First-principles study of the electronic and molecular structure of protein nanotubes

Hajime Okamoto
Department of Materials Science and Engineering, Waseda University, Tokyo 169-8555, Japan

Kyozaburo Takeda
Department of Materials Science and Engineering and Advanced Research Institute for Science and Engineering, Waseda University, Tokyo 169-8555, Japan

Kenji Shiraishi
Institute of Physics, University of Tsukuba, Ibaraki 305-8577, Japan

Received 14 November 2000; revised manuscript received 5 April 2001; published 31 August 2001

The electronic and molecular structures of protein nanotubes (PNT’s) have been investigated theoretically by first-principles electronic structure calculations. The results have been discussed in comparison to those of the polypeptide open chains (POC’s) and polypeptide closed rings (PCR’s) in order to give a systematic understanding. Focusing on the intra-ring and inter-ring hydrogen bonds (HB’s), we also investigate the PCR stacking mechanism. The present calculation reveals that PNT’s are semiconductors and that an extra proton in the tube interior has the potential to be an electron acceptor.

DOI: 10.1103/PhysRevB.64.115425

I. INTRODUCTION

The recent advances towards the development of self-assembling polypeptide closed rings (PCR’s) made it possible to produce an entirely new class of artificial proteins. These nanoscopic proteins are referred to as the protein nanotubes (PNT’s) because they have open-ended hollow tubular structures (Fig. 1). Ghadiri et al. have succeeded in synthesizing many types of PNT’s by varying the number \( n \) and choosing amino acid groups in the PCR’s.\(^1\) The interest of the curious hollow core structures found in those PNT’s are not limited to the use of transmembrane ionic channels, e.g., but rather the space of their applications are unfathomable.

Apart from the actual applications, these artificial PNT’s would provide us with deep understandings of the “electronic roles and functions” of the peptides as well as the amino acids. The reason is that these PNT’s have a periodic structure. For many years, quantum biologists in the protein science have tried to elucidate the electronic roles of the individual amino acids as well as those functions of their integrated peptide forms. However, the vastness, complexity, and randomness of proteins have always confronted us, and prevented us from performing detailed quantum-mechanical calculations. Now, the periodicity, i.e., the structural characteristic found in the PNT’s, would allow us the straightforward use of the traditional quantum calculations, and with which some of the long-standing questions can be solved. The present results obtained for “periodic” PNT’s are strictly accurate, but are limited to the PNT’s themselves. However, one can deduce several conclusions which should be true in the proteins of actual organisms. This deduction seems to be one of two conclusions that the success in the PNT synthesis inherently includes.

Several ab initio theoretical works have been carried out for PNT’s: Lewis et al., have investigated the stable molecular structures and the energetics of cyclo[(D-Ala-Glu-D-Ala-Gln)]\(_{m=1-4}\) PCR’s and their PNT’s.\(^2\) Carloni et al., have studied the electronic structures of those aggregated PNT system,\(^3\) and Jishi et al., have discussed the influence of the guest atoms in PNT’s.\(^4\) These pioneering works give several important features on the electronic and molecular structures. Nevertheless, we start to reinvestigate the polypeptide open chains (POC’s), because we provide a systematic understanding of the POC’s, PCR’s, and PNT’s. Detailed investigations of their electronic and molecular structures should deepen our understanding. Moreover, we discuss the self-stacking mechanism of the PCR’s to form the PNT’s. We

FIG. 1. Illustration of the protein nanotube (PNT) (a) and polypeptide closed ring (PCR).
FIG. 2. Illustration of the peptide skeleton folding (a), the $\beta$-pleated sheets formation (b), and the membered ring (MR) formation (c).

FIG. 3. The change in the electronic band structure of the TP-POC’s of $\Theta=0^\circ$ (a), and other POC’s (b) of $30^\circ$ (12POC), $36^\circ$ (10POC), $45^\circ$ (8POC), and $60^\circ$ (6POC).
II. CALCULATIONAL METHODOLOGY

In order to give the systematic understanding on the electronic and molecular structures of these POC’s, PCR’s, and PNT’s, the calculations on the electronic structures should be performed by the parameter-free method. We, therefore, use the following two methods in accordance with the problems. For finite systems such as PCR’s, we perform ab initio molecular orbital (MO) calculations based on the Hartree-Fock (HF) approach with the employment of the GAUSSIAN98 program. For infinite systems such as POC’s and PNT’s, the crystal orbital calculation is performed, because of their structural periodicity.

In the former, a large number of the constituting atoms, included in PCR’s and PNT’s, limits the actual calculations by the Slater-type orbital 3G basis set for the geometrical optimization. However, the details in those electronic structures and intra- and inter-ring hydrogen bonds (HB’s) have been investigated by the 6-31G** basis set for a comparison. In the latter, the first-principles band calculations have been carried out using the local-density functional (LDF) approximation under the Hohenberg-Kohn-Sham scheme. The normconserving pseudopotential introduced by Bachelet, Hamann, and Schluter is also used, in which the core basis functions are ignored in the calculation. The Bloch functions of the corresponding crystal orbitals are expressed by the linear combination of atomic orbitals (LCAO) form of the pseudoatomic wave functions, which are obtained numerically in terms of the sum of the (Gaussian-type orbitals) GTO’s. The exchange-correlation energy is approximated by the functional form of Ceperley and Alder’s potential, parametrized by Perdew and Zunger. An outline of the present calculations, as well as the details in the calculation parameters, have been reported in our previous work.

III. OPEN CHAIN SYSTEM

A. Trans-planar-Polypeptide Open chains

The skeletal form of polypeptide chains is characterized by three rotation angles φ, ψ, and ω of the internal bonds, N–Cα, Cα–C, and C–N, respectively. In the fully extended (trans-planar, TP) skeleton form, all those values are 180°, and characteristic amide planes are formed between Cα-Cα atoms (ω=180°) in the peptide sequence. This plane is maintained until any intrapeptide HB’s occur. Typical secondary forms of the β-pleated sheet structure are produced by folding the peptide skeleton and maintaining this amide plane by varying the remaining internal rotation angles, φ and ψ. Thus, here, we characterize the internal bond rotations as the skeleton folding by using the folding angle Θ between the neighboring amide planes (Fig. 2).

The most extended POC is a trans-planar (TP), one giving the folding angle of Θ=0°. Although any folded POC’s are produced by varying Θ, one should focus on those folded POC’s having Θ=30°, 36°, 45°, and 60°, because they have...
a potential to produce closed chains of 12, 10, 8, and 6 membered ring (MR) forms, respectively, (Fig. 2); we therefore call these folded POC's having $\Theta = 30^\circ$, $36^\circ$, $45^\circ$, and $60^\circ$ as 12POC, 10POC, 8POC, and 6POC, respectively. According to this definition, the parallel $\beta$-sheet ($\beta_p$) is a parallel aggregation of those 6POC's ($\Theta = 60^\circ$), and the anti-parallel $\beta$-sheet ($\beta_a$) is one of the antiparallel 8POC's ($\Theta = 45^\circ$). Thus, by using the folded POC's, we can systematically discuss the POC's, their aggregated forms ($\beta$-pleated sheets), and also peptide closed rings PCR's.

How does the skeleton folding change the electronic structure? Before we discuss the change in the electronic structures of the folded POC, let us summarize the characteristic features of the TP-POC, briefly. This is because the TP-POC has a unique molecular plane corresponding to its own amide planes, and their orbitals are crucial for those in the PCR's as well as in the folded POC's.

Figure 3 (a) shows the calculated electronic structure of the TP-POC. Four characteristic bands are found at the upper part of the valence band (VB), while two bands are at the bottom part of the conduction band (CB). Based on the unique molecular plane of the TP-POC, these eigenstates should be classified into two types: One is the $\pi$ state, whose orbital lobs are standing vertically to the POC's plane and the other is the $\sigma$ state having in-planar orbital lobs. The above resulting states are assigned according to energy as $\pi$, $\sigma$, $\pi$, and $\sigma$ for the VB states, and as $\pi$ for the CB states, respectively. Thus, the TP-POC produces the $\pi$ band-edge states for both the highest occupied valence band (HOVB) and the lowest unoccupied conduction band (LUCB). The resulting $\epsilon - k$ dispersion also reveals that holes at the HOVB state tend to be slightly delocalized along the peptide skeleton but electrons at the LUCB state are localized.

B. Skeleton Folding

Now let us discuss the energetics and the electronic structures of the folded single POC's. Figure 4(a) shows the change in the total energy with varying POC folding angles $\Theta$. With increasing folding angle $\Theta$, the total energy increases. Thus, an important result is that the TP-POC is most stable energetically among POC's.

Figure 3(b) shows the change in the band-edge structure by varying the peptide skeleton folding angle $\Theta$. The characteristic feature in the VB-edge states is an interchange of the orbital character. With increasing skeleton folding, the HOVB $\pi$ state ($B_2$ state in the TP-POC) stabilizes energetically while the HOVB-1 $\pi_x$ band ($A_1$ state) destabilizes. When the folding angle is $\Theta = 45^\circ$ (8POC), these two states of $B$ and $A$ bands accidentally “degenerate,” and a level interchange occurs when $\Theta$ is $60^\circ$ (6POC).
No interchange among the two band states (A and B) occurs in the LUCB state, but a large amount of dispersion is found in the \( \epsilon - k \) relation with the skeleton folding (Fig. 3). Since the LUCB state is basically the \( \sigma \sigma \pi^* \) antibonding state, the skeleton folding induces \( \sigma \sigma \pi \sigma \) bonding couplings among the skeleton atoms. Mulliken’s analysis well reveals that this change in the hybridization also distributes electrons uniformly to each skeleton atom. Thus, with increasing folding of the POC’s, electrons at the LUCB state delocalize toward the peptide axis with a stabilization of its energy eigenstate; the calculated effective-mass ratio is reduced to 3.11 for the 8POC and to 1.51 for the 6POC.

**C. Condensation Energy of \( \beta \)-Pleated Sheets**

As mentioned before, the parallel \( \beta \)-pleated sheet (\( \beta_6 \)) is produced by stacking 6POC’s in parallel, and the stacking of 8POC’s in antiparallel produces the antiparallel \( \beta \)-pleated one (\( \beta_a \)). Therefore, the electronic features of these \( \beta \)-pleated sheets can be well discussed in terms of those composing 6POC and 8POC. We first obtain the optimized inter-chain distance and estimate the condensation energy of the \( \beta_p \) and \( \beta_a \) forms. Figure 4(b) shows the change in the total energy (per residue) with varying inter-chain distance. The optimized inter-chain distance is 4.72 Å for \( \beta_p \) and 4.73 Å for \( \beta_a \). These values are well consistent with the experimental values of 4.85 Å and 4.75 Å for \( \beta_p \) and \( \beta_a \), respectively.18

Figure 4(b) also reveals the condensation energy of 13.6 kcal/mol for \( \beta_6 \) and 8.4 kcal/mol for \( \beta_a \). These values are consistent with the other theoretical results by Suhai.19 The noticeable point is that the resulting \( \beta_a \) form is energetically more stable than the \( \beta_p \) form by 3.7 kcal/mol, although the larger condensation energy is found in the \( \beta_p \) form. This is because the energy is destabilized by the strong folding in the 6POC (\( \Theta = 60^\circ \)), as compared with that of the 8POC (8.9 kcal/mol), in the isolated single chain form.20

**IV. CLOSED RING SYSTEM**

**A. Circular Rings and Membered Rings**

The simplest PCR is a precise circular ring (CR). These type of PCR’s are formed by bending and cyclizing the TP-POC roundly [Fig. 5(a)]. Therefore, \( C^\alpha-C^\alpha \) amide planes of \( \omega = 180^\circ \) are broken and the highly strained PCR’s result. One can also produce the other types of PCR’s, by rotating the internal rotation angles \( \phi \) and \( \psi \) equally but oppositely (Table I). In these PCR’s, the amide planes are maintained to be \( \omega = 180^\circ \), and have a MR form [Fig. 5(b)]. When the composing amino acids are unified to have the same circular dichroism (i.e., LL or DD sequences), each MR should be the folded POC having the corresponding folding angle; i.e., \( nMR \) should be \( nPOC \). Thus, the MR can be characterized by using the folding angle \( \Theta \) defined in the folded POC’s.

Figure 6 shows the relationship among the total energy (per residue) and the amino residue number \( n \), of the CR’s and MR’s. With decreasing \( n \), a significant destabilization is found in the CR system. As discussed in the POC’s [Fig. 4(a)], the full extension of the \( \pi \) electrons produces the most energetically stable peptide skeleton; alternatively speaking, the peptide skeleton tends to form the trans-planar backbone in which the all peptide planes are in the unique molecular plane. Thus, it is easily understood that the CR’s are energetically more unstable than the MR’s because of their largely distorted amide planes. Such distortion increases in the smaller CR’s, and this is the reason why the smaller CR’s are more destabilized than the larger CR’s.21

In the case of \( n \) being infinite, the structures of both the MR and CR are precisely equal to that of the TP-POC. The increase in the \( n \) number lets the structures of both the MR and CR approach the unique form of this TP-POC. One should remember that the TP-POC is the most stable form. Thus, both CR’s and MR’s energetically stabilize with in

<table>
<thead>
<tr>
<th>PCR</th>
<th>CR</th>
<th>MR</th>
<th>NR</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>( \phi )</td>
<td>( \psi )</td>
<td>( \omega )</td>
</tr>
<tr>
<td>12</td>
<td>(-161^\circ)</td>
<td>(+168^\circ)</td>
<td>(-161^\circ)</td>
</tr>
<tr>
<td>10</td>
<td>(-157^\circ)</td>
<td>(+167^\circ)</td>
<td>(-157^\circ)</td>
</tr>
<tr>
<td>8</td>
<td>(-151^\circ)</td>
<td>(+165^\circ)</td>
<td>(-151^\circ)</td>
</tr>
<tr>
<td>6</td>
<td>(-142^\circ)</td>
<td>(+159^\circ)</td>
<td>(-142^\circ)</td>
</tr>
</tbody>
</table>

**FIG. 6.** Comparison of the calculated total energies among CR’s, MR’s, and NR’s with varying \( n \) number and with the inverse square of the radius \( 1/r^2 \).
creasing $n$ number, and the energy differences between those CR's and MR's should be decreased (Fig. 6).

### B. Electronic Structure

Based on the symmetry consideration described in Appendix A, let us discuss the electronic structure of the PCR’s. Figure 7(a) shows the resulting electronic states near the HOMO and LUMO states of the MR’s by varying the ring number $n$ ($n = 6 \sim 12$). The solid lines in the figure correspond to those $\sigma$ and $\pi$ states in the TP-POC. The upper part of the valence states is formed by two characteristic states of the $\pi$ and $\sigma$ orbitals. This feature is basically similar to that of the TP-POC, but the following are different (Appendix B). One is an existence of the doubly degenerated states due to the rotatory reflection of the PCR’s. Although those doubly degenerated states are sandwiched by the $\sigma$ or $\pi$ orbitals,

---

**FIG. 7.** Change in the electronic structure near the HOMO-LUMO gap of MR’s with varying $n$ number (a). Doubly degenerated states are represented by the splitted lines. We also illustrate the relative energetic position of the $\pi$ and $\sigma$ states of the valence band of the TP-POC, folded POC’s and PCR’s (b).
their orbitals characters are indistinguishable because of their orbital degeneracy. The other is the relative energetical position among $p$ and $s$ orbitals, being different from that in the POC’s. In the POC’s, the $p$ and $s$ states produce the nested bands, whereas the $p$ and $s$ states are completely separated in the MR’s. We summarize the relative position of the $p$ and $s$ states in the TP-POC, the folded POC’s and the PCR’s in Fig. 7.

C. Novel Rings

It is easily understood that the highly strained distortion of the CR’s peptide planes lets every CR be energetically more unstable than the MR having the same ring number, $n$ (Fig. 6). One should, however, notice that those MR’s are not the fully optimized PCR’s but the transition ones. The normal vibrational analysis searches out for the existence of energetically optimized stable PCR’s. Figure 5(c) shows the resulting structures, which give novel ring (NR) forms. In Table I, we also tabulate the details of their internal bond rotation angles $\phi$, $\psi$, and $\omega$.

The most characteristic feature found in these fully optimized NR structures is that the C$\alpha$-C$\alpha$ amide planes of $\omega=180^\circ$ are broken but the novel planes of $\psi\sim180^\circ$ are commonly produced between N atoms (N-N plane). Let us compare the stabilizing energy from MR’s to NR’s (Fig. 6). The corresponding value is 11.95 kcal/mol per residue for the 6NR, and the 12NR is more stable than 12MR by 10.60 kcal/mol. Although the stabilization energy changes with the $n$ number, the absolute values of the resulting NR’s hardly change with $n$ number. Those values rather give the “minimum” at $n=8$. This feature is noticeable because those energies of both CR’s and MR’s decrease monotonously with increasing $n$ number.

1. Amide Planes

What causes the structural change from the MR’s to the NR’s, and what breaks the normal C$\alpha$-C$\alpha$ amide planes, and also how are the novel N-N amide planes produced? First, we decompose the total energy into several energy terms; the repulsive part of the inter-atoms and inter-electrons Coulomb energies, the sum of the HF MO energies, and the kinetic part. In Fig. 8, these energy terms are represented by the difference between those of the MR and NR having the same $n$ number. With increasing $n$ number, these differences are converged to be zero, because the MR’s and NR’s should approach the unique form of the TP-POC when the $n$ number is infinite. Figure 8 also reveals the following feature; the repulsive part of the intercore and interelectron terms itself prefers the normal C$\alpha$-C$\alpha$ amide planes, while the attractive term due to the sum of the HF MO energy eigenvalues does the novel N-N amide planes. However, the latter attractive term overcomes the former repulsive term, and the NR form results. Thus, it can be concluded that any PCR’s wish to stabilize its orbital energies to form some peptide planes in the backbone. This feature is consistent with the idea that the planar peptide skeleton is formed through the ener-

![FIG. 8. Change in the energy difference between MR’s and NR’s with varying $n$ number. Values of the energy difference are divided into several energy terms of the total energy (TE), kinetic energy, attractive term, and repulsive term.](image1)

![FIG. 9. Comparison of the molecular structures of the 6MR and 6NR, and two characteristic intra-ring HBs of HB1 and HB2. The N-N planes are shown as the hatched area.](image2)
getical stabilization due to one-electron orbital energies via the $\pi$ electrons delocalization, as found in the TP-POC. In this meaning, any other peptide planes would be, however, possibly produced in the PCR’s. In the following, we mention that these N-N peptide planes are produced by the intra-ring HB’s.

### 2. Intra-ring hydrogen bonds

Let us reconsider the atomic geometry of the PCR’s. The coronal cyclization inevitably generates the two types of the intra-ring HB’s as shown in Fig. 9. The first type of HB’s are those between O⋯H-N (HB1). The structural optimization shortens the values of the HB1 by about 10%–11%. The strengthening in the HB1’s can be confirmed by an increase in their overlap population (OP) values (≈50%). Thus, it can be concluded that the NR’s are produced by strengthening the HB1’s, i.e., the novel N-N planes are caused by the shortening of the HB1’s, as described in the following.

How do HB1’s form N-N planes? Figure 10 reveals that the bond lengths of the amino acids in the resulting NR’s hardly change with varying $n$ number. As shown in Fig. 9, HB1’s in MR’s are formed among the neighboring amide planes. Therefore, in order to shorten these HB1’s under maintaining the amino acid bond lengths (BL’s), the individual HB1 should be in the own atomic plane. For this purpose, the ordinary C$^\alpha$.C$^\alpha$ planes are broken, and the novel peptide plane is generated between N atoms by changing the internal rotation angle $\omega$. This feature is commonly found in all the NR’s, and the resulting HB1 monotonously decreases with the change in $n$ number.

Now, we discuss the role of the other type of the intra-ring HB’s, C$^\alpha$.H$^\alpha$⋯O (HB2), which can be found on the opposite side of every HB1 (Fig. 9). All C$^\alpha$ atoms have an H atom (H$^\alpha$) beside the substitution group ($R$), and the alternating sequence of the L- and D-type amino residues has these H$^\alpha$ atoms, being out of the molecular plane vertically. The noticeable point is that those HB2’s give the minimum value when the ring number $n$ is 8 (Fig. 11). This feature can be geometrically understood as follows. The formation of the N-N planes due to the HB1 shortening lets the HB2’s bridge beyond the neighboring N-N planes. Therefore, when the plane determined by C, O, and H$^\alpha$ atoms coincides with the N-N plane (Fig. 9), these HB2’s give the minimum BL value. According to Fig. 11(a), this structural situation occurs when $n$ is 8. This atomic configuration gives the maximum overlap population for the HB2 [Fig. 11(b)]; i.e., the 8NR causes

![FIG. 10. Illustration of the inclined N-N plane against the $S_n$ axis direction (a) and changes in the inclination angle $\theta$ and the skeleton bond lengths of the NR’s with varying $n$ number (b).](image-url)

![FIG. 11. Change in the BL, bond angle BA (a) and OP and TE (b) of the HB2 with varying $n$ number. The bond angle is determined by the three atoms of C$^\alpha$.H$^\alpha$⋯O.](image-url)
the strongest HB2's. This is the reason why the 8NR gives the minimum value in the total energy.

To sum up, the following is concluded: Two types of the intra-ring HB's (HB1's and HB2's) perform important roles for the energetically stable NR formation. HB1's produce the novel N-N planes in the coronal PCR's, and the HB2's determine the optimal n number among these stable nNR's. This theoretical tendency is consistent with the experimental results, where the first synthesized PNT's are those 8PCR's.

3. Orientation of N-H bonds

All the resulting NR's have the $S_n$ main axis at the center of the individual ring plane. The other characteristic feature of the novel N-N planes is an inclination of these planes against this $S_n$ axis, and also its inclination ($\theta$) decreases with increasing n number (Fig. 10).

However, a point to be noticed, is that all the resulting N-H bonds are oriented almost parallel to the $S_n$ axis, independent to the n number as well as to the inclination of N-N planes. This unique N-H orientation is an origin of the self-stacking of PCR's to form the PNT structure, because all types of the present PCR's (CR's, MR's, and NR's) arrange the individual N-H bonds parallel to the $S_n$ axis. Thus, basically, any PCR's would be stacked via the inter-ring hydrogen bonds of N-H...O-C.

V. PROTEIN NANOTUBE

A. Energetics and Inter-ring Hydrogen Bonds

In this section, we will discuss the optimized structure of the protein nanotube (PNT). For this purpose, we carried out the total-energy calculations in the following manner. We started to obtain an optimal inter-ring distance by using the model tube, constructed by stacking three NR's in parallel.\(^{25}\) The minimum energy value is found at 4.92 Å under the restriction of conserving the individual NR form (Table II). Following this inter-ring optimization, the molecular structure of three NR's is fully optimized by freezing the above inter-ring distance. Next, we pick out the center PCR and reform the PNT by stacking these PCR's. Then, we recalculate the optimal inter-ring distance (4.72 Å). Finally, at this reoptimized inter-ring distance, we optimize the molecular structure of the three PCR's, and determine the "optimized" PNT form by rechoosing the center PCR [hereafter called a tubing ring, (TR)] and restacking them in parallel to form the PNT (TR-PNT).

We show the resulting structure in Fig. 12. The PNT formation due to the PCR parallel stacking, rebreaks the N-N bonds.
novel plane found in the isolated PCR. It reproduces the Cα-Cα amide planes found in the MR, while the internal bond rotation of \( \omega \) is \(-171^\circ\). In this respect, the Cα-Cα amide planes (\( \omega=180^\circ \)) are not perfectly restored. The other bond rotation angles (\( \phi \) and \( \psi \)) and the bond lengths also reveal that the form of the TR takes the middle position between the NR and MR.

Two types of the inter-ring HB’s (HB3 and HB4) are found in PNT’s (Fig. 12). Here, we consider the role of these HB’s in order to investigate the structural change in PCR’s. A significant shortening (~0.9 Å) is found in the HB3 (N-H•••O) when the PNT changes from the NR type to the TR one. This bond length shortening is caused by the restoring of the peptide plane from the N-N novel type (NR) to the Cα-Cα normal one (MR). Thus, the inter-ring HB3’s work the structural change from the NR-PNT to the TR-PNT. The overlap population analysis also reveals that the inter-ring HB4’s function to form the TR-PNT structure, while the intra-ring HB1’s prefer to form the NR-PNT structure. This “surviving” function of the HB1 lets the TR revert to that of the MR imperfectly.

Finally, we estimate the stabilization energy due to the PCR’s stacking. Figure 13 shows the calculated total-energy relationship among the single PCR’s (6MR, 6NR, and 6TR) and also the resulting stabilization energy via the PCR stacking. All values are based on one amino acid (Gly) residue. In the isolated PCR, the 6NR is energetically more stable than the 6MR by 11.95 kcal/mol, but the 6TR is more unstable than the 6NR by 2.10 kcal/mol. However, the larger condensation (stacking) energy (9.19 kcal/mol) lowers the 6TR-PNT, being energetically more stable than the isolated 6NR (7.09 kcal/mol). The first-principles LDF calculation also gives a similar result where the 6TR-PNT is energetically more stable than the 6NR-PNT by 5.31 kcal/mol.

**B. Electronic Band Structures**

In the above section, we found the important role of the inter-ring HB’s for the PCR’s stacking and the PNT formation. Can electrons or holes delocalize along the tube via these inter-ring HB’s? In this section, we calculate the first-principle electronic structures of the PNT’s and discuss whether carriers have a potential to cause the band conduction toward the tube axis. We will start to discuss the electronic structure of the 6MR-PNT, since the energetically
stable PNT is the TR-PNT whose PCR’s are rather similar to the MR’s. A parallel stacking of 6MR’s is also assumed for the 6MR-PNT, because the parallel β-pleated sheet (βp) corresponds to this 6MR-PNT, and then we can compare those results.

In Fig. 14, we show the calculated band structures of the 6MR-PNT and that of the βp (toward the inter-chain y direction). Similar electronic structures are found below those lower than −3 eV. These are the σ states, which are significantly influenced by the inter-chain (βp) or the inter-ring (6MR-PNT) interaction. Therefore, both states cause the well ε − k dispersion. We also show how the two types of the inter-ring HB’s (HB3 and HB4) are induced in those states.

A significant difference is found among the four band states near the VB edge. The HOVB state of the βp has the σx character (parallel to the peptide main axis), while the HOVB state of the 6MR-PNT has the π character. This is because, as mentioned in the POC’s, the increase in the folding destabilizes the skeleton σ state and the HOVB state interchanges from the skeleton π to the σ one when 6POC is produced. On the contrary, the coronal cyclization does not interchange the eigenstates among the skeleton π and σ orbitals, and therefore, the 6MR maintains the skeleton π HOVB state (Fig. 7). Since H’s 1s atomic orbitals (AO’s) are not hybridized with these inplanar π states due to its orthogonal character, the HOVB state in the 6MR-PNT is
strongly localized toward the tube axis. The strong localized nature found in the LUCB state is also caused for the same reason.

The structure of the TR is basically similar to that of the MR, except for the restoration of the perfect $C^\alpha-C^\alpha$ peptide plane ($\omega \sim 171^\circ$ not 180$^\circ$). Considering this structural similarity, the resulting electronic structure (Fig. 15) can be well understood by the results for the above 6MR-PNT. The strong localization is found at both the band-edge states, because of the orthogonal atomic configuration between the in-planar band-edge states and H and/or H$^+$’s AO’s. Thus, the band conduction toward the tube axis hardly occurs as long as the TR or MR form is maintained in the PNT.

C. Electronic Delocalization in PNTs

In order to delocalize electrons and holes (at the band-edge states) along the tube axis, the inter-ring HB’s should be mixed with the HOV or LUCB band-edge states. However, considering that the HOV and LUCB states are both in-planar $\pi$ states, the above orbital hybridization hardly occurs.

Let us investigate the model PNT, which is formed by stacking 6CR’s (6CR-PNT). Since the CR has the circular ring skeleton, two types of hydrogen atoms, H$^+$ and H$^-$, are settled equally against the peptide plane (Fig. 15). This atomic configuration is the same as those of the TP-POC, except for the highly distorted peptide plane. Thus, these two types of H atoms cause the inter-ring HB’s (HB4’s) with the O’s nonbonding lone pair (2p like) states. However, the resulting band structure reveals that the hybridization via these inter-ring HB4’s are still so weak that the effective mass of the HOV or LUCB states are not reduced; the calculated effective mass (ratio) is, with difficulty, 2.06 for the electrons at the LUCB state.

How do other types of PNT’s function? Figure 15 also shows the resulting band structure of the 6NR-PNT. Basically, the band-edge states are the in-planar $\pi$ state, and the orbital mixing of H AO’s is not significant via the inter-ring HB’s. However, the inclined peptide plane (N-N plane) against the tube axis produces an efficient orbital hybridization via the inter-ring HB3’s (Fig. 15). This hybridization comparably appears more in the LUCB state, and the well dispersive $\epsilon-k$ relation results along the tube axis. Considering the calculated value of $m^*=0.75m_0$, the band conduction toward the tube axis would be possible for electrons when a type of the 6NR-PNT is produced.

Based on the resulting energetics (Fig. 13), the above NR-PNT is more unstable than the TR-PNT by 5–7 kcal/mol. Considering that the electronic delocalization along the tube axis strongly depends on how the N-N planes incline, the future problem for designing the biomediated PNT’s would be how the energetic instability could be evaded.

VI. PROTONATION

When H$^+$ migrates in the PNT’s, how does it change the electronic structure? Figure 16(a) shows the calculated electronic structure when H$^+$ is at the center of the molecular plane of the 6CR (Appendix C). Since this proton arrangement does not change the symmetry of the system ($S_n$), the migrated H$^+$ should be mixed only with the states having the total symmetry, i.e., $A_g$ symmetry states of the $n$PCR’s ($n$
having symmetry, or A symmetry states of the nPCR′s (n = 4k having no i symmetry). As discussed before (Fig. 7), the 6CR causes the in-planar HOMO state having an $A_g$ symmetry. The migrated $H^+$ should then be hybridized with this HOMO state via the $sp\sigma$ bonding coupling. Thus, one can find an unoccupied localized impurity level (ULIL) in the HOMO-LUMO band gap, and also the energetical stabilization of the in-planar $A_g$ state changes itself into the HOVB-4 state.

The Mulliken charge analysis reveals the strongly localized nature of this level. An important point is that the unoccupied localized level created by the migrated proton is the bound state in the ground state. Therefore, some charge transfers are available into this level. These electronic features are found in MR′s having $n$ of 6–12 (Fig. 17) and NR′s. Thus, it can be concluded that the migrated proton in the PCR′s causes the “acceptor”-like unoccupied localized level in the HOMO-LUMO gap.

How does the hydroxyl ($OH^-$) ion change the electronic structure? For this future consideration, we start to investigate by using the $H^-$ ion. This is because the $H^-$ ion is the “simplest” negative ion and it has the same total symmetry as the proton. The comparison between the $H^+$ and $H^-$ ions leads to the systematic understanding of ion migration in this system (Appendix C).

According to Fig. 16(c), the $H^-$ ion creates the occupied level in the HOMO-LUMO gap. This level is placed energetically far from the HOMO state but near to the LUMO state, and is fully occupied by two electrons (OLIL). The Mulliken charge analysis reveals over 96% of those doubly occupied electrons being strongly localized at the migrated $H^-$ ion site. Although the unstable nature of the $H^-$ ion lets the created level not be a “bound” state, the efficient charge transfer would occur from this occupied level to the conduction band (unoccupied levels) when the system includes some other impurities causing stable occupied states. Thus, this type of negative ion species theoretically forecasts the possibility of the “donor”-level creation.

Finally, we expand our discussion to the PNT system, in which the proton or the $H^-$ ion is migrated. For the purpose of designing the biomediated electronic materials, we should focus on the NR-PNT, because this type of the PNT delocalizes electrons at the LUCB state most significantly (Fig. 15).

As we expected, the migrated proton creates the ULIL state just above the HOVB state (a), and the migrated $H^-$ ion creates the OLIL (b) (Fig. 18). Thus, some type of the PNT could produce “semiconducting properties” in their electronic structures.

FIG. 17. Calculated molecular orbital states of MR′s having $n$ of 6–12. The migrated proton is settled at the center of the molecular plane. The symbol ULIL means the unoccupied localized state originated from the migrated proton.
The fully optimized nPCR’s give the common structures, in which the Cα-Cα planes of ω=180° are broken, but the novel planes of φ=180° are produced between N atoms (N-N plane). The resulting total energy of the nNR’s hardly changes with n number.

Two types of the intra-ring HB’s of HB1 and HB2 appear in the PCR’s. The novel “secondary” forms of the NR’s are formed by strengthening these HB1’s. On the contrary, HB2’s give the optimal ring number n of 8.

Even in the single PCR, all the resulting N-H bonds are oriented parallel to the Sα axis, independent to n number. This unique orientation of N-H bonds is responsible for the stacking of PCR’s. The hollow tubular structure is then formed via these oriented inter-ring hydrogen bonds of N-H···O-C.

- The inter-ring interaction rebreaks the N-N novel peptide plane found in the isolated PCR, and reproduces the Cα-Cα amide planes. Consequently, the inclination of the peptide planes should appear. The electronic delocalization of electrons and holes is tunable by this inclination of the peptide plane.
- The unoccupied localized impurity level due to the foreign proton, is created just above the HOVB (HOMO) state, when the foreign proton is migrated and captured by the PNT (PCR). According to the resulting small value of the corresponding activation energy, the captured proton is expected to function as the thermal acceptor for the PNT’s valence electrons.

ACKNOWLEDGMENTS

The authors would like to express their thanks to Katsuhiro Fukasaku for his stimulating discussion and to Norihiko Takahashi for his advice in the use of the super computer. Parts of the present calculations have been performed at the Super Computer Center of NTT and at the Institute for Molecular Science (IMS), Okazaki National Research Institutes. One of the authors (KT) would like to express his thanks to the Iketani Science Foundation.

APPENDIX A: SYMMETRY CONSIDERATION AMONG POLYPEPTIDE OPEN CHAINS AND POLYPEPTIDE CLOSED RINGS

We discuss here how the symmetry changes when the folded POC’s form PCR’s? Figure 19 illustrates the change in the symmetry between the folded POC’s and the PCR’s. We also show the typical σ and π orbitals and the connectivity among the irreducible representations. POC’s have the main rotation axis of C2 toward the peptide backbone, while the PCR’s do that of Sn, which stands vertically on the molecular plane. Thus, the doubly degenerated eigenstates should appear in the latter PCR system. An interesting point is that the nondegenerated states of the nPCR’s should be further classified into A and A', or A and B, in accordance with the n number of 4k+2 (k:integer) or 4k. This is because the PCR’s whose n number being n=4k+2, have an inverse symmetry, while the PCR’s having n number of n = 4k, break such an inverse symmetry. Thus, the symmetrical analysis reveals that the irreducible representations of the π orbitals, A and B, in the folded POC’s (C2 symmetry) should be oppositely noted as B and A in the PCR’s (Sn symmetry). On the contrary, the same notification are maintained for the irreducible representations of the σ orbitals among the folded POC’s and PCR’s.
APPENDIX B: CHARACTERISTICS IN ELECTRONIC STRUCTURES OF POLYPEPTIDE CLOSED RINGS

With increasing $n$ number such as 6, 8, 10, and 12, a number of these degenerated orbitals increases as 2, 3, 4, and 5, respectively. Since the unit cell of the TP-POC includes two amino acid residues, two $\pi$ states and two $\sigma$ states are produced at the individual wave number. Thus, in the TP-POC, the upper part of VB is composed of those $\pi$ and $\sigma$ bands.\(^{11}\) Now, let us consider the PCR’s of $n$ being 6, for instance. It involves three units of the TP-POC, and therefore six states should be produced from these two $\pi$ and $\sigma$ bands. The ring formation having an $S_6$ symmetry, however, degenerates four of them doubly. Thus, two doubly degenerated states ($E_u$ and $E_g$) result, and the remaining two nondegenerated ($A_g$ and $A_u$) should appear in both of the upper and lower VB parts. When $n$ of the PCR’s is 8, three doubly degenerated states ($E_1$, $E_2$, and $E_3$) should appear, and be sandwiched by the two nondegenerated ($A$ and $B$) states. With increasing $n$ number, these apertures are filled up by the doubly degenerated states, and finally the periodicity in the TP-POC of $n$ being infinite causes the dispersive $\varepsilon-k$ relation.

As discussed in the POC system, the HOMO state interchanges from the $\pi$ state to the $\sigma$ state when the folding angle $\Theta$ becomes more than 45° $\sim$ POC. However, in the MR system, the corresponding $\sigma$ state (having $A_g$ symmetry for $n=4k+2$ or $A$ symmetry for $n=4k$) is settled lower than the two $\pi$ states. Moreover, it stabilizes with decreasing $n$ number (i.e., increasing folding angle). Thus, no interchange occurs, and the two characteristic states of the $\pi$ and $\sigma$ orbitals remain perfectly separated (Fig. 7).

APPENDIX C: HYDROGEN IONS

In Fig. 16, we show how the electronic charging changes the electronic structure of the PCR. For this purpose, the proton [H$^+$ (a)], the neutral H atom [H (b)] or the singly ionized [H$^-$ (c)] ion is migrated at the center of the 6CR. Both of the cases in which ions (H$^+$ and H$^-$) are migrated, have been investigated by the restricted Hartree-Fock calculation because of their closed-shell nature. On the contrary, the CR having an H atom gives the open shell, and the corresponding electronic structure has been calculated by the unrestricted Hartree-Fock scheme. This is why the resulting
electronic structure of the CR including an H atom (b) gives the doubly eigenstates (of $\alpha$ or $\beta$ spin) in the figure.

The calculated total charges describe well the nature of the impurity levels. When a neutral H atom is migrated, the present HF calculation gives the total charge value of “1.00” for this H atom (Fig. 16b). This value means that “no” charge transfer occurs among the H atom and the 6CR. Therefore, the ground state of the H atom intrudes in the occupied states of the 6CR without a disturbance (hybridization). Most electrons hardly flow out even when this H atom is singly charged up ($H^-\)$. Thus, the strong inter-electron interaction occurs locally at the $H^-$ ion site. This is the reason why the OLIL is strongly destabilized and is lifted just below the LUMO state. On the contrary, parts of electrons (64%) flow into the proton site when the H atom is positively charged up ($H^+$). For a proton, its “one-electron eigenstate” would be a vacuum, i.e., zero, because the proton has no electron. The resulting energetical position of the ULIL is settled by two thirds between those of the vacuum and the neutral H atom (0.5 Hartree) levels. This position is consistent with the above calculated charge-transfer ratio. Thus, the present HF calculations qualitatively describe the characteristic feature of the ion migrated PCR system.

14. In the present calculations for POC’s, the $x$ axis is setto the main axis of the POC, and the $y$ axis is set so that the $xy$ plane is equal to the skeleton plane of the TP-POC. Therefore, the $z$ axis is toward the out-of-skeleton plane.
15. Their irreducible representation (at the point $G$) is $B_1$, $A_1$, $A_2$, and $B_1$ for the VB states, and $A_2$ and $B_2$ for the CB states, respectively.
16. The skeleton folding disappears the mirror plane of the TP-POC, and the line group (isomorphic to the point group $C_{2v}$ for the TP-POC) should be changed into that of $C_2$ for the folded POC’s. Therefore, the resulting irreducible representations should be classified into $A$ or $B$ as shown in Fig 20.
17. This orbital interchange can be well understood by considering the orbital nature. The HOVB state of the TP-POC ($B_2$) is a $pp\pi^*$ antibonding one among the neighboring skeleton atoms. Therefore, the skeleton folding induces the $pp\sigma$ bonding nature, which stabilizes the corresponding $B$ state in the folded POC’s. On the contrary, with strengthening of the backbone folding, the hybridization of oxygen’s atomic orbital (O’s AO) decreases while that of N’s AO increases in the HOVB-1 ($A$) state. This change in the orbital hybridization releases electrons localized at O atoms (TP-POC) to the N atoms (folded POC’s), and induces the $pp\pi^*$ antibonding couplings among the N-C$\alpha$ atoms, where the $A$ state should be destabilized.
20. The mechanism of the larger condensation found in the $\beta_p$ is well understood by considering the two types of the interchain hydrogen bonds (HB’s), which are found in their optimized structures [Fig. 4(c)]; one is that of the neighboring N-H-O (called hereafter as type HB3), and the other is that of the neighboring C$\alpha$-H$\cdot\cdot\cdot$O (type HB4). The HB3 is a common and primary type both in the $\beta_p$ and $\beta_d$ forms. The optimized structures result in the similar HB3’s both in the $\beta_d$ and $\beta_p$ forms; the interatomic distances (1.79 Å for $\beta_d$ and 1.68 Å for $\beta_p$) and the interatomic angle ($\angle$NHO = 174.8° for $\beta_d$ and $\angle$NHO = 176.9° for $\beta_p$). This means that the HB3 produces similar sp$\sigma$ couplings, both in the $\beta_d$ and $\beta_p$, and that the condensation in these systems is controlled by the second type of HB’s, i.e., the HB4. The optimized $\beta_p$ produces remarkably shortened HB4’s. The calculated Mulliken charge also reveals a 28% increase in the positive charging at the $H^+$ atoms within the $\beta_p$ form. Thus, the second HB’s in the $\beta_p$ form causes the additional Coulomb attractive force via the interchain HB’s, and causes the larger condensation energy found in the $\beta_p$ form.
21. The continuum elastic medium model well describes this feature. It is well known that the flexion energy of the rigid body is inversely proportional to the squared value of the bending curvature radius $r$. In Fig. 6, we insert the resulting total energies of the CR’s and MR’s vs $1/r^2$. A linear relationship is well found in the CR’s but not in the MR’s. Thus, the first-principle energy calculation reveals that, once when the planar peptide skeleton is produced, its distortional mechanism is approximated by the flexion stress for the rigid body, i.e., the distortion of the $\pi$ orbitals in the amide planes is simply described by a “rigid bar.”
22. In the present paper, the optimization for $\beta$-pleated sheets has been carried out by varying the inter-ring distance under the restriction of frozen internal coordinates. The fully optimization has been performed to obtain the ring structures (NR’s).
23. The HF calculation with STO-3G basis set gives that the 6NR is energetically more stable than 6NR only by 0.2 kcal/mol per
residue. This value should be limited in the present basis set, and the value depends on the choice of the basis set. Delicate problems are included. Similarly, the details in the NR structures, deduced by the STO-3G calculations, are also delicate. This basis set does not include the polarization portion of the orbitals, which play an important role in the hydrogen bonding (HB). The underestimation of the polarized nature of HB might produce an “over-rotation” in $\alpha$. The optimization by using the 6-31G** basis sets results in the other “NR” structure, which is rather similar to the MR form. However, the important point is that all the resulting N-H bonds in the latter NR form are also oriented parallel to the $S_n$ axis.

Figure 11(a) also shows the change in the bond angle ($\angle C^{\alpha}HO$) of the three atoms $C^{\alpha}$-$H^{\alpha}$-$O$. The maximum bond angle is found when the bond length of the HB2 is at a minimum, i.e., in the 8NR. Thus, in the resulting 8NR, three atoms $C^{\alpha}$-$H^{\alpha}$-$O$ are arranged most linearly. One should, however, pay attention that this interpretation is standing on the simple $sp\sigma$ coupling and should be treated with caution. Although HB2 plays an important role to produce the optimal ring number, more precise interpretation on the HB’s requires more accurate basis including diffused and polarized parts.

A parallel stacking of the NR’s is also assumed, because the NR’s have the N-H bonds parallel to the $S_n$ axis and even a simple parallel stacking causes the inter-ring condensation via these N-H $\cdots$ O hydrogen bonds by 2.64 kcal/mol for the 6NR-PNT. An antiparallel stacking is also available. However, that stacking is excluded to save the CPU from loading, because the antiparallel stacking lets the size of the unit cell be twice as large than the parallel one.

A weak dispersion found at the LUCB state of the TR produces the indirect-type band structure.

In the present paper, we employed two total-energy calculation methods, the first-principles LDF band calculation, and the $ab\ initio$ HF MO calculation. Each of them has its merits and demerits. Both of them give the correct electronic structures of the conduction band (unoccupied states) as well as the valence band (occupied states). However, the LDF band calculations tend to underestimate the values of the band gap, but the HF MO approach overestimates them. Therefore, the numerical value of the LDF band gap, 0.3 eV, is underestimated. Similarly, those of the ULIL’s (3.9, 2.8, 2.3, and 1.9 eV for 6MR, 8MR, 10MR, and 12MR) also are overestimated because the ULIL is unoccupied by electrons. Those values should be treated with caution and may be changed by more accurate calculations.