J|A|C|S
ARTICLES
Published on Web 06/21/2002

A Quantum Chemistry Study of Binding Carotenoids in the Bacterial Light-Harvesting Complexes

Yanli Wang and Xiche Hu*

Contribution from the Department of Chemistry, University of Toledo, Toledo, Ohio 43606

Received January 16, 2002

Abstract: Carotenoids play the dual function of light harvesting and photoprotection in photosynthetic organisms. Despite their functional importance, the molecular basis for binding of carotenoids in the photosynthetic proteins is poorly understood. We have discovered that all carotenoids are surrounded either by aromatic residues or by chlorophylls in all known crystal structures of the photosynthetic pigment−protein complexes. The intermolecular $\pi-\pi$ stacking interactions between carotenoids and the surrounding aromatic residues in the light-harvesting complex II (LH-II) of *Rhodospirillum molischianum* were analyzed by high level ab initio electronic structure calculations. Intermolecular interaction energies were calculated with the second-order Møller−Plesset perturbation method (MP2) using the modified 6-31G*(0.25) basis set with diffuse d-polarization by Hobza and co-workers. The MP2/6-31G*(0.25) calculations yield a total stabilization energy of −15.66 kcal/mol between the carotenoid molecule and the four surrounding aromatic residues (α-Trp-23, β-Phe-20, β-Phe-24, β-Phe-27). It is thus concluded that $\pi-\pi$ stacking interactions between carotenoids and the aromatic residues play an essential role in binding carotenoids in the LH-II complex of *Rhodospirillum molischianum*. The physical nature of the $\pi-\pi$ stacking interactions was further analyzed, and the dispersion interactions were found to be the dominant intermolecular attraction force. There is also a substantial electrostatic contribution to the overall intermolecular stabilization energy.

1. Introduction

Life as we know it today exists because of photosynthesis, the process through which light energy is converted into chemical energy by plants, algae, and photosynthetic bacteria.[1−4] The molecular machine responsible for the initial light conversion is the photosynthetic unit (PSU). In the PSU, thousands of pigment molecules, mainly chlorophylls (Chl) and carotenoids, are noncovalently bound to proteins to form the so-called pigment−protein complexes, such as the light-harvesting complexes (LHs) and the photosynthetic reaction center (RC). The primary processes of photosynthesis consist of light absorption by the LHs and the subsequent excitation transfer to the RC for the primary charge separation.[5−10] Bulk chlorophylls in the LHs are responsible for the initial capture of sunlight, while a pair of specialized chlorophylls in the RC, the so-called reaction center special pairs, are directly involved in the primary charge separation.[11−13] Carotenoids play the dual function of light harvesting and photoprotection. The carotenoids absorb energy in a spectral region complementary to that of chlorophylls and, most importantly, function as a photoprotective agent that quenches the excited triplet state of chlorophylls. The latter state would otherwise be long-lived and could readily react with molecular oxygen to generate singlet oxygen, which is extremely reactive and destructive.[14,15]

Tremendous progress in our understanding of photosynthesis has been achieved in the last two decades with the determination of the atomic structures of the bacterial photosynthetic reaction center, followed by two high-resolution crystal structures of the bacterial light-harvesting complexes.[4] Structures of the RC for *Rhodopseudomonas (Rps.) viridis*[11] as well as for *Rhodobacter (Rb.) sphaeroides*[16,17] were determined to atomic resolution by X-ray crystallography. High-resolution crystal structures of the light-harvesting complexes from two species (*Rps. acidophila* and *Rhodospirillum (Rs.) molischianum*) have been resolved.[18,19] Most recently, a 2.5 Å resolution crystal structure of photosys-

* To whom correspondence should be addressed. E-mail: xhu@utoledo.edu.

(1) Clayton, R. K., Sistrom, W. R., Eds. *The Photosynthetic Bacteria*; Plenum: New York, 1978.
(2) Blankenship, R. E.; Madigan, M. T.; Bauer, C. E. *Anoxygenic Photosynthetic Bacteria*; Kluwer Academic Publishers: Dordrecht, 1995.
(3) Ort, D. R.; Yocum, C. F. *Oxygenic Photosynthesis: The Light Reactions*; Kluwer Academic Publishers: Dordrecht, 1996.
(4) Hu, X.; Ritz, T.; Damjanović, A.; Autenrieth, F.; Schulten, K. *Q. Rev. Biophys.* **2002**, *35*, 1.
(5) Hu, X.; Damjanović, A.; Ritz, T.; Schulten, K. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *35*, 1.
(6) Hu, X.; Schulten, K. *Phys. Today* **1997**, *50*, 28.
(7) van Grondelle, R.; Dekker, J.; Gillbro, T.; Sundstrom, V. *Biochim. Biophys. Acta* **1994**, *1187*, 1.
(8) Fleming, G.; van Grondelle, R. *Phys. Today* **1994**, *47*, 48.
(9) Knox, R. S. In *Primary Processes of Photosynthesis*; Barber, J., Ed.; Elsevier: Amsterdam, 1977; p 55.
(10) Sauer, K. In *Bioenergetics of Photosynthesis*; Govindjee, Ed.; Academic Press: New York, 1975; p 115.
(11) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. *Nature* **1985**, *318*, 618.
(12) Lancaster, C. R. D.; Ermler, U.; Michel, H. In *Anoxygenic Photosynthetic Bacteria*; Blankenship, R. E., Madigan, M. T., Bauer, C. E., Eds.; Kluwer Academic Publishers: Dordrecht, 1995; p 503.
(13) Zuber, H.; Brunisholz, R. A. In *Chlorophylls*; Scheer, H., Ed.; CRC Press: Boca Raton, 1991; p 627.
(14) Cogdell, R.; Frank, H. *Biochim. Biophys. Acta* **1987**, *895*, 63.
(15) Nilsson, R.; Merkel, P. B.; Kearns, D. R. *Photochem. Photobiol.* **1972**, *16*, 117.
(16) Allen, J. P.; Yeates, T. O.; Komiya, H.; Rees, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6162.
(17) Ermler, U.; Fritzsch, G.; Buchanan, S. K.; Michel, H. *Structure* **1994**, *2*, 925.

tem I (PS-I) from the cyanobacterium *Synechococcus (S.) elongatus* has been reported.[20] These crystal structures provide the detailed knowledge of the organization of pigment molecules in the photosynthetic membrane necessary for in depth theoretical and experimental studies of structure−function relationships.

Although carotenoids play a very important role in the photoprotection of (bacterio)chlorophylls, little is known about binding of carotenoids in the photosynthetic pigment−protein complexes. It was suggested that carotenoids are bound to proteins by hydrophobic interactions.[20] However, nonspecific hydrophobic interactions alone fail to account for the fact that all carotenoids are specifically bonded in the pigment−protein complexes. Recently, possible intermolecular hydrogen bond interactions (C−H···O−C) between carotenoid and the B800 bacteriochlorophyll (BChl) in the LH-II complex from the purple bacterium *Rps. acidophila* have been proposed on the basis of the Hartree−Fock (HF) level electronic structure calculations.[21] It should be noted that such hydrogen bonds, if they exist as suggested, are limited to the LH-II complex from *Rps. acidophila* only, because in the LH-II complex from *Rs. molischianum* the B800 bacteriochlorophyll adopts a different orientation.[19] On the basis of careful examination of all known crystal structures of photosynthetic proteins,[11,16,18−20,22] we have discovered that all carotenoids are surrounded either by aromatic residues or by chlorophylls (see below); the latter are molecules with conjugated π systems. We hypothesize that $\pi-\pi$ stacking interactions are the molecular forces that bind carotenoids in the photosynthetic pigment−protein complexes. In this article, the strength of these $\pi-\pi$ stacking interactions in the LH-II complex of *Rs. molischianum* is characterized by means of high level ab initio electronic structure calculations.

$\pi-\pi$ stacking interactions play an important role in a large number of biological and chemical systems, including base-stacking in DNA, molecular recognition, aromatic crystal packing, and biomolecular self-aggregation. They have been the subject of great theoretical interest ever since the early days of London.[23−32] One of the most widely studied systems is the benzene dimer, which serves as the prototype for aromatic $\pi-\pi$ stacking. $\pi-\pi$ stacking interactions of the benzene dimer have been studied at several levels of ab initio theory.[33−38] The intricate interplay of $\pi-\pi$ stacking and hydrogen bonding in DNA base pairing has been extensively studied by Hobza and co-workers.[30,39]

Weakly bonded interactions are essentially a juxtaposition of several elements, including electrostatic interactions, exchange-repulsion interactions, induction, and dispersion forces. Of these, dispersion forces constitute the dominant attractive forces between neutral molecules.[24−26] Dispersion forces arise from the mutual correlation of electrons that belong to interacting monomers (intermolecular correlation effects); the correlation energy is typically on the same order of magnitude as the intermolecular interaction energy. Consequently, inclusion of electron correlation is important in any accurate ab initio electronic structure calculation of weakly bonded complexes.

There are three principal methods that include the correlation correction: (a) configuration interaction (CI) methods; (b) coupled cluster (CC) methods; and (c) many-body perturbation theory (MBPT), also known as Møller−Plesset perturbation theory (MP). The full CI expansion is only of theoretical value because of its prohibitive computational intensity. Other variants of CI methods do not satisfy the necessary requirement of size consistency for treating intermolecular complexes. Coupled cluster methods, in particular, the coupled cluster method with single, double, and perturbative triple excitations (CCSD(T)), have been successfully applied to weakly bonded complexes of small molecules. Hydrogen-bonded and stacked formamide···formamide and formamidine···formamidine dimers were characterized at the CCSD(T) level by Hobza and Sponer.[40] Tsuzuki et al. performed a CCSD(T) calculation of $\pi-\pi$ stacking in the ethylene dimer.[41] The largest intermolecular complexes studied at the CCSD(T) level so far are the benzene dimer[35,38] and the naphthalene dimer.[38] However, CCSD(T) is very demanding in computational resources in terms of the CPU speed, size of the core memory, and capacity of the hard disk. It is impractical to apply the CCSD(T) method to large biomolecular systems. A popular and feasible way to include the correlation effects is the second-order Møller−Plesset perturbation theory (MP2), which usually covers a significantly large part of the correlation energy. The MP2 method has been applied to a wide variety of weakly bonded complexes, including $\pi-\pi$ stacking and hydrogen bonding in DNA, van der Waals complexes of atoms and molecules, etc.[30,31,39,42] One of the largest aromatic dimer systems studied at the MP2 level of theory to date is the MP2/6-31+G calculation of naphthalene and anthracene by Gonzalez and Lim.[43]

In this article, we implement the second-order Moller−Plesset perturbation method to calculate the strength of $\pi-\pi$ stacking interactions between carotenoids and the aromatic residues in the LH-II complex of *Rs. molischianum*.[19] The rest of the article is organized as follows. In section 2, we report structural details of three photosynthetic pigment−protein complexes, with special emphasis on the surroundings of carotenoids. Detailed implementation of the MP2 method, along with the choice of basis set, is described in section 3. Section 4 presents the results of

(18) McDermott, G.; Prince, S.; Freer, A.; Hawthornthwaite-Lawless, A.; Papiz, M.; Cogdell, R.; Isaacs, N. *Nature* **1995**, *374*, 517.
(19) Koepke, J.; Hu, X.; Münke, C.; Schulten, K.; Michel, H. *Structure* **1996**, *4*, 581−597.
(20) Jordan, P.; Fromme, P.; Witt, H. T.; Klukas, O.; Saenger, W.; Krauss, N. *Nature* **2001**, *411*, 909.
(21) He, Z.; Sundstrom, V.; Pullerits, T. *FEBS* Lett. **2001**, *496*, 36−39.
(22) Hofmann, E.; Wrench, P.; Sharples, F.; Hiller, R.; Welte, W.; Diederichs, K. *Science* **1996**, *272*, 1788−1791.
(23) Eisenschitz, R.; London, F. *Z. Phys.* **1930**, *60*, 491.
(24) Hobza, P.; Zahradnik, R. *Chem. Rev.* **1988**, *88*, 871−897.
(25) Chalasinski, G.; Gutowski, M. *Chem. Rev.* **1988**, *88*, 943.
(26) Buckingham, A. D.; Fowler, P. W.; Hutson, J. M. *Chem. Rev.* **1988**, *88*, 963.
(27) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.
(28) Chalasinski, G.; Szczesniak, M. M. *Chem. Rev.* **1994**, *94*, 1723.
(29) Chalasinski, G.; Szczesniak, M.; Cybulski, S. In *Pauling's Legacy: Modern Modelling of the Chemical Bond*; Maksic, Z., Orville-Thomas, W., Eds.; Elsevier: Amsterdam, 1999; p 665.
(30) Sponer, J.; Berger, I.; Spackova, N.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **2000**, *S2*, 383.
(31) Kim, K. S.; Tarakeshwar, P.; Lee, J. Y. *Chem. Rev.* **2000**, *100*, 4145.
(32) Chalasinski, G.; Szczesniak, M. M. *Chem. Rev.* **2000**, *100*, 4227.
(33) Karlstrom, G.; Linse, P.; Wallqvist, A.; Jonsson, B. *J. Am. Chem. Soc.* **1983**, *105*, 3777−3782.
(34) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Am. Chem. Soc.* **1994**, *116*, 3500−3506.
(35) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 48, 18790.
(36) Tsuzuki, S.; Uchimaru, T.; Mikami, M.; Tanabe, K. *Chem. Phys. Lett.* **1996**, *252*, 206.
(37) Jaffe, R. L.; Smith, G. D. *J. Chem. Phys.* **1996**, *105*, 2780−2788.
(38) Tsuzuki, S.; Uchimaru, T.; Matsumura, K.; Mikami, M.; Tanabe, K. *Chem. Phys. Lett.* **2000**, *319*, 547−554.
(39) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117−135.
(40) Hobza, P.; Sponer, J. *THEOCHEM* **1996**, *388*, 115−120.
(41) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *Chem. Phys. Lett.* **1998**, *287*, 202.
(42) Del Bene, J.; Shavitt, I. In *Molecular Interactions*; Scheiner, S., Ed.; Wiley: Chichester, 1997; p 157.
(43) Gonzalez, C.; Lim, E. C. *J. Phys. Chem. A* **2000**, *104*, 2953.

intermolecular interaction strengths, as well as an analysis of the physical origin of intermolecular forces. The biological significance of our findings is discussed in section 5, followed by a brief summary (section 6).

2. Binding Environment of Carotenoids

Carotenoids are molecules that contain conjugated C=C π bonds. The degree of π conjugation is typically 8−11 double bonds, but can range all the way from 5 to 15 double bonds. More than 600 naturally occurring carotenoids are known.[44] All known crystal structures of photosynthetic pigment−protein complexes contain one or more carotenoids,[11,16,18−20,22] the identity of which varies from species to species. Here, we analyze the binding environment of carotenoids in three representative pigment−protein complexes: the LH-II complex of *Rs. molischianum*,[19] the RC from *Rb. sphaeroides*,[17] and the PS-I from *S. elongatus*.[20]

Bacterial Light-Harvesting Complex-II. The LH-II complex functions as a light-harvesting antenna in the bacterial PSU, absorbing photons and transferring the excitation energy to the photosynthetic reaction center.[4] Lycopene has been found to be the major carotenoid in the LH-II complex of *Rs. molischianum*.[45] As revealed by its crystal structure,[19] the LH-II from *Rs. molischianum* is an octameric aggregate of $\alpha\beta$-heterodimers; the latter, as depicted in Figure 1, contains a pair of short peptides (α- and β-apoproteins) noncovalently binding three bacteriochlorophyll molecules and one lycopene. As seen in Figure 1, four aromatic residues (α-Trp-23, β-Phe-20, β-Phe-24, β-Phe-27) are within van der Waals contact (<5.0 Å) of the lycopene, forming a binding pocket for the chromophore.

Bacterial Photosynthetic Reaction Center. The bacterial photosynthetic reaction center is a specialized pigment−protein complex that initiates the primary charge separation process in photosynthesis.[4,11] The crystal structure of the RC from *Rb. sphaeroides*[17] contains three protein subunits, known as L (light), M (medium), and H (heavy), respectively. The L and M subunits are homologous and are related with a pseudo 2-fold circular symmetry. Multiple pigment molecules (cofactors) are bound to the L and M subunits and are arranged accordingly in two symmetric branches, commonly referred to as A branch and B branch: two BChls that form a strongly interacting dimer constituting the so-called special pair (P_A, P_B), two accessory BChls in close proximity to the special pair (B_A, B_B), two bacteriopheophytins (H_A, H_B), and a pair of quinones (Q_A, Q_B).[16,17] In addition, one carotenoid, a spheroidene molecule, is in close proximity to the M subunit of the RC. As shown in Figure 2, spheroidene in the RC from *Rb. sphaeroides* is surrounded by one BChl molecule (the accessory BChl molecule connected to the M subunit, commonly referred to as B_B) and multiple aromatic residues on the M subunit (M-W66, M-F67, M-F68, M-F74, M-F85, M-W115, M-F120, M-F123, M-W157, M-F162, M-W171).

Cyanobacterial Photosystem-I. In contrast to the bacterial photosynthetic reaction center, which collects light energy through separate membrane-intrinsic light-harvesting complexes, PS-I features a core integral antenna system. The 2.5 Å resolution crystal structure of PS-I from the cyanobacterium *S. elongatus*[20] reveals that it is a gigantic complex assembly of 12 protein subunits (named PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, PsaI, PsaJ, PsaK, PsaL, PsaM, and PsaX) and 127 cofactors. The latter comprise 22 carotenoids, 96 chlorophylls,

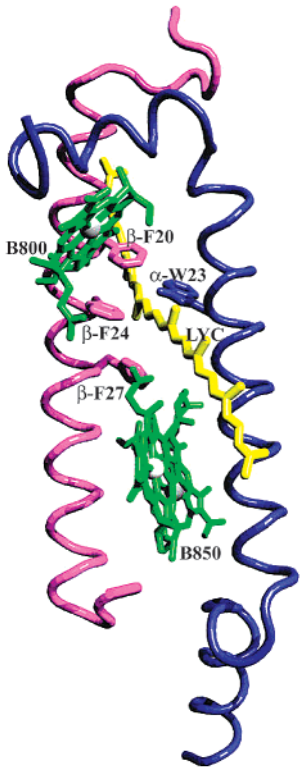

Figure 1. Binding environment of carotenoids (lycopenes) in the LH-II complex from *Rs. molischianum*[19] (PDB accession number 1LGH). Shown here is one of the $\alpha\beta$-dimers in the octameric LH-II complex.[19] The polypeptides are represented as α-carbon traces with the α-apoprotein and β-apoprotein in blue and magenta, respectively. The lycopene molecule is in yellow, surrounded by four aromatic residues: α-W23, β-F20, β-F24, and β-F27. Here, the prefixes α and β denote amino acids belonging to the α-apoprotein and the β-apoprotein, respectively. Bacteriochlorophylls (BChls) are colored green with the phytyl tail truncated for clarity, and the central Mg atoms of BChls are shown as silver spheres. (This figure is produced with the program VMD.[46])

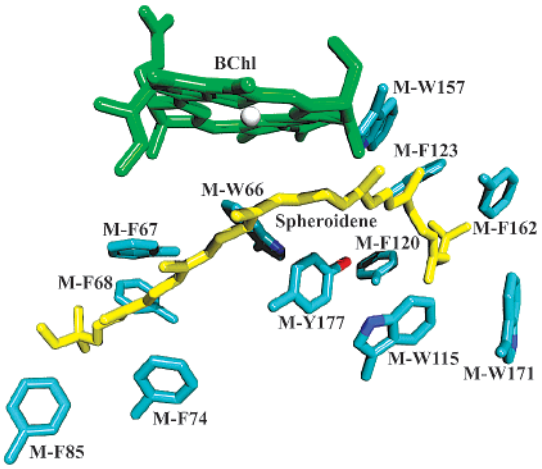

Figure 2. Binding environment of carotenoid in the bacterial photosynthetic reaction center (RC) of *Rb. sphaeroides*[17] (PDB accession number 1PCR). The sole carotenoid in the RC is a spheroidene molecule associated with the M subunit. Spheroidene is surrounded by one BChl molecule (the accessory BChl molecule connected to the M subunit) and multiple aromatic residues (M-W66, M-F67, M-F68, M-F74, M-F85, M-W115, M-F120, M-F123, M-W157, M-F162, M-W171) within a distance of 5 Å. The prefix M- indicates the M subunit of the RC.

(44) Goodwin, T. W. *Annu. Rev. Nutr.* **1986**, *6*, 273.
(45) Germeroth, L.; Lottspeich, F.; Robert, B.; Michel, H. *Biochemistry* **1993**, *32*, 5615.

Table 1. Carotenoids (β-Carotenes) and Their Surrounding Aromatic Groups in Photosystem I of *S. elongatus*[a]

β-carotene[b]	chlorophylls within 5 Å[b]	aromatic residues within 5 Å[b]
BCR-4001	Chl: 1113, 1118, 1120	
BCR-4002	Chl: 1112	
BCR-4003	Chl: 1127	
BCR-4004	Chl: 1212, 1217, 1218	PsaB-F224
BCR-4005	Chl: 1225	
BCR-4006	Chl: 1211	
BCR-4007	Chl: 1122	PsaA-F415
BCR-4008	Chl: 1124, 1133	
BCR-4009	Chl: 1220	PsaB-F318
BCR-4010	Chl: 1222, 1231	PsaB-F390
BCR-4011	Chl: 1126	PsaA-F681, PsaA-W744
BCR-4012	Chl: 1230	
BCR-4013	Chl: 1101, 1302	PsaA-W118
BCR-4014	Chl: 1229, 1301	
BCR-4015	Chl: 1229, 1235, 1303	PsaB-F431
BCR-4016	Chl: 1228, 1701	
BCR-4017	Chl: 1206, 1239	
BCR-4018	Chl: 1132, 1204, 1207	
BCR-4019	Chl: 1201, 1502	PsaI-F31
BCR-4020	Chl: 1131, 1207, 1502	PsaI-W20
BCR-4021	Chl: 1201	
BCR-4022		PsaL-F125, PsaL-F133

[a] There exists a total of 22 carotenoids, that is, β-carotenes (BCR), in PS-I of *S. elongatus*.[20] Chlorophylls (Chls) and aromatic residues within 5 Å of carotenoids are listed here. [b] The naming convention for peptide chains, chlorophylls, and carotenoids follows ref 20, and the residue identification numbers are in accord with the PDB file for the PS-I (accession number 1JB0). The prefixes PsaA-, PsaB-, PsaI-, and PsaL- indicate various subunits of PS-I.

two phylloquinones, three Fe_4S_4 clusters, and a putative Ca^{2+} ion. All carotenoids in PS-I were identified as β-carotene.[20] A systematic examination of the binding pockets of all 22 β-carotenes showed that all of the carotenoids are in van der Waals contact (<5 Å) with chlorophylls or aromatic residues. Table 1 lists all 22 β-carotenes and their surrounding aromatic groups.

The binding environment of one of the 22 β-carotenes, β-carotene-4004, is representative. As shown in Figure 3, β-carotene-4004 is surrounded by three Chls (Chl-1212, Chl-1217, and Chl-1218) and one aromatic residue (PsaB-F224). Here, all of the residue identification numbers, such as 4004, 1212, etc., are in accord with the PDB file for the crystal structure of PS-I[20] (PDB accession number 1JB0). Chl-1217 is in extended $\pi-\pi$ stacking contact with the β-carotene, and the other two Chls contact the ends of the β-carotene.

In all three pigment−protein complexes described above, all carotenoids are surrounded by (bacterio)chlorophylls and/or aromatic residues. On the basis of this discovery, we hypothesized that $\pi-\pi$ stacking interactions between carotenoids and these aromatic groups are responsible for binding carotenoids in the photosynthetic pigment−protein complexes. High level ab initio electronic structure calculations were then carried out to characterize the strength of these $\pi-\pi$ stacking interactions in one of the three pigment−protein complexes, the LH-II complex of *Rs. molischianum*.

3. Methods

The strength of $\pi-\pi$ stacking interactions between carotenoids and the aromatic residues in the LH-II complex of *Rs. molischianum* was characterized by means of ab initio electronic structure calculations at the MP2 level. Intermolecular $\pi-\pi$ stacking interaction energies between the lycopene molecule and each of the four surrounding aromatic residues α-Trp-23, β-Phe-20, β-Phe-24, β-Phe-27 (see Figure 1) were calculated in a pairwise manner.

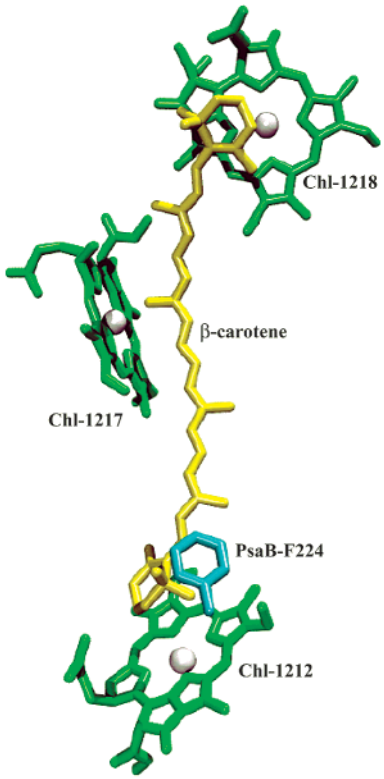

Figure 3. Binding environment of carotenoids in PS-I of the cyanobacterium *S. elongatus*.[20] The PS-I contains a total of 22 carotenoids (see Table 1). One of the carotenoids, β-carotene-4004, and its surrounding aromatic groups are shown here. Three chlorophylls (Chl-1212, Chl-1217, Chl-1218) and one phenylalanine residue (PsaB-F224) are within 5 Å of β-carotene-4004. The naming convention is the same as in Table 1.

As in all other quantum mechanical calculations, the quality of calculated results depends on the choice of basis set. It has been shown that for a proper treatment of $\pi-\pi$ stacking interactions, inclusion of diffuse basis sets is required.[36,39] These diffuse basis sets are localized sufficiently far from the atomic nuclei and thus fill the empty space between two interacting monomers. The latter is where a substantial portion of correlation energy originates. At the MP2 level, Dunning's correlation consistent basis sets (cc-pVXZ, X = D, T, Q, and 5) and the augmented aug-cc-PVXZ basis sets are desirable and have been applied to both $\pi-\pi$ stacking and hydrogen-bonding complexes of small molecules.[41,47] However, such huge basis sets are not computationally feasible for the large system of interest here. A more feasible choice for our system is a medium-sized basis set, such as the polarization augmented double-ζ 6-31G* basis set. In a series of studies of DNA base stacking, Hobza and co-workers employed a modified 6-31G* basis set with diffuse (momentum-optimized, dispersion energy-optimized) d-polarization at the MP2 level of theory.[30,39,48,49] In the conventional 6-31G* basis set, the d-polarization functions for non-hydrogen atoms (C, N, and O atoms) are energy-optimized with an exponent of 0.8. In the modified basis set, an exponent of 0.25 is used for the d-polarization functions of C, N, and O atoms, instead. Following the author's convention,[30,50] the modified basis set is designated 6-31G*-(0.25). It has been shown that the inclusion of more diffused d-polarization functions in the 6-31G*(0.25) basis set improves the electron correlation stabilization energy of stacked DNA base dimers

(46) Humphrey, W. F.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
(47) Tarakeshwar, P.; Choi, H. S.; Kim, K. S. *J. Am. Chem. Soc.* **2001**, *123*, 3323.
(48) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590.
(49) Kratochvil, M.; Sponer, J.; Hobza, P. *J. Am. Chem. Soc.* **2000**, *122*, 3495.
(50) Hobza, P.; Sponer, J.; Polasek, M. *J. Am. Chem. Soc.* **1995**, *117*, 792.

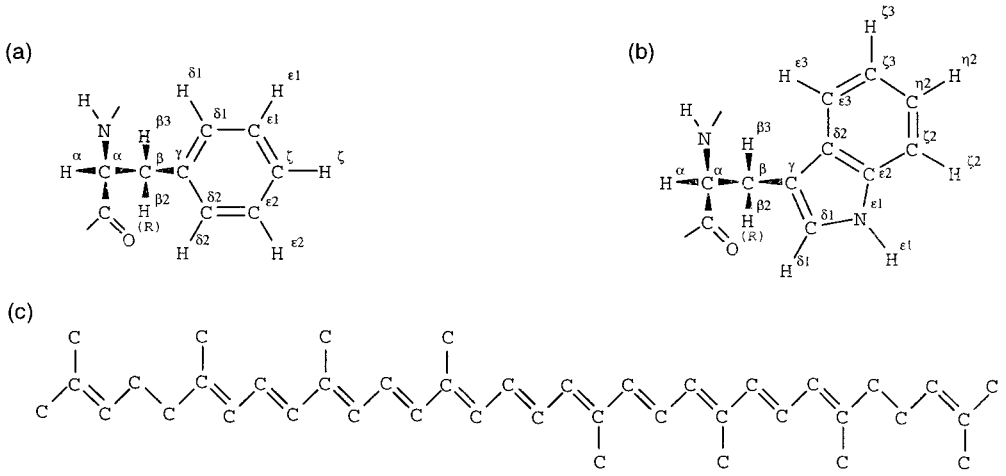

Figure 4. Molecular structures of lycopene and its intermolecular interacting partners: (a) phenylalanine; (b) tryptophan; and (c) lycopene with hydrogen atoms omitted.

substantially.[30,51] The 6-31G*(0.25) basis set is used in all of our calculations.

The quantum chemical program package Gaussian 98[52] was employed for all of the calculations. The intermolecular interaction energy was calculated at the MP2/6-31G*(0.25) level using the supermolecular approach. The energy of interaction between molecules A and B is defined as the difference between the energy of the interacting dimer E_{AB} and the energies of the monomers E_A and E_B:

$$E_{int} = E_{AB} - E_A - E_B \quad (1)$$

In all of our calculations, the coordinates of the carotenoid molecule and its interacting partners were extracted from the X-ray crystal structure. Therefore, the internal coordinates of the monomers used in computing E_A and E_B are the same as within the dimer AB.

In practical applications of the supermolecular approach, one often encounters the so-called basis set superposition error (BSSE) because of the use of incomplete basis sets. BSSE results from nonphysical lowering of the monomer's energy in the dimer's calculation because of the "borrowing" of the basis set from the interacting partner. There exist standard procedures to correct BSSE;[53] we used the Boys and Bernardi Counter Poise method.[54]

In addition to the intermolecular interaction energy, we are also interested in determining the physical origin of intermolecular interactions. In other words, we want to determine the contributions of all of the force components to the overall intermolecular interaction for a particular $\pi-\pi$ stacking interaction. Partitioning of intermolecular interaction energy into its constituting components is best defined in the framework of exchange or symmetry adapted perturbation theory.[27] The intermolecular interaction energy is expressed as a sum of contributions from at least the first- and second-order perturbation terms.[24] The first-order perturbation consists of electrostatic interactions and exchange-repulsion forces. The second-order perturbation contains dispersion forces, induction, as well as exchange-induction and exchange-dispersion forces. Electrostatic forces originate from permanent electric multipole interactions. Exchange-repulsion forces are a quantum effect that results from the Pauli exclusion principle, which forbids the electrons of one monomer from penetrating into the occupied space of its partners. Induction energy results from the interaction of the induced electric moments of one monomer with the permanent charge distribution of its partners. Dispersion forces arise from the mutual correlation of electrons that belong to interacting monomers (intermolecular correlation effects). In the variational supermolecular approach adopted here, the dispersion interaction energy corresponds primarily to the correlation energy, which is simply the difference between the MP2 intermolecular interaction energy and the HF intermolecular interaction energy. The HF intermolecular interaction energy itself is practically identical to the sum of four perturbation terms: electrostatic interaction, exchange-repulsion force (first-order terms), induction, and exchange-induction (second-order terms).[24,55] The electrostatic interaction energies can be analyzed by means of the distributed multipole method of Stone et al., as implemented in the program Orient 3.2.[56] Distributed multipoles[57] themselves are evaluated from the Gaussian 98 output wave functions by means of the GDMA 1.0 program.[56] To a first-order approximation, the contribution of the exchange-repulsion force can be roughly estimated as the difference of the HF energy and the electrostatic interaction energy.

4. Results

Intermolecular $\pi-\pi$ stacking interactions between the lycopene molecule and each of the four surrounding aromatic residues (α-Trp-23, β-Phe-20, β-Phe-24, β-Phe-27) are treated in a pairwise manner. Figure 4 depicts the structural formula for the lycopene molecule and two amino acids (phenylalanine and tryptophan). Only the side-chain atoms of the amino acids were included in the calculation; the α-carbon atom and its associated main-chain groups were excluded (see Figure 4). The α-carbon itself was replaced by a hydrogen atom. For both lycopene and the amino acids, the coordinates of non-hydrogen atoms were taken directly from the X-ray crystal structure of LH-II from *Rs. molischianum*.[19] The positions of all hydrogen atoms were placed by ab initio geometry optimization at the HF/6-31G* level with all of the heavy (non-hydrogen) atom positions fixed. It is worth mentioning a few words about the size of the current complexes. The Lycopene···Phe and Lycopene·

(51) Hobza, P.; Sponer, J.; Leszczynski, J. *J. Phys. Chem. B* **1997**, *101*, 8038.
(52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.
(53) van Duijneveldt, In *Molecular Interactions*; Scheiner, F. S., Ed.; Wiley: Chichester, 1997; p 81.
(54) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(55) Gutowski, M.; Kakol, M.; Piela, L. *Int. J. Quantum Chem.* **1983**, *23*, 1843.
(56) Stone, A.; Dullweber, A.; Hodges, M.; Wales, P. *Orient: A Program for Studying Interactions between Molecules*, version 3.2; University of Cambridge, 1995.
(57) Stone, A. J. *Mol. Phys.* **1985**, *56*, 1065−1082.

Table 2. MP2/6-31G*(0.25) and HF/6-31G*(0.25) Pairwise Intermolecular Interaction Energies of the Intermolecular Complex

intermolecular pair	E_{HF} (kcal/mol)[a]	E_{MP2} (kcal/mol)[a]
Lyc···β-F20	2.10	−3.64
	(2.00)	(3.70)
Lyc···β-F24	3.04	−2.75
	(1.98)	(3.82)
Lyc···β-F27	−0.14	−2.28
	(0.96)	(1.40)
Lyc···α-W23	0.97	−6.99
	(2.50)	(5.42)

[a] Intermolecular interaction energies at both the HF (E_{HF}) and the MP2 (E_{MP2}) levels are corrected for BSSE. The BSSE value for each complex is shown in parentheses.

··Trp pairs contain a total of 111 and 115 atoms, respectively. With the 6-31G*(0.25) basis set, the Lycopene···Phe complex consists of 346 electrons with a total of 833 basis functions (1572 primitive Gaussians); the Lycopene···Trp complex contains 366 electrons with a total of 880 basis functions (1660 primitive Gaussians).

Table 2 presents results of MP2/6-31G*(0.25) energy calculations for all pairwise intermolecular interactions between lycopene and the four aromatic residues. The magnitude of the BSSE correction for each interaction is also listed in parentheses. The stabilization energy of −6.99 kcal/mol between lycopene and α-Trp-23 is the strongest among all four pairs. This is not unexpected, because the tryptophan residue contains a larger two-ring π-system, while phenylalanine contains only one phenyl ring. The stabilization energy between lycopene and the three phenylalanine residues ranges from −2.28, −2.75 to −3.64 kcal/mol for β-Phe-27, β-Phe-24, and β-Phe-20, respectively. This is largely determined by the extent of $\pi-\pi$ stacking, which in turn depends on the distance and orientation of the two interacting π systems. Indeed, β-Phe-20 is much closer to the conjugated π-bonds of lycopene geometrically than are the other two phenylalanine residues. Summarizing all four pairwise interactions, a total of −15.66 kcal/mol of stabilization energy arises.

Also listed in Table 2 are intermolecular interaction energies at the HF/6-31G*(0.25) level. With the exception of the Lyc··β-Phe-27 pair, HF treatment results in positive intermolecular interaction energies (i.e., unstable complexes) for the pairwise intermolecular interactions. This further underscores the point made earlier about the necessity of including correlation correction when dealing with weakly bonded complexes, which is consistent with observations on many other $\pi-\pi$ stacking and hydrogen-bonded complexes.[30,31,39,58]

The basis set superposition error (BSSE), as determined by the Boys and Bernardi Counter Poise method, is shown in Table 2. The BSSE is found to be substantial for the current system. At the MP2/6-31G*(0.25) level of theory, BSSE correction ranges from 1.40 kcal/mol for the Lyc···β-F27 pair to 5.42 kcal/mol for the Lyc···α-W23 pair. Because the fundamental cause of BSSE is basis set incompleteness, this calls for larger and better basis sets to be used in the future as the computational power evolves for treating large dimers as reported here. Dunning's correlation consistent basis sets (cc-pVXZ, X = D, T, Q, and 5) and the augmented aug-cc-PVXZ basis sets have been shown to be adequate for treating $\pi-\pi$ stacking and hydrogen-bonding complexes of small molecules.[41,47] In terms of suitable basis sets for carotenoid molecules, the 6-31++G** basis set has been employed in the framework of the time-dependent density functional theory (TDDFT) to treat excited states of carotenoids, resulting in a much better prediction of the spectroscopic properties of carotenoids than the 6-31G** basis set gave.[59,60] It is conceivable that the 6-31++G** basis set will perform equally well for treating the ground state of carotenoids. With the above observation, we consider the basis set dependence of the calculated dimerization energies for the current system a subject that merits further investigation.

Table 3. Elements of the Pairwise Intermolecular Interaction Energies (in kcal/mol)

intermolecular pair	E_{MP2}[a]	$E_{MP2}-E_{HF}$[b]	$E_{electrostatic}$[c]	$E_{repulsion}$[d]
Lyc···β-F20	−3.64	−5.75	−0.57	2.67
Lyc···β-F24	−2.75	−5.79	−0.94	3.98
Lyc···β-F27	−2.28	−2.15	−0.64	0.50
Lyc···α-W23	−6.99	−7.96	−5.38	6.35

[a] E_{MP2}: Total intermolecular interaction energies calculated at the MP2/6-31G*(0.25) level with BSSE correction. [b] $E_{MP2} - E_{HF}$: Difference between MP2 and HF energies, both calculated with the 6-31G*(0.25) basis set (see text) with BSSE correction. [c] $E_{Electrostatic}$: Electrostatic interaction energies calculated on the basis of a multipole analysis of the MP2/6-31G*(0.25) wave functions using the Orient 3.2 program.[56] [d] $E_{repulsion}$: Exchange-repulsion interaction energies estimated as the difference between the HF energy and the electrostatic interaction energy.

To understand the physical nature of the intermolecular interaction, we decomposed the overall MP2/6-31G*(0.25) stabilization energy into its constituting components. Table 3 lists components of the intermolecular interaction energy for all four pairwise $\pi-\pi$ stacking interactions. The correlation energy (i.e., the difference between the MP2 energy and the HF energy) corresponds primarily to the dispersion interaction energy for the stacking complex. The dispersion energy ranges from −2.15 kcal/mol for the Lyc···β-Phe-27 pair to as high as −7.96 kcal/mol for the Lyc···α-Trp-23 pair. The electrostatic interaction energies are determined by the Orient 3.2 program using multipoles extracted from the MP2/6-31G*(0.25) wave functions of the interacting complex as mentioned in the Methods section. It should be mentioned that calculation of the second-order correction wave functions is required for such a analysis, which increases the computational intensity of the MP2 calculation substantially. For all four pairwise intermolecular interactions, the electrostatic interaction between lycopene and the aromatic residue is attractive. Overall, the electrostatic interaction contributed a total of −7.53 kcal/mol to the stabilization energy between lycopene and the four aromatic residues. The contribution of the exchange-repulsion force was roughly estimated as the difference between the HF energy and the electrostatic interaction energy as described in the Methods section. As expected, the contribution of the exchange repulsion force to the intermolecular interaction is repulsive.

5. Discussion

Photosynthetic membranes contain thousands of pigment molecules, mainly (bacterio)chlorophylls and carotenoids. The latter are noncovalently bound to proteins to form well-organized pigment−protein complexes.[4,13,61,62] Over the past few decades,

(58) Tsuzuki, S.; Uchimaru, T.; Matsumura, K.; Mikami, M.; Tanabe, K. *J. Chem. Phys.* **1999**, *110*, 11906,
(59) Hsu, C. P.; Walla, P. J.; Head-Gordon, M.; Fleming, G. R. *J. Phys. Chem. B* **2001**, *105*, 11016.
(60) Hsu, C. P.; Hirata, S.; Head-Gordon, M. *J. Phys. Chem. A* **2001**, *105*, 451.
(61) Zuber, H. *Trends Biochem. Sci.* **1986**, *11*, 414.
(62) Zuber, H.; Cogdell, R. J. In *Anoxygenic Photosynthetic Bacteria*; Blankenship, R. E., Madigan, M. T., Bauer, C. E., Eds.; Kluwer Academic Publishers: Dordrecht, 1995; p 315.

extensive biochemical and spectroscopic studies of photosynthetic systems have revealed that the pigment−protein complexes are organized in the form of PSU. In the PSU, the photosynthetic reaction center is surrounded by an array of light-harvesting antenna. The light-harvesting antenna itself is composed of multiple light-harvesting complexes with varying spectral characteristics and a particular structural organization in the whole antenna.[4] In most purple bacteria, for example, the PSUs contain two types of light-harvesting complexes, commonly referred to as light-harvesting complex-I (LH-I) and LH-II.[13] LH-I is found surrounding directly the RCs,[63] while LH-II is not in direct contact with the reaction center, but transfers energy to the reaction center via LH-I.[4,7] There exists a pronounced energetic hierarchy in the light-harvesting system. Pigments of outer light-harvesting complexes absorb light at a higher energy than do the inner ones. For example, the LH-II complex, which surrounds LH-I, absorbs maximally at 800 and 850 nm; LH-I, which in turn surrounds the RC, absorbs at a lower energy (875 nm).[13] This energy cascade serves to funnel electronic excitations from the LH-IIs through LH-I to the RC. For such an energy cascade to function, (bacterio)chlorophylls and carotenoids need to be located in certain well-organized positions and orientations.[4,5] It is remarkable that LH-I and LH-II complexes actually result from the self-aggregation of a large number of identical, noncovalently bonded transmembrane helices, BChls, and carotenoids.

How are the pigments, that is, carotenoids and chlorophylls, bonded to proteins to form the sophisticated pigment−protein complexes? To address this question, one needs to understand the intermolecular forces that govern pigment−protein interactions. (Bacterio)chlorophylls are normally bound to protein by metal ligation bonds between the central Mg atoms and the peptide side-chain atoms. For carotenoids to function as photoprotection agents of (bacterio)chlorophylls, they have to be bound structurally to the protein in geometrical proximity to the (bacterio)chlorophylls. Until now, little has been known about carotenoid binding in the photosynthetic pigment−protein complexes. The present work shows that intermolecular $\pi-\pi$ stacking interactions between carotenoids and the protein surroundings are responsible for binding carotenoids.

This work opens the door for further theoretical and experimental investigations of carotenoid binding in the LH-II complex of *Rs. molischianum*, in particular, and in the photosynthetic pigment−protein complexes, in general. The work reported here should stimulate more focused biochemical studies on probing the binding forces of carotenoids. For example, the theoretically predicted stabilization role of aromatic residues calls for experimental verification by means of site-directed mutagenesis of the relevant residues. This iterative interplay of theory and experiment will eventually lead to a firm understanding of carotenoid binding forces.

Ab initio calculation of weakly bonded interactions of large biomolecular systems remains a great challenge. On one hand, high level methods with correlation correction are needed to properly treat dispersion interactions;[28,30−32,39,58] on the other hand, the computational intensity of all post HF methods increases with the very high exponential power of the system size. The choice of MP2 theory and the modified 6-31G*(0.25) basis set represents a compromise between accuracy and feasible turn-around time. Although the absolute magnitude of the $\pi-\pi$ stacking interaction is inevitably dependent on the level of theory used, the attractive nature of the $\pi-\pi$ stacking interaction between lycopene and the surrounding aromatic residues in the LH-II complex of *Rs. molischianum*, as well as its substantial magnitude, is firmly established. It should also be born in mind that the MP2/6-31G*(0.25) method has previously been extensively applied to $\pi-\pi$ stacking of DNA bases.[30,39] According to Hobza and co-workers,[51] the MP2/6-31G*(0.25) method provides a good estimate of the stacking stabilization energy because the MP2 method recovers a substantial portion of the electron correlation energy. It has been observed that the MP2 method has a tendency to overestimate the dispersion energy when compared to the more rigorous CCSD(T) method with the same basis sets.[35,38,64] It is also well known that the medium-sized basis set, such as 6-31G*(0.25), typically underestimates dispersion interactions as compared to a much larger basis set at the same level of theory.[36,38,51] Thus, a potential cancellation of two offsetting errors was at work.

Although the density functional approach (DFT) is gaining popularity for treating large biomolecules because of its low computing cost for including the correlation effect, it is unfortunately inadequate for treating weakly bonded intermolecular complexes because it does not include the dispersion effect.[32,39]

6. Conclusion

Carotenoids have the vital dual function of light harvesting and photoprotection in the photosynthetic pigment−protein complexes. For carotenoids to function, they have to be bound structurally to the protein. This work shows that the intermolecular interactions between carotenoids and the protein surroundings provide a stabilizing effect for the formation of the overall pigment−protein complex. The strength of $\pi-\pi$ stacking interactions between carotenoids and the aromatic residues in the LH-II complex of *Rs. molischianum* has been characterized by means of the supermolecular approach at the MP2 level. The modified 6-31G*(0.25) basis set with diffuse d-polarization by Hobza and co-workers was adopted here.[50] The MP2/6-31G*(0.25) calculations yielded a total stabilization energy of −15.66 kcal/mol between the lycopene molecule and the four surrounding aromatic residues (α-Trp-23, β-Phe-20, β-Phe-24, β-Phe-27). Thus, it is concluded that $\pi-\pi$ stacking interactions between carotenoids and the aromatic residues play an essential role in binding carotenoids in the LH-II complex of *Rs. molischianum*.

On the basis of the MP2/6-31G*(0.25) wave functions, the physical nature of the intermolecular interactions between the lycopene molecule and the surrounding aromatic residues has been further analyzed. Dispersion interactions are found to be the dominant intermolecular attraction forces, although there is a substantial contribution of electrostatic attraction to the overall intermolecular stabilization energy.

Acknowledgment. The authors would like to thank The University of Toledo for providing the initial financial support for setting up a state-of-the-art computational chemistry laboratory with high performance computing and visualization workstations on which the ab initio calculations reported here were carried out. This work is supported in part through a grant from the deArce Fund for Medical Research and Development.

JA025618I

(63) Walz, T.; Ghosh, R. *J. Mol. Biol.* **1997**, *265*, 107.

(64) Sponer, J.; Hobza, P. *Chem. Phys. Lett.* **1997**, *267*, 263.