of the low-energy absorbance intensity to higher energies is consistent in both the *ab initio* and experimental data (Figures 3 and 5), though the direct band gap of each is nearly identical, owing to the lower slope of the absorbance edge in (AT-GC)n. Two competing effects may explain this observation. DNA with GC bases has been known to exhibit a lower band gap than DNA consisting solely of AT bases,[59] as seen in our *ab initio* and experimental spectra. Conjugation of electronic states—delocalization of the electron wave function in real space—is aided by runs of similar material, which results in a shift of transitions to lower energies.[18,59] Conversely, alternation of bases shifts the energies of these onset absorptions to slightly higher energies. This effect is evident in the onset absorptions $e_1^{(AT\text{-}GC)n}$ and $e_2^{(AT\text{-}GC)n}$ compared to $e_1^{(AT)n(GC)n}$ and $e_2^{(AT)n(GC)n}$ in the *ab initio* spectra. Absorption intensity is reduced and shifted higher in energy in the alternating (AT-GC)n DNA when compared to the (AT)n(GC)n DNA, where conjugation between neighboring like bases may be present. This is further corroborated by the experimental spectrum, wherein absorption intensity at low energies (below 4.5 eV) is shifted to higher energies in (AT-GC)n in comparison to (AT)n(GC)n.

Comparison of Experimental Spectra

The broadening of the primary absorption feature in region I of the experimental spectrum trends across all three DNA oligonucleotides with increasing complexity. The (AT)n has the lowest complexity, and subsequently experiences the least amount of inhomogeneous broadening from the superposition of overlapping electronic transitions. The (AT)n(GC)n duplex shows a slightly broader spectrum due to the GC section of the molecule contributing more overlapping transitions, especially at lower energies, indicated most visibly by the lower slope of the onset edge and the reduction of the A_{260}/A_{230} ratio. These comparisons suggest that the slope of the absorption onset, in conjunction with the

A_{260}/A_{230} ratio, may prove to be useful metrics in predicting the composition of DNA.

Comparison with Previous Literature

Previous studies of the electronic structure of duplex DNA list a variety of interpretations for the valence band and conduction band edge assignments from the PDOS with conflicting conclusions. Most commonly the valence band edge is assigned to guanine where present[59–65] or adenine where guanine is not present.[59,66] However, the HOMO state is highly sensitive to the counterions in the proximity of the PO_4^{3-} group. In the present OLCAO calculation, we found the HOMO to originate from the (PO_4^{3-}--Na^+) complex. Conduction band edge states are often assigned to either cytosine[61,63,65] or an Na^+/PO_4^{3-} antibonding state,[62,64] whereas in our calculation, all lower CB states are anti-bonding states from base pairs. While contributions to the band edges from all these states are present in our PDOS, the strong contribution of the PO_4^{3-} group to the valence band edge is a new observation. Because of the strong charge localization in PO_4^{3-} in other network-forming phosphates like $AlPO_4$[15] or apatites[28] where the charge compensations occur with different types of positively charged ions, this influence of the phosphate group on valence band structure and subsequently on the observed band gap and optical absorbance follows reasonably from previous results.

We note also that thus far the only related existing *ab initio* optical property calculations we are aware of are related specifically to isolated, single nucleotides.[26,29] Such calculations do not account even approximately for the intermolecular base-pair interactions and cannot be compared with experimental data nor our DNA calculations.

DFT calculations have the potential also to provide insight into the spatial extent of the electron wave functions within the oligonucleotide, elucidating both the degree of wave function delocalization that results from conjugation between like base pairs, and the extent of interaction between dissimilar base pairs at the interface between the two segments and beyond. It is also possible to divide the PDOS by a different set of constituent features, such as by atomic species or individual atoms. Different ways of resolving the PDOS might

provide other insights into the origin of spectral features. The optical properties presented here (Figures 3 and 5) take into account the dipole matrix transition elements (Equation 1). Transition assignments were based on *ab initio* electron states with the help of PDOS for interpretation. Omitted from the scope of this study is the effect on the optical absorption that occurs when the purines or pyrimidines sit exclusively on separate strands; that is, whether A and G (or T and C) share the same strand, as they do in our model, or occupy both strands. This comparison would help elucidate the spatial effects of orbital overlap, such as conjugation of states between structurally similar bases, which affect the energies of available states and thus observed electronic transitions.

Experimental analyses of DNA optical properties commonly consider only A260 or purity ratios such as A_{260}/A_{230} and A_{260}/A_{280}. These have already been shown to neglect influential parameters such as pH and ionic concentration of solvent.[1,56] These simple metrics furthermore neglect the movement of spectral features that result from changes in the composition or stacking sequence, potentially leading to error or misinterpretation. These observations highlight the importance of considering whole spectra when analyzing DNA, but also suggest the possibility of new metrics, such as the slope of the 260 nm absorption peak or the A_{260}/A_{230} ratio, as possible indicators of DNA structural complexity and degree of base pair alternation.

Conclusions

We have studied the effects of composition and stacking sequence on the UV-Vis optical properties of three DNA oligonucleotides via decadic molar absorption spectroscopy and *ab initio* calculations, noting the significant correlations between the two. Heretofore no optical calculations on large DNA double helix models have ever been attempted and in this respect, despite the approximations involved, our work presents a significant advance.

These new results highlight the benefits of full spectral analysis for DNA, which accounts for the movement of spectral features that result from changes in the composition or stacking sequence.

Acknowledgements

This work is supported by U.S. DOE, BES-DMSE-BMM under Award DE-SC0008064 and DE-SC008176. This research used the resources of the National Energy Research Scientific Computing Center supported by DOE under Contract No. DE-AC03-76SF00098.

Author contributions

J.B.S. and K.E.K. performed experimental optical measurements. J.B.S., Y.M., D.M.D, R.H.F., and N.F.S. analyzed the data. L.P. and W.Y.C. performed *ab initio* quantum mechanical calculations. K.E.K. performed melting curve calculations. L.K.D. contributed new code for the calculation of decadic molar absorption coefficients. R.P. and V.A.P. assisted in data interpretation and comparison to literature. All authors contributed to the discussion of the paper, and all authors edited the manuscript.

Bibliographic References

1. Bloomfield, V. A., Crothers, D. M. & Tinoco Jr., I. *Nucleic Acids: Structures, Properties, and Functions*. 2000, (University Science Books).
2. Chen, Q. *et al.* 3D Motion of DNA-Au Nanoconjugates in Graphene Liquid Cell Electron Microscopy. *Nano Lett.* 2013, 13, 4556–4561.
3. Liao, X. *et al.* Desktop nanofabrication with massively multiplexed beam pen lithography. *Nat. Commun*., 2013, 4, 1–7.
4. Liu, N., Hentschel, M., Weiss, T., Alivisatos, A. P. & Giessen, H. Three-Dimensional Plasmon Rulers. *Science,* 2011 332, 1407–1410.
5. Omabegho, T., Sha, R. & Seeman, N. C. A Bipedal DNA Brownian Motor with Coordinated Legs. *Science,* 2009, 324, 67–71.
6. Young, K. L. *et al.* Using DNA to Design Plasmonic Metamaterials with Tunable Optical Properties. *Adv. Mater.*, 2014, 26, 653–659.
7. Zhang, C. *et al.* A general approach to DNA-programmable atom equivalents. *Nat. Mater.,* 2013, 12, 741–746.
8. Zheng, J. *et al.* From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature*, 2009, 461, 74–77.
9. Boon, E. M. & Barton, J. K. Charge transport in DNA. *Curr. Opin. Struct. Biol.*, 2002, 12, 320–329.
10. Sontz, P. A., Muren, N. B. & Barton, J. K. DNA Charge Transport for Sensing and Signaling. *Acc. Chem. Res.,* 2012, 45, 1792–1800.
11. Gutierrez, R. & Cuniberti, G. Modeling Charge Transport and Dynamics in Biomolecular Systems. *J. Self-Assem. Mol. Electron.,* 2013, 1, 01.
12. De Pablo, P. J. *et al.* Absence of dc-Conductivity in λ-DNA. *Phys. Rev. Lett.,* 2000, 85, 4992.
13. Kutnjak, Z., Filipic, C., Podgornik, R., Nordenskiöld, L. & Korolev, N. Electrical Conduction in Native Deoxyribonucleic Acid: Hole Hopping Transfer Mechanism? *Phys. Rev. Lett.*, *2003,* 90, 098101.
14. Cuniberti, G., Fagas, G. & Richter, K. *Introducing molecular electronics*. (Springer, 2005).
15. Dryden, D. M., Tan, G. L. & French, R. H. Optical Properties and van der Waals-London Dispersion Interactions in Berlinite Aluminum Phosphate from Vacuum Ultraviolet Spectroscopy. *J. Am. Ceram. Soc.,* 2014, 97, 1143–1150.
16. Rulis, P., Ouyang, L. & Ching, W. Electronic structure and bonding in calcium apatite crystals: Hydroxyapatite, fluorapatite, chlorapatite, and bromapatite. *Phys. Rev. B,* 2004, 70, 155104.
17. French, R. H., Meth, J. S., Thorne, J. R. G., Hochstrasser, R. M. & Miller, R. D. Vacuum ultraviolet spectroscopy of the optical properties and electronic structure of seven poly(di-alkylsilanes). *Synth. Met.,* 1992, 50, 499–508.
18. French, R. *et al.* Novel hydrofluorocarbon polymers for use as pellicles in 157 nm semiconductor photolithography: fundamentals of transparency. *J. Fluor. Chem.,* 2003, 122, 63–80.
19. Hemminger, J. *From Quanta to the Continuum: Opportunities for Mesoscale Science*. (U.S Department of Energy Basic Energy Sciences Advisory Committee, 2012).
20. Taniguchi, M. & Kawai, T. DNA electronics. *Phys. E Low-Dimens. Syst. Nanostructures,* 2006, 33, 1–12.
21. Rajter, R. F., French, R. H., Ching, W. Y., Podgornik, R. & Parsegian, V. A. Chirality-dependent properties of carbon nanotubes: electronic structure, optical dispersion properties, Hamaker coefficients and van der Waals-London dispersion interactions. *RSC Adv.* 2012, 823–842. doi:10.1039/C2RA20083J
22. Voet, D., Gratzer, W. B., Cox, R. A. & Doty, P. Absorption spectra of nucleotides, polynucleotides, and nucleic acids in the far ultraviolet. *Biopolymers* 1963, 1, 193–208.

23. Callis, P. R. Electronic States and Luminescence of Nucleic Acid Systems. *Annu. Rev. Phys. Chem.,* 1983, 34, 329–357.
24. Omerzu, A., Mihailovic, D., Anzelak, B. & Turel, I. Optical spectra of wet and dry M-DNA. *Phys. Rev. B*,2007, 75, 121103(R).
25. Zalar, A., Tepfer, D., Hoffmann, S. V., Kenney, J. M. & Leach, S. Directed exospermia: I. Biological modes of resistance to UV light are implied through absorption spectroscopy of DNA and potential UV screens. *Int. J. Astrobiol.,* 2007, 6, 229.
26. Silaghi, S. D. *et al.* Dielectric functions of DNA base films from near-infrared to ultra-violet. *Phys. Status Solidi B,* 2005, 242, 3047–3052.
27. Sonmezoglu, S. & Ates Sonmezoglu, O. Optical and dielectric properties of double helix DNA thin films. *Mater. Sci. Eng. C,* 2011, 31, 1619–1624.
28. Ching, W.-Y. & Rulis, P. *Electronic Structure Methods for Complex Materials: The Orthogonalized Linear Combination of Atomic Orbitals*. (Oxford University Press, 2012).
29. Preuss, M., Schmidt, W. G., Seino, K., Furthmüller, J. & Bechstedt, F. Ground- and excited-state properties of DNA base molecules from plane-wave calculations using ultrasoft pseudopotentials. *J. Comput. Chem*., 2004, 25, 112–122.
30. Poudel, L., Rulis, P., Liang, L. & Ching, W.-Y. Partial Charge and Hydrogen Bonding in Four Periodic B-DNA Models. *Phys. Rev. E*., 2014, ***90****, 022705*.
31. Cornell, W. D. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc*., 1995, 117, 5179–5197.
32. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.,* 1996, 6, 15–50.
33. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B,* 1996*,* 54, 11169–11186.
34. Baroni, S., de Gironcoli, S., Dal Corso, A. & Giannozzi, P. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.*, 2001, 73, 515–562.
35. Rohlfing, M. & Louie, S. G. Electron-hole excitations and optical spectra from first principles. *Phys. Rev. B*, 2000, 62, 4927–4944.
36. Rocca, D., Lu, D. & Galli, G. Ab initio calculations of optical absorption spectra: Solution of the Bethe–Salpeter equation within density matrix perturbation theory. *J. Chem. Phys.*, 2010, 133, 164109.
37. Marini, A., Hogan, C., Grüning, M. & Varsano, D. yambo: An ab initio tool for excited state calculations. *Comput. Phys. Commun.*, 2009, 180, 1392–1403.
38. Runge, E. & Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.*, 1984, 52, 997–1000.
39. Onida, G., Reining, L. & Rubio, A. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.*, 2002, 74, 601–659.
40. Ullrich, C. *Time-dependent density-functional theory: concepts and applications*. (Oxford University Press, 2012).
41. Casida, M. E. & Huix-Rotllant, M. Progress in Time-Dependent Density-Functional Theory. *Annu. Rev. Phys. Chem.*, 2012, 63, 287–323.
42. Hirata, S. & Head-Gordon, M. Time-dependent density functional theory within the Tamm–Dancoff approximation. *Chem. Phys. Lett.*, 1999, 314, 291–299.
43. Chantzis, A., Laurent, A. D., Adamo, C. & Jacquemin, D. Is the Tamm-Dancoff Approximation Reliable for the Calculation of Absorption and Fluorescence Band Shapes? *J. Chem. Theory Comput.*, 2013, 9, 4517–4525.
44. Ratcliff, L. E., Hine, N. D. M. & Haynes, P. D. Calculating optical absorption spectra for large systems using linear-scaling density functional theory. *Phys. Rev. B*, 2011, 84, 165131.
45. Ratcliff, L. E. & Haynes, P. D. Ab initio calculations of the optical absorption spectra of C60-conjugated polymer hybrids. *Phys. Chem. Chem. Phys.*, 2013, 15, 13024–13031.

46. Foresman, J. B., Head-Gordon, M., Pople, J. A. & Frisch, M. J. Toward a systematic molecular orbital theory for excited states. *J. Phys. Chem.*, 1992, 96, 135–149.
47. Liang, L., Rulis, P., Kahr, B. & Ching, W. Y. Theoretical study of the large linear dichroism of herapathite. *Phys. Rev. B,* 2009, 80, 235132.
48. Kahr, B., Freudenthal, J., Phillips, S. & Kaminsky, W. Herapathite. *Science*, 2009, 324, 1407–1407.
49. Liang, L., Rulis, P., Ouyang, L. & Ching, W. Y. Ab initio investigation of hydrogen bonding and network structure in a supercooled model of water. *Phys. Rev. B,* 2011 83, 024201.
50. Adhikari, P. *et al.* Electronic Structure, Dielectric Response, and Surface Charge Distribution of RGD (1FUV) Peptide. *Nature Sci. Rep.* 2014, **4**, 5605.
51. Lee, B. & Richards, F. M. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.*, 1971, 55, 379–IN4.
52. French, R. H. Electronic band structure of Al_2O_3, with comparison to AlON and AlN. *J. Am. Ceram. Soc.,* 1990, 73, 477–489.
53. John SantaLucia, Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. PNAS, 1998, 95 (4), 1460-1465.
54. Fox, J. *Applied regression analysis and generalized linear models*. (Sage, 2008).
55. Fox, J. *An R companion to applied regression*. (SAGE Publications, 2011).
56. Wilfinger, W. W., Mackey, K. & Chomczynski, P. Effect of pH and Ionic Strength on the Spectrophotometric Assessment of Nucleic Acid Purity. *BioTechniques* 22, 474–481 (1997).
57. Owen, T. Fundamentals of modern UV-visible spectroscopy, https://www.chem.agilent.com/Library/primers/Public/59801397_020660.pdf, accessed 7/29/2014.
58. Largy, E., Granzhan, A., Hamon, F., Verga, D. & Teulade-Fichou, M.-P. in *Quadruplex Nucleic Acids* , eds. Chaires, J. B. & Graves, D., Springer Berlin Heidelberg, 2012, 330, 111–177.
59. Xie, P., Teramae, H., Liu, K. & Aoki, Y. Electronic states of mixed base pairs systems of dna and the effect of base composition and sequences on the band structures using screw axis translational symmetry. *Int. J. Quantum Chem.*, 2013, 113, 489–496.
60. Shapir, E. *et al.* Electronic structure of single DNA molecules resolved by transverse scanning tunnelling spectroscopy. *Nat. Mater.*, 2007, 7, 68–74.
61. Orimoto, Y., Gu, F. L., Imamura, A. & Aoki, Y. Efficient and accurate calculations on the electronic structure of B-type poly(dG)-poly(dC) DNA by elongation method: First step toward the understanding of the biological properties of aperiodic DNA. *J. Chem. Phys.*, 2007, 126, 215104.
62. Gervasio, F., Carloni, P. & Parrinello, M. Electronic Structure of Wet DNA. *Phys. Rev. Lett.*, 2002, 89, 108102.
63. Rengifo, E. A. & Murillo, G. DFT-Based Investigation of the Electronic Structure of a Double-Stranded AC B-DNA Dimmer at Different Levels of Theory. *Revista de Ciencias,* 2012, 16, 117.
64. Taniguchi, M. & Kawai, T. Electronic structures of A- and B-type DNA crystals. *Phys. Rev. E,* 2004, 70, 011913.
65. Wang, H., Lewis, J. P. & Sankey, O. F. Band-Gap Tunneling States in DNA. *Phys. Rev. Lett.*, 2004, 93, 016401.
66. Lewis, J. P., Cheatham, T. E., Starikov, E. B., Wang, H. & Sankey, O. F. Dynamically Amorphous Character of Electronic States in Poly(dA)-Poly(dT) DNA. *J. Phys. Chem. B*, 2003, 107, 2581–2587.

Table 1 *Ab initio* HOMO-LUMO gaps and experimentally determined direct band gaps for each DNA oligonucleotide sequence. For the specific n value for each oligonucleotide in *ab initio* calculation and experiment, please refer to the text.

Sequence	HOMO-LUMO Gap [eV] (*ab initio)*	Band Gap [eV] *(experimental)*	Experimental Fitting Range [eV]
(AT)n	3.12	4.30	4.33 – 4.40
(AT)n(GC)n	2.52	4.24	4.25 – 4.39
(AT-GC)n	2.14	4.23	4.28 – 4.32

Table 2: Electronic transition assignments from the PDOS of each oligonucleotide. Features are divided into regions I, II, and III, individually indexed, correlated where possible to an equivalent experimental feature. Transition assignments are given from an occupied state, indicated by a capital letter, to an unoccupied state, given by a lower-case letter, according to the indexing of significant spectral features in the PDOS (Figure 2). For the specific n value for each oligonucleotide in *ab initio* calculation and experiment, please refer to the text.

(AT)n

Region	Feature Index	Feature Energy (eV)	Feature Identification	Equivalent Experimental Feature (eV)
I	$e_1^{(AT)n}$ $e_2^{(AT)n}$	3.12 3.67 3.83	HOMO-LUMO Gap T1→(t1,a2), P2→(t1,a2) A1→a1, P1→(t2,a3) *(260 nm peak)*	4.30 4.64 4.79
II	$m_1^{(AT)n}$ $e_3^{(AT)n}$	4.14 4.45	*(230 nm trough)* A1→(t2,a3)	5.17 5.41
III	$e_4^{(AT)n}$ $e_5^{(AT)n}$	4.78 5.00 >5.00	(S3,P5)→a1, P1→t3 A2→(t1,a2), (T1,P2)→t3 Continuum of overlapping transitions	5.70 -- --

(AT)n(GC)n

Region	Feature Index	Feature Energy (eV)	Feature Identification	Equivalent Experimental Feature (eV)
I	$e_1^{(AT)n(GC)n}$ $e_2^{(AT)n(GC)n}$ $e_3^{(AT)n(GC)n}$ $e_4^{(AT)n(GC)n}$	2.52 3.07 3.25 3.56 3.88	HOMO-LUMO Gap(C1,T1)→t1, P2→g1 G1→(t2,c2,g2) C2→c1, A1→(t1,g1), (T1,C1,P2)→t3, P1→t4 *(260 nm peak)* C2→(t2,c2,g2), A1→a1, P2→t4, G1→a2, P1→g3*(260 nm peak)*	4.24 Broadening Broadening 4.57 4.79
II	$m_1^{(AT)n(GC)n}$ $e_5^{(AT)n(GC)n}$	4.14 4.48	*(230 nm trough)* (S3, P5)→(c1, t1), G2→g1, C2→a2, A1→(a2,c3)	5.21 5.41
III	$e_6^{(AT)5(GC)n}$	4.93 to 5.24	Continuum of overlapping transitions	-- --

(AT-GC)n

Region	Feature Index	Feature Energy (eV)	Feature Identification	Equivalent Experimental Feature (eV)
I	$e_1^{(AT-GC)n}$ $e_2^{(AT-GC)n}$ $e_3^{(AT-GC)n}$ $e_4^{(AT-GC)n}$	2.14 3.2 3.37 3.59 3.81	HOMO-LUMO Gap P1→g2, G1→g1, T1→(a1,c1,g2) P2→(a1,c1), G1→(g3) P1→t1, P3→g1, (A1,C1)→g2 (P3,A1,C1)→(a1,c1), (P2,G1)→t1 *(260 nm peak)*	4.23 Broadening Broadening 4.56 4.80
II	$m_1^{(AT-GC)n}$ $e_5^{(AT-GC)n}$	4.14 4.48	*(230 nm trough)* Continuum of overlapping transitions	5.21 5.41
III	$e_6^{(AT-GC)n}$	4.76 to 5.21	Continuum of overlapping transitions	-- --

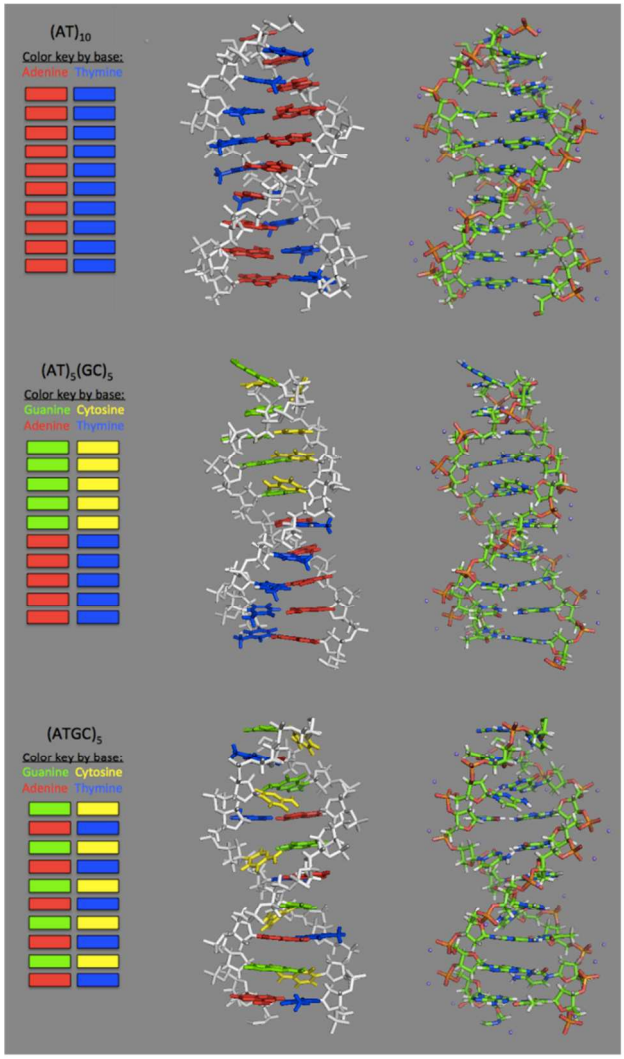

Figure 1: Molecular models of ten-base-pair oligonucleotide duplexes a) (AT)10, b) (AT)5(GC)5, and c) (AT-GC)5, given as schematic and 3D models. Base pairs are given in the standard color-coding: blue is thymine (T), yellow is cytosine (C), green is guanine (G), and red is adenine (A). Individual atoms are color coded in the rightmost column: carbon is green, hydrogen is white, nitrogen is blue, oxygen is red, phosphorus is orange, and sodium is purple. Images were created using PyMol software.

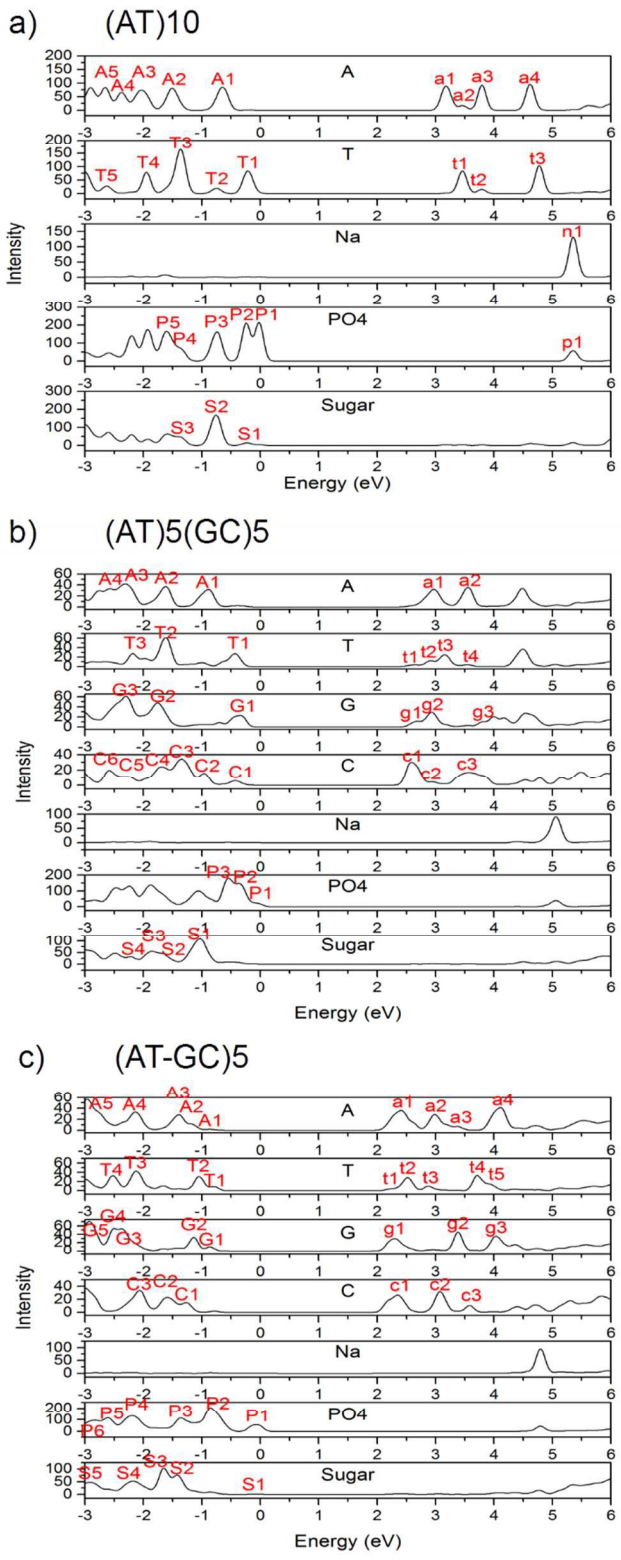

Figure 2: Partial Densities of States (PDOS) for a) (AT)10 , b) (AT)5(GC)5, and c) (AT-GC)5

from OLCAO-DFT *ab initio* calculation. The PDOS are sectioned by structural group, given as adenine (A), thymine (T), guanine (G), cytosine (C), Na^{+} counterion (NA), PO_4^{3-} complex anion in the backbone (PO4), and deoxyribose sugar in the backbone (Sugar). Occupied state features are identified with a capital letter and number (e.g. P1) and unoccupied state features are given by a lower-case letter and number (e.g. p1). In (AT)10 the HOMO is defined by a doublet feature originating from the phosphate group (P1 and P2), and the LUMO results from a state associated with the adenine bases (a1). In (AT)5(GC)5 the HOMO is defined by a feature originating from the phosphate group (P1), and the LUMO results from a state associated with the cytosine bases (c1). In (AT-GC)5 the HOMO is defined by a feature originating from the phosphate group (P1), and the LUMO results from a state associated with the cytosine and guanine bases (c1 and g1).

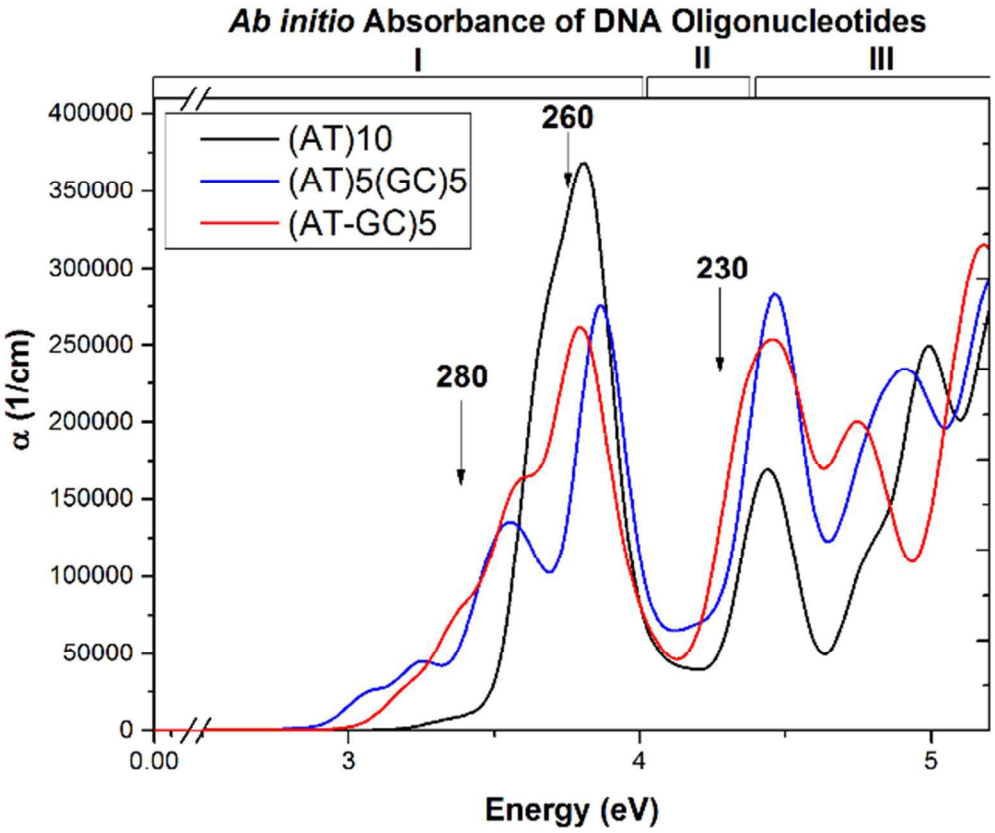

Figure 3: *Ab initio* absorption coefficient spectra for (AT)10 (black line), (AT)5(GC)5 (blue line), and (AT-GC)5 (red line). Regions I, II, and III are marked, as are the features corresponding to the 280 nm, 260 nm, and 230 nm features in experimental spectra.

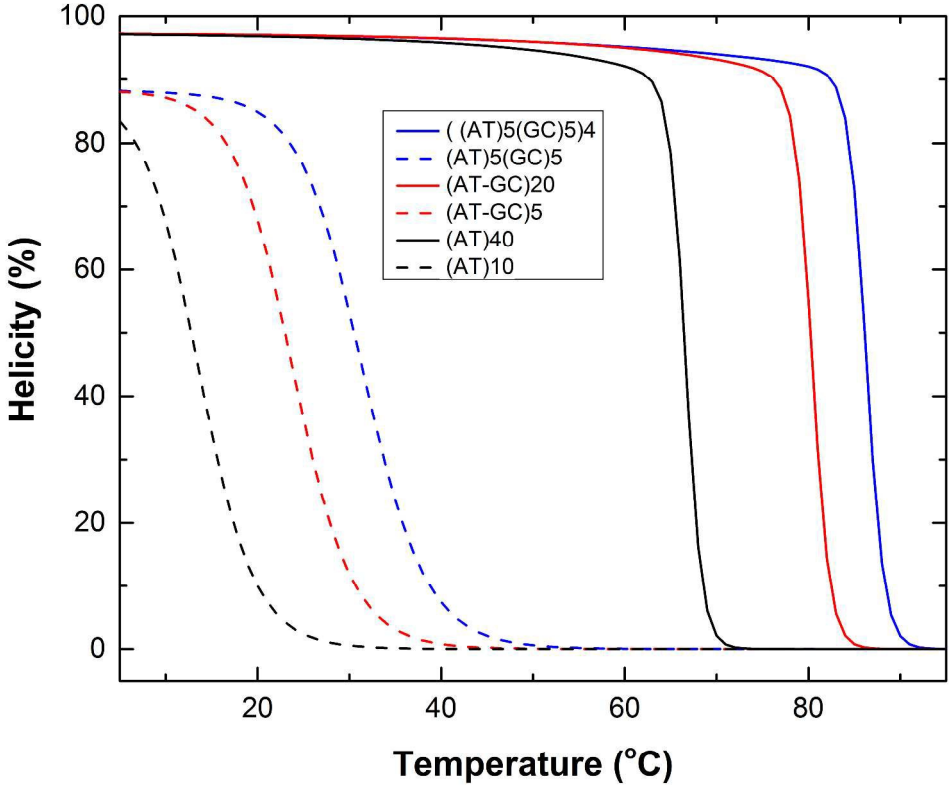

Figure 4: Calculated melting curves of 10 and 40 base-pair (AT)10, (AT)40, (AT)5(GC)5, ((AT)5(GC)5)4, (AT-GC)5, and (AT-GC)20.

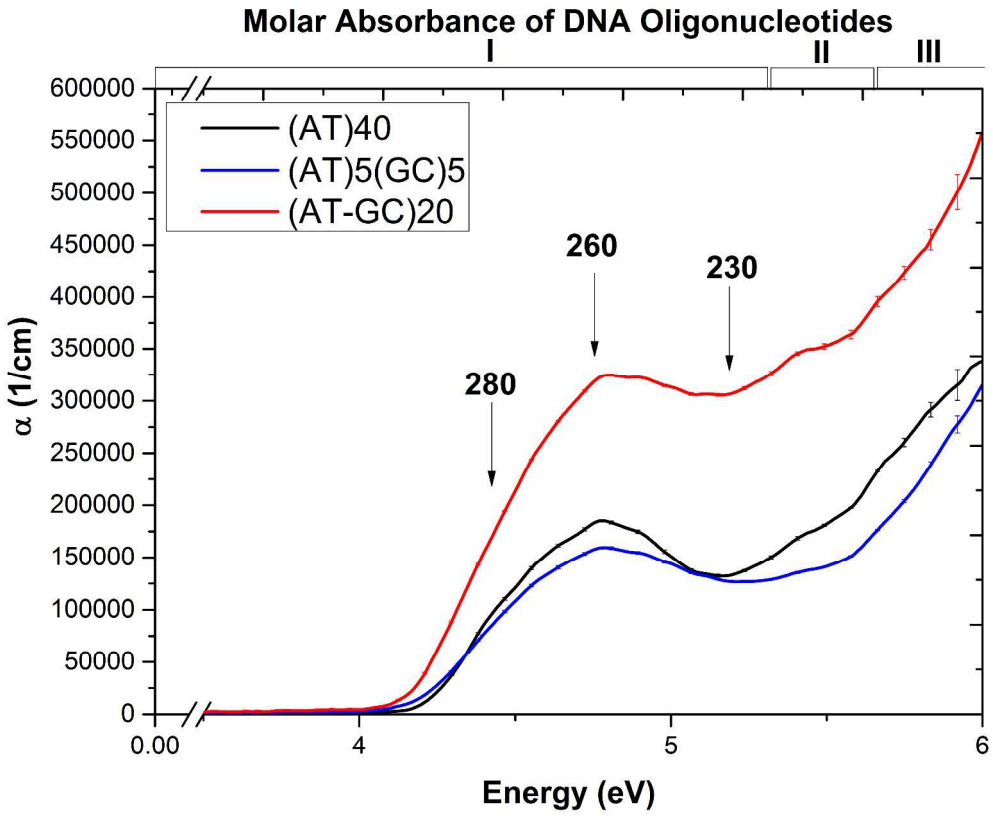

Figure 5: Experimental molar absorption coefficient spectra for (AT)40, (AT)5(GC)5, and (AT-GC)20. Regions I, II, and III are marked, as are the 280 nm, 260 nm, and 230 nm features. Error bars shown are the heterogeneous error from the fitting of the concentration data to the Beer-Lambert law.

Table of contents graphic