

Calculation of the total energy per unit cell and of the band structures of the five nucleotide base stacks using the local-density approximation

M.-L. Zhang, M. S. Miao, V. E. Van Doren, J. J. Ladik, and J. W. Mintmire

Citation: *The Journal of Chemical Physics* **111**, 8696 (1999); doi: 10.1063/1.480209

View online: <http://dx.doi.org/10.1063/1.480209>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/111/18?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[Ab initio calculation of the CdSe/CdTe heterojunction band offset using the local-density approximation-1/2 technique with spin-orbit corrections](#)

J. Appl. Phys. **111**, 073708 (2012); 10.1063/1.3699054

[Local-density approximation for orbital densities applied to the self-interaction correction](#)

J. Chem. Phys. **128**, 044105 (2008); 10.1063/1.2822113

[Investigation of the energy band structure of orthorhombic BaSi 2 by optical and electrical measurements and theoretical calculations](#)

Appl. Phys. Lett. **81**, 1032 (2002); 10.1063/1.1498865

[Comment on "Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory treatment" \[*J. Chem. Phys.* **114**, 5149 \(2001\)\]](#)

J. Chem. Phys. **116**, 11039 (2002); 10.1063/1.1480873

[Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment](#)

J. Chem. Phys. **114**, 5149 (2001); 10.1063/1.1329889



AIP | APL Photonics

APL Photonics is pleased to announce
Benjamin Eggleton as its Editor-in-Chief



Calculation of the total energy per unit cell and of the band structures of the five nucleotide base stacks using the local-density approximation

M.-L. Zhang, M. S. Miao, and V. E. Van Doren

Dept. of Physics, University of Antwerp (RUCA), Groenenbergenlaan 171, B-2020 Antwerpen, Belgium

J. J. Ladik

Theoretical Chemistry and Laboratory of the National Foundation for Cancer Research, Friedrich-Alexander University Erlangen-Nürnberg, Egerlandstr. 3, D-91058 Erlangen, Germany

J. W. Mintmire

Chemistry Division, Naval Research Laboratory, Washington, DC 20375-5342

(Received 16 July 1999; accepted 18 August 1999)

All-electron first-principles total energy electronic structure calculations were carried out for single chains of four nucleotide base stacks (composed of adenine, thymine, guanine, and cytosine, respectively) in the DNA *B* conformation (3.36 Å stacking distance and 36° screw angle θ) using the local-density approximation (LDA) within a helical band structure approach. A uracil stack was also computed in the DNA *B* conformation and compared with the results obtained for the four DNA base stacks. The total energies per unit cell as a function of the stacking distance (at fixed screw angle $\theta=36^\circ$) and of the screw angle (at $d=3.36$ Å) show in most cases rather good agreement with the experiment. As expected with LDA calculations, the band gaps were underestimated by nearly 50% compared to experimentally suggested values. Finally, some suggestions are given for the improvement of the band structures of the nucleotide base stacks. © 1999 American Institute of Physics. [S0021-9606(99)30342-1]

I. INTRODUCTION

Early Hartree–Fock (HF) calculations of the homopolynucleotides (poly-*XSP*, where *X* is a base, *S* the sugar-ring in DNA, and *P* stands for its phosphate group) showed that the calculated band structures of helical DNA *B* conformations (stacking distance $d=36^\circ$, screw angle $\theta=36^\circ$) of the homopolynucleotides are given to a very good approximation by the superposition of the band structures of the base stacks and those of the sugar-phosphate chain.^{1,2} HF band structure calculations of a cytosine stack in the presence of water clusters³ further showed that the presence of water has little influence on the band structure. Additional calculations suggest that counterions should have minimal effect because of the screening by alternating negative and positive charges on the phosphate group, on the sugar ring, and on the bases.⁴ For all these reasons the calculation of the base stacks alone still gives rather good information of the electronic structure of the periodic base stacks in DNA *B*.

In a previous paper, the total energy per unit cell and the band structure of polyparaphenylene were calculated using an all-electron first-principles total energy method using the local-density approximation (LDA) with a helical band structure approach.^{5,6} In this paper, we present the total energies per unit cell of the five nucleotide base stacks adenine (A), thymine (T), guanine (G), cytosine (C), and finally of uracil (U). We note that uracil does not occur in DNA, but rather in RNA with different helical conformations. For the sake of comparison we have considered a uracil stack with the same helical structure as the other four nucleotide base stacks. Be-

cause there are correlation-corrected Hartree–Fock band structure calculations (using the inverse Dyson equation in the diagonal approximation⁷ with a Møller–Plesset⁸ self energy in the framework of the electronic polaron model⁹) on the base stacks,¹⁰ it seemed interesting and useful to compare these band structures with the LDA ones. Further, having already performed LDA calculations on different chemical bond chains,^{6,11,12} we wanted to see how the LDA program works for periodic molecular stacks in a helical arrangement.

In reality, DNA is composed of an aperiodic sequence of the nucleotides, so that a more accurate treatment of the negative factor counting (NFC) method in its matrix block form should be used.¹³ For the approximation and test of the electronic structure of aperiodic DNA, the first step is necessarily the calculation of the band structures of the base stacks. Further, one should point out that according to statistical analyses of base sequences there is a preference in native DNA to have the same bases repeated several times (in some cases up to 30). On the other hand, the homopolynucleotides always containing the same base have been synthesized and they are commercially available. Finally, one should point out that the NFC method can be easily formulated also in the framework of the density functional method.

II. METHODS

In this work we use the local-density functional approach using a linear combination of Gaussian-type orbitals (LCGTO) originally developed for two-dimensionally periodic systems¹⁴ and chain polymers,¹⁵ to calculate the total

energies and electronic structures of helical chain polymers.^{5,11} The local density approximation is used with a Kohn–Sham exchange term¹⁶ and a Ceperley–Alder correlation term,¹⁷ using the Perdew–Zunger analytic fit.¹⁸ The helicity of the polymer system is described by a screw operator S , with the one-electron wave functions transforming under S according to Bloch's theorem. These wave functions are constructed from a linear combination of products of nuclear-centered Gaussian-type orbitals with real solid spherical harmonics. Furthermore, the exchange-correlation potential and the charge density in the Coulomb potential are fitted with a linear combination of Gaussian-type functions.^{5,6} Details of the theory can be found in the references given above.

For the conformations of the base stacks, we have taken these in the helical DNA B conformation (for the details for the molecular geometries of the five bases see Ref. 19). In a helical chain (or stack) we have instead of the simple translation d in the z direction, the helix operation: translation+rotation. One can define a screw operator $\hat{S}(d, \theta)$, where θ is the screw angle of the right-handed rotation (in DNA $B+36^\circ$) about the z axis.^{5,20}

$$\hat{S}(d, \theta)(x, y, z) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta, z + d). \quad (1)$$

The symmetry group generated by the screw operation $\hat{S}(d, \theta)$ is isomorphic with the one-dimensional translational group. Therefore the one-electron wave function (crystal orbitals) $\Psi(\mathbf{r}, K)$ will transform under the operation of $\hat{S}(d, \theta)$ as

$$\hat{S}^m \Psi(\mathbf{r}, K) = e^{imK} \Psi(\mathbf{r}, K). \quad (2)$$

Here m is any integer, K ($-\pi < K < \pi$) is a good quantum number (quasimomentum) similar to k in the case of a simple translation, if $\theta=0^\circ$, $K=k_z d$.

The basis sets applied have in Huzinaga's notation²¹ the contraction scheme for oH (centered on the nuclei) (41/1*), oC(7111/411/1*), oN(7111/411/1*), and oO(7111/411/1*). Further, as usual auxiliary Gaussians were used to express the density, the exchange and correlation terms were aH(3,1;3,1) aC(4,4;4,4) aN(4,4;4,4) aO(4,4;4,4). These contracted Gaussians and auxiliary functions were optimized for local density functional (LDF) calculations by Godbout *et al.*²² In the case of the auxiliary (a) functions, four s functions and four sets of (another) s and d functions, all of them with the same exponents, were applied (4,4; 4,4). In the case of the DNA- B base stacks (A , T , G , and C), this basis was applied. For the uracil stack the smaller basis oH(4,1), aH(4,4); oC(721/51/1*), aC(4,3;4,3); oN(711/51/1*), aN(4,3; 4,3), and finally oO(721/51/1*), aO(4,3;4,3) was used. The number of K points in the half Brillouin zone $[0, \pi]$ was 11.

III. RESULTS AND DISCUSSION

We have carried out different sets of calculations where we have varied the stacking distance and the screw angle independently. In Figs. 1–3 we give the total energy per unit cell (in Hartree atomic units) of the five stacks either as a function of their stacking distance (keeping the screw angle constant at 36° and using the position of the screw axis as in

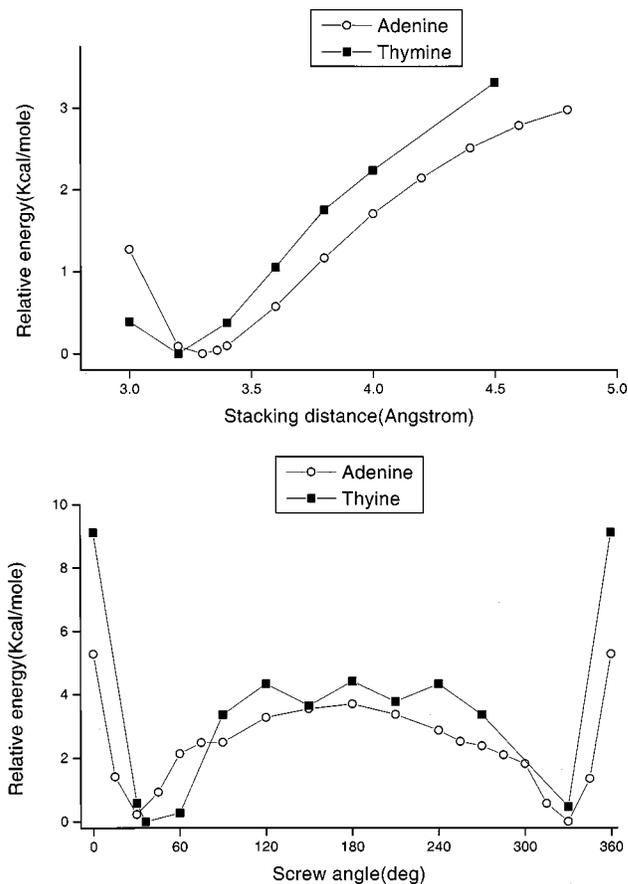


FIG. 1. (a) Total energy per unit cell as a function of stacking distance at a screw angle of 36° , of an adenine stack (open circle), and a thymine stack (solid square)—screw axis as in DNA B ; (b) Total energy per unit cell as a function of screw angle at a stacking distance of 3.36 Å of an adenine stack (open circle) and a thymine stack (solid square)—screw axis as in DNA B .

DNA B) and as a function of the screw angle (keeping the stacking distance d constant at its experimental value in DNA B , 3.36 Å).¹⁹ In the case of the A stack, the total energy per stack, E/N , has its minimum at $d_{\min}=3.30$ Å [$E(d=3.30$ Å, $\theta=36^\circ)=-458.96237$ Hartree, Fig. 1(a)]. This is, however, not the case at the T stack, where $d_{\min}=3.20$ Å [$E(d=3.20$ Å, $\theta=36^\circ)=-446.25278$ Hartree, Fig. 1(a)] and at the G stack where $d_{\min}=3.20$ Å [$E(d=3.20$ Å, $\theta=36^\circ)=-533.14456$ Hartree, Fig. 2(a)]. In the case of the C stack, $d_{\min}=3.22$ Å [$E(d=3.22$ Å, $\theta=36^\circ)=-388.01728$ Hartree, Fig. 2(a)] and for the U stacks, $d_{\min}=3.20$ Å [$E(d=3.20$ Å, $\theta=36^\circ)=-407.76481$ Hartree, Fig. 3(a)]. It should be pointed out that the experimental conformation of a U stack in the DNA geometry is quite different than those of the DNA base stacks (the bases in RNA are not all perpendicular to the screw axis). This means that the reason for the completely different conformation of RNA as compared to DNA B is first of all not the absence of the methyl group in U , but most probably the presence of ribose in RNA instead of deoxyribose in DNA. Of course, only calculations on whole polynucleotides in the two different systems could give more insight into this important problem.

For the A stack we found the deepest minimum at a screw angle of 330° [$E(d=3.36$ Å, $330^\circ)$

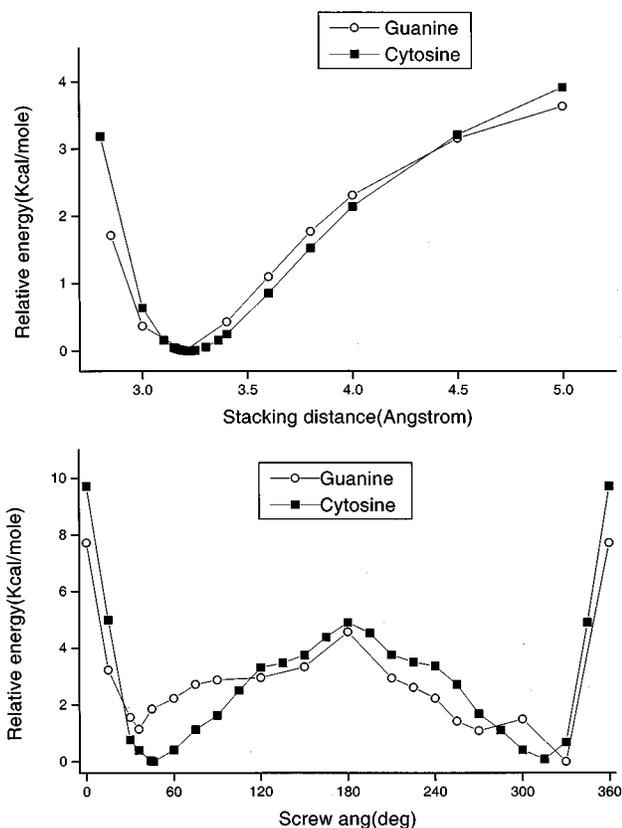


FIG. 2. (a) Total energy per unit cell as a function of stacking distance at a screw angle of 36° , of a guanine stack (open circle), and a cytosine stack (solid square)—screw axis as in DNA B. (b) Total energy per unit cell as a function of screw angle at a stacking distance of 3.36 \AA of a guanine stack (open circle) and a cytosine stack (solid square)—screw axis as in DNA B.

$= -458.96295$ Hartrees] and a second one at 30° lying only 0.02 eV higher [see Fig. 1(b)]. Because a screw angle of 330° in a right-handed helix is equivalent to 30° in a left-handed one, this very small difference shows an energetic near equivalence of the right-handed and left-handed helices (which is to be expected in the case of a monomer with inversion or reflection symmetry). The whole curve is nearly symmetric about 180° .

The situation is similar in the case of the *T* stack, with the energetically similar minima at about 36° [$E(d = 3.36 \text{ \AA}, \theta = 36^\circ) = -446.25236$ Hartree] and 330° . The curve is again nearly symmetric (with the exception of the neighborhood of the minimum at 330° ; see Fig. 1(b)) with respect to reflection about 180° .

In the case of the *G* stack there are again two energetically similar local minima [see Fig. 2(b)] at 36° and 270° , respectively, and a deeper global minimum at

TABLE I. The valance and conduction bands of an adenine stack, $d = 3.36 \text{ \AA}$, $\theta = 36^\circ$, larger basis set DNA B geometry (in eV).

	Lower Limit	Upper Limit	Width
Valence band	$-4.221(K=0)^a$	$-3.637(K=\pi)$	0.584
Conduction band	$-0.4878(K=0)$	$-0.211(K=\pi)$	0.277
Gap		3.733	

^a k is the quasi-wave-vector.

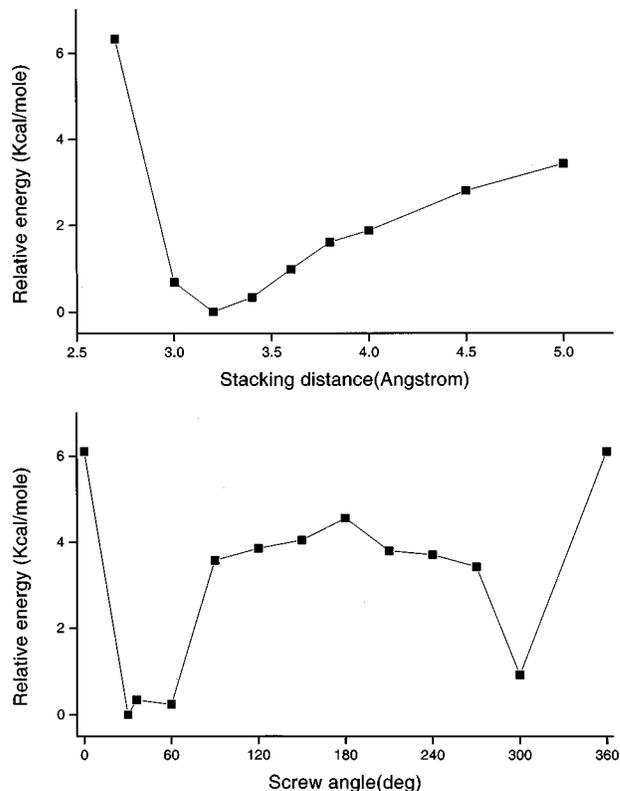


FIG. 3. (a) Total energy per unit cell as a function of stacking distance at a screw angle of 36° , of a uracil stack—screw axis as in DNA B. (b) Total energy per unit cell as a function of stacking distance at a screw angle of 36° , of a uracil stack—screw axis as in DNA B.

330° [by about $0.002 \text{ a.u.} \approx 0.05 \text{ eV}$, $E(d = 3.36 \text{ \AA}, \theta = 330^\circ) = -533.14588$ Hartree]. In this case the curve does not show a symmetry for a reflection about 180° . On the other hand, the *C* stack has two equivalent minima at 46° [$E(d = 3.36 \text{ \AA}, \theta = 46^\circ) = -388.01767$ Hartree] and 315° , and shows perfect symmetry for reflection at 180° [Fig. 2(b)]. Finally, the *U* stack [Fig. 3(b)] has a deeper minimum at 30° [$E(d = 3.36 \text{ \AA}, \theta = 30^\circ) = -407.76499$ Hartree] and two shallow ones at 60° and 300° , respectively. The *E/N* curve is again symmetric for reflection through a plane perpendicular to the curve (plane of the paper) at 180° .

If one compares Fig. 1(b) with Fig. 3(b) (*T* and *U* stacks, respectively), one can see that the forms of the shallow curves are strongly similar, only in the case of the *T* stack due to the substitution by the CH_3 group of an H atom of *U*, the *E/N* values are shifted to larger negative values by 39 a.u. s (the same is true for the curve which gives the stacking distance difference at 36° screw angle).

Turning to the band structures of the five stacks given in Tables I–V, one can notice that in all of the five cases the

TABLE II. The valance and conduction bands of a thymine stack, $d = 3.36 \text{ \AA}$, $\theta = 36^\circ$, larger basis set DNA B geometry (in eV).

	Lower Limit	Upper Limit	Width
Valence band	$-4.643(K=0)$	$-4.356(K=\pi)$	0.287
Conduction band	$-1.175(K=\pi)$	$-1.021(K=0)$	0.154
Gap		3.127	

TABLE III. The valence and conduction bands of a guanine stack, $d = 3.36 \text{ \AA}$, $\theta = 36^\circ$, larger basis set DNA B geometry (in eV).

	Lower Limit	Upper Limit	Width
Valence band	$-3.934(K=0)$	$-3.652(K=\pi)$	0.282
Conduction band	$-0.462(K=0)$	$-0.270(K=\pi)$	0.192
Gap		3.190	

main features are the same: (1) the valence (highest filled bands) lies between -3.6 and -4.6 eV; (2) the position of the conduction (lowest unfilled) bands are between -0.2 and -1.2 eV; (3) the bandwidths are between 0.6 and 0.03 eV. These values lie essentially higher for the valence bands and lower for the conduction bands, respectively, than those obtained previously using the Hartree-Fock (HF) method and correcting the band structures at the MP2 level for correlation using the electronic polaron model (see the Introduction in Ref. 10). The bandwidths obtained are not essentially different in average (though not in the individual cases).

Due to the higher lying conduction bands and lower lying valence bands in the case of HF+MP2, the fundamental gap is the too large with values between 8.6 and 9.5 eV. The present density functional calculation, as is usual in the case of LDA computations, has given too small gap values (values between 3.1 and 3.7 eV). The estimated experimental gap values [which are not equal with the onset of the ultraviolet (UV) spectra] of the 4 DNA base stacks are between 5.0 and 6.0 eV.

IV. CONCLUSION

This first LDA calculation of the nucleotide base stacks has shown mostly the right behavior of the total energy per unit cell as a function of the stacking distance d and the screw angle θ .

On the other hand, the band structure calculations of the stacks have given less satisfactory results. Though the widths of the valence and conduction bands are similar to the HF+MP2 bandwidths, their positions are not satisfactory. The valence bands lie too high and the conduction bands too low, resulting in fundamental gap values nearly by a factor of 2 less than the estimated experimental values.

To improve the LDA band structures, several options could be attempted. In the HF and MP2 calculation, the gap dropped from 8.7 to 6.6 eV, when two sets of non-atom-centered p orbital basis functions were included equidistant between the stacks.¹⁰ One could try this procedure also in the LDA case with supplemented auxiliary functions (in the case of a stack there is a large empty space between the bases

TABLE IV. The valence and conduction bands of a cytosine stack, $d = 3.36 \text{ \AA}$, $\theta = 36^\circ$, larger basis set DNA B geometry (in eV).

	Lower Limit	Upper Limit	Width
Valence band	$-3.897(K=0)$	$-3.859(K=\pi)$	0.038
Conduction band	$-0.522(K=\pi)$	$-0.487(K=0)$	0.035
Gap		3.337	

TABLE V. The valence and conduction bands of a uracil stack, $d = 3.36 \text{ \AA}$, $\theta = 36^\circ$, larger basis set DNA B geometry (in eV).

	Lower Limit	Upper Limit	Width
Valence band	$-4.609(K=0)$	$-4.551(K=\pi)$	0.058
Conduction band	$-1.205(K=\pi)$	$-0.950(K=0)$	0.255
Gap		3.346	

which is not described by Gaussians on the real molecule. An attempt to put different d functions on the bases hardly improved the gap¹⁰.

Another possibility would be to use a more sophisticated form of the density functional method developed to describe excited states, such as the optimal effective potential (OEP) method.²³ Of course, it is not clear at all whether this method would really improve significantly the band structure, and with it, the gap.

Certainly different additional calculations are needed to obtain a realistic band structure for complicated molecular stacks as the nucleotide base stacks.

ACKNOWLEDGMENTS

We should like to express our gratitude to Drs. F. Bartha and F. Bogar for fruitful discussions. We are further indebted to the Flemish Science Foundation (Grant No. G 2131,94) and to the Concerted Action of the University of Antwerp (Grant No. GOA-BOF-UA 23). J.W.M. acknowledges support by the Office of Naval Research (ONR) through the Naval Research Laboratory and directly through the ONR Chemistry-Physics and Materials Divisions.

¹J. J. Ladik and S. Suhai, Int. J. Quantum Chem., Quantum Biol. Symp. **7**, 181 (1980).

²P. Otto, E. Clementi, and J. Ladik, J. Chem. Phys. **78**, 4547 (1983).

³P. Otto, J. Ladik, G. Corongiu, S. Suhai, and W. Förner, J. Chem. Phys. **77**, 5026 (1982).

⁴Y.-J. Ye, unpublished result.

⁵J. W. Mintmire, in Density Functional Methods in Chemistry, edited by J. Labanowski and J. Andzelm (Springer, New York, 1991, p. 125); J. W. Mintmire, B. I. Dunlap, and C. T. White, Phys. Rev. Lett. **68**, 631 (1992); J. W. Mintmire, D. H. Robertson, and C. T. White, J. Phys. Chem. Solids **54**, 1835 (1993).

⁶M. S. Miao, P. E. Van Camp, V. E. Van Doren, J. J. Ladik, and J. W. Mintmire, Phys. Rev. B **54**, 10430 (1996).

⁷L. S. Cederbaum and W. Domcke, Adv. Chem. Phys. **36**, 205 (1977); C.-M. Liegener, J. Phys. C **18**, 6011 (1985).

⁸C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

⁹Y. Toyozawa, Prog. Theor. Phys. **12**, 422 (1954); S. Suhai, Phys. Rev. B **27**, 3506 (1983); see also, J. J. Ladik, Quantum Theory of Polymers as Solids (Plenum, New York, 1988), Chap. 5.

¹⁰F. Bogar and J. J. Ladik, Chem. Phys. **237**, 273 (1998).

¹¹M. S. Miao, P. E. Van Camp, V. E. Van Doren, J. J. Ladik, and J. W. Mintmire, J. Chem. Phys. **109**, 9623 (1998).

¹²M. S. Miao, P. E. Van Camp, V. E. Van Doren, J. J. Ladik, and J. W. Mintmire, Int. J. Quantum Chem. **64**, 243 (1997).

¹³For a detailed description of these methods and for references, see J. J. Ladik, Quantum Theory of Polymers as Solids (Plenum, New York, 1988), Chap. 4.

¹⁴J. W. Mintmire, J. R. Sabin, and S. B. Trickey, Phys. Rev. B **26**, 1743 (1982).

¹⁵J. W. Mintmire and C. T. White, Phys. Rev. Lett. **50**, 101 (1983); Phys. Rev. B **28**, 3283 (1983); J. W. Mintmire, *ibid.* **39**, 13350 (1989).

¹⁶W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).

¹⁷D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

¹⁸J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

- ¹⁹R. J. Feldman, in *Atlas of Macromolecular Structure on Microfiche (AM-SOM)*, National Institutes of Health, Bethesda, Washington, Md. (1976).
- ²⁰A. Imamura, *J. Chem. Phys.* **52**, 3168 (1970); A. Blumen and C. Merkel, *Phys. Status Solidi B* **83**, 425 (1977); see also, J. Ladik, in *Excited States in Quantum Chemistry*, edited by C. A. Nicolaides and D. R. Beck (Reidel, Dordrecht), 1979, p. 495.
- ²¹*Gaussian Basis Sets for Molecular Calculations*, edited by S. Huzihaga (Elsevier, New York, 1984).
- ²²N. Godbout, D. R. Salahub, J. Andzelm, and E. Wimmer, *Can. J. Chem.* **70**, 560 (1992).
- ²³J. Ladik, A. Sujianto, and P. Otto, *J. Mol. Struct.: THEOCHEM* **228**, 271 (1991).