Electronic Structure and Stability of Semiconducting Graphene Nanoribbons

Verónica Barone, Oded Hod,* and Gustavo E. Scuseria

Department of Chemistry, Rice University, Houston, Texas 77005-1892

Received July 21, 2006; Revised Manuscript Received November 2, 2006

ABSTRACT

We present a systematic density functional theory study of the electronic properties, optical spectra, and relative thermodynamic stability of semiconducting graphene nanoribbons. We consider ribbons with different edge nature including bare and hydrogen-terminated ribbons, several crystallographic orientations, and widths up to 3 nm. Our results can be extrapolated to wider ribbons providing a qualitative way of determining the electronic properties of ribbons with widths of practical significance. We predict that in order to produce materials with band gaps similar to Ge or InN, the width of the ribbons must be between 2 and 3 nm. If larger bang gap ribbons are needed (like Si, InP, or GaAs), their width must be reduced to 1−2 nm. According to the extrapolated inverse power law obtained in this work, armchair carbon nanoribbons of widths larger than 8 nm will present a maximum band gap of 0.3 eV, while for ribbons with a width of 80 nm the maximum possible band gap is 0.05 eV. For chiral nanoribbons the band gap oscillations rapidly vanish as a function of the chiral angle indicating that a careful design of their crystallographic nature is an essential ingredient for controlling their electronic properties. Optical excitations show important differences between ribbons with and without hydrogen termination and are found to be sensitive to the carbon nanoribbon width. This should provide a practical way of revealing information on their size and the nature of their edges.

Low-dimensional carbon structures have been the focus of extensive research since the discovery of fullerenes1 and carbon nanotubes2 (CNTs). These novel forms of carbon present unique opportunities to study low-dimensional physical phenomena. The peculiar electronic properties of carbon nanotubes arise from the characteristic band structure of graphene (a single infinite sheet of graphite). Depending on their chirality and diameter, CNTs present either a semiconducting or metallic character.3 Due to their special electronic and structural properties, such systems are considered as promising candidates to build new technologies with applications ranging from nanoelectronics, nanosensors, photonic and nanomechanical devices to active material encapsulation, drug delivery, and surface functionalization, among others.4

Recently, a related type of carbon based quasi-one-dimensional systems, referred to as carbon nanoribbons (CNRs), has been synthesized experimentally.5 Their synthesis was followed by experimental evidence of coherent electron transport on large length scales measured at relatively high temperatures.6 These findings open exciting opportunities for the design of novel electronic devices and interconnects.

CNRs are basically elongated stripes of single layered graphene with a finite width. Due to their structural resemblance to carbon nanotubes and to the quantum confinement effect, CNRs are expected to present electronic properties similar to those of CNTs. Nevertheless, recent experimental progress5 indicates that CNRs can be produced in a highly controllable manner offering the opportunity to take advantage of the unique and well-understood electronic properties of CNTs while circumventing their growth control problems which have been one of the major obstacles toward their wide commercial use.

Theoretical studies mainly based on tight-binding approximations have shown remarkable similitude between nanoribbons and single-walled carbon nanotubes (SWNTs).7−10 Graphene nanoribbons can also be either metallic or semiconducting depending on the crystallographic direction of the ribbon axis. Similar to armchair nanotubes, it has been predicted that ribbons with zigzag edges are all metallic7 and may present interesting magnetic properties.11,12 Band gap oscillations have also been predicted for semiconducting narrow armchair ribbons as a function of their width.13 These oscillations introduce the possibility of tailoring their electronic structure.

A deep understanding of the physical parameters governing the electronic structure and thermodynamic stability of CNRs is necessary before considering them as building blocks in the design of future nanodevices. Since the main difference (apart from the strain energy) between CNTs and CNRs is the existence of sharp edges in CNRs, the majority of the current literature focuses on the edge effects on the electronic and magnetic properties of CNRs.7−10 Furthermore, most of these studies are based on tight-binding approximations which are known to give qualitatively good results for the electronic structure of CNTs but fail to quantitatively reproduce experimental results.
Some fundamental questions regarding several important aspects of CNR systems remain open. These include the ability to control the electronic structure of CNRs with realistic widths (~80 nm), the band gap dependence on the crystallographic orientation of the CNR axis, the electronic structure dependence on the nature of the edge passivation, among other important issues such as the effect of structural defects on the electronic and thermodynamic properties of CNRs.

It is the purpose of the current study to offer an answer to some of these questions by presenting a detailed analysis of the electronic structure and thermodynamic stability of semiconducting CNRs using density functional theory (DFT). First, we systematically study the band gap dependence on the width of CNRs, discussing the possibility of tailoring their electronic properties. This is followed by a prediction of some unique features in their optical spectra which may be utilized as a characterization tool. A qualitative estimation of the relative stability of free-standing CNRs with different edge nature is given thereafter.

The CNRs considered in the present study are obtained by “unfolding” infinite periodic CNTs and cutting (or extending) the resulting sheet to the desired width, while eliminating dangling bonds. We follow the notation used in the literature in which a zigzag CNT unfolds into an armchair CNR (i.e., a CNR with armchair edges). In Figure 1 we present a set of typical hydrogen-terminated CNRs obtained by unfolding CNTs of different chiralities. Two numbers are assigned to each CNR, the chiral angle \( \phi \) and the width \( W \). The chiral angle is related to the corresponding CNT indices, \( (n,m) \), via \( \tan(\phi) = (3^{1/2}m/(2n + m)) \) and represents the crystallographic direction of the axis of the CNR. To calculate the width, the positions of all carbon atoms are projected onto a line perpendicular to the CNR axis. The width is taken as the maximum distance between any two carbon atoms along this line.

Calculations reported in this paper have been carried out using the development version of the Gaussian suite of programs. Bloch functions are expanded in terms of atomic Gaussian-type orbitals, and the Khon–Sham (KS) equations are solved self-consistently in that basis set. The optimized geometries and the electronic structure of each ribbon have been obtained using the 6-31G* Gaussian basis set and two different functionals, the PBE realization of the generalized gradient approximation and the screened exchange hybrid density functional, HSE. The latter functional has been tested in a wide set of materials and was shown to accurately reproduce experimental band gaps. Furthermore, HSE has also proven to deliver accurate first and second optical excitation energies in metallic and semiconducting SWNTs.

We start by studying the electronic properties of armchair (\( \phi = 0^\circ \)) CNRs. In Figure 2 we present our results for the band gap dependence on the width of bare-edged (left panel) and hydrogen-terminated (right panel) armchair CNRs as calculated using the PBE and HSE functionals. Even though a quantitative difference exists, as expected, between the band gaps calculated by the two functionals, qualitatively, both predict similar band gap oscillations as a function of the CNR’s width. We find that hydrogen termination enhances the amplitude of the band gap oscillations for narrow ribbons (\( W < 1.5 \) nm) and that this enhancement vanishes for larger widths. Similar band gap oscillations were previously predicted using tight binding calculations. Nevertheless, a major difference is identified. While in the tight-binding calculations, at certain widths, armchair CNRs appear to become metallic, similar to the \( \mod(3) \) rule for zigzag CNTs, our DFT calculations predict all armchair CNRs studied to be semiconducting.

The periodicity of 3 obtained for the band gap oscillations as a function of the width of the system can be related to the electronic properties of the corresponding CNT systems. A simple picture may rationalize the results as follows. Consider the direction perpendicular to the periodic axis of an armchair nanoribbon. One can imagine the addition of one \( C_2 \) unit, defined in ref 7 as adding a single site to a
zigzag chain of carbon atoms in that direction. Since the \( \pi \) system of the ribbon is composed of a contribution of a single \( p_z \) electron from each atomic site, the Fermi wavelength along this chain is approximately four atomic sites. Therefore, the addition of three sites to such a chain will add a full wavelength and result in a 3-fold periodic pattern.

For the range of widths covered by our DFT calculations \((L < 3 \text{ nm})\) we do not observe a significant quenching of the energy gap oscillations. However, CNRs are expected to reach the graphene limit of zero band gap for sufficiently large widths. Therefore, the question arises of how these oscillations behave for widths larger than 3 nm. To study this, we note that the armchair CNRs band gaps presented in Figure 2 can be separated into three groups, namely, the points forming the envelope of the maxima of the oscillations, those forming the envelope of the minima of the oscillations, and the remaining intermediate points. It is possible now to extrapolate the behavior of each of these subgroups to larger widths. In Figure 3 such an extrapolation is presented using an inverse power law with two fitting parameters. This simple rule presents the correct asymptotic behavior and provides qualitative information of the electronic structure of ribbons beyond the range of widths studied by our first-principles calculations. Some points are worth noting. The same type of inverse power law fits well all the six different sets of data, including bare and H-terminated CNRs. However, an important difference arises between passivated and nonpassivated CNRs. While for the bare CNRs (left panel of Figure 3) the intermediate set lies very close to the lower envelope of the oscillations, the case for the hydrogen-terminated CNRs (right panel of the figure) is the opposite, and the intermediate set is found closer to the upper envelope. This implies that if one considers CNRs of width \( L = 8 \text{ nm} \), for instance, produced in ultrahigh vacuum (therefore having bare edges), we expect \( \sim 67\% \) of them to present a band gap of about 0.1 eV and \( \sim 33\% \) a band gap of 0.3 eV. For H-terminated ribbons, \( \sim 67\% \) will present an energy gap of 0.2 eV while \( \sim 33\% \) of them will have a band gap close to zero. If we consider ribbons ten times wider (\( L \))

**Figure 2.** Dependence of the band gap on the ribbon width for bare (left panel) and hydrogen-terminated (right panel) armchair CNRs.

**Figure 3.** Extrapolation of the envelope of the HSE/6-31G* band gap oscillation to large CNRs width for bare (left panel) and hydrogen-terminated (right panel) CNRs. The inverse power law used for the extrapolation has two parameters \( a \) and \( b \) according to \( E_g(L) = aL^{-b} \).
dependence on the edge nature and on the crystallographic different CNRs presented in Figure 1. The different panels correspond to the passivated chiral CNRs. The different panels correspond to the different ribbon chiralities. This is in contrast to the band gap oscillations as the chiral angle increases. This trend implies that in order to have control on the band gap of the ribbons, it is necessary to control not only their width but also their chiral angle. It is interesting to note that the band gap minima do not occur at the same width for different ribbon chiralities. This is in contrast to the band gap commensuration with respect to different ribbon types obtained by Ezawa.13

Another parameter that may help to control the electronic properties of CNRs is the crystallographic direction of their main axis. This has an equivalent rule as that of the chiral angle in CNTs and is expected to be important in CNRs electronic structure design. As a result of the complicated nature of the edges formed when unrolling a chiral CNT and due to the sensitivity of the electronic structure of CNRs on the nature of the edges, this effect is expected to have emphasized importance in the case of CNRs. In Figure 4 we present a plot of the band gap as a function of the width of different chiral hydrogen-terminated CNRs. It is evident that there is a progressive decrease of the amplitude of the band gap oscillations as the chiral angle increases. This finding is in accordance to the fact that at the limit of $\phi = 30^\circ$ (zigzag ribbons) all CNRs are predicted to be metallic.7 What is most surprising is that for a chiral angle as low as $\phi \sim 9^\circ$ the band gap oscillations almost completely vanish. This trend implies that in order to have control on the band gap of the ribbons, it is necessary to control not only their width but also their chiral angle. It is interesting to note that the band gap minima do not occur at the same width for different ribbon chiralities. This is in contrast to the band gap commensuration with respect to different ribbon types obtained by Ezawa.13

The predicted band gap oscillations and their strong dependence on the edge nature and on the crystallographic orientation of the CNR axis suggest that, similar to CNTs,26–29 optical spectra can be proven as a powerful characterization tool for CNRs. We have thus calculated the optical spectra of armchair CNRs employing the first-order noniterative random phase approximation expression for the imaginary part of the dielectric function $\epsilon$.30

$$\text{Im}(\epsilon) = \frac{1}{\omega^2} \sum_k \sum_{\alpha, \sigma} |\langle \psi^k_\alpha | \mathbf{p} | \psi^k_\sigma \rangle|^2 \delta(\epsilon^k_\alpha - \epsilon^k_\sigma - \omega)$$

Here, $\omega$ is the excitation photon energy, $\mathbf{p} = -i\hbar \nabla$ is the linear momentum operator, $\psi^k_\alpha$ and $\psi^k_\sigma$ stand for occupied and unoccupied Bloch orbitals with energies $\epsilon^k_\alpha$ and $\epsilon^k_\sigma$, respectively, and $\mathbf{k}$ stands for the reciprocal space vector.

In Figure 5 the first ($E_{11}$) and second ($E_{22}$) optical excitation energies for armchair CNRs calculated at the HSE/6-31G* level of theory are presented. As expected, we observe that the first excitation energy, $E_{11}$, presents oscillations that correspond to the band gap oscillations discussed above. More surprising is the fact that the second excitation energy, $E_{22}$, presents width-dependent oscillations as well.

An interesting point to notice is the fact that for ribbons with hydrogen termination (right panel of Figure 5) there is a phase shift of $\pi$ between the oscillation of the first and second optical transition energies such that the local maxima of $E_{11}$ coincide with the local minima of $E_{22}$ for the whole range of widths studied. This is expected to give rise to a doublet in the optical spectrum. A similar doublet is expected to appear for bare CNRs with $L < 1.2$ nm (see the left panel of the figure). Nevertheless, for larger widths, the band gap oscillations of the first and second optical transition energies of bare ribbons are in phase and the doublet is expected to disappear.

The predicted optical spectra for three CNRs with consecutive widths corresponding to one period of the band gap oscillations are presented in Figure 6. A doublet clearly appears in the upper panel of the figure which corresponds to the local maximum of $E_{11}$ (minimum of $E_{22}$) for this period. This holds true for both bare and hydrogen-terminated CNRs.

It is interesting to further note that for the consecutive CNRs series presented in Figure 6 the passivation produces a red shift of the $E_{11}$ peak for a CNR width of $L = 0.86$ nm (lower panel of Figure 6) and a blue shift of this peak at a width of $L = 0.96$ nm (middle panel of the figure). In the upper panel one can see that the lower peak of the doublet red shifts while the higher peak blue shifts, causing a peak separation upon hydrogen passivation. It is suggested that both the appearance of the doublet in the optical spectra and the peak shifts upon passivation may provide a means for experimentally determining the edge nature of CNRs.

Up to now we have considered the effect of the edge nature and the crystallographic construction of CNRs on their electronic structure. Another important aspect is the influence
of the edge nature on the thermodynamic stability of such systems. It is intuitively expected that upon “unfolding” a CNT, the loss of the chemical bonds experienced by the edges of the obtained CNR will cause it to be less thermodynamically stable than the corresponding CNT. Passivating the bare edges with proper termination groups will stabilize the structure. To quantify these conjectures it is necessary to evaluate the relative stability of CNRs with and without edge passivation. As these ribbons have different chemical compositions, the cohesive energy per atom does not provide a suitable way to compare their relative stability. Therefore, we adopt the approach customary used in binary phase thermodynamics to account for chemical composition and utilized previously to analyze the relative stability of endohedral silicon nanowires.\textsuperscript{31} Within this approach one defines a molar Gibbs free energy of formation $\delta G$ for a CNR as

$$
\delta G(x) = E(x) + x_H \mu_H + (1 - x_H) \mu_C \quad (2)
$$

where $-E(x)$ is the cohesive energy per atom of the CNR, $x_H$ is the molar fraction of hydrogen atoms, and $\mu_H$ and $\mu_C$ are the chemical potentials of the constituents at a given state. We choose $\mu_H$ as the binding energy per atom of the $H_2$ molecule and $\mu_C$ as the cohesive energy per atom of a single graphene sheet. This definition allows for a direct energy comparison between passivated and nonpassivated nanoribbons with different compositions.

In Figure 7 we present a plot of the relative stability of different hydrocarbon systems as a function of the hydrogen molar fraction. According to the definition given in eq 2, stable structures with respect to the constituents present a negative value of $\delta G$ while metastable structures present a positive $\delta G$ value. For bare armchair CNRs the relative stability comparison is straightforward; the wider the ribbon is, the more stable it becomes (lower cohesive energy) approaching the calculated graphene limit of $-7.70$ eV, as can be seen in the inset of the figure. For comparison purposes we included the relative stability of a pristine\textsuperscript{5,5} SWNT which despite the considerable strain energy is still more stable ($E_c = -7.51$ eV) than any of the nonpassivated armchair CNRs studied here. The situation changes upon H-termination. As described above, the highly reactive edges of the CNRs stabilize by chemically absorbing hydrogen atoms, lowering their free energy. A striking feature is evident where all hydrogen-passivated CNRs studied (armchair and chirals) present Gibbs free energies lower than $0.05$ eV. This indicates a thermodynamic stability higher than that of nonpassivated CNRs and comparable to that of a graphene sheet. The crystallographic direction of the ribbon axis has little influence on their relative stability; in other words, chiral and armchair-passivated CNRs are predicted to be equally stable in our analysis. It should be noted that the relative
and the nature of the edge of CNRs. Finally, we have shown that a careful design of their crystallographic nature is an important ingredient in their electronic properties control. All stability considerations we are presenting here are mainly qualitative and that we are neglecting thermal effects, zero point energy corrections, and substrate effects.

In summary, we have presented a detailed DFT analysis of the electronic properties and relative stabilities of several types of semiconducting carbon nanoribbons with and without hydrogen termination and widths up to 3 nm. These results can be extrapolated to wider ribbons providing a qualitative way of determining the electronic properties of ribbons with widths of practical significance. In order to produce an armchair CNR with a band gap similar to that of Ge or InN, we predict that its width should be between 2 and 3 nm. If a larger band gap material is needed, with a gap comparable to that of Si, InP, or GaAs, for example, our calculations indicate that the width should be reduced to as low as 1–2 nm. According to the extrapolated inverse power law obtained in this work, CNRs with widths larger than 8 nm will present a maximum band gap of 0.3 eV while for ribbons with \( L = 80 \) nm the maximum possible band gap is 0.05 eV. For chiral CNRs, the band gap oscillations rapidly vanish as a function of the chiral angle indicating that a careful design of their crystallographic nature is an essential ingredient in their electronic properties control. All these features are identified as highly important for tailoring the electronic properties of CNRs and for the design of future nanoelectronic devices. It is suggested that edge termination with other functional groups or chemical and physical modification of the carbon network may serve as a further control parameter for the amplitude and extent of the band gap oscillations. Research in that direction is currently underway. Optical excitations show important differences between ribbons with and without hydrogen terminations and are found to be sensitive to the CNR width. This should provide a practical way of revealing information on the size and the nature of the edge of CNRs. Finally, we have shown that as free-standing materials, edge passivation is highly stable.

Acknowledgment. This work was supported by NSF Award Number CHE-0457030 and the Welch Foundation. Calculations were performed in part on the Rice Terascale Cluster funded by NSF under Grant EIA-0216467, Intel, and HP. O.H. would like to thank the generous financial support of the Rothschild and Fulbright Foundations.

References

(16) This basis set consists of (3szp1d) contracted Gaussian functions.

NL0617033