Quantum Mechanics/Molecular Mechanics Calculation of the Raman Spectra of the Phycocyanobilin Chromophore in α-C-Phycocyanin

Maria Andrea Mroginski,* Franz Mark,† Walter Thiel,‡ and Peter Hildebrandt*

*Technische Universität Berlin, Institut für Chemie, Max-Volmer-Laboratorium für Biophysikalische Chemie, Sekr. PC 14, D-10623 Berlin, Germany; †Max-Planck-Institut für Bioanorganische Chemie, D-45470 Mülheim, Germany; and ‡Max-Planck-Institut für Kohlenforschung, D-45470 Mülheim, Germany

ABSTRACT We have established a quantum mechanics (QM)/molecular mechanics (MM) hybrid method for calculating the Raman spectra of protein-bound cofactors using the α-subunit of C-phycocyanin containing a phycocyanobilin (PCB) chromophore as a test case. The PCB cofactor was described with density functional theory, whereas the protein matrix was treated with the CHARMM force field. The Hessian matrix of the QM region was built by taking into account bonded and nonbonded interactions with the protein environment and projected onto the internal coordinate space. Force constants were scaled with a global set of scaling factors, and the Raman intensities were computed using a finite-field method combined with a fourth-order differentiation algorithm for the calculation of the polarizability derivatives. In general, the QM/MM results provided a substantially improved description of the experimental resonance Raman (RR) spectra of the protein-bound cofactor compared to QM calculations of isolated PCB models in vacuo. The results allow the assessment of the effect of the protein-cofactor interactions on the RR spectra and reveal the potential and limitations of QM calculations on isolated tetrapyrroles for determining the chromophore structures in the various species and states of phytochromes for which three-dimensional structures are not available.

INTRODUCTION

Phytochromes constitute a class of sensor photoreceptors found in plants, fungi, and bacteria (1). The protein carries an open-chain methine-bridged tetrapyrrole as the light-sensing cofactor which upon photoexcitation initiates a reaction cascade to switch the receptor between the inactive and the active state. The photocycle involves a sequence of chromophore structural changes which are finally coupled to a major rearrangement of the protein matrix. However, the underlying molecular events are still not fully understood; so the elucidation of the photoactivation mechanism remains one of the key issues in phytochrome research.

Although the recently published three-dimensional structure of the parent state of a bacterial phytochrome represents a breakthrough (2), further progress in this field has to rely upon spectroscopic methods, specifically resonance Raman (RR) spectroscopy (3).This technique selectively probes the chromophore and thus may provide essential information about the structure of the tetrapyrrole cofactor and its interactions with the protein environment. However, extracting the information from the RR spectra still represents a considerable challenge.

In our previous studies, we have employed density functional theory (DFT) calculations to guide the spectra interpretation (4). Using a set of global force field scaling factors calibrated for a series of training molecules (5,6), Raman spectra have been calculated for various tetrapyrrole geometries. Upon comparison with the experimental RR spectra, it was possible to identify the gross chromophore structures of several states in the photocycle of phytochrome (7). However, this approach is associated with principal limitations since the calculations refer to the isolated chromophore in vacuo. Although the protein matrix is largely invisible in the RR spectra, steric and electrostatic interactions with the protein environment may cause significant perturbations of the geometry and the electron density distribution of the tetrapyrrolic chromophore and thus may have a pronounced effect on the RR spectra.

To overcome these restrictions, hybrid methods that are based on the combination of quantum mechanics (QM) calculations with molecular dynamics (MD) or molecular mechanics (MM) simulations represent a promising conceptual approach (8). QM/MM methods have been increasingly used to study proteins and enzymes by treating the active center quantum mechanically and by employing an empirical force field for describing the molecular environment (9). So far, most of the studies have been directed toward energies, geometries, transition states, and reaction pathways (9), whereas only a few investigations have focused on the calculation of vibrational spectra (10–14). Thus, the main goals of this work are to establish a QM/MM methodology for the calculation of Raman spectra of protein-bound cofactors and to assess the effect of the protein environment on the vibrational spectra of the prosthetic group.

A good candidate for testing the performance of theoretical approaches is the α-subunit of C-phycocyanin (α-CPC). C-phycocyanin is a phycobiliprotein that serves as an antenna pigment in photosynthetic cyanobacteria. α-CPC is a sufficiently small protein of 163 amino acids for which the three-dimensional structure is known (15,16) and RR spectra

Submitted March 14, 2007, and accepted for publication May 9, 2007.

Address reprint requests to Maria Andrea Mroginski, Tel.: 49-30-314-21584; Fax: 49-30-314-21122; E-mail: andrea.mroginski@tu-berlin.de.

Editor: Steven D. Schwartz.

0006-3495/07/09/1885/10 $2.00

doi: 10.1529/biophysj.107.108878

are available (17). Furthermore, α-CPC contains a covalently bound phycocyanobilin chromophore (PCB) which is structurally very similar to the cofactors in phytochromes.

COMPUTATIONAL METHOD

The simulation system

The initial structural model was set up by adding missing hydrogen atoms and solvent water to the experimental x-ray structure of the α-subunit of CPC from *Spirulina platensis* (18). The model system was then relaxed by performing a series of force field energy minimizations followed by MD simulations (see Supplementary Material) using the CHARMM22 force field (19). Unless noted otherwise, the optimized van der Waals (vdW) parameters from the CHARMM22 force field (set A) were employed in the calculations. For some tests we also used extreme values of the vdW parameters on the PCB molecule (set B taken from Table 1 of Riccardi et al. (20)), which were designed to maximize the hard sphere radius and to minimize the well depth of the Lennard-Jones potential function (20).

From the equilibration trajectory we extracted the snapshot corresponding to the conformation with the lowest energy. The differences between the crystallographic structure and the 'relaxed' structure resulting from the MD simulations are minimal. The root mean-square deviations (RMSD) between these two conformations are 0.14 Å for the backbone and 0.56 Å for the remainder of the protein. In the vicinity of the chromophore the most significant changes are the reorientation of the Arg-79 (~30°) and Lys-83 (~10°) side chains toward PCB, strengthening the hydrogen-bond (H-bond) interaction with the PCB carboxyl groups. The position of Asp-87, which forms H-bonds with the two NH groups of the tetrapyrrolic inner rings, remains basically unchanged (Fig. 1).

QM/MM setup

The geometry optimization of PCB and its protein environment involved a total of 1593 atoms contained in a 15-Å sphere centered at the PCB's center of mass and was performed using the QM/MM approach with the modular program package ChemShell (21,22). The QM fragment was described with DFT at the B3LYP/6-31G* level of theory, whereas the MM part was treated with the empirical CHARMM22 force field. The coupling between the QM and MM regions was handled by the electrostatic embedding model using the charge-shift scheme. This model includes the electrostatic interactions between QM and MM regions by incorporating the classical MM point charges in the core Hamiltonian. Bonding and vdW interactions across the QM/MM division were calculated using the MM force field. The covalent bonds cut at the QM/MM border were saturated by hydrogen-link atoms. The geometry of the QM/MM system was optimized using a limited memory quasi-Newton L-BFGS algorithm working in hybrid delocalized internal coordinates (HDLC) (23). Further technical details about the MM and QM/MM setup are given as Supplementary Material.

To check the influence of the link atom and the size of the QM region on the QM/MM results, we tested three different QM/MM partition schemes: an 86-atom QM fragment (M86) consisting of the entire PCB molecule and the side chain of Cys-84 to which the chromophore is covalently bound; a 93-atom QM fragment (M93) containing the PCB molecule and the side chains of Cys-84 and the Asp-87 residue, which forms strong H-bonds with the inner rings of PCB; and finally an 83-atom QM fragment (M83) consisting of the backbone of the PCB molecule together with the Cys-84 and Asp-87 side chains (Fig. 2). Model M86 requires a single hydrogen link atom to terminate the broken bond, whereas two and four link atoms are needed in models M93 and M83, respectively. Model M86B is identical to M86 except for the use of different vdW parameters (set B, see above).

QM model compound

For comparison with the QM/MM models, DFT calculations of the isolated chromophore were carried out for the dimethylester derivative of PCB

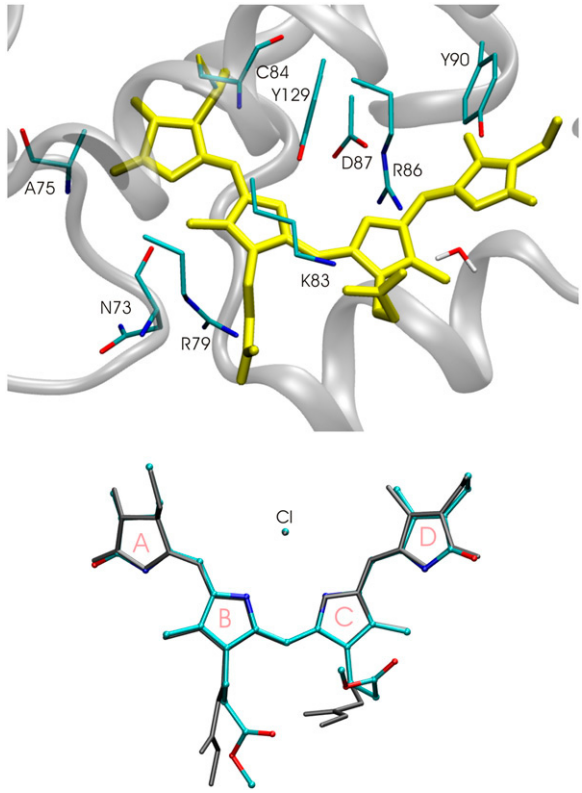

FIGURE 1 (*Top*) Chromophore binding pocket of α-CPC with the PCB cofactor drawn in yellow. (*Bottom*) Superposed structures of the PCBE-HCl models (QM, QM-strained) used in the QM calculations.

(PCBE) in the same protonated *ZZZasa* configuration as adopted by the protein-bound cofactor, using a chloride counterion to mimic the interactions with the carboxylate side chain of Asp-87. The geometries of several PCBE molecules with different conformations of the propionic side chains were generated in a systematic way by performing a conformational search of various side-chain conformations. Among all possible conformations of the propionic side chains, we only considered those where the propionic side

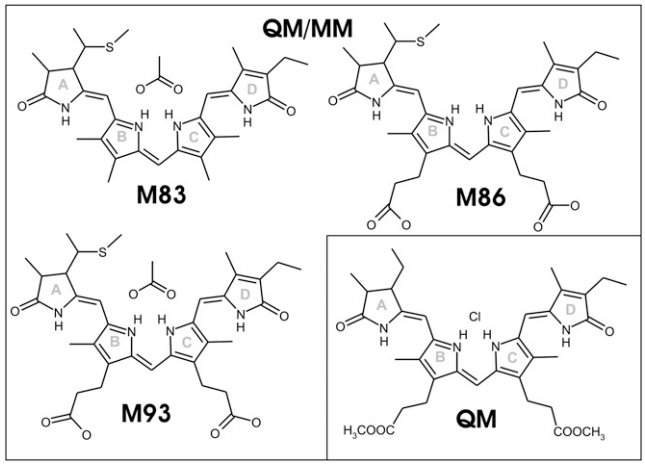

FIGURE 2 Definitions of the QM regions of models M83, M86, and M93 and of the pure QM model.

chain of ring *B* is pointing downward with respect to the plane formed by rings *B* and *C* and the propionic side chain on ring *C* is oriented upward as observed in the x-ray structures of the α84 chromophore in *Mastigocladus laminosus* (15) and *Spirulina platensis* (18) (Fig. 1). Geometries were subsequently optimized on the B3LYP level by means of Gaussian03 (24) using the 6-31G* basis sets except for Cl, for which a 6-31 + G* basis set was adopted. The isolated PCBE-HCl adduct, corresponding to the lowest energy downward-upward conformer structure that is calculated quantum mechanically, is denoted the QM model (Fig. 2). The geometry of the tetrapyrrolic backbone of this model compound is very similar to that of the crystal structure except for the dihedral angles at the methine bridges *A-B* and *C-D* (Table 1). In an attempt to force these dihedral angles toward the crystallographic values, we found a second conformer, which displays a slightly different vibrational spectrum. This conformer adopts a strained geometry that lies in a very shallow local minimum of the potential energy surface 3.41 kcal/mol above the QM model conformer, thus we denote it the QM-strained model.

Vibrational frequencies

The vibrational frequencies and normal modes of the QM fragment were obtained through diagonalization of the mass-weighted Hessian matrix. The

TABLE 1 Selected experimental and calculated bond lengths (Å), bond angles (degree), and dihedral angles (degree) of protonated PCB molecule complexed with Cl^- and with aspartate residue

		QM		QM/MM			
	Exp*	QM	QM-strained	M93	M86	M86B	M83
PCB†							
Bond length							
C=C (AB)	1.405	1.356	1.355	1.361	1.361	1.363	1.360
C-C (AB)	1.423	1.443	1.444	1.448	1.446	1.449	1.449
C=C (BC)	1.411	1.392	1.396	1.389	1.387	1.390	1.390
C-C (BC)	1.421	1.402	1.397	1.400	1.399	1.400	1.399
C=C (CD)	1.416	1.361	1.360	1.359	1.359	1.361	1.359
C-C (CD)	1.415	1.441	1.441	1.442	1.440	1.442	1.442
Bond angles							
CCC (AB)	133.8	129.2	128.9	131.2	131.3	130.9	131.3
CCC (BC)	133.9	133.5	133.8	134.8	134.7	134.1	134.9
CCC (CD)	129.4	130.1	127.5	126.2	126.1	126.9	125.9
$CC_{\beta}C$ (pps B)	110.3	112.9	113.4	114.8	114.9	113.6	112.9
$CC_{\alpha}C$ (pps B)	114.1	112.0	112.3	113.3	113.2	113.8	114.0
OCO (pps B)	124.3	123.6	123.5	125.6	125.5	125.7	121.4
$CC_{\beta}C$ (pps C)	113.5	114.5	113.7	113.3	113.6	112.7	114.4
$CC_{\alpha}C$ (pps C)	110.8	113.1	112.6	111.6	111.8	110.0	112.5
OCO (pps C)	123.1	123.5	123.6	124.4	124.4	124.6	120.6
Dihedral							
N(A)CCC (AB)	20.8	−3.1	−3.1	0.8	1.2	1.4	0.7
CCCN(B) (AB)	131.6	148.4	147.4	152.3	152.3	154.8	152.5
N(B)CCC (BC)	4.3	3.4	1.5	−0.1	−0.6	0.9	0.1
CCCN(C) (BC)	5.3	3.6	2.7	8.3	9.2	6.2	8.2
N(C)CCC (CD)	141.1	152.4	144.6	143.7	144.6	148.1	143.3
CCCN(D) (CD)	1.1	−2.2	−2.9	−4.5	−4.9	−3.2	−4.6
$CC_{\alpha}C_{\beta}C$ (pps B)	179.5	−171.7	139.4	174.3	174.9	173.5	175.7
$C_{\alpha}C_{\beta}C$-O (pps B)	158.7	−160.3	143.8	165.0	164.3	167.8	163.6
$C_{\alpha}C_{\beta}C$=O (pps B)	−21.6	21.0	−38.2	−16.8	−17.3	−13.8	−15.5
$CC_{\alpha}C_{\beta}C$ (pps C)	61.7	78.6	41.4	55.5	54.9	52.1	53.6
$C_{\alpha}C_{\beta}C$-O (pps C)	75.1	−151.6	67.8	53.1	54.1	75.5	55.5
$C_{\alpha}C_{\beta}C$=O (pps C)	−104.2	30.6	−110.7	−126.4	−125.2	−101.4	−124.1
PCB-Asp							
N(B)...O1(Asp)	2.87	3.12	3.11	2.86	2.67	3.15	2.85
(N(B)...Cl)							
N(B)...O2(Asp)	4.23			4.17	4.12	4.3	4.12
N(C)...O1(Asp)	2.82	3.13	3.12	2.78	2.67	3.24	2.79
(N(C)...Cl)							
N(C)...O2(Asp)	3.26			3.24	3.09	3.39	3.24
N(B)-H..O1		170.6	172.2	168.6	169.8	144.1	169.1
(N(B)-H..Cl)							
N(C)-H..O1		170.1	171.4	177.6	170.6	165.9	177.8
(N(C)-H..Cl)							

*Experimental data are taken from Wang et al. (18).

†AB, BC, and CD denote the methine bridges between ring *A* and ring *B*, ring *B* and ring *C*, and ring *C* and ring *D*, respectively; pps B and pps C stand for propionic side chain on ring *B* and on ring *C*, respectively.

Hessian matrix is built from the second derivatives of the total energy with respect to the nuclear displacements of the QM atoms. Contribution to the Hessian matrix arising from electrostatic interactions between the QM and MM parts were computed quantum mechanically, whereas contributions resulting from vdW and bonding interactions between these two parts were calculated analytically based on the molecular mechanical CHARMM22 force field. In the calculations of the QM/MM Hessian, the presence of the link atoms introduces additional degrees of freedom so that their contributions need to be projected out. In agreement with the observations by Cui and Karplus (13) we have noticed that even with this projection the link atoms lead to artifacts in the vibrational spectra for those modes that involve the boundary QM atoms. Despite these artifacts, the present approximation is sufficient for comparison with the experimental spectra in the frequency range between 400 and 4000 cm^{-1} (see Table S3 of Supplementary Material to assess the similarity of QM/MM frequencies computed for three different QM regions with different numbers of link atoms; see also Eichinger et al. (12) for further analysis).

The force fields computed for the QM and QM/MM models were corrected by a set of global scaling factors determined previously for a series of model compounds (5,6,25) (see Supplementary Material). This method partially compensates for the errors in the calculated force fields resulting from deficiencies of the quantum mechanical method and from the harmonic approximation, thereby affording an accuracy of $\sim\pm 11$ cm^{-1} for the calculated frequencies (5).

Raman intensities

Raman intensity calculations were restricted to nonresonant conditions, which provide a satisfactory description of the experimental intensities obtained with preresonance excitation (4). For an experimental arrangement with a backward scattering geometry and off-resonance excitation (ν_0), the Raman intensity $I_{Ra,s}$ of a normal mode s with frequency ν_s is given by Eq. 1, where α' and β' are the derivatives of the mean polarizability α and its anisotropy β with respect to the normal mode s (26):

$$I_{Ra,s} \propto (45(\alpha')^2 + 7(\beta')^2)(\nu_0 - \nu_s)^4 \left[\nu_s\left(1 - \exp\left(-\frac{h\nu_s}{kT}\right)\right)\right]^{-1}. \quad (1)$$

Using the electric field method, the derivatives α' and β' were evaluated by a twofold numerical differentiation of the forces acting on the atoms with respect to an applied field (27). To obtain numerically accurate Raman intensities, the derivatives were computed using a fourth-order differentiation algorithm that effectively minimizes the errors associated with the numerical differentiation (5). Calculations were performed using an electric field strength of 0.002 a.u.

RESULTS AND DISCUSSION

Geometries

The geometry optimization of the M83, M86, and M93 QM/MM models yields very similar minimum energy structures of PCB despite the different QM fragments. The RMSDs with respect to an average structure are <1 mÅ for the bond lengths and 1° for the bond angles and dihedral angles. Comparison of the x-ray structure (18) with these optimized geometries reveal that the relaxation of the PCB chromophore inside the protein pocket is associated with structural changes which particularly take place at the methine bridge connecting ring *A* with the remainder of the chromophore and at both propionic side chains. The dihedral angles at the C-C and C=C bonds at the methine bridge between ring *A* and ring *B* (*A-B* methine bridge) are twisted by $\sim$20° in opposite directions, whereas the positions of the two rings remain practically fixed, leading to a tilt of the C-H group with respect to the starting structure. This distortion results from an attempt of the *A-B* methine hydrogen to maximize the interaction with the carbonyl group of Tyr-129 by forming an intermolecular H-bond.

During the QM/MM geometry optimization, the distance between the *A-B* methine carbon atom and Tyr-129 carboxyl oxygen decreases drastically from 3.60 Å (x-ray structure) to 2.67 Å (M93). Moreover, the fact that the quantum mechanical calculations of the isolated chromophore yield similar values for the *A-B* methine bridge dihedral angles as those obtained for the QM/MM models indicates that the *A-B* bridge in the crystal structure exhibits a slightly strained conformation. Unlike the *A-B* methine bridge, changes of the torsional angles at the *B-C* and *C-D* methine bridges are calculated to be $<5°$. The positions of the four pyrrole rings are stabilized by the immediate protein matrix, which remains practically unchanged compared to the crystallographic structure. Furthermore, we observed that the decrease of the tetrapyrrolic torsional angles is accompanied by an increase of the single/double bond length alternation at the *A-B* and *C-D* methine bridges together with the closing of the corresponding C=C-C angles. The propionic acid side chain of ring *B* is only slightly rotated by $\sim$5°, whereas the propionic side chain of ring *C* is rotated by 20° toward the charged residues of neighboring amino acids.

H-bond interactions between the chromophore and the apoprotein require special attention. The crystal structure of α-CPC shows that four charged amino acids are located at a distance of <3.5 Å to the PCB cofactor (Fig. 1). The carboxylic end of the Asp-87 resides in the vicinity of the inner rings of the tetrapyrrolic chromophore and is forming H-bonds to the NH groups of the corresponding pyrrolic rings. This amino acid is responsible for the nearly coplanar arrangement of rings *B* and *C*. One lysine (Lys-83) and two arginines (Arg-79, Arg-86) have their terminal groups pointing toward the oxygens of the propionate side chains. The distances between the corresponding side-chain amino groups and the carboxylate oxygens decreases to 2.7 Å upon geometry optimization, reflecting the formation of strong H-bonds. Besides these residues, Ala-75 and Asn-73 play an important role in stabilizing ring *A* by forming weak H-bonds with the keto and N-H groups of the chromophore. Thus the PCB cofactor forms at least six H-bonds with the protein matrix. Among them, the interactions with Asp-87 are particularly important, since they have a strong influence on the electron density distribution in the chromophore and consequently on its vibrational spectra. A sound description of the H-bond interaction, therefore, requires the quantum mechanical treatment of the hydrogen donor and of the acceptor partners.

In the models M83 and M93 that include the Asp-87 side chain in the QM fragment, this requirement is fulfilled for the

NH···O bonds formed with the inner rings *B* and *C* of the chromophore (Fig. 2 and Table 1). For these two models, the hydrogen donor-acceptor [N(*B*) ···O1, N(*C*)···O1] distances are in very close agreement with the crystallographic values (0.01–0.04 Å). If the H-bond is cut by the QM/MM boundary, the distance between donor and acceptor strongly depends on the MM force field parameters used in the calculations. When the standard vdW parameters (set A) are assigned to the QM fragment of the M86 model, the N···O distances are significantly reduced by ~0.1–0.2 Å compared to the geometry of the pure QM model. The use of set B vdW parameters in M86B, on the other hand, induces an elongation of the H-bonds by ~0.3–0.4 Å (Table 1). Significant geometrical differences resulting from the choice of vdW parameters for the QM fragment are also noted for the propionate side chains of PCB. Here, for example, the torsion of the terminal carboxylate group of the side chain attached to ring *C* differs by 22° between the models M86 and M86B. The geometry of the tetrapyrrolic backbone, however, is less sensitive to the numerical values of vdW parameters.

The structural differences between the geometries of the two models for the isolated chromophore, namely models QM and QM-strained, are confined preferentially at the *C-D* methine bridge and at the propionic side chains. In the 'relaxed' QM model, ring *D* is twisted by 11° with respect to the crystallographic structure, whereas in the strained conformation (QM-strained) this difference is reduced to only 4°. The torsion of the *C-D* methine bridge and, to a lesser extent, the torsions on the *A-B* and *B-C* bridges are correlated with the orientations and conformations of the two propionic side chains. In the strained model, the CC_α-$C_\beta C$ dihedral angle of the propionic side chain on ring *B* is reduced from 180° (x ray) to 139° (Table 1). This geometric change leads, on the one hand, to a 40° deviation of the methylene groups from a well-staggered conformation and, on the other hand, to an increase of the C=O···H-C interaction between the terminal carboxylate oxygens and the *B-C* methine group. In the QM model, the carboxylate group of the propionic side chain on ring *B* is pointing away from the tetrapyrrolic backbone as observed in the crystal structure (Fig. 1).

Structural changes of the propionic side chain on ring *C* in the QM-strained model are more pronounced than those observed on ring *B*. The corresponding CC_α-$C_\beta C$ dihedral angle differs with respect to the crystallographic values in 17° and in 20° for the QM and for QM-strained model, respectively (Table 1). In the QM model, the propionic side chain on ring *C* is oriented in a similar way as found in the x-ray structure of oligopyrroles (28,29) with the C=O group forming a weak H-bond with the *B-C* methine bridge hydrogen. The large twist of the propionic side chain predicted for the strained model leads to a conformation in which the carboxylate group lies above pyrrole ring *C* at a distance of 3.45 Å with respect to the nearest carbon on the ring and points toward the N-H group of ring *D*. The stability of this strained structure can be understood in terms of sizeable vdW interactions between the C=O π-bond and the tetrapyrrolic π-system. These interactions force the twist of the *C-D* methine bridge toward a strained conformation of the tetrapyrrolic backbone as found in the crystal. Thus, despite its high energy, the QM-strained model of the isolated PCBE-HCl molecule provides a better description for the tetrapyrrolic backbone conformation of the protein-bound chromophore (vide infra).

There are only slight geometrical differences at the level of the tetrapyrrolic backbone between the structures optimized using the hybrid QM/MM method and that resulting from the quantum mechanical geometry optimization of an isolated PCBE-HCl with a *ZZZasa* conformation (QM-strained). This finding, a posteriori, seems to justify the strategy to use quantum mechanical calculations of the isolated chromophore as a first approximation for calculations of protein-bound cofactors.

Raman spectra

The experimental Raman spectra of α-CPC in H_2O and in D_2O were recorded at low temperature using 1064-nm excitation as described elsewhere (30). Although the excitation line is red shifted with respect to the first electronic transition, the (pre)resonance enhancement of the vibrational modes of the chromophore is still sufficiently high for effective discrimination of the Raman bands of the protein matrix. In addition, the relative intensities of the Raman bands of linear tetrapyrroles at 1064-nm excitation are very similar to those obtained under rigorous resonance conditions (31,32). Thus we will refer to the experimental Raman spectra of the PCB cofactor of α-CPC as RR spectra. In general, the main features of the experimental RR spectra are well described by both the QM/MM-calculated spectrum (M93) of α-CPC and the QM-calculated spectra of the isolated PCB models (PCBE-HCl) (Fig. 3); however, the hybrid method affords a distinct improvement in reproducing specific details.

Protonation marker bands

Characteristic marker bands for the protonated (cationic) state of tetrapyrroles are two bands between 1500 and 1600 cm^{-1} which appear with medium (1590–1550 cm^{-1}) and low (1530–1510 cm^{-1}) intensity in the experimental RR spectra (4). These bands, which have been assigned to the N-H in-plane bending (ip) modes of the inner rings *B* and *C*, disappear in D_2O and are replaced by a band between 1080 and 1050 cm^{-1} (N-D ip). In the experimental RR spectra of α-CPC, however, the expected bands appear as doublets at 1590/1567 and 1521/1511 cm^{-1} in H_2O (Fig. 4) and at 1073/1061 cm^{-1} in D_2O (Fig. 5), indicating that the PCB chromophore in α-CPC most likely exists in two different states.

The QM calculations of the two PCBE-HCl models afford geometries in which the chloride counterion is equidistant to

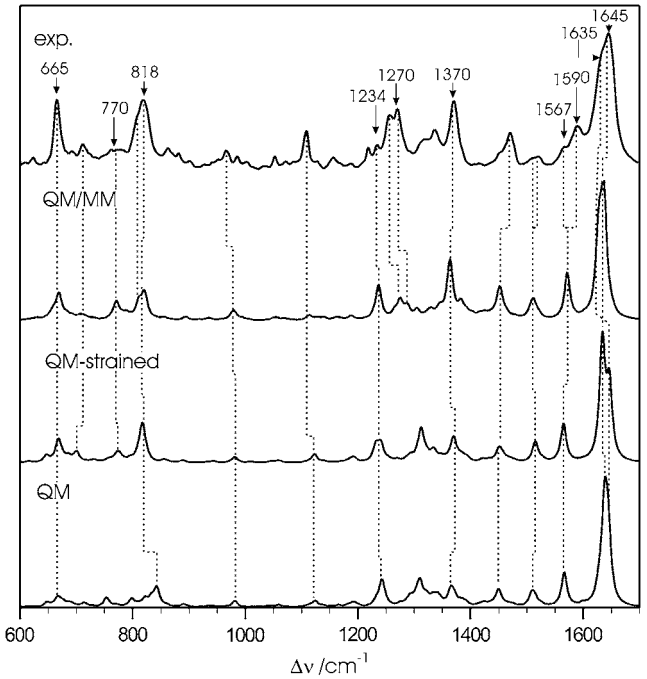

FIGURE 3 Experimental RR spectra of α-CPC in H_2O (*top*), calculated QM/MM Raman spectra of α-CPC (model M93), and calculated Raman spectra of PCBE-HCl in vacuo for the QM-strained model and the 'relaxed' QM model.

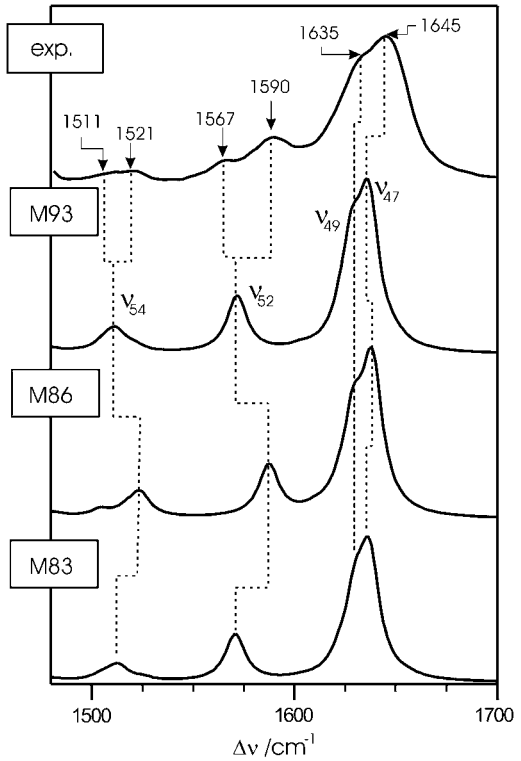

FIGURE 4 High frequency range (1500–1700 cm^{-1}) of the experimental (*top*) and calculated Raman spectra of α-CPC in H_2O using M83, M86, and M93 QM/MM partition schemes.

the N-H hydrogens of rings *B* and *C*. Correspondingly, structure and electron density distribution of these rings as well as the two bond lengths of the *B-C* methine bridges are virtually identical (Table 1). Thus, each of the two N-H ip modes, calculated at 1565 and 1514 cm^{-1}, include nearly equal contributions from the N-H ip coordinates of rings *B* and *C*. In the QM/MM calculations, this quasidegeneracy is removed by the interactions with the surrounding protein matrix, specifically with Asp-87 that is in H-bond distance to the ring *B* and *C* N-H groups. For each QM/MM model, the calculations reveal a bond length alternation for the *B-C* methine bridge with an ~0.005 Å increase and decrease of the bond lengths for C(ring *C*)-C(10) and C(ring *B*)-C(10), respectively (Table 1). Furthermore, the electron density distribution calculated for each hybrid model and, consequently, the corresponding vibrational spectra sensitively depend on the QM/MM boundary.

Whereas for M93 the N-H ip modes of the inner rings are calculated to be at 1572 (ν_{52}) and 1512 cm^{-1} (ν_{54}), the corresponding frequencies obtained for M83 (including the Asp side chain in the QM fragment) are at 1571 and 1513 cm^{-1}. In the case of M86, where the QM/MM boundary intersects the H-bond with Asp-87, the choice of the vdW force field parameters for the H-bond donor and acceptor plays an important role. The vibrational spectra calculated using the set A of vdW parameters for the chromophore (M86) display distinctly higher frequencies for both modes (1587 and 1524 cm^{-1}). Using soft vdW parameters (M86B), however, leads to physically meaningless underestimation of the frequencies by ~30 cm^{-1}. Except for M86B, all other models afford a contribution of ~40% N-H ip (ring *C*) and ~20% N-H ip (ring *B*) to the mode ν_{52}, whereas the mode ν_{54} is largely localized on ring *B* with only a small contribution from the ring *C* N-H ip coordinate.

The decreasing frequency of mode ν_{52} in the order M86 > M83 > M93 > QM is roughly correlated with the decreasing N(*B*)–Asp-87(O1)/chloride and N(*C*)–Asp-87(O1)/chloride distances and the increasing electron charge densities on the ring *B* and *C* nitrogens. A similar albeit less pronounced tendency is noted for ν_{54}. On the basis of these correlations one may assign the high frequency components of the ν_{52} and ν_{54} doublets in the experimental RR spectra of α-CPC (1590 and 1521 cm^{-1}) to a PCB species with relatively short N(*B*)/N(*C*)–Asp-87(O1) distances, whereas the corresponding low frequency components (1567 and 1511 cm^{-1}) reflect a PCB state with larger distances between proton donor and acceptor. The same interpretation may hold for the N-D ip doublet in the experimental RR spectrum (Fig. 5). It is important to emphasize that the static QM/MM approach cannot account for the coexistence of the two PCB forms, but the individual models that were chosen provide a description for one or the other form.

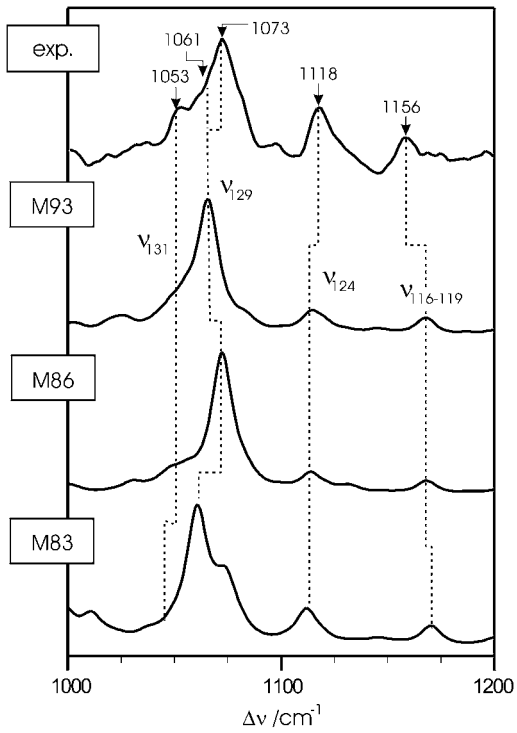

FIGURE 5 Middle frequency range (1000–1200 cm^{-1}) of the experimental (*top*) and calculated Raman spectra of α-CPC in D_2O using M83, M86, and M93 QM/MM partition schemes.

Marker bands for the methine bridge geometry

The modes dominated by the C=C stretching coordinates of the methine bridges and the individual pyrrole rings have been shown to display a band pattern that is characteristic for the configurations and conformations of the methine bridges in protonated tetrapyrroles (7). These modes are found in the region between 1550 and 1650 cm^{-1}, where the experimental RR spectrum of α-CPC in H_2O displays a strong and asymmetric peak with two components at 1645 and 1635 cm^{-1}. Both components are well reproduced by the QM/MM calculations, which afford two modes at 1637 (ν_{47}) and 1628 cm^{-1} (ν_{49}). The size of the QM fragment does not affect the frequencies ($\pm$1 cm^{-1}) and the relative intensities as well as the normal modes compositions, which indicate that C=C stretching coordinates of the *A-B* and *C-D* methine bridges prevail in the modes ν_{49} and ν_{47}, respectively.

In the QM calculations of the isolated PCBE-HCl, however, the order of both modes is reversed (Fig. 3) since the C=C (*A-B*) stretching mode is calculated at a higher frequency. This discrepancy can be attributed to the effect of the protein environment. In the QM/MM calculations, the protein-cofactor interactions reduce the *A-B* methine bridge torsion and thus cause a gain in conjugation, which may account for the 18-cm^{-1} downshift of the C=C (*A-B*) stretching mode as compared to the isolated PCBE-HCl models. The sensitivity of the C=C stretchings toward methine bridge torsions is also illustrated by comparing the spectra of the QM and QM-strained models. The 10° difference in the N_CC-CC dihedral angle at the *C-D* methine bridge is reflected by a 4-cm^{-1} frequency difference for the corresponding C=C stretching mode. Further modes between 1600 and 1650 cm^{-1}, including the C=C stretching of ring *D* and of the *C-B* methine bridge, are predicted with relatively low Raman intensity (Supplementary Material).

In previous QM studies of isolated tetrapyrroles (7), further support for the determination of the methine bridge geometry was derived from the comparison of the spectrally more crowded region between 1200 and 1500 cm^{-1}, which includes a variety of vibrational modes involving methyl deformations (~1370–1500 cm^{-1}), methylene wagging, and twisting deformations (~1200–1350 cm^{-1}), C-O stretchings of the propionic side chains (at 1393 and 1383 cm^{-1}), as well as C-H ip and N-H ip (rings *A* and *D*) deformations.

According to the QM/MM calculations, the N-H ip coordinates of the two outer rings *A* and *D* contribute to vibrational modes predicted between 1350 and 1450 cm^{-1}. These coordinates are mixed with C-C and C-N stretchings as well as with C-H ip deformations of the methine bridges. The amount by which these coordinates contribute to the potential energy distributions (PEDs) of the normal modes in this narrow frequency window varies strongly for the three different QM/MM models. However, the corresponding Raman intensity patterns predicted in this region do not significantly change. All QM/MM calculations give rise to an intense band at ~1363 cm^{-1} including the N-H ip bending of ring *A* and the C-H ip bending of the methine bridge *A-B*. This mode is assigned to the experimental band observed at 1370 cm^{-1} (Fig. 3). A further mode of relatively strong intensity originates from the C-H ip deformation of the *C-D* methine bridge. It is calculated at 1237 cm^{-1} and thus readily attributed to the experimental RR band at 1234 cm^{-1}. Also most of the other bands in this region are well reproduced by the QM/MM calculations, specifically for the models M93 and M86, which include the propionate side chains in the QM fragment. A distinctly poorer description is achieved by QM calculations of the isolated PCBE-HCl models.

Marker bands for torsions of the tetrapyrrole skeleton

The modes between 600 and 1000 cm^{-1} include the N-H and C-H out-of-plane (oop) coordinates and torsional coordinates of the tetrapyrrole skeleton, such that this spectral range is expected to contain information about structural details of tetrapyrroles beyond the level of methine bridge configurations and conformations. This region of the experimental RR spectrum of α-CPC is well reproduced by the M93 and M86 models so that nearly all of the experimentally observed bands can be assigned. Consequently, the strongest band in the experimental RR spectrum at 818 cm^{-1}, which is nearly insensitive to H/D exchange at the pyrrole nitrogens

(Figs. 6 and 7), and its low-frequency shoulder at 808 cm^{-1} are attributed to the modes ν_{161} (calculated at 820 cm^{-1}) and ν_{162} (calculated at 811 cm^{-1}). Both modes are predicted to include a significant (20%) C-H oop (*A-B* and *C-D* methine bridge) character. All other C-H oop-containing modes are calculated with very low Raman activity. This is also true for the mode ν_{167} (771 cm^{-1}), which is dominated by the N-H oop coordinates of rings *A* and *D*. This mode is assigned to the weak Raman band at 770 cm^{-1} that disappears in D_2O. A small contribution of the N-H oop coordinate of ring *B* is also calculated for the mode ν_{184} (663 cm^{-1}) which, like the adjacent mode ν_{183} (669 cm^{-1}), is dominated by the methine wagging of ring *D*. Both modes give rise to the relatively intense band at 665 cm^{-1}.

In the region between 900 and 600 cm^{-1}, the performance of the calculations is quite different for the QM and QM-strained models. Whereas the calculated spectrum of the QM-strained model provides a satisfactory description of the experimental spectrum, we note distinct discrepancies for the QM model (Fig. 3). Specifically, we note that the calculated C-H oop frequency of the *C-D* methine bridge (843 cm^{-1}) is distinctly higher than the calculated frequency of the QM-strained (826 cm^{-1}) and M93 (820 cm^{-1}) models as well as the experimental RR band frequency (818 cm^{-1}). This finding can be rationalized in terms of the structure of the rings *C* and *D*, which adopt a largely coplanar orientation in the QM model. In the QM-strained model the *C-D* methine bridge is twisted by 10° and thus is quite similar to the experimental crystal structure and the structure optimized by the QM/MM models M93 and M86. Thus, the C-H oop frequencies can be taken as a spectral marker for the torsions of the respective methine bridges.

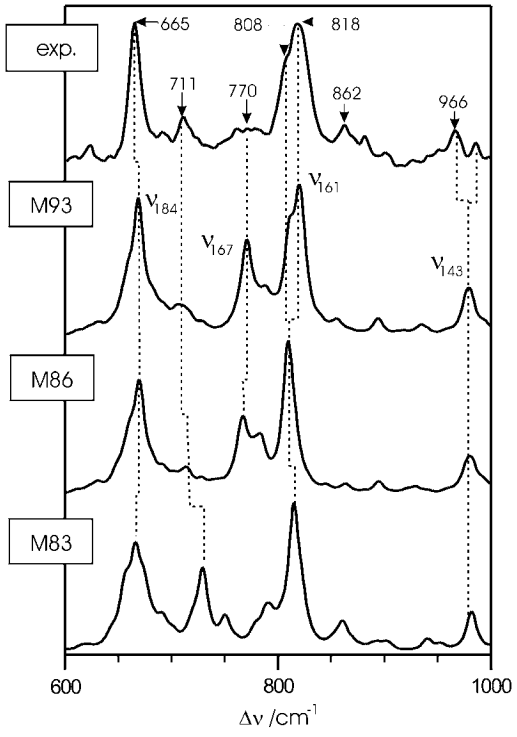

FIGURE 6 Low frequency range (600–1000 cm^{-1}) of the experimental (*top*) and calculated Raman spectra of α-CPC in H_2O using M83, M86, and M93 QM/MM partition schemes.

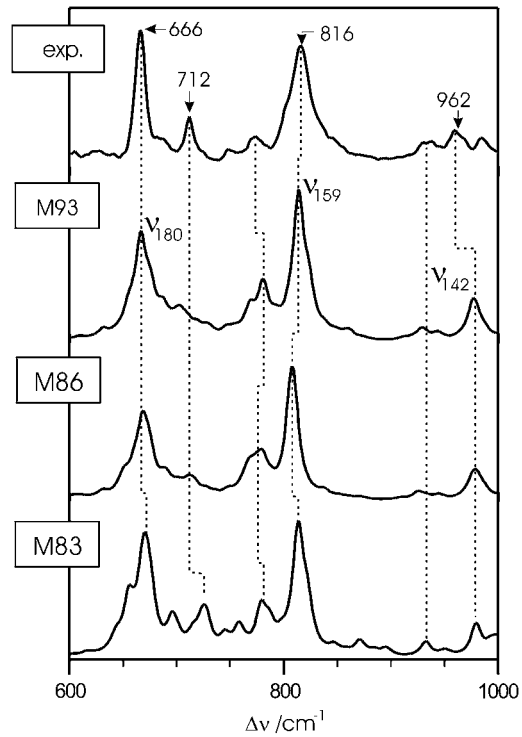

FIGURE 7 Low frequency range (600–1000 cm^{-1}) of the experimental (*top*) and calculated Raman spectra of α-CPC in D_2O using M83, M86, and M93 QM/MM partition schemes.

Influence of the propionic side chains on the vibrational spectra

Modes that are dominated by internal coordinates of the propionate side chains exhibit usually only relatively low intensity in the experimental RR spectra since these substituents are not part of the chromophoric site. The only exceptions refer to the modes ν_{105} (1288 cm^{-1}) and ν_{119} (1161 cm^{-1}), which include large contributions of the propionate methylene twisting and rocking, respectively. These modes are predicted to exhibit significant Raman intensity and may be attributed to the relatively intense bands at 1270 and 1156 cm^{-1} in the experimental spectrum (Figs. 3 and 5). Smaller but nonnegligible contributions of the propionate coordinates are found for a large number of modes below 1000 cm^{-1}, which, hence, should be sensitive to the conformations of these side chains. It is, therefore, not surprising that unlike the M93 and M86 models, the QM/MM calculations of the M83 model, which does not include the propionates in the QM fragment, do not satisfactorily reproduce the experimental RR spectra of α-CPC in H_2O and D_2O specifically in the low frequency region.

CONCLUSIONS

1. We have demonstrated that a hybrid QM/MM can be employed to calculate the Raman spectra of a relatively large cofactor in proteins. The application of this approach to α-CPC provides a considerable improvement compared to pure QM calculations of the isolated tetrapyrroles.
2. Size reduction of the QM fragment by omitting nonconjugated side chains has significant effects on the calculated Raman spectra and thus leads to a poorer reproduction of the experimental RR spectra, specifically in the frequency region below 1000 cm^{-1}.
3. H-bond interactions with the tetrapyrrole backbone require special attention since they have a pronounced effect on the electron density distribution of the chromophore and thus on the vibrational spectra. Conversely, proper choice of the QM fragment may provide further insight into the protein-cofactor interactions.
4. Compared to the in vacuo calculations, energy-optimized tetrapyrrole structures for the QM/MM models reveal structural changes that do not only refer to the side chains but also include the chromophore backbone. These geometric differences as well as the electrostatic interactions with the protein matrix cause changes in the entire spectral range (400–1700 cm^{-1}) that is usually probed in RR experiments.
5. The alterations also refer to the C=C stretching region, which is used for identifying the methine bridge configuration and conformation in tetrapyrroles. Thus, solely the comparison of the experimental RR spectra of the protein-bound chromophore with calculated spectra of isolated cofactors may not always lead to unambiguous conclusions. It is, therefore, required to enlarge the set of experimental data, for example, by extending the RR studies to protein adducts with isotopically labeled chromophores.
6. The good performance of the QM/MM approach here constitutes a promising basis for extending the studies to the various states of bacterial and plant phytochromes for which the three-dimensional structures are not known.

SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

The work was financed by the Deutsche Forschungsgemeinschaft (SFB 498, B9).

REFERENCES

1. Briggs, W. R., and J. L. Spudich. 2005. Handbook of Photosensory Receptors. Wiley, Weinheim, Germany.
2. Wagner, J. R., J. S. Brunzelle, K. T. Forest, and R. D. Vierstra. 2005. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature*. 438:325–331.
3. Mroginski, M. A., D. H. Murgida, and P. Hildebrandt. 2007. The chromophore structural changes during the photocycle of phytochrome: a combined resonance Raman and quantum chemical approach. *Acc. Chem. Res.* 40:258–266.
4. Kneip, C., P. Hildebrandt, K. Németh, F. Mark, and K. Schaffner. 1999. Interpretation of the resonance Raman spectra of linear tetrapyrroles based on DFT calculations. *Chem. Phys. Lett.* 311:479–484.
5. Magdó, I., K. Németh, F. Mark, P. Hildebrandt, and K. Schaffner. 1999. Calculation of vibrational spectra of linear tetrapyrroles. 1. Global sets of scaling factors for force fields derived by ab initio and density functional theory. *J. Phys. Chem. A*. 103:289–303.
6. Mroginski, M. A., K. Németh, T. Bauschlicher, W. Klotzbücher, R. Goddard, O. Heinemann, P. Hildebrandt, and F. Mark. 2005. Calculation of vibrational spectra of linear tetrapyrroles. 3. Hydrogen-bonded hexamethylpyrromethene dimers. *J. Phys. Chem. A*. 109: 2139–2150.
7. Mroginski, M. A., D. H. Murgida, D. von Stetten, C. Kneip, F. Mark, and P. Hildebrandt. 2004. Determination of the chromophore structures in the photoinduced reaction cycle of phytochrome. *J. Am. Chem. Soc.* 126:16734–16735.
8. Shurki, A., and A. Warshel. 2003. Structure/function correlations of proteins using MM, QM/MM, and related approaches: methods, concepts, pitfalls, and current progress. *Adv. Protein Chem.* 66:249–313.
9. Senn, H. M., and W. Thiel. 2007. QM/MM methods for biological systems. *Top. Curr. Chem.* 268:173–290.
10. Nonella, M., G. Mathias, M. Eichinger, and P. Tavan. 2003. Structures and vibrational frequencies of the quinones in Rb. sphaeroides derived by a combined density functional/molecular mechanics approach. *J. Phys. Chem. B*. 107:316–322.
11. Klähn, M., G. Mathias, C. Kötting, M. Nonella, J. Schlitter, K. Gerwert, and P. Tavan. 2004. IR spectra of phosphate ions in aqueous solution: predictions of a DFT/MM approach compared with observations. *J. Phys. Chem. A*. 108:6186–6194.
12. Eichinger, M., P. Tavan, J. Hutter, and M. Parrinello. 1999. A hybrid method for solutes in complex solvents: density functional theory combined with empirical force fields. *J. Chem. Phys.* 110:10452–10467.
13. Cui, Q., and M. Karplus. 2000. Molecular properties from combined QM/MM methods. I. Analytical second derivative and vibrational calculations. *J. Chem. Phys.* 112:1133–1149.
14. Woo, T. K., L. Cavallo, and T. Ziegler. 1998. Implementation of the IMOMM methodology for performing combined QM/MM molecular dynamics simulations and frequency calculations. *Theor. Chem. Acc.* 100:307–313.
15. Duerring, M., R. Huber, W. Bode, R. Ruembeli, and H. Zuber. 1990. Refined three-dimensional structure of phycoerythrocyanin from the cyanobacterium *Mastigocladus Laminosus* at 2.7 Å. *J. Mol. Biol.* 211:633–644.
16. Schirmer, T., W. Bode, and R. Huber. 1987. Refined three-dimensional structures of two cyanobacterial C-phycocyanins at 2.1 and 2.5 Å resolution. *J. Mol. Biol.* 196:677–695.
17. Hildebrandt, P., K. Németh, C. Kneip, and F. Mark. 1998. Calculations of vibrational spectra of linear tetrapyrroles. A key for interpreting the resonance Raman spectra of phytochrome. *Biophys. J.* 74:A133. (Abstr.)
18. Wang, X. Q., L. N. Li, W. R. Chang, J. P. Zhang, L. L. Gui, B. J. Guo, and D. C. Liang. 2001. Structure of C-phycocyanin from *Spirulina platensis* at 2.2 Å resolution: a novel monoclinic crystal form for phycobiliproteins in phycobilisomes. *Acta Crystallogr. Sect. D*. 57: 784–792.
19. MacKerell, A. D. Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. I. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe,

J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.

20. Riccardi, D., G. H. Li, and Q. Cui. 2004. Importance of van der Waals interactions in QM/MM simulations. *J. Phys. Chem. B.* 108:6467–6478.
21. Sherwood, P., A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjovoll, A. Fahmi, A. Schafer, and C. Lennartz. 2003. QUASI: a general purpose implementation of the QM/MM approach and its application to problems in catalysis. *J. Mol. Struct. THEOCHEM.* 632:1–28.
22. www.chemshell.org.
23. Billeter, S. R., A. J. Turner, and W. Thiel. 2000. Linear scaling geometry optimization and transition state search in hybrid delocalised internal coordinates. *Phys. Chem. Chem. Phys.* 2:2177–2186.
24. Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyouta, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramilo, R. Gomperts, R. E. Stratman, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople. 2004. Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT.
25. Rauhut, G., and P. Pulay. 1995. Transferable scaling factors for density functional derived vibrational force fields. *J. Phys. Chem.* 99:3093–4000.
26. Polavarapu, P. L. 1990. Ab initio vibrational Raman and Raman optical activity spectra. *J. Phys. Chem.* 94:8106–8112.
27. Komornicki, A., and J. W. McIver Jr. 1979. An efficient *ab initio* method for computing infrared and Raman intensities: application to ethylene. *J. Chem. Phys.* 70:2014–2016.
28. Beckmann, S., T. Wessel, B. Franck, W. Hönle, H. Borrmann, and H. G. von Schnering. 1990. [22]Coproporphyrin II for photodynamic therapy. *Angew. Chem. Int. Ed. Engl.* 29:1395–1397.
29. Wyllie, G. R. A., and W. R. Scheidt. 2003. NO orientation and tilting in (nitrosyl)iron(II) deuteroporphyrin IX. *Inorg. Chem.* 42:4259–4261.
30. Kneip, C., A. Parbel, H. Foerstendorf, H. Scheer, F. Siebert, and P. Hildebrandt. 1998. Fourier transform near-infrared resonance Raman spectroscopic study of the alpha-subunit of phycoerythrocyanin and phycocyanin from the cyanobacterium *Mastigocladus laminosus*. *J. Raman Spectrosc.* 29:939–944.
31. Kneip, C., P. Hildebrandt, W. Schlamann, S. E. Braslavsky, F. Mark, and K. Schaffner. 1999. Protonation state and structural changes of the tetrapyrrole chromophore during the Pr → Pfr phototransformation of phytochrome: a resonance Raman spectroscopic study. *Biochemistry.* 38:15185–15192.
32. Fodor, S. P. A., J. C. Lagarias, and R. A. Mathies. 1990. Resonance Raman analysis of the Pr and Pfr forms of phytochrome. *Biochemistry.* 29:11141–11146.