Ab initio calculations of the electronic structure of helical polymers

ARTICLE *in* JOURNAL OF COMPUTATIONAL CHEMISTRY · SEPTEMBER 2004

Impact Factor: 3.59 · DOI: 10.1002/jcc.540050606

Ab Initio Calculations of the Electronic Structure of Helical Polymers*

J. M. André† and D. P. Vercauteren†

IBM Corporation, Dept. D55, Bldg. 701-2, P.O. Box 390, Poughkeepsie, New York 12602 and National Foundation for Cancer Research, 7315 Wisconsin Ave., Bethesda, Maryland 20814

V. P. Bodart and J. G. Fripiat

Facultès Universitaires de Namur, Laboratoire de Chimie Thèorique Appliquée, B-5000 Namur, Belgium
Received 21 December 1983; accepted 10 April 1984

An *ab initio* self-consistent-field (SCF) algorithm taking into account all the features of the one-dimensional translational periodicity and the helical symmetry is presented. This algorithm includes the long-range correction to the Coulomb potential and is designed to calculate the band structure of periodic one-dimensional polymers (planar or helical). Its efficiency in terms of computing time and numerical accuracy is tested via applications on a $(LiH)_n$ chain, polyethylene, and four conformers of polypropylene.

I. INTRODUCTION

In this article, we describe a self-consistent-field (SCF) *ab initio* program for calculating band structures of helical polymers using Cartesian Gaussian basis sets and taking into account long-range interactions. Standard programs are new routinely available for handling linear polymers and are currently applied in several groups like Budapest,[1] Vienna,[2] Erlangen,[3] Poughkeepsie,[4] and Namur.[5]

A recent and significant improvement in dealing with the huge number of two-electron integrals has been described previously[6] and implemented into a computer program.[7] It turns out that taking into account the only translational symmetry, it is now possible to compute Hartree–Fock (HF) energy band structures of periodic polymers with a high degree of numerical accuracy. In practice, one calculates explicitly all electronic integrals over a short-range domain (i.e., 9–15 repeating units) depending on the system and the basis set delocalization; and the long-range contributions are evaluated via a multipole expansion.[8–10]

Obviously, this approach is only applicable to systems having small unit cells and would become almost prohibitively expensive when applied to polymers of chemical interest in their actual conformation; for instance, isotactic polypropylene $(CH_2—CH(CH_3))_n$ does not assume a zig-zag conformation but instead is a 2*3/1 helix; thus the translational unit cell contains three monomeric units (9 carbons and 18 hydrogens). Even, in a nearest neighbor cell approximation, an *ab initio* calculation would be a formidable task in such conditions.

The natural way of reducing this numerical effort is to consider explicitly the helical symmetry as proposed by Imamura and Fujita,[11] and Blumen and Merkel.[12] In this way, one considers *a priori* the asymmetric unit as the relevant unit for the calculation. The successive units are then obtained by screw displacements of the asymmetric unit, which, at the same time, involves the rotation of the atomic orbital basis sets.

It is thus obvious that the simplest numerical way of solving the problem of helical polymers was to use the floating spherical Gaussian orbitals (FSGO),[13] or the Gaussian lobe function (GLF) approach.[14] In both approaches, only 1*s* Gaussian-type orbitals (GTO) are used. Thus the rotation of the atomic orbitals basis set only involves the calculation of the centers of the *s*-type GTOs; since these are spherically symmetric, no integral transformation is required. This approach for helical polymers is used in some of the previously cited groups.[2,5,12] The *s*-multiple expansion, as modified by the screw symmetry operations, has also been analyzed previously.[15] If conceptually simple, this approach does not lead to the most

*Original proofs for this paper were lost in the mail.
†On leave from Facultés Universitaires de Namur, Laboratoire de Chimie Théorique Appliquée, B-5000 Namur, Belgium.

Journal of Computational Chemistry, Vol. 5, No. 6, 535–547 (1984)

CCC 0192-8651/84/060535-13$04.00

efficient implementation, mainly by the fact that it does not contain naturally the "shell" structure of p orbitals (i.e., same center, same radial dependence but different orientations). Considering the extension to p and d orbitals, it is thus necessary to change the methodology and explicitly consider a Cartesian Gaussian basis.

Two attempts are being made in Poughkeepsie. In the first one,[4] the explicit rotation of p orbitals is performed; the program supplies the transformation of p orbitals in the next-nearest neighbor approximation, but lacks the consideration of more far away interactions and particularly of long-range interactions. A generalization is actually being made.[16]

The other approach is the subject of this article. It combines in a single computer program, applicable to helical polymers, the techniques developed during the last few years in the Namur group. In particular, it performs the general rotation of p orbitals in the integral calculation; it extends the multipole expansion to the helical case and makes use as well of other interesting features such as the use of compact Coulomb integrals[6,7] (drastically reducing the storage problem), the numerically integrated density matrices using 12-point Gaussian–Legendre quadrature,[17] the direct solution of the complex eigenvalue problem using the Householder similarity transformation, the QR algorithm[18] (thus completely avoiding time-consuming $S^{-1/2}$ transformations[19]), and the direct calculation of the first derivative of the energy band[20] required for getting the correct indexing and labelling of bands in the density of states (DOS) calculation.[21] Two versions of our program are actually available upon request: a first one, adapted on an IBM-4341 computer, from Dr. E. Clementi at IBM Poughkeepsie, and a second version, adapted on a DEC 20/60 computer, from our group in Namur.

II. HARTREE–FOCK EQUATIONS FOR HELICAL SYSTEMS

A. General Equations

The common helix designation[22] L^*M/t specifies the number (L) of skeletal atoms in the asymmetric unit of the chain and the number of such asymmetric units (M) per t turns of the helix in one translational repeat. The translational period is denoted by a, and $\tau = a/M$ is the length of the reduced (helical) period which we hereafter call reduced unit cell or simply cell. The origin unit cell O contains Ω nuclei of positions $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\Omega$ and ω atomic basis functions, $\{\chi_p(\mathbf{r} - \mathbf{P})\}$; $1 \le p \le \omega$, specified in a right-handed coordinate system with the Z axis in the direction of periodicity. The whole polymer is generated from this unit by applying screw symmetry operations $S^j(\mathbf{r}) = D(j\alpha)(\mathbf{r}) + j\tau$ where $\alpha = 2\pi t/M$ (or $-2\pi t/M$) for a right- or left-handed helix. $D(\alpha)$ is a matrix representing a clockwise rotation around the Z axis by the angle α; τ is a vector of component τ directed along Z.

In the LCAO theory of helical polymers, the polymeric orbitals $\Psi(k, \mathbf{r})$ for a given k point of the one-dimensional extended Brillouin zone ($-\pi/\tau \le k \le +\pi/\tau$) are expressed as linear combinations of ω atomic Bloch functions, $\Phi(k, \mathbf{r})$:

$$\Psi(k, \mathbf{r}) = \Phi(k, \mathbf{r})\mathbf{C}(k) \tag{1}$$

where a component $\Phi_p(k, \mathbf{r})$ of $\Phi(k)$ is explicitly

$$\begin{aligned}\Phi_p(k, \mathbf{r}) &= \sum_{j=-N}^{+N} \exp(ikj\tau) \\ &\quad \times \chi_p\{D(-j\tau)[\mathbf{r} - S^j(\mathbf{P})]\} \\ &= \sum_j \exp(ikj\tau)\chi_p^j\end{aligned} \tag{2}$$

The set of optimal one-electron polymer orbitals $\Phi(k, \mathbf{r})$ is obtained in the usual way by solving the eigenvalue equations:

$$\mathbf{F}(k)\mathbf{C}(k) = \mathbf{S}(k)\mathbf{C}(k)\mathbf{E}(k) \tag{3}$$

where $\mathbf{E}(k)$ defines the band structures of the polymer, and the Fock and overlap matrices between Bloch orbitals $\mathbf{F}(k)$ and $\mathbf{S}(k)$ are defined by the following equations:

$$\mathbf{F}(k) = \sum_j \exp(ikj\tau)\mathbf{F}(j) \tag{4}$$

$$\mathbf{S}(k) = \sum_j \exp(ikj\tau)\mathbf{S}(j) \tag{5}$$

Let us note from eqs. (4) and (5) that the Fock and overlap matrices between Bloch orbitals $\{\mathbf{F}(k), \mathbf{S}(k)\}$ are the Fourier transforms of the Fock and overlap matrices of the atomic basis set

defined as

$$S_{pq}^{j} = \int \chi_p^0(\mathbf{r})\chi_q^j(\mathbf{r})\,dv \tag{6}$$

$$F_{pq}^{j} = \int \chi_p^0(\mathbf{r})F(\mathbf{r})\chi_q^j(\mathbf{r})\,dv$$

$$= T_{pq}^{j} - \sum_{h=-N}^{N}\sum_{A} Z_A V_{pq}^{j}|_A^h + \sum_{h=-N}^{N}\sum_{l=-N}^{N}\sum_{r}\sum_{s} D_{rs}^{hl}\left[2\left(_{pq}^{j}|_{rs}^{hl}\right) - \left(_{pr}^{h}|_{qs}^{jl}\right)\right]$$

$$N \to \infty;\ 1 \le r, s \le \omega \tag{7}$$

In eq. (7), T_{pq}^{j}, $V_{pq}^{j}|_A^h$, and $\left(_{pq}^{j}|_{rs}^{hl}\right)$ are the kinetic, nuclear attraction, and electron repulsion integrals, respectively, defined as:

$$T_{pq}^{j} = \int \chi_p^0(\mathbf{r})\left(-\tfrac{1}{2}\nabla^2\right)\chi_q^j(\mathbf{r})\,dv \tag{8}$$

$$V_{pq}^{j}|_A^h = \int \chi_p^0(\mathbf{r})|\mathbf{r} - S^h(\mathbf{A})|^{-1}\chi_q^j(r)\,dv \tag{9}$$

$$\left(_{pq}^{j}|_{rs}^{hl}\right) = \iint \chi_p^0(\mathbf{r}_1)\chi_q^j(\mathbf{r}_1)|\mathbf{r}_1 - \mathbf{r}_2|^{-1} \times \chi_r^h(\mathbf{r}_2)\chi_s^l(\mathbf{r}_2)\,dv_1\,dv_2 \tag{10}$$

The LCAO density matrix $D_{rs}^{hl} = D_{rs}^{l-h}$ is defined by the integration over the extended Brillouin zone (BZ):

$$D_{pq}^{j} = L_{\mathrm{RC}}^{-1}\int_{\mathrm{BZ}} \sum_{n=1}^{n_0} C^*_{np}(k)C_{nq}(k)\exp(ikj\tau)\,dk$$

where L_{RC} is the length $2\pi/\tau$ of the BZ and $2n_0$ is the number of electrons per unit cell.

B. The Cutoff Procedure

The summations appearing in equations such as eq. (7) are, in principle, infinite. In practice one is not able to compute an infinite number of polycenter integrals, and there is a need for limiting N to some tractable value N. A crucial point is thus the way in which those lattice summations are truncated. Different strategies are applied in most existing programs; they constitute the basic differences between them and are responsible for different numerical outputs. Note also that the cutoff procedure directly determines the number of two-electron integrals to be computed and thus directly influences the computation time required by a given application.

In the framework of a Nth neighbor approximation, the summations over j in eqs. (4) and (5), obviously, run from $-N$ to $+N$. This corresponds, in the Coulomb part of eq. (7), to take into account an exponential (Gaussian) decrease of the distribution:

$$P_{pq}^{j}(\mathbf{r}_1) = \chi_p^0(\mathbf{r}_1)\chi_q^j(\mathbf{r}_1)$$

symmetrically centered around the origin unit cell (O). Following the same argument, the summation over l would run from $h - N$ to $h + N$ to allow for a symmetric decrease of the distribution:

$$P_{rs}^{hl}(\mathbf{r}_2) = \chi_r^h(\mathbf{r}_2)\chi_s^l(\mathbf{r}_2)$$

around the hth unit cell.

A correct handling of the h summation is more difficult to find; we propose to cut it into three parts:

(*i*) *From* $-N$ *to* $+N$. This determines a short-range domain where all the two-electron integrals are exactly calculated. The largest separation between two atomic orbitals involved in a two-electron integral occurs when $j = N$, $h = -N$, and $l = h - N = -2N$. This short-range domain thus covers $3N + 1$ unit cells.

(*ii*) *From* $-N'$ *to* $-(N + 1)$ *and from* $(N + 1)$ *to* N'. For being consistent with the previous arguments on the decrease of the charge distributions, the actual absolute value of N' must be equal or greater than N. It defines an intermediate-range region where computation of the electrostatic terms can be simplified by using faster asymptotic approximations of the F_ν error functions, involved in the calculations of two-electron integrals between Gaussian functions.

(*iii*) *From* $-\infty$ *to* $-(N' + 1)$ *and from* $(N' + 1)$ *to* ∞. In this long-range region, we make use, as explained in the next section, of a multipole expansion to exactly compute up to infinity the electrostatic repulsion terms between the nonoverlapping charge distributions P_{pq}^{j} and P_{rs}^{hl}. This constitutes an important need and affords properly stabilized and, where required, symmetric matrix elements within a Nth neighbor approximation.

C. The Multipole Expansion for Long-Range Coulomb Interactions in Helical Polymers

As noted previously, in actual polymer calculations, it is obviously impossible to deal with very large values of the number of unit cells since the

two-electron part of an LCAO-CO calculation involves an enormous number of integrals. There is a need for limiting that number of unit cells to some amenable value N. Long-range interactions behave like conditionally and slowly convergent series, and actually are significantly contributing far beyond this number N. An analysis to this problem has solved this question by multipole expansion in the helical case.[15] F_{pq}^{j} [eq. (7)] is cast in the form

$$F_{pq}^{j} = T_{pq}^{j} + \sum_{h=-N'}^{N'} \left\{ \sum_{l=h-N}^{h+N} D_{rs}^{hl}\left(_{pq}^{j}|_{rs}^{hl}\right) - \sum_{A} Z_A V_{pq}^{j}|_{A}^{h} \right\} + L_{pq}^{j}(N') - \tfrac{1}{2} \sum_{h} \sum_{l} \sum_{r} \sum_{s} D_{rs}^{hl}\left(_{pr}^{h}|_{qs}^{jl}\right) \quad (11)$$

with

$$L_{pq}^{j}(N') = \left(\sum_{h=-\infty}^{\infty} - \sum_{h=-N'}^{N'} \right) \times \left\{ - \sum_{A} Z_A V_{pq}^{j}|_{A}^{h} + \sum_{l=h-N}^{h+N} \sum_{r,s} D_{rs}^{hl}\left(_{pq}^{j}|_{rs}^{hl}\right) \right\}$$

The detailed deduction of the multipole expansion gives

$$L_{pq}^{j}(N') = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} U_{pq}^{j(k,l)}(N') \tau^{-(k+l+1)} \quad (12)$$

where

$$U_{pq}^{j(k,l)}(N') = \sum_{m} 2(-1)^{l+m}(k+l)!\left[(k+|m|)!(l+|m|)!\right]^{-1} \times w(k+l+1)\Delta_{k+l+1}(N', m, \alpha) \times M_{p\,q}^{*\,j(k,m)}(0) M^{(l,m)}(0) \quad (13)$$

In the above expressions, $w(k+l+1)$ takes the value 1 for $k+l+1$ odd; $+i$ for $k+l+1$ even and right-handed helix; $-i$ for $k+l+1$ even and a left-handed helix. The functions $\Delta_{k+l+1}(N, m, \alpha)$ are defined by

$$\Delta_{k+l+1}(N, m, \alpha) = \zeta(k+l+1, m\alpha) - \sum_{h=1}^{N} \cos(hm\alpha) h^{-(k+l+1)}$$

if $k+l+1$ is odd

or

$$= \zeta(k+l+1, m\alpha) - \sum_{h=1}^{N} \sin(hm\alpha) \times h^{-(k+l+1)} \text{ if } k+l+1 \text{ is even} \quad (14)$$

and

$$\zeta(k+l+1, m\alpha) = \sum_{h=1}^{\infty} \cos(hm\alpha) h^{-(k+l+1)}$$

if $k+l+1$ is odd

or

$$= \sum_{h=1}^{\infty} \sin(hm\alpha) h^{-(k+l+1)} \text{ if } k+l+1 \text{ is even} \quad (15)$$

The functions $\Delta_{k+l+1}(N, m, \alpha)$ have to be evaluated once since the arguments are only related to the type of helix and to the conditions of the calculations (i.e., the input number of cells N). Capital letter $M^{(l,m)}$ refers to the mth component of the 2^l electric pole related to the charge distributions associated with the orbital product $P_{pq}^{j}(\mathbf{r})$:

$$M_{pq}^{j(k,m)}(0) = \langle \chi_p^0 | r_0^k P_k^{|m|}(\cos\theta_0) \exp(im\Phi_0) | \chi_q^j \rangle \quad (16)$$

or to the total charge (electrons + nuclei) associated with each translational unit:

$$M^{(k,m)} = - \sum_{A} Z_A r_A^k P_k^{|m|}(\cos\theta_A) \exp(im\Phi_A) + \sum_{j=-N}^{N} \sum_{p,q} D_{pq}^{j} \langle \chi_p^0 | r^k P_k^{|m|}(\cos\theta) \times \exp(im\Phi) | \chi_q^j \rangle \quad (17)$$

Explicit formulas for the expansion up to $k+l+1=3$ are given in section III.

D. The Use of Compacted Coulomb Integrals

It has been shown in previous articles[6,7,23,24] that important simplifications can be made on the structure of two-electron matrices by taking into account the full translational symmetry of polymer density matrices. In particular, the Coulomb part of eq. (7) can be shown to be

equivalently rewritten as

$$\sum_{h=-N}^{+N}\sum_{l=h-N}^{h+N}\sum_{r}\sum_{s} D_{rs}^{hl}\left(\begin{smallmatrix} j & hh+l \\ pq & rs \end{smallmatrix}\right) = \sum_{l=-N}^{+N}\sum_{r}\sum_{s} D_{rs}^{l}\sum_{h=-N}^{N}\left(\begin{smallmatrix} j & hl \\ pq & rs \end{smallmatrix}\right) \tag{18}$$

It is striking to note that the h summation does not evolve with the iterations and can be computed and stored at once, reducing the number of stored terms by a factor of $(2N + 1)$, and thus minimizing I/O operations. Furthermore, as will be developed in the next section, the computation of such summations is algorithmically easy due to the presence of many common terms since the charge distributions P_{rs}^{hh+l} are invariant upon translation. Let us note, however, that such summations are divergent and that the divergencies must be removed by proper combinations of electron–nuclei and electron–electron terms. A detailed analysis of the properties of compact summations like eq. (18) have been given elsewhere.[6]

III. SPECIFIC PROGRAMMING CONSIDERATIONS

A. Two-Electron Package

The programming difficulties encountered in the calculations of the electronic properties of polymers are the same as in molecular calculations; examples are the use of contracted gaussians, shells, and fast algorithms for computing two-electron integrals. Elegant solutions to those problems are popularized by programs like IBMOL[25] or GAUSSIAN,[26] and are an integrand part of the coding of polymer programs.

We discuss hereafter specific solutions of the problems encountered in polymer calculations.

The two-electron redundancy is not only orbital-index dependent but also cell-index dependent and is summarized by the following always valid identities:

$$\left(\begin{smallmatrix} j & hl \\ pq & rs \end{smallmatrix}\right) = \left(\begin{smallmatrix} j & lh \\ pq & sr \end{smallmatrix}\right) = \left(\begin{smallmatrix} j & h-jl-j \\ qp & r\ \ s \end{smallmatrix}\right) = \left(\begin{smallmatrix} l-j & h-j \\ qp & s\ \ r \end{smallmatrix}\right)$$
$$= \left(\begin{smallmatrix} l-h & \bar{h}j-h \\ rs & pq \end{smallmatrix}\right) = \left(\begin{smallmatrix} l-h & j-h\bar{h} \\ rs & q\ \ p \end{smallmatrix}\right)$$
$$= \left(\begin{smallmatrix} h-l & \bar{l}\ \ j-l \\ sr & pq \end{smallmatrix}\right) = \left(\begin{smallmatrix} h-l & j-l\bar{l} \\ sr & q\ \ p \end{smallmatrix}\right) \tag{19}$$

so that only one of those is, in principle, to be computed. Furthermore, as seen in section IID, they are used as compacted summations:

$$\delta_{pqrs}(j,l) = \sum_{h}\left(\begin{smallmatrix} j & hh+l \\ pq & rs \end{smallmatrix}\right) \tag{20}$$

for which the following equalities hold (if the h summation is dealt up to infinity):

$$\begin{aligned}\delta_{pqrs}(j,l) &= \delta_{pqsr}(j,-l) = \delta_{qprs}(-j,l)\\ &= \delta_{qpsr}(-j,-l) = \delta_{rspq}(l,j)\\ &= \delta_{rsqp}(l,-j) = \delta_{srpq}(-l,j)\\ &= \delta_{srqp}(-l,-j)\end{aligned} \tag{21}$$

The respective extent of j, h, or l summations has been discussed in section IIB; in particular, the h summation is extended up to infinity by adding, to the short-range domain, an intermediate-range region where fast expansions of F_ν functions are used, and a long-range region where the multipole expansion is applied. In such conditions, the choice taken in our computer program is to only compute and store the integrals needed for the Coulomb and exchange contributions

$$\left(\begin{smallmatrix} j & hh+l \\ pq & rs \end{smallmatrix}\right) \quad \text{and} \quad \delta_{pqrs}(j,l)$$

for which

$$p \geq q \quad \text{and} \quad r \geq s \quad \text{if } p > r$$

or

$$p \geq q \quad \text{and} \quad q \geq s \quad \text{if } p = r$$

and

$$-N \leq j, h, l \leq +N$$

It is important to stress the fact that the accumulation of two-electron integrals given by eq. (20) leads to a very fast algorithm in the two-electron package. Indeed, each two-electron integral can be written as a sum of products of three factors: the first one depending on the charge distribution $\chi_p^0\chi_q^j$, the second one depending only on the charge distribution $\chi_r^0\chi_s^l$, and the last one being the only one to refer to the four-center indices. For example, in the case of s-type orbitals, we have the explicit expression:

$$\left(\begin{smallmatrix} j & hh+l \\ pq & rs \end{smallmatrix}\right) = S_{pq}^{\ j} S_{rs}^{\ l}\left\{\frac{4(\alpha_p+\alpha_q)(\alpha_r+\alpha_s)}{\pi(\alpha_p+\alpha_q+\alpha_r+\alpha_s)}\right\}^{1/2} \times F_0\left\{\frac{(\alpha_p+\alpha_q)(\alpha_r+\alpha_s)}{(\alpha_p+\alpha_q+\alpha_r+\alpha_s)}\left(\mathbf{G}_{pq}^{\ j} - \mathbf{G}_{rs}^{hh+l}\right)^2\right\} \tag{22}$$

In eq. (22), the αs are the Gaussian exponents, F_0 is a reduced form of the incomplete gamma function $[F_0(t) = (1/t^{1/2})\int_0^{t^{1/2}} \exp(-u^2)\,du]$,[27] and $\mathbf{G}_{pq}^{j}$, $\mathbf{G}_{rs}^{hh+l}$, are the barycenters of the respective distributions $\chi_p^0\chi_q^j$ and $\chi_r^h\chi_s^{h+l}$.

It is straightforward to note that the h summation of the compacted Coulomb integrals [eq. (20)] only applies to the F terms; in the previous example of s-type integrals:

$$\delta_{pqrs}(j,l) = \sum_h \left(\begin{smallmatrix} j \\ pq \end{smallmatrix}\middle|\begin{smallmatrix} hh+l \\ rs \end{smallmatrix}\right)$$

$$= S_{pq}^{j} S_{rs}^{l} \left\{ \frac{4(\alpha_p + \alpha_q)(\alpha_r + \alpha_s)}{\pi(\alpha_p + \alpha_q + \alpha_r + \alpha_s)} \right\}^{1/2}$$

$$\times \sum_h F_0\left\{ \frac{(\alpha_p + \alpha_q)(\alpha_r + \alpha_s)}{(\alpha_p + \alpha_q + \alpha_r + \alpha_s)} \left(\mathbf{G}_{pq}^{j} - \mathbf{G}_{rs}^{hh+l}\right)^2 \right\} \tag{23}$$

Taking into account the discussion of Section IIB, eq. (23) leads to a very fast algorithm since the h summation is divided into three parts:

(i) a very short-range region where the F_0 function is calculated by usual interpolation,

(ii) an intermediate region where the F_0 function is precisely approximated by its asymptotic form[28]:

$$F_0(x) = \tfrac{1}{2}(\pi/x)^{1/2} \qquad \text{if } x \geq 16$$

It can be easily shown (see Appendix 1) that with usual basis sets this approximation is already very precise for small distances. For instance, in the case of Clementi's $7s/3p$ basis set[29] for carbon, the lowest exponent is 0.167, and thus, in this less favorable case, the approximation starts to be correct at distances as short as 5.2 Å.

As already pointed out, the equalities [eq. (21)] strictly hold only for infinite summations and thus, if the long-range interactions are included, the explicit form of the first terms (limited to $k + l + 1 = 3$ in order to keep the invariance with respect to the origin[8]) can be deduced from eq. (12) and gives for the explicit working expansion of eq. (20):

$$\delta_{pqrs}(j,l) = \sum_{h=-N}^{N} \left(\begin{smallmatrix} j \\ pq \end{smallmatrix}\middle|\begin{smallmatrix} hh+l \\ rs \end{smallmatrix}\right) \quad \text{short-range region}$$

$$+ \left(\sum_{-N'}^{-N-1} + \sum_{N+1}^{N'} \right) \left(\begin{smallmatrix} j \\ pq \end{smallmatrix}\middle|\begin{smallmatrix} hh+l \\ rs \end{smallmatrix}\right)$$

intermediate-range region

$$+ \Delta_3(N',0,\alpha)/\tau^3 \left\{ S_{pq}^{j} M_{rs}^{l(2)}[3z^2 - r^2] + S_{rs}^{l} M_{pq}^{j(2)}[3z^2 - r^2] - 4M_{pq}^{j(1)}[z]\,M_{rs}^{l(1)}[z] \right\}$$

$$+ \Delta_3(N',1,\alpha)/\tau^3 \left\{ 2M_{pq}^{j(1)}[x]\,M_{rs}^{l(1)}[x] + 2M_{pq}^{j(1)}[y]\,M_{rs}^{l(1)}[y] \right\}$$

long-range region (24)

when $\Delta_3(N',m,\alpha)$ is defined by eq. (14) and $M_{pq}^{j(k)}$ or $M_{rs}^{l(k)}(\Sigma_{\alpha\beta\gamma} a_{\alpha\beta\gamma} x^\alpha y^\beta z^\gamma;\ \alpha+\beta+\gamma = k)$ are the 2^k-pole moments of the $\chi_p^0\chi_q^j$ or $\chi_r^0\chi_s^l$ distributions. The last point worth mentioning is that some electron integrals have to be rotated to take into account the helical symmetry as seen from eq. (2). The s- and p_z-type orbitals are invariant under the helical rotation. The $\bar{p}_x$- and $\bar{p}_y$-rotated orbitals are obtained from cartesian p_x and p_y ones by the following relations:

$$\bar{p}_x^j = p_x^j \cos(j\alpha) + p_y^j \sin(j\alpha)$$

$$\bar{p}_y^j = -p_x^j \sin(j\alpha) + p_y^j \cos(j\alpha) \tag{25}$$

Rotated one- and two-electron integrals are then obtained as linear combinations of Cartesian ones.

B. SCF Package

The complex eigenvalue problem [eq. (3)] is solved at 12 nonequidistant k points in half the Brillouin zone. Those 12 k points are selected by the 12-points Gauss–Legendre quadrature to numerically integrate the density matrices elements

$$D_{pq}^{j} = L_{\mathrm{RC}}^{-1} \int_{\mathrm{BZ}} \sum_{n=1}^{n_0} C_{np}^{*}(k) C_{nq}(k) \exp(ikj\tau)\,dk$$

$$= \sum_{l=1}^{12} A_l f(x_l) \tag{26}$$

The 12 nodes x_l are the roots of the Legendre polynomial of degree 12. The weights $\{A_l\}$ and the roots $\{x_l\}$ are listed for half the Brillouin zone in Appendix 2. The diagonalization procedure is made by using the efficient and self-contained CBORIS routine.[18] This routine diagonalizes complex hermitian matrices using the Householder similarity transformation and the QR algorithm. It works in a way that completely avoids time consuming $S^{-1/2}$ transformations.

Once self-consistency is reached, the complex eigenvalue equations are solved at 21 equidistant k points in half the Brillouin zone. For each energy value, $E_n(k)$, the first derivative $E_n'(k)$ is directly and locally computed according to a previously developed algorithm.[20]

$$E_n'(k) = \mathbf{C}_n^{\dagger}(k)[\mathbf{F}'(k) - E_n(k)\mathbf{S}'(k)]\mathbf{C}_n(k) \tag{27}$$

$\mathbf{F}'(k)$ and $\mathbf{S}'(k)$ are the first k derivative of the Hamiltonian and overlap matrices, respectively; they are easily computed since they are expressed as trigonometric series [see eqs. (4) and (5)]

$$\mathbf{F}'(k) = \frac{d}{dk}\mathbf{F}(k) = i\tau\sum_j j\exp(ikj\tau)\mathbf{F}(j)$$

$$\mathbf{S}'(k) = \frac{d}{dk}\mathbf{S}(k) = i\tau\sum_j j\exp(ikj\tau)\mathbf{S}(j) \tag{28}$$

This is related to the band indexing problem. Due to the large computing time associated with diagonalizations, the energy bands are known at discrete wave number values only, $E_n(k)$ (as already mentioned, 21 k points in our case). Bands are usually given an index, n, increasing in order of increasing energies, but difficulties appear because the rule breaks down at degeneracy points where crossing could arise. For k points close in energy but somewhat distant in the reciprocal space, it is sometimes difficult to make a choice between a crossing or noncrossing situation. Obviously, the knowledge of the derivatives of the energy bands helps to solve this question.

As soon as the correct indexing and labeling of the electronic energy bands is obtained, the density of electronic states (DOS) is calculated and plotted following the methodology developed in our group by Delhalle and Delhalle.[21] Examples of such plotting results are given in section IVC.

IV. TEST APPLICATIONS

A. Model $(LiH)_n$

The *ab initio* program has been tested in the case of a model chain of a linear arrangement of LiH molecules [r(LiH) = 4 a.u., r(HLi) = 6 a.u., τ = 10 a.u.] using simple three-Gaussian expansions of s and p orbitals. The LiH unit cell is a standard testing case since it has a strongly enough polar unit cell to check the capability of the multipole expansion used and, furthermore, has a simple enough unit cell to allow a direct comparison with molecular results. Table I gives the numerical results as obtained from the IBMOL Poughkeepsie program on molecular clusters of $(LiH)_n$, for odd ns ranging from 1 to 29, compared with the results of our infinite polymer *ab initio* program. Examination of this table calls for the following comments. As expected, the relative cluster energies show very poor convergence due to the role of the end effects. The difference of relative energies between $n = 27$ and $n = 29$ is still of the order of 10^{-4} a.u. The energy increment is more rapidly converged; it is 2.10^{-6} a.u. between $n = 27$ and $n = 29$. The energy and charge results (the H atomic net charge presented in the case of the molecular clusters correspond to the more internal of the chain) of our polymer program fit perfectly well with the extrapolated molecular results. It must be stressed at this point that the polymer results are obtained in a few seconds of CPU time on an IBM 4341 computer while the molecular calculation for the 29 units cluster requires more than 4 h of CPU time on the same computer.

In Table II the convergence properties of the polymer program, with respect to the sizes of the short-range region (N) and of the intermediate region (N') including the $k + l + 1 = 3$ multipole corrections, are given. It is observed that the convergence is extremely fast. For the sake of comparison, we compare in Figure 1 our convergence rate with other published results[2] presented at the same relative scales (the basis sets being not the same). This convergence rate turns out to be an excellent achievement of the computer strategy we have developed.

B. Polyethylene

The *ab initio* program has also been tested in the case of linear polyethylene in its experimental stable geometry [r(CC) = 1.09 a.u., r(CH) = 1.54 a.u., $\hat{CCC}$ = 109.5°, $\hat{HCH}$ = 109.5°]. The interest of polyethylene is that it can be considered as a zig-zag linear polymer with a $(CH_2—CH_2)$ unit cell or as a 1*2/1 helix with an elementary asymmetric unit (CH_2). The basis set used is Clementi's $7s/3p$ for carbon and $4s$ for hydrogen, detailed in the Appendix 1. Table III provides the dependence of the energy results on the sizes of the short-range region (N) and of

Table I. Energy results and H atomic net charge for clusters of LiH (in a.u.).

n	E_n	E_n/n	$\Delta E_{n+1}=E_{n+1}-E_n$	q_H
Molecular Clusters				
1	-7.817840	-7.817840		0.0104
3	-23.493153	-7.831051	-7.837657	0.0474
5	-39.174549	-7.834910	-7.840698	0.0543
7	-54.856869	-7.836696	-7.841160	0.0563
9	-70.539469	-7.837719	-7.841300	0.0571
11	-86.222190	-7.838381	-7.841361	0.0574
13	-101.904974	-7.838844	-7.841392	0.0577
15	-117.587793	-7.839186	-7.841410	0.0578
17	-133.270635	-7.839449	-7.841421	0.0579
19	-148.953491	-7.839657	-7.841428	0.0579
21	-164.636357	-7.839827	-7.841433	0.0580
23	-180.319229	-7.839966	-7.841436	0.0580
25	-196.002107	-7.840084	-7.841439	0.0580
27	-211.684987	-7.840185	-7.841440	0.0581
29	-227.367871	-7.840271	-7.841442	0.0581
Infinite Polymer				
∞		-7.841449		0.0581

Table II. Convergence properties of the energy per unit cell with respect to the sizes of the short-range region (N unit cells) and of the intermediate region (N' unit cells); $k + l + 1 = 3$ multipole corrections included. Linear chain of $(LiH)_n$. Results in a.u.

$N \backslash N'$	5	9	13	17	25	49	99
3	-7.841404			-7.841450	-7.841451	-7.841451	-7.841451
5		-7.841445		-7.841449	-7.841449	-7.841449	-7.841449
7			-7.841448	-7.841449	-7.841449	-7.841449	-7.841449
9				-7.841449	-7.841449	-7.841449	-7.841449

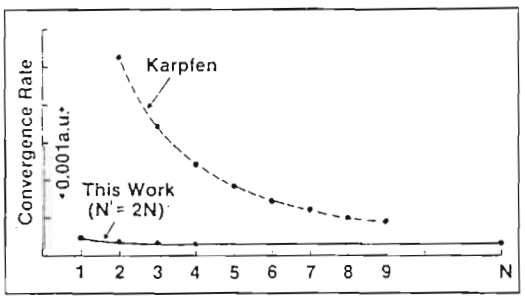

Figure 1. Comparison of our convergence rate of the total energy of a $(LiH)_n$ chain, in function of the number of cells (N), with other published results.

the intermediate-range region (N'), in both the linear and helical assumptions. Figure 2 compares the rate of convergence with respect to the number of cells (N) obtained by our scheme and by other polymer programs.[2] It is clear from that figure that our scheme is a net improvement for getting accurate numerical results without requiring prohibitive computation times. The results obtained by the two different methodologies, the polyethylene considered as a zig-zag polymer or as a helix, are in complete agreement. It is to be noted that an agreement to the last

Table III. Energy results for polyethylene (a.u.).

Unit cell: (CH_2-CH_2)

N	N'	E per $(CH_2)_2$	E per (CH_2)
5	9	-77.759621	-38.879810
5	25	-77.759609	-38.879804
7	25	-77.759546	-38.879773
9	25	-77.759544	-38.879772
11	25	-77.759544	-38.879772

Unit cell: (CH_2)

N	N'	E per (CH_2)
5	9	-38.866876
9	17	-38.879876
11	21	-38.879806
11	37	-38.879786
11	49	-38.879784
13	49	-38.879775
17	49	-38.879773
19	37	-38.879775
19	49	-38.879773

digit cannot be obtained. Firstly, since the sizes of the short-range and of the intermediate-range regions must cover an odd number of unit cells, and since the size of the reduced helical cell is just half the size of the linear one, we are unable to cover exactly the same domain in the integral expansion. Furthermore, the multipole ability will also be different since the $k + l + 1 = 3$ multipole expansion is used to represent in the first case a $(CH_2—CH_2)$ electron density of twice the size of the (CH_2) density of the helical case. The band structures of polyethylene are represented in Figure 3 for the linear case (upper part of the figure) and for the helical case (lower part). Again, the two band structures are in complete agreement; the linear case is treated with a unit cell of twice the size of the helical case, so that, in the reciprocal space, the linear polyethylene has a reciprocal unit cell and hence a first Brillouin zone of half the size of the helical case. The agreement is better seen by

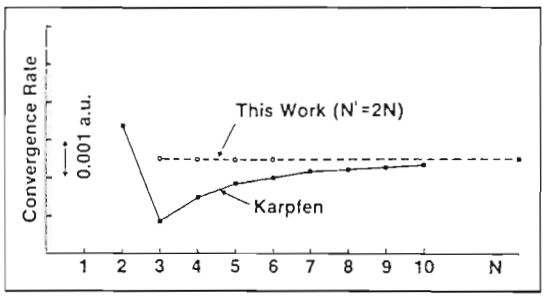

Figure 2. Comparison of our convergence rate of the total energy of polyethylene, in function of the number of cells (N), with other published results.

unfolding, in the linear case, the energy bands according to the symmetry requirements.

C. Polypropylene

The previous sections show the capability of our methodology and of our computer program to produce in reasonable computing time *ab initio* band structures of linear and helical polymers.

We were thus led to investigate a real case which was not affordable by previously available *ab initio* polymer programs. We have chosen the case of the single-methyl substituted polyethyl

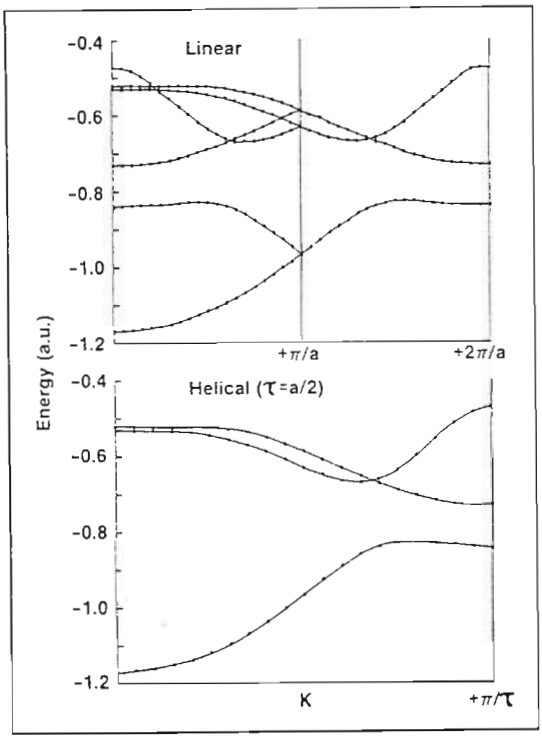

Figure 3. Band structure of polyethylene considered as (a) (upper part) a zig-zag linear chain with a $(CH_2—CH_2)$ unit cell, and (b) (lower part) a helical $1^*2/1$ polymer with a (CH_2) unit cell.

ene, polypropylene, which is known to exist in iso- and syndiotactic arrangements. In the case of an isotactic polypropylene, the geometrical arrangement is that of an isoclined 2*3/1 helix.[30,31] In the case of syndiotactic forms, both a zig-zag[32] and a two-order helical form[33] are experimentally observed; the zig-zag form being the most stable obtained after dissolving the helical one. Let us note here that the isotactic helical case with three [CH_2—$CH(CH_3)$] units in the unit cell was a case out of reach from previous programs.

The calculations have been made using Clementi's $7s/3p$ basis for carbon and $4s$ for hydrogen (see Appendix 1). In the syndiotactic case, both for the zig-zag and the helical conformations, the reduced unit cell contains two [CH_2—$CH(CH_3)$] units; the short-range region and the intermediate one were composed of three and five unit cells, respectively. All two-electron integrals have thus been calculated between 10 propylene units (30 carbon atoms + 60 hydrogens). In the isotactic case, the reduced unit contains only a single [CH_2—$CH(CH_3)$] and the short- and intermediate-range regions were composed of five and nine cells, respectively (thus corresponding to calculations over 27 carbon atoms and 54 hydrogens).

The energy results {per [CH_2—$CH(CH_3)$] unit} and the relative stabilization of each species are presented in Table IV. They are in close agreement with chemical intuition and experimental observations; the syndiotactic forms are less stable than the helical isotactic one due to less steric interactions, the zig-zag conformation being stabilized with respect to the helical form by 7.6 kcal/mol per [CH_2—$CH(CH_3)$] unit. For the isotactic forms, the three-order helix is the stable form; due to strong steric interactions between the methyl groups, the zig-zag form is very unstable {125.4 kcal/mol per [CH_2—$CH(CH_3)$] unit}. The density of states of the four forms are presented in Figure 4. Assuming comparable contributions of the $2s$- and $2p$-atomic function in the one-electron states for the various conformations, convoluted density of states could be considered as a reasonable approximation of the possible UPS spectrum. A

Table IV. Energy results and relative stabilities for syndio- and isotactic forms of polypropylene per [CH_2—$CH(CH_3)$] unit.

	E(a.u.)	ΔE(kcal/mol)
Isotactic helix 3	-116.662298	0.00
Syndiotactic zig-zag	-116.648104	8.90
Syndiotactic helix 2	-116.635947	16.52
Isotactic zig-zag	-116.462304	125.39

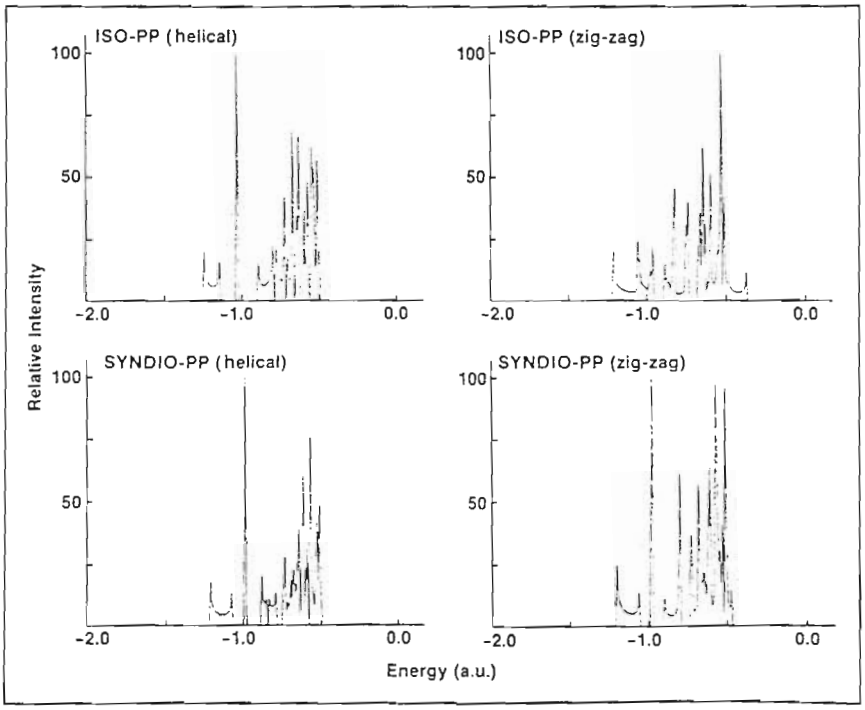

Figure 4. Density of states of four forms of polypropylene: (a) (upper part) isotactic in its helical or zig-zag conformation, and (b) (lower part) syndiotactic in its helical or zig-zag conformation.

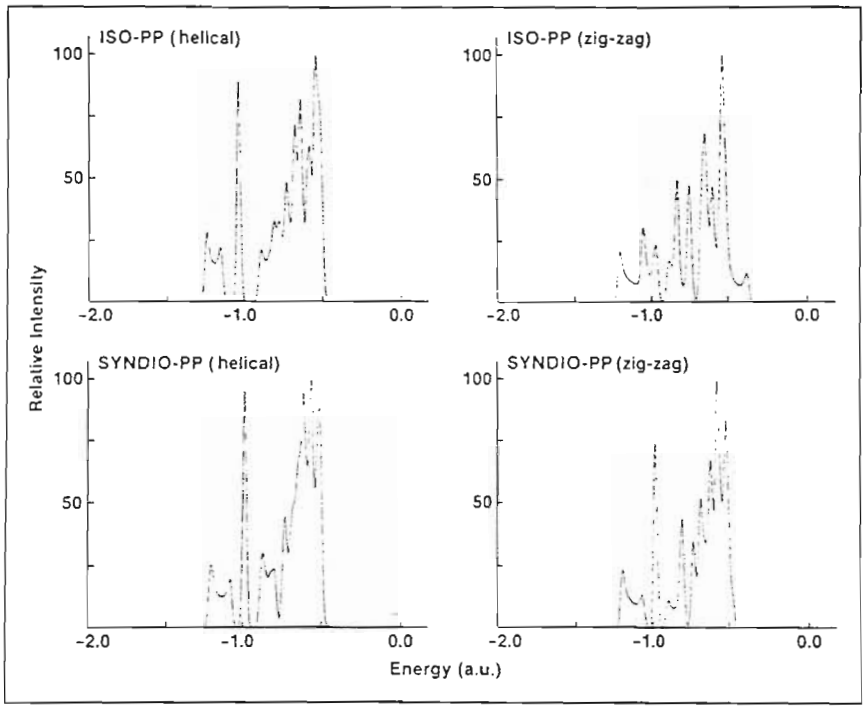

Figure 5. Convolution of the density of states of the four forms of polypropylene presented in Figure 4: (a) (upper part) isotactic in its helical or zig-zag conformation, and (b) (lower part) syndiotactic in its helical or zig-zag conformation.

density of states convoluted by a Gaussian of 0.7 eV is presented in Figure 5. Actually, as far as we know, no UPS spectra of polypropylene are available; only, some XPS spectra of isotactic polypropylene is actually published.[34] The XPS spectra differs from the UPS spectra only by the fact that the $2p$-dominated peaks (top of valence bands) are much lower in intensity due to weaker cross sections of the $2p$ orbitals compared with the $2s$ ones ($\sigma_{2s}/\sigma_{2p} \simeq 12$). The agreement between our density of states and the experimental spectrum is excellent. The calculations confirm the idea that conformational changes produce significant modifications in the DOS spectrum, and that the theoretical study of valence bands is an excellent complement to XPS and UPS spectroscopies.

CONCLUSIONS

In this article we have shown that the methodology developed in the last 15 years in several groups for computing electronic properties of linear and helical polymers is now implemented in efficient computer programs, and that polymers of true chemical or physical interest are now open to routine theoretical investigations.

We have presented examples on lithium hydride chain, polyethylene, and polypropylene. The program has also been applied successfully to the electronic structure of polypyrrole in relation with its high conductivity properties upon doping[35] and to some polyglycines.[36] As the program includes the long-range correction to the Coulomb potential, it is designed to treat polar systems as well and not restricted to hydrocarbon polymers.

Our calculations also demonstrate the actual interest of band structure calculations; the influence of slight conformational changes in the DOS of polypropylenes clearly shows that chemical intuition is not enough for getting a good idea of the electronic structure of polymers but that explicit rigorous calculations are required.

V.P.B. would like to thank the Belgian Institute for Scientific Research in Industry and Agriculture (IRSIA) for a research fellowship. J.M.A. is indebted to the Belgian National Fund for Scientific Research (FNRS) for financial support. J.M.A. and D.P.V. wish to thank the National Foundation for Cancer Research (NFCR) for the financial support of this work, and IBM-Poughkeepsie and IBM-Belgium for making possible their stay in the States. They also would like to express their sincere gratitude to Dr. E. Clementi for fruitful discussions and for allowing them to attend the stimulating atmosphere of Department D55. J.M.A. acknowledges also the support of NATO contract 130.82 on molecular design of materials for nonlinear optics.

APPENDIX 1

Clementi's minimal basis for carbon and hydrogen.[29] Exponents (α), coefficients (c), and minimal distances (d) where the asymptotic approximation of the F_γ function can be turned on (see section IIIA).

		α(a.u.$^{-2}$)	c	d(Å)
C	1s	1267.183972	0.005489	0.06
		190.603978	0.040625	0.15
		43.247698	0.181793	0.32
		11.964903	0.458558	0.61
		3.663116	0.457314	1.11
	2s	0.539158	0.410472	2.88
		0.167130	0.645265	5.18
	2p	4.187345	0.112119	1.03
		0.854053	0.468083	2.17
		0.199770	0.620758	4.73
H	1s	13.013372	0.019678	0.59
		1.962496	0.137952	1.51
		0.444569	0.478313	3.17
		0.121953	0.501131	6.06

APPENDIX 2

Weights and roots of the Legendre polynomial of degree 12 as needed for numerical integration in half the Brillouin zone.

x_1(radians)	A_1
0.0289 6448 7993	0.0235 8766 8193
0.1506 1226 1496	0.0534 6966 2997
0.3614 3603 4181	0.0800 3916 4272
0.6482 3944 1541	0.1015 8371 3361
0.9930 0795 9289	0.1167 4626 8269
1.3740 8014 8713	0.1245 7352 2906
1.7675 1250 4877	0.1245 7352 2906
2.1485 8469 4301	0.1167 4626 8269
2.4933 5321 2049	0.1015 8371 3361
2.7801 5661 9408	0.0800 3916 4272
2.9909 8039 2093	0.0534 6966 2997
3.1126 2816 5596	0.0235 8766 8193

References

1. M. Kertesz, *Acta Phys. Acad. Sci. Hung.*, **41**, 127 (1976).
2. A. Karpfen, *Int. J. Quantum Chem.*, **19**, 1207 (1981).
3. S. Suhai and J. Ladik, *Solid State Commun.*, **22**, 227 (1977).
4. P. Otto, E. Clementi, and J. Ladik, *J. Chem. Phys.*, **78**, 4547 (1983).
5. J. L. Brédas, J. M. André, J. G. Fripiat, and J. Delhalle, *Gazz. Chim. Ital.*, **108**, 307 (1978).
6. J. M. André, D. P. Vercauteren, and J. G. Fripiat, *J. Comput. Chem.*, in press.
7. J. M. André, V. P. Bodart, J. L. Brédas, J. Delhalle, and J. G. Fripiat, in *Quantum Theory of Polymers, Solid State Aspects*, J. M. André and J. Ladik, Eds., D. Reidel, 1984; see also *Documentation for an ab initio polymer program (PLH)*, IBM Technical Report POK-28, 1983.
8. L. Piela and J. Delhalle, *Int. J. Quantum Chem.*, **13**, 605 (1978).
9. L. Piela, in *Recent Advances in the Quantum Theory of Polymers*, J. M. André, J. L. Brédas, J. Delhalle, J. Ladik, G. Leroy, and C. Moser, Eds., Lecture Notes in Physics, Vol. 113, Springer-Verlag, Heidelberg, 1980, p. 104.
10. J. Delhalle, L. Piela, J. L. Brédas, and J. M. André, *Phys. Rev. B*, **22**, 6254 (1980).
11. A. Imamura and H. Fujita, *J. Chem. Phys.*, **61**, 115 (1974).
12. A. Blumen and C. Merkel, *Phys. Status Solidi*, **B83**, 425 (1977).
13. A. A. Frost, *J. Chem. Phys.*, **47**, 3707 (1967).
14. J. Whitten, *J. Chem. Phys.*, **39**, 349 (1963).
15. L. Piela, J. M. André, J. L. Brédas, and J. Delhalle, *Int. J. Quantum Chem.*, **S14**, 405 (1980).
16. P. Otto, personal communication.
17. J. M. André, J. Delhalle, C. Demanet, and M. E. Lambert-Gerard, *Int. J. Quantum Chem.*, **S10**, 99 (1976).
18. J. Zupan and E. Zakrajšek, *Ann. Soc. Scient. Brux.*, **89**, 337 (1975).
19. J. Del Re, J. Ladik, and G. Biczo, *Phys. Rev.*, **155**, 997 (1967).
20. J. M. André, J. Delhalle, G. Kapsomenos, and G. Leroy, *Chem. Phys. Lett.*, **14**, 485 (1972).
21. J. Delhalle and S. Delhalle, *Int. J. Quantum Chem.*, **11**, 349 (1977).
22. B. Wunderlich, *Macromolecular Physics*, Academic, New York, 1973, Vol. I.
23. C. Pisani and R. Dovesi, *Int. J. Quantum. Chem.*, **17**, 501 (1980).
24. R. Dovesi, C. Pisani, and C. Roetti, *Int. J. Quantum Chem.*, **17**, 517 (1980).
25. E. Ortoleva, G. Castiglione, and E. Clementi, *Comput. Phys. Commun.*, **19**, 337 (1980).

26. J. S. Binkley, R. A. Whiteside, R. Krishman, R. Seeger, D. J. Defrees, H. B. Schlegel, S. Topiol, L. R. Kahn, and J. A. Pople, *Quantum Chemistry Program Exchange*, **13**, 406 (1981).
27. I. Shavitt, in *Methods of Computational Physics*, B. Adler, S. Fernbach, and M. Rotenberg, Eds., Academic, New York, 1963, Vol. 2, p. 1.
28. L. J. Schaad and G. O. Morrell, *J. Chem. Phys.*, **54**, 1965 (1971).
29. L. Gianolo, R. Pavani, and E. Clementi, *Gazz. Chim. Ital.*, **108**, 181 (1978).
30. G. Natta, P. Corradini, and P. Ganis, *Makromol. Chem.*, **39**, 238 (1960).
31. G. Natta and P. Corradini, *Nuovo Cimento*, **15**, 40 (1960).
32. G. Natta, M. Peraldo, and G. Allegra, *Makromol. Chem.*, **75**, 215 (1964).
33. G. Natta, I. Pasquon, and A. Zambelli, *J. Am. Chem. Soc.*, **84**, 1488 (1962).
34. J. M. André, J. Delhalle, and J. J. Pireaux, *ACS Symposium Series*, **162**, 151 (1981).
35. J. M. André, D. P. Vercauteren, G. B. Street, and J. L. Brédas, *J. Chem. Phys.*, **80**, 5643 (1984).
36. J. M. André, D. P. Vercauteren, and E. Clementi, to be submitted.