The Infrared Spectra of the Retinal Chromophore in Bacteriorhodopsin Calculated by a DFT/MM Approach

G. Babitzki, G. Mathias, and P. Tavan*

Theoretische Biophysik, Lehrstuhl für Biomolekulare Optik, Ludwig-Maximilians-Universität Oettingenstr. 67, 80538 München, Germany

Received: March 18, 2009; Revised Manuscript Received: May 25, 2009

In the preceding paper (DOI 10.1021/jp902428x), we have derived the polarized force field "PBR" for bacteriorhodopsin (BR) using hybrid methods which combine density functional theory (DFT) with molecular mechanics (MM) models. This polarized force field has enabled extended molecular dynamics (MD) simulations of BR's chromophore binding pocket which closely preserve the experimentally well-known structure. Here, we employ the PBR-MD trajectories obtained for the conformational substates prevalent at physiological temperatures as material for the DFT/MM computation of the chromophore's vibrational spectra. By comparison with DFT results on the structure and vibrational spectra of an isolated chromophore, we identify the structural and spectral changes induced by the protein environment. Comparisons with the wealth of experimental data available in the literature on the chromophore's vibrational spectra yield estimates on the accuracy of the DFT/MM descriptions. We discuss why highly accurate DFT/MM descriptions are expected to become a decisive tool for solving the long-standing enigma of how the light-driven proton pump BR actually works.

Introduction

In the introductory remarks to the preceding paper,[1] which we will call "part I" of this work (DOI 10.1021/jp902428x), we have shortly reviewed the physiological role of the retinal protein bacteriorhodopsin (BR) (see ref 2 for a more extended review) and have indicated the reasons why this light-driven proton pump has served as a standard test system for first applications of new biophysical techniques. Clearly, one could percieve also our reported efforts[1] to generate a computer model of the chromophore binding pocket of BR remaining stable in quite extended molecular dynamics (MD) simulations as a purely methods oriented contribution. In view of the widespread jaundice that the mechanism of proton pumping in BR has been thoroughly clarified in the past, such a misperception would seem quite natural.

One of the sources of the quoted prejudice is a "News and Views" article that appeared nearly a decade ago in Nature under the heading "Bacteriorhodopsin—the movie".[3] Here, we read in the abstract: "For 30 years and more, the mechanism of a microbial proton pump has been subject to increasingly sophisticated analysis. The full picture of how the pump operates is now emerging." As we will argue below, there is no such "full picture" up to now and there are very good reasons to put the issue of the BR pump mechanism, to which one of the authors contributed a speculative suggestion three decades ago,[4,5] once again on the agenda. The corresponding analysis will then additionally answer the question why efforts to compute the vibrational spectra of the BR chromophore as accurately as possible are still of vital scientific interest.

This paper is organized as follows. First, we will shortly analyze why the tremendous efforts to clarify the proton pump mechanism in BR have not yet rendered a solution. Thereby, we will motivate our own efforts and, hopefully, inspire others. Subsequently, we will sketch the computational methods employed by us for the computation of the chromophore's vibrational spectra. The corresponding results will be presented and discussed on the background of the available wealth of spectroscopic data.[6−8] A summary and an outlook conclude the paper.

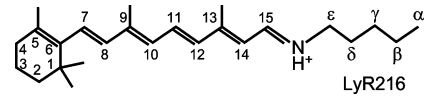

Figure 1. In the light-adapted initial state BR_{568} of the photocycle, the chromophore is the protonated Schiff base of all-*trans* retinal (all-*trans* $RSBH^+$) covalently attached to the apoprotein through Lys216.

What is the Chromophore Structure in L? Many experts in the field of BR are convinced that an answer to the above question would immediately lead to a deep and thorough understanding of the proton pump mechanism acting in BR (see, e.g., refs 4, 5, and 9−16). Here, "L" designates an early intermediate of the BR photocycle, which appears a few microseconds after the all-*trans* → 13-*cis* photoisomerization[17] of the BR_{568} chromophore (see Figure 1 for the chromophore's chemical structure in the light-adapted initial state BR_{568}). The key role of the L intermediate for proton pumping immediately becomes apparent if one considers the fact that the first proton transfer during the proton pump cycle occurs in the L to M transition. Here, the proton originally attached to the $RSBH^+$ is transferred to a nearby aspartate (Asp85) and the key mechanistic questions are (i) why do the chromophore and/or Asp85 drastically change their pK_a's in the BR_{568} to L transition and (ii) why is the subsequent proton transfer irreversible?[4,5,18]

Despite the fact that many X-ray structures[16,19−22] allegedly describing the L intermediate have been published, the detailed structure of this intermediate and particularly that of the chromophore have remained elusive up to now. The reason for this lack of knowledge is quite subtle:

All quoted attempts to characterize the L structure by crystallographic means were based on the same low-temperature

* Corresponding author. E-mail: tavan@physik.uni-muenchen.de. Phone: +49-89-2180-9220. Fax: +49-89-2180-9202.

10.1021/jp902432e CCC: $40.75
Published on Web 07/06/2009

trapping technique. Similarly, also recent solid-state NMR data on the chromophore structure in L were obtained for a structure trapped at 150 K.[23] If the trapped low-temperature L state would represent a reasonable model for the functional L intermediate of the room-temperature proton pump cycle, then this approach would be valid. However, already some time ago Rödig et al.[14] have shown by carefully comparing time-resolved Fourier transform infrared (FTIR) spectra with low-temperature ones that the structure of the chromophore is different in the two cases. The L intermediate occurring in the photocycle at physiological temperatures shows the so-called 15-Hoop band (where "15-Hoop" denotes the hydrogen out-of-plane vibration at C_{15}) with a sizable intensity, indicating substantial twists around the $C_{14}-C_{15}$ single bond of the $RSBH^+$,[24] which finding has important implications concerning the pump mechanism.[9−12] In contrast, the corresponding low-temperature spectrum exhibits a vanishing intensity of the 15-Hoop indicating that, in the cryogenically trapped L state, the chromophore is essentially planar in the region of the $C_{14}-C_{15}$ bond. Furthermore, in the L intermediate of the photocycle, the $C_{14}-C_{15}$ stretching vibration is found at a different spectral location than in the trapped so-called L state.[14]

The spectra of Rödig et al.[14] have recently been confirmed by Lórenz-Fonfría et al.[25] applying the same time-resolved step-scan FTIR technique to a wider spectral region (800−4000 cm^{-1}). Concentrating on vibrational bands associated with internal water molecules, these authors additionally found that the distribution of internal water molecules in the L intermediate of the photocycle substantially differs from that in the trapped low-temperature L state.

These results show that low-temperature X-ray structures of trapped L can neither provide correct information on the chromophore structure nor provide correct information on the distribution of internal waters. Moreover, such X-ray structures cannot give information on the pumping mechanism because the trapped L state predominantly shows a shunt reaction back to BR_{568} upon heating, as recently rediscovered by Dioumaev et al.[26] Thus, the low-temperature L, instead of proceeding to the M state through the deprotonation of the 13-*cis* $RSBH^+$, i.e., instead of pumping protons, prefers to return directly to the initial state BR_{568} through a thermal 13-*cis* to all-*trans* isomerization of the $RSBH^+$.[27]

In summary, the trapped low-temperature "L" state does not represent a model for the L intermediate of the photocycle because it has a distinctly different IR spectrum and because it does not pump protons. Concerning the open question on the proton pump mechanism, it is therefore not too important that all existing X-ray models of L[16,19−22] are at variance with each other concerning the structure of the chromophore and concerning the locations of the internal water molecules. Thorough discussions of the experimental and refinement problems, which cause these differences and similarly apply also to X-ray models of other cryogenically trapped BR intermediates, are given in refs 15 and 16. Note that theoretical treatments, which selected one of these trapped "L" structures as their starting point for simulations, identified some of the X-ray models for the chromophore as incompatible with quantum chemistry.[18,28] Note furthermore that other theoretical treatments[29] took the low-temperature X-ray structures of the various intermediates for granted and derived a mechanistic explanation of the proton pumping from these structures.

We conclude that only time-resolved techniques addressing the BR photocycle at physiological temperatures can provide a valid access to the key mechanistic question of the chromophore structure(s) in L. These techniques must be capable to cope with the likely conformational heterogeneities discussed in part I of this work[1] (DOI 10.1021/jp902428x) which suggest (at least) two different L states.

Up to now, the only detailed information on the chromophore structure in the L intermediate of the BR photocycle comes from time-resolved vibrational spectroscopy.[10,14,25,30] Therefore, currently, the only access to the chromophore structure in L (and, thus, to an insight into the physical mechanism of proton pumping) can be gained through a high-quality theoretical analysis of vibrational spectra. This view of the issue has inspired this and previous[31] work.

Computing the Vibrational Spectra of the Chromphore. As explained in the Introduction to part I (DOI 10.1021/jp902428x), a vibrational analysis, which aims at a reliable decoding of the observed FTIR and resonance Raman (RR) spectra in terms of the underlying chromophore structures, is a particularly hard task as long as the chromophore is protonated[32] like it is, e.g., in BR_{568} and in the early K and L intermediates. Therefore, early attempts[33−35] to model the intramolecular force field of an $RSBH^+$ empirically with fitted force constants, whose initial values were derived for unprotonated and strongly alternating polyenes, are expected to be error-prone. Nevertheless, all early (and nowadays still accepted) conclusions on the isomeric state of the BR chromophore in the various photocycle intermediates were derived from such analyses (see, e.g., ref 6). Other early descriptions, which were obtained by combining semiempirical quantum chemistry for the $RSBH^+$ with simplified one-counterion models for the protein environment,[36] suffered from the insufficient accuracy of the employed Hartree−Fock approximation and from the crudeness of the one-counterion model. These difficulties led to unresolved conflicts[9−12] in the structural interpretation of the spectra.

Today, however, DFT/MM techniques are available, which enable much more accurate descriptions of chromophore vibrational spectra (for a review, see ref 37), if the associated protein structures are precisely known. For light-adapted BR_{568}, the X-ray structure 1C3W has been gained by Luecke et al.[38] from crystals scattering to a resolution of 1.55 Å. This cryogenic structure is unaffected by the problems of partial occupancy shedding doubts on the quality of intermediate structures.[15] As shown in part I of this work (DOI 10.1021/jp902428x), this reliable starting structure 1C3W has enabled the generation of a room-temperature ensemble of BR structures by MD simulation. Furthermore, one may speculate whether the progress of computer technology and of computational methods can enable a simulation-based construction of accurate models for the K and L intermediates in the near future. Then, one could compute the associated vibrational spectra and compare with observations to conclude on structures. The reliability of such conclusions critically depends on the quality of the employed computational methods, of course.

This is the motivation why we wanted to check whether current DFT/MM hybrid methods can describe the vibrational spectra of the BR_{568} chromophore *in situ* with an accuracy that could suffice to decide such intricate questions like that of the chromophore structure in the L intermediate. Furthermore, the DFT/MM results will allow us to judge the quality of earlier attempts for a vibrational analysis of the BR_{568} chromophore, which were either based on empirical[33−35] or on simple quantum chemical[36] methods. Because the generally accepted concepts on the isomeric state of the chromophore in the various intermediate states of BR (or of other retinal proteins) originally

derive from the quoted empirical analyses, a corresponding reevaluation seems appropriate.

Methods

IR line spectra were calculated for the BR chromophore by conventional DFT/MM normal-mode analyses using the all-atom model 1C3W/MD[1] derived from the 100 K X-ray structure 1C3W.[38] In these calculations, three different partitions of the simulation system into DFT and MM fragments were used. Figure 2 depicts and labels these partitions. In the partitions S and M, the two fragments are connected by only one covalent C−C linkage which is treated by the "scaled position link atom method (SPLAM)".[31] In L, two SPLAM cuts are applied to the $C_{\alpha,215}-C'_{215}$ and $C_{\alpha,216}-C'_{216}$ bonds within the backbone. An additional DFT normal-mode analysis was applied to the isolated DFT fragment of partition M.

All DFT calculations were carried out with the plane-wave code CPMD[39] using the gradient-corrected exchange functional of Becke,[40] the correlation functional of Perdew,[41] and the norm-conserving pseudopotentials of Troullier and Martins.[42] The rectangular box containing the grid for the plane-wave expansion of the Kohn−Sham orbitals was placed around the respective DFT fragment in such a way that no atom of the fragment came closer than 3 Å to one of the faces and that the box volume became minimal. The size of the basis set was then defined by a plane-wave cutoff of 70 Ry. This particular DFT approach will be denoted as "MT/BP".

Inhomogeneously broadened IR spectra were calculated in the DFT/MM setting by applying the protocol for instantaneous normal-mode analyses (INMA) described in ref 43 to many snapshots taken from the MD trajectories discussed in part I of this work (DOI 10.1021/jp902428x).[1] These DFT/MM computations were carried out with the program combination EGO/CPMD originally introduced in ref 31. If not stated otherwise, the MM fragment was described with the PBR force field.[1]

For comparisons with experimental fundamentals, all harmonic frequencies calculated by MT/BP in the DFT and DFT/MM settings were scaled by a factor of 1.005. This factor is somewhat smaller than the factor of 1.0122 suggested in ref 44 for comparing harmonic MT/BP descriptions of quinone spectra with observed fundamentals.

We will characterize the geometry of the polyene chain within the $RSBH^+$ by three local observables, which we call the alternation α_i, the curvature κ_i, and the helicity η_i. Here, i is the number of the carbon atom C_i within the $RSBH^+$ (cf. Figure 1). The local alternation

$$\alpha_i = (-1)^{i-1}(\lambda_i - \bar{\lambda}), \quad i \in \{6, \ldots, 15\} \tag{1}$$

measures how much the length λ_i of the $C_{i-1}-C_i$ bond deviates from the average C−C bond length

$$\bar{\lambda} = \frac{1}{10}\sum_{i=6}^{15}\lambda_i \tag{2}$$

For a normal polyene or an unprotonated retinal Schiff base (RSB), α_i is positive and assumes everywhere similar values, whereas, for a $RSBH^+$, it should decrease from the cyclohexene ring toward the Schiff base.[45] The local curvature

$$\kappa_i = (-1)^i(\varphi_i - 120°) \tag{3}$$

gives the deviation of the bond angle φ_i between the atoms $C_{i-1}-C_i-X_{i+1}$ from the value of 120° (X ∈ {C, N}). For regularly extended polyenes, κ_i vanishes everywhere, whereas a total positive curvature

$$\kappa_{tot} = \sum_{i=6}^{15}\kappa_i \tag{4}$$

is associated with the well-known banana shape of retinal chromophores. Finally, the local helicity

$$\eta_i = (-1)^i(\theta_i - 180°) \tag{5}$$

measures the deviation of the dihedral angle θ_i at the C_i-X_{i+1} bond from the value of 180°, signifying planarity. The polyene helicity

$$\eta_{pol} = \sum_{i=7}^{15}\eta_i \tag{6}$$

then provides a rough measure of how much the plane of the C═N group is twisted with respect to that of the C_7═C_8 group attached to the cyclohexene ring.

In the BR chromophore, a positive unit charge is delocalized over the molecule. A clear representation of this delocalization is obtained, if the atomic ESP charges[31,46] resulting from DFT/MM calculations for the respective DFT fragment are summed up over the atom groups j defined by Figure 3. The total charge in the Schiff base region of the molecule is then given by the quantity

$$q_{SB} = q_{13-14} + q_{15-N} + q_{Lys} \tag{7}$$

Results and Discussion

Before comparing experimentally observed spectra with the inhomogeneously broadened IR spectra predicted for the BR_{568} chromophore at room temperature by the INMA approach in

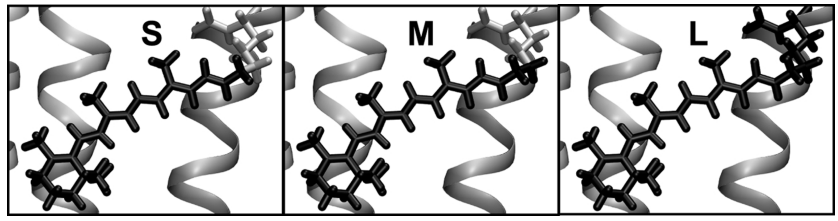

Figure 2. Three partitions S, M, and L of the BR simulation model defining differently extended DFT (black) and MM fragments (light gray).

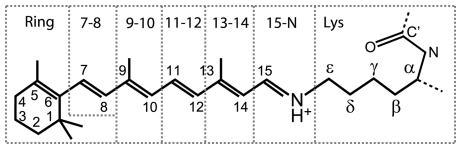

Figure 3. The total charges q_j within the seven indicated atom groups $j \in \{\text{Ring, 7-8, 9-10, 11-12, 13-14, 15-N, Lys}\}$ are used to characterize the charge delocalization predicted by DFT or DFT/MM treatments for the chromophore; q_j are the sums of the ESP charges[31,46] at the DFT atoms within the respective groups.

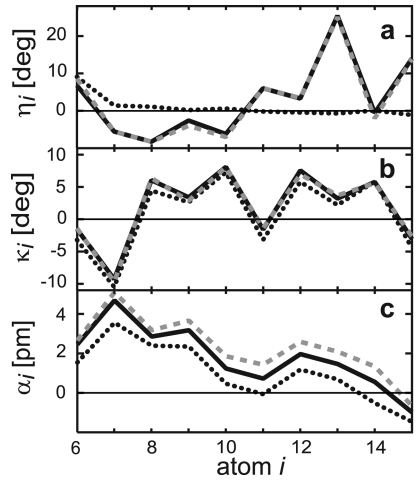

Figure 4. DFT and DFT/MM results explaining the effects of the apoprotein (structure 1C3W/MD, force fields PBR and C22, DFT fragment M) on the chromophore's geometry. Shown are the local (a) helicity η_i (eq 5), (b) curvature κ_i (eq 3), and (c) alternation α_i (eq 1) along the polyene chain. Different apoprotein models are distinguished by line style: VAC, black dotted; PBR, black solid; C22, gray dashed.

the DFT/MM setting, we will first show how the descriptions change with increasing complexity and realism of the modeling. Instead of the large structural ensembles resulting from MD simulations at 300 K required for INMA, we will use a single BR structure for these purposes. Thus, we will apply conventional DFT/MM normal-mode analyses to our all-atom model 1C3W/MD[1] of the 100 K X-ray structure 1C3W.[38]

How the Protein Modifies the RSBH$^+$. To obtain a first insight into the effects of the apoprotein on the properties of the chromophore, we carried out three different normal-mode analyses. The first calculation (labeled as VAC) treated the isolated DFT fragment of the DFT/MM partition M (cf. Figure 2), i.e., the all-*trans* ethyl-RSBH$^+$ in the vacuum. The two other calculations were executed in the DFT/MM setting for partition M and, correspondingly, included the apoprotein. As a molecular model, we took in both cases the all-atom structure 1C3W/MD, but in one case (labeled as C22), we employed the CHARMM22 partial charges[47] for the apoprotein and in the other case (labeled as PBR) those of the polarized force field PBR.

Chromophore Geometry. Figure 4 compares the helicity, curvature, and bond alternation obtained from the VAC, PBR, and C22 calculations. According to the dotted line in Figure 4a, the helicity of the isolated RSBH$^+$ vanishes nearly everywhere, indicating that this molecule has an almost planar π electron system. The single nonzero value η_6 marks a 10° torsion around the C_6—C_7 single bond connecting the cyclohexene ring with the planar ($\eta_{pol} = 0°$) polyene moiety of the chromophore. This twist is caused by sterical interactions of the two methyl groups at C_1 with the hydrogen atom at C_8 (cf. Figure 1). As witnessed by the almost identical C22 and PBR helicities, the embedding of the RSBH$^+$ into the protein causes twists around nearly all bonds within the polyene moiety which are essentially independent of the detailed electrostatics model. The largest twists are assigned to the C_{13}═C_{14} and C_{15}═N double bonds, resulting in a positive overall helicity of the chromophore's polyene chain ($\eta_{pol} \approx 24°$). Solely at the C_{14}—C_{15} single bond, the carbon chain is predicted as locally planar. This result seems to contradict the solid-state NMR data of Lansing et al.,[48] who determined for the H—C_{14}—C_{15}—H dihedral angle the small but nonzero value of 16 ± 4°. However, additionally extracting this dihedral angle from the C22 and PBR calculations yields the nonzero values of 13 and 11°, respectively. The smallness of the twist around the C_{14}—C_{15} bond fits well with the small intensity of the 15-Hoop band at 976 cm^{-1} in the IR spectrum of BR_{568} in D_2O.[12]

Thus, the various twists around the C—C bonds are mainly caused by steric interactions of the RSBH$^+$ with its binding pocket. In contrast, the net curvature of the polyene chain predicted by the curves in Figure 4b is mainly caused by intramolecular steric interactions of the two bulky methyl groups at C_9 and C_{13} with the neighboring hydrogen atoms. Correspondingly, the local curvatures are particularly large at $i \in \{8, 9, 10\}$ and $i \in \{12, 13, 14\}$. Furthermore, they nicely display the local symmetry of the chromophore structure defined by a mirror plane vertically crossing the polyene chain at C_{11} (cf. Figures 4b and 1). Within the protein, the total curvatures κ_{tot} measure 18.4° for C22 and 18.6° for PBR. Thus, they are independent of the electrostatics model. These total curvatures are much larger than the 6.9° found for the isolated chromophore, showing that steric protein—chromophore interactions additionally contribute to κ_{tot}.

The bond alternation defined by eq 1 is a very important quantity because it correlates with the stiffness of the various C—C bonds along a polyene chain with respect to torsion and stretching. Here, large α_i indicate stiff double and floppy single bonds, whereas vanishing α_i point to bonds of intermediate stiffness. Because the IR spectra of polyene derivatives are dominated by the C—C stretching vibrations, the inspection of a graph like Figure 4c can give a first hint on the relative sizes of the local stretching force constants entering corresponding normal modes.[36]

Figure 4c shows the decrease of the bond alternation α_i from the cyclohexene ring toward the Schiff base which is to be expected for protonated chromophores.[45] The average alternation is smallest for the isolated dye, becomes larger upon transfer into the protein (black solid and gray dashed lines), and sensitively depends on the details of the electrostatics model because for C22 the α_i are sizably larger than for PBR. In all cases, the local alternation α_i becomes slightly negative at C_{14}, indicating that all three chromophore models tend to assign a somewhat larger double bond character to the C_{14}—C_{15} bond than to the formal C_{13}═C_{14} double bond. This enhanced double bond character can partially explain why the twists induced by sterical strains into the chromophore are larger at the C_{13}═C_{14} than at the C_{14}—C_{15} bond (cf. Figure 4a).

The DFT/MM calculations start at the chromophore geometry given by the X-ray structure 1C3W.[38] Subsequently, they minimize the total energy function of the hybrid model for BR by slightly varying all atomic coordinates. Because of the limited resolution of the X-ray data and of the high quality of DFT/MM structure calculations, the thus obtained chromophore geometry should be much more accurate than the X-ray model.

Figure 5 illustrates the geometry changes resulting from a DFT/MM energy minimization. The largest changes are ob-

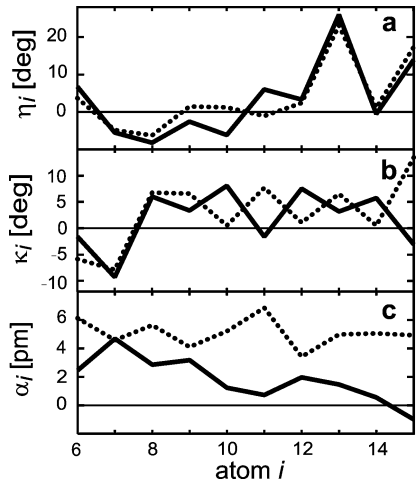

Figure 5. X-ray structure 1C3W[38] of the chromophore (dotted lines) compared with the DFT/MM prediction (solid lines, copied from Figure 4) on the chromophore's geometry: (a) helicity; (b) curvature; (c) alternation.

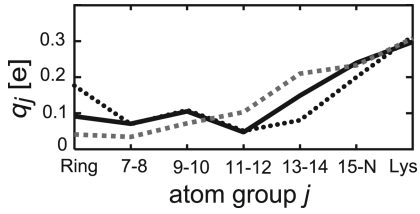

Figure 6. Effects of the apoprotein on the delocalization of the positive charge within the chromophore as measured by the group charges q_j defined in Figure 3. See the caption to Figure 4 for further information.

served for the bond alternation in Figure 5c. Whereas the X-ray model (dotted curve) assigns uniformly a large and positive alternation to the polyene chain, the DFT/MM result (solid curve) yields the expected decrease of alternation toward the Schiff base. Concerning the alternation, the X-ray model reflects the properties of the standard MM force field employed during refinement which imposes short lengths (1.35 Å) to formal double and large lengths (1.46 Å) to formal single bonds. The X-ray curvature in Figure 5b neither shows the banana shape nor the local mirror symmetry at C_{11} imposed by the methyl groups at C_9 and C_{13}. The corresponding errors of the X-ray model are about 3% and are of the same order of magnitude as the errors of the bond lengths. Solely for the helicity (Figure 5a), the differences are smaller. We suggest that the noted shortcomings of the X-ray structure reflect the prejudices contained in the MM force field used for refinement.

Charge Distribution within the Chromophore. Up to this point of the presentation, the bond alternation was the only chromophore property for which we have identified a strong dependence on external electric fields. A second and intimately related property of this kind is the distribution of the positive charge within the chromophore displayed by Figure 6.

Like the local alternations α_i, also the q_j show strong variations not only upon transfer of the $RSBH^+$ into the protein but also upon small changes of the local electrostatics. According to Figure 6, the delocalization of the charge toward the ring is strongest in the vacuum case (black dotted curve) and is successively reduced for the PBR (black solid curve) and C22 (gray dashed curve) models of the electrostatics. Conversely, the charge q_{SB} remaining in the Schiff base region (cf. eq 7) is smallest for VAC (0.6 *e*), intermediate for PBR (0.7 *e*), and

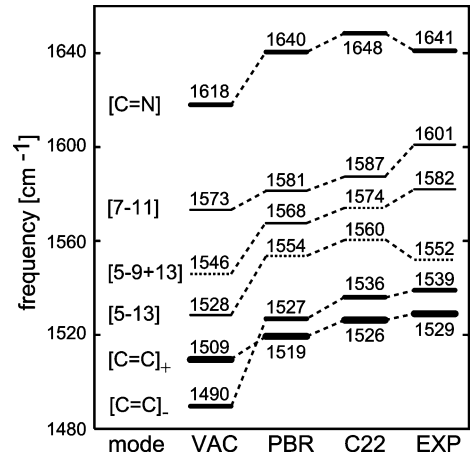

Figure 7. Effect of the apoprotein on the C═C and C═N stretch frequencies of the $RSBH^+$. The results of the DFT calculation VAC on the isolated $RSBH^+$ are compared with the DFT/MM calculations C22 and PBR on the BR chromophore and with experimental data (EXP) obtained for BR_{568} by Alshuth[49] and by Gerwert.[8] For further explanations, see the text and the caption to Figure 4.

largest for C22 (0.8 *e*). A comparison with Figure 4c now shows that q_{SB} is actually correlated with the size of the bond alternation.

Vibrational Spectra of the Chromophore. Because the bond alternation α_i represents a measure for the C—C stretching force constants in retinal chromophores, the prominent C—C stretching bands in the vibrational spectra of these dyes are expected to show a similarly strong field dependence. Figure 7 clearly verifies this expectation for the high-frequency spectral range characteristic for the C═C and C═N double bond stretches. As compared to the experimental band positions determined by RR spectroscopy for BR_{568} at physiological temperatures,[49] the C═C and C═N stretch frequencies calculated for the isolated chromophore show an average redshift of 31.6 cm^{-1}. The inclusion of the protein electrostatics reduces this root-mean-square deviation to 11.8 cm^{-1} for PBR and to 8.1 cm^{-1} for C22, respectively. The quoted deviations depend, of course, on the assignment of observed RR band positions to normal-mode frequencies which is given in Figure 7 by dashed lines. The foundation of this assignment will be discussed further below using data on isotopically labeled chromophores. Note that the increase of the double bond stretching frequencies displayed by Figure 7 for the sequence VAC, PBR, and C22 nicely correlates with the correspondingly increasing bond alternation and localization of positive charge near the Schiff base discussed further above.

Besides frequencies, Figure 7 contains two types of additional information: (I) The thickness of the lines marking the vibrational levels coarsely measures the associated IR intensity. The experimental intensity information required for the drawing of column EXP has been extracted from a FTIR difference spectrum (BR-M) published by Gerwert.[8] (II) The symbols [...] in the first column schematically characterize the compositions of the normal modes resulting from the DFT and DFT/MM treatments. For instance, the symbol [5-13] characterizes a normal mode predominantly composed of an out-of-phase combination of the C_5═C_6 and C_{13}═C_{14} stretches. Similarly, the symbol $[C═C]_+$ denotes a strongly delocalized, in-phase stretching motion of all C═C double bonds in the BR chromophore.[36] Figure 17 in the Supporting Information (SI) explains the given mode symbols through a graphical representation of the various normal modes. According to our experience, the

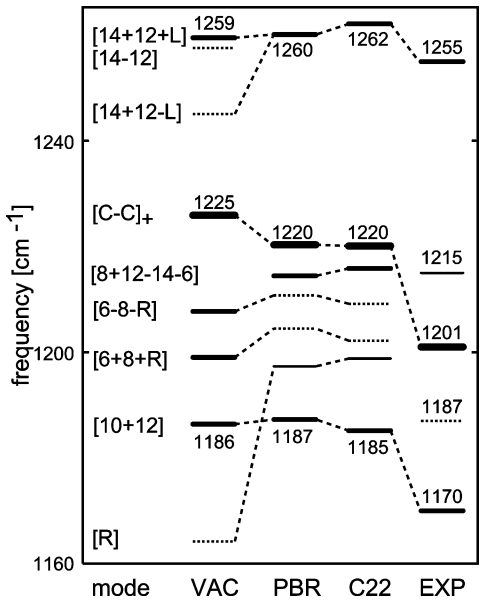

Figure 8. Effect of the apoprotein on the frequencies of C−C single bond stretches of the $RSBH^+$ (cf. the caption to Figure 7).

mode compositions of the double bond stretches are very stable properties and hardly change for different environments.

With these explanations, the experimental data and the DFT/MM results collected in Figure 7 clearly illustrate the well-known fact that the double bond stretching region in BR's IR spectrum is dominated by three intense bands. The most intense band at 1529 cm^{-1} belongs to the totally symmetric ethylenic stretch $[C{=}C]_+$. This band has a large shoulder which is centered at 1539 cm^{-1} and belongs to the totally antisymmetric mode $[C{=}C]_-$. According to the DFT/MM calculations, the intensity of the $[C{=}C]_-$ mode is by about a factor of 2 smaller than that of the $[C{=}C]_+$ mode. Quite intense is also the band of the C=N stretch at 1641 cm^{-1}. Smaller IR intensities are predicted for the normal modes [7-11] and [5-9+13] assigned to RR bands at 1601 and 1582 cm^{-1}, and a very small IR intensity for the mode [5-13] assigned to a RR band at 1552 cm^{-1}.[49]

According to the computational results in Figure 7, the transfer of the $RSBH^+$ from the vacuum into a protein environment does not only cause an overall blueshift of the frequencies but also has one additional consequence: Whereas in the isolated cation the less intense $[C{=}C]_-$ mode is predicted as the lowest C=C stretch, the DFT/MM calculations consistently claim that in BR_{568} the more intense symmetric combination $[C{=}C]_+$ should be found at a lower frequency. This claim agrees with the available experimental evidence.[35]

As noted above, the frequencies of the double bond stretches exhibit a clear correlation with the field-dependent bond alternation. Thus, one would naively expect an anticorrelation for the frequencies of the single bond stretches. Figure 8 compares the results of the calculations VAC, PBR, and C22 with experimental data[8,49] for the spectral region between 1160 and 1260 cm^{-1} which is associated with C−C single bond stretches and harbors the so-called fingerprint bands of the BR_{568} chromophore. In the IR spectrum, the fingerprint of the BR_{568} chromophore consists of a very intense band at 1201 cm^{-1} and of two intense bands at 1170 and 1255 cm^{-1}, respectively. In general agreement with the FTIR data, also the calculations predict one very intense and two intense C−C single bond stretches which are thus assigned to the observed fingerprint bands. Figure 18 in the SI graphically explains the mode compositions of these intense C−C stretches.

Among the C−C stretches generating the fingerprint motif, only the frequency of the most intense mode $[C{-}C]_+$ (which is a delocalized in-phase combination of all C−C stretches within the polyene moiety of the $RSBH^+$) exhibits the expected anticorrelation with the bond alternation. Correspondingly, the $[C{-}C]_+$ frequency is redshifted with increasing alternation by 5.6 and 0.2 cm^{-1} in the transitions VAC to PBR and PBR to C22, respectively. Although bringing the computational results closer to the observations, the DFT/MM calculations still overestimate the C−C (Figure 8) and underestimate the C=C frequencies (Figure 7), indicating that the DFT/MM models still underestimate the bond alternation of the chromophore in BR_{568}. In the chromophore of BR_{568}, the positive charge seems to be even more strongly localized toward the Schiff base than in the DFT/MM models discussed so far.

The frequencies of the other C−C stretches do not show a likewise simple dependence on the bond alternation. In the VAC to PBR transition, for instance, the frequency of the mode [14+12+L] increases by 1 cm^{-1}, although the $C_{14}{-}C_{15}$ and $C_{12}{-}C_{13}$ bonds become elongated in this transition, indicating smaller stretching force constants. Such a more complicated behavior has to be expected for a congested spectral region like the one covered by Figure 8, because here the C−C stretches are strongly coupled to many higher and lower frequency modes (including the in-plane bending modes of the hydrogen atoms attached to the polyene chain). On the one hand, this coupling can strongly diminish the effects of altered C−C stretching force constants. On the other hand, in the case of nearly degenerate levels, even small changes of the geometry or of external fields can easily induce large changes of mode compositions and of spectral locations. For instance, in the transition from VAC to PBR, one sees a large blueshift of a C−C stretching mode localized in the cyclohexene ring (mode [R]) and the appearance of a new mode [8+12-14-6] spectrally near the $[C{-}C]_+$ mode which was absent in VAC. The complexity of the coupling patterns in the densely populated spectral region of the C−C single bond stretches is the reason why we will mainly look at the much simpler high-frequency region of the C=C and C=N double bond stretches in our continued assessment of different theoretical descriptions.

Intermediate Summary. The comparison of the single point calculations VAC, PBR, and C22 with experimental data has shown that the inclusion of the apoprotein not only enables a highly precise determination of the chromophore geometry but also generally brings the calculated frequencies much closer to the experimental values. On the other hand, the DFT/MM calculations seem to overestimate the delocalization of the positive charge into the polyene moiety. In this respect, the C22 force field apparently performs better than PBR which is an unexpected finding. However, the PBR force field has been developed for a different partitioning (L) of the system into DFT and MM fragments. Thus, we now must check how the choice of the DFT fragment affects the descriptions.

Effects of the DFT Fragment Size. For this study, we have chosen the three partitionings S, M, and L of the system into DFT and MM fragments (cf. Figure 2), because we wanted to keep the DFT fragment for reasons of computational feasibility as small as possible. Here, the apoprotein was modeled by PBR[1] and the water molecules by Jorgensen's transferable three-point potential (TIP3P).[50]

As demonstrated by Figure 19 in the SI, the variation of the DFT fragment size has a negligible effect on the curvature of

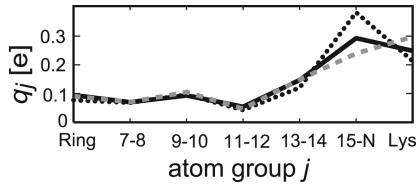

Figure 9. Effect of the DFT fragment size (cf. Figure 2) on the charge distribution q_j within the RSBH$^+$ (DFT/MM results, structure 1C3W/MD force field PBR): S, black dotted; M, gray dashed; L, black solid.

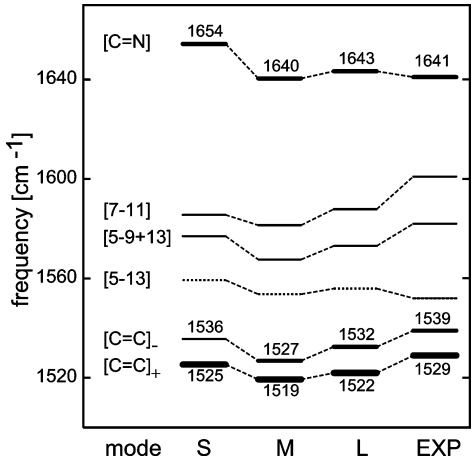

Figure 10. Effect of fragment size S, M, or L on the C=C and C=N stretch frequencies of the BR chromophore from DFT/PBR computations on the structure 1C3W/MD. Like in Figure 7, the computed line spectra are compared with experimental data.

the polyene moiety and causes only small changes of the dihedral angles and of the bond alternation near the Schiff base. In good correlation with the charge distribution shown in Figure 9, the bond alternation is only slightly larger for partition S than for M or L. In fact, for S, the total charge q_{SB} near the Schiff base (cf. eq 7) is only by 3% larger than that for M or L. Figure 9 confirms that the variation of the DFT fragment changes the charge distribution solely near the C=N group.

The slightly larger value of q_{SB} leads for model S to the expectation of somewhat stiffer C=C and C=N double bonds. Figure 10 demonstrates that this is actually the case. Compared to M and L, the double bond stretching frequencies obtained for S are shifted to the blue by averages of 8 and 4 cm^{-1}, respectively. However, as compared to model M, also for L the frequencies are shifted toward the blue. As a result, the root-mean-square deviation between the experimental and calculated frequencies is reduced from 11.8 cm^{-1} for M to 9.3 cm^{-1} for S and to 7.8 cm^{-1} for L. The latter deviation is even a little smaller than the value of 8.1 cm^{-1} noted above for combining the DFT fragment M with the C22 force field. Correspondingly, it seems that the PBR force field in combination with the largest DFT fragment L provides the best description of the experimental data.

Note, however, that the model S with its larger alternation and charge q_{SB} performs better than L when one exclusively considers the C=C stretch frequencies. It is only the 13 cm^{-1} overestimate of the C=N stretch frequency found for S which excludes this most simple modeling. Replacing the lysine chain by a methyl group, as one does in the DFT treatment with fragment S, apparently places the boundary between the DFT and MM regions too close to the Schiff base group for a correct evaluation of the C=N force constant. Furthermore, the suboptimal performance of model M is caused by the choice

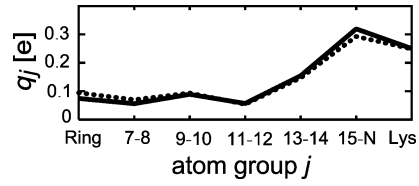

Figure 11. Distribution of charges q_j within the chromophore for conformation C_1 and DFT/MM partition L: The 300 K ensemble average (solid curve) is compared with the 100 K result (dotted curve). See the text for an explanation.

of the electrostatics for the lysine chain in the PBR force field favoring a somewhat too strong delocalization of the positive charge. As a result, the computationally most expensive variant L among the studied alternatives for DFT/MM partitioning had to be chosen for the subsequent study of temperature effects.

Temperature Effects. The above descriptions of chromophore structures and IR spectra were obtained for an all-atom version of the X-ray structure 1C3W.[38] Correspondingly, these computational results pertain to a system at 100 K, whereas the experimental spectra employed for comparisons were recorded at room temperature.

In part I of this work (DOI 10.1021/jp902428x), we succeeded to generate a 300 K conformational ensemble of the chromophore binding pocket by MD simulations.[1] Due to the use of the polarized force field PBR, these simulations preserved the crystallographic structure over extended time spans (>50 ns). Using a specially designed sampling technique, the 300 K ensemble was shown to be composed of (at least) two, and possibly even three, distinct and slowly interconvertible conformational substates called C_1, C_2, and C_3, respectively. The increase of the temperature from 100 to 300 K turned out to have two effects: (i) Within a given substate, the amplitudes of the thermal fluctuations become larger. (ii) The relative populations of the substates C_i are modified. Whereas the crystallographic conformation C_1 is predominantly populated at 100 K, conformation C_2 gains a considerable statistical weight at higher temperatures. For the rare conformation C_3, the limited statistics sampling precluded an accurate prediction of its weight at 300 K. We expect that the effects (i) and (ii) will modify the average structure and the IR spectra of the chromophore.

Effects of Thermal Fluctuations. First, we will address point (i) for conformation C_1. Our all-atom model 1C3W/MD obviously represents a 100 K snapshot of C_1. Furthermore, a PBR-MD simulation starting at 1C3W/MD has rendered for C_1 a 300 K trajectory covering 50 ns. Applying the INMA approach (cf. Methods) to this trajectory enables the computation of room-temperature IR spectra. Thus, for state C_1, one can identify the effects of increasing thermal fluctuations by comparing average 300 K properties of the chromophore with data from the 100 K snapshot. At 100 K, the thermal fluctuations are small and, therefore, a single minimized structure suffices.

Figure 11 compares for partition L and conformation C_1 the average distribution of charges q_j within the chromophore at 300 K with the distribution within the 100 K snapshot (dotted curve, identical to the solid curve in Figure 9). The 300 K average has been calculated from 25 snapshots selected from the central part of the 50 ns PBR-MD trajectory at 5 ps temporal increments. This snapshot ensemble has also been used for the construction of all other DFT/MM results that will be presented for C_1 at 300 K.

Figure 11 shows for state C_1 that elevating the temperature from 100 to 300 K leads within the chromophore to a slight shift of positive charge toward the Schiff base. Correspondingly, the value of q_{SB} increases by 3% upon heating and becomes as

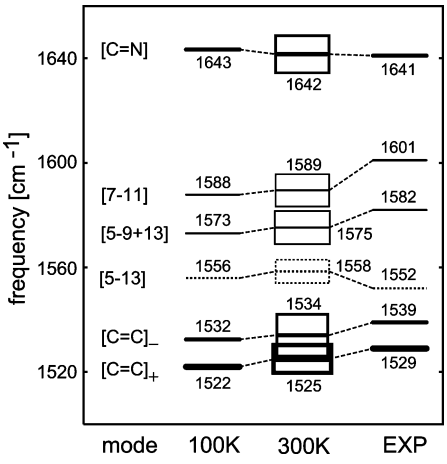

Figure 12. Average C═C and C═N stretch frequencies and corresponding band widths in conformation C_1 at 300 K compared with the 100 K results (structure 1C3W/MD, partition L; see also Figure 10). The band widths are indicated by heights of boxes measuring the standard deviations of the underlying 300 K frequency ensembles. Experimental data (cf. Figure 7) are given as a reference.

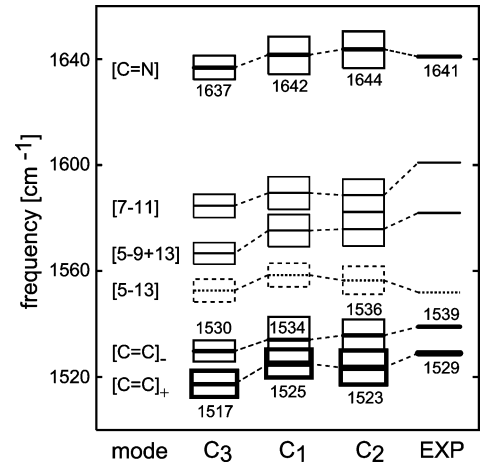

Figure 13. Average C═C and C═N stretch frequencies and standard deviations in the 300 K conformations C_1, C_2, and C_3. See the caption to Figure 12 for further explanations.

large as in the low-temperature fragment S. In line with the behavior of q_{SB}, also the average bond alternation increases a little (0.1 pm) upon heating. Figure 20 in the SI documents that these and other structural changes induced by heating are actually quite small. The noted increases of q_{SB} and of the bond alternation entail for state C_1 the prediction that heating should shift the C═C stretch frequencies to the blue.

Figure 12 confirms this prediction. Upon heating, all C═C frequencies are shifted nearly uniformly by about 2 cm^{-1} to the blue. In contrast, the C═N stretch is shifted by 1 cm^{-1} to the red. Correspondingly, the deviation between the calculated and observed double bond stretching frequencies decreases from 7.8 to 6.6 cm^{-1}. Despite the quite close matching of calculated and experimental frequencies, which meanwhile has been achieved through a successively more realistic and complex modeling of the experimental situation, the C═C stretching force constants still seem to be a little too small. Thus, even our current chromophore model still features a charge delocalization that is apparently a little too large.

Effects of Conformational Transitions. To characterize the three conformational substates identified in part I (DOI 10.1021/jp902428x), we have selected, in addition to the 25 C_1 snapshots, 20 snapshots from each of the extended 300 K MD trajectories sampling the conformational states C_2 and C_3.

Figure 21 and the associated text in the SI present, compare, and discuss associated DFT/MM results that explain the influence of the respective conformational state on the structure of the chromophore. Inspecting these data, one concludes that the sterical strain exerted by the binding pocket on the chromophore decreases in the sequence $C_1 \rightarrow C_2 \rightarrow C_3$. For instance, in state C_3, the chromophore has a reduced curvature just like the model VAC discussed further above. Furthermore, in C_2 and C_3, the polyene moiety of the $RSBH^+$ exhibits much smaller twists than in the sterically restrained but enthalpically favored[1] conformation C_1. On the other hand, the localization of positive charge near the Schiff base and the bond alternation are seen to increase in the sequence $C_3 \rightarrow C_1 \rightarrow C_2$ such that the C═C stretching frequencies should exhibit blueshifts in that sequence.

Figure 13 now shows that the C═C stretching frequencies of the states C_1 and C_2 are actually blueshifted with respect to those of state C_3. However, for the transition from C_1 to C_2, no overall blueshift is seen, indicating that the spectral positions of the C═C stretches do not exclusively depend on the degree of charge delocalization but also on the chromophore geometry (as mentioned above, in C_1, the chromophore exhibits larger twists than in C_2). Compared with the experimental double bond stretch frequencies, the calculated values show root-mean-square deviations of 11.1 cm^{-1} for C_3 and of 6.6 cm^{-1} for C_1 and C_2 (the latter two spectra deviate from each other by 1.6 cm^{-1}).

Considering the band widths, which are measured by the heights of the boxes surrounding the frequency levels, one concludes that the IR bands associated with the three conformational states will strongly overlap. Therefore, particularly the IR spectra of C_1 and C_2 should be very similar in the high-frequency region (and as shown by Figure 22 in the SI, such a similarity also holds in the fingerprint region). The structural differences between C_1 and C_2, which are mainly localized at the aliphatic chain connecting the $RSBH^+$ with the polypetide backbone and at the backbone near Lys216, are thus predicted to be spectroscopically silent.

Total Spectral Effect of Heating. According to the results of part I (DOI 10.1021/jp902428x), the conformations C_1 and C_2 should dominate the 300 K conformational ensemble, whereas C_3 should occur in much smaller or even negligible amounts. Thus, we now assume that the crystallographic conformation C_1 is exclusively present at 100 K and that a 42% admixture of C_2 appears at 300 K. Then, the DFT/MM normal-mode analysis of the structure 1C3W/MD can be taken as our description of the IR spectrum of the BR chromophore at cryogenic temperatures (cf. column "L" in Figure 10 or column "100 K" in Figure 12). Furthermore, we can construct an INMA prediction for the room-temperature IR spectrum of the BR chromophore by choosing 25 snapshots from the C_1 trajectory and 18 snapshots from the C_2 trajectory at 300 K to model a 58:42 conformational mixture. Next, one can compare this 300 K spectrum with its 100 K relative to study the combined effects of enhanced thermal fluctuations and changed conformational composition.

The graph thus obtained for the C═C and C═N stretching frequencies would be nearly identical to the one shown in Figure 12, because the frequencies calculated for the 300 K mixture differ from those shown in Figure 12 for the pure conformation C_1 at 300 K by at most 1 cm^{-1}. Because of this similarity, the 300 K mixture data are not documented at this point. Instead, they will be shown and discussed in a different context further below (cf. Figure 15).

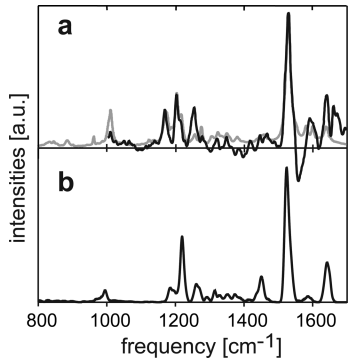

Figure 14. IR spectrum (b) of the BR chromophore at 300 K as calculated by the INMA procedure in the DFT/MM setting for a 58:42 mixture of conformations C_1 and C_2 compared with experimental spectra. The black curve (a) is a BR-M FTIR difference spectrum extracted from the thesis of Gerwert.[8] The gray curve in the background of part a is a RR spectrum of BR_{568} taken from the thesis of Alshuth.[7] All spectra have been normalized to identical heights of the dominating peak of the ethylenic stretch.

Interestingly, however, the 42% admixture of C_2 sizably decreases the differences between the 100 and 300 K data shown in Figure 12 (which refer to pure C_1). As a result, in the high-frequency region, the band positions are predicted to shift by at most 2 cm^{-1} upon heating. Furthermore, the 42% admixture of C_2 reduces the root-mean-square deviation of the calculated from the experimental double bond stretching frequencies slightly from 6.6 cm^{-1} (pure C_1) to 6.5 cm^{-1} (mixture). These numbers have to be compared with the deviation of 7.8 cm^{-1} identified for the cryogenic model and suggest that our 300 K mixture model should describe the vibrational spectra of the BR chromophore reasonably well.

The IR Spectrum of the BR Chromophore at 300 K. To check this expectation on the quality of our 300 K model, we construct from our INMA data on the conformations C_i a prediction for the absolute IR spectrum $I(\nu)$ of the BR chromophore. For this purpose, we center at each INMA frequency a normalized Gaussian with a standard deviation of 5 cm^{-1} (here, the width has been chosen such that the resulting superposition is sufficiently smooth, despite the limited size of the INMA snapshot ensemble, while still allowing the distinction of neighboring bands). Weighting this Gaussian by the intensity calculated for the associated normal mode, we determine the IR spectrum $I(\nu|C_i)$ of C_i by summing over all Gaussian bands which can be derived from the associated INMA data set. For C_1 and C_2, the INMA data sets contained 25 and 20 line spectra, respectively. The predicted spectrum $I(\nu)$ is then given as the mixture $0.58I(\nu|C_1) + 0.42I(\nu|C_2)$.

Figure 14b shows the resulting prediction for the IR spectrum of BR_{568} in the frequency range from 800 to 1700 cm^{-1}. For visual comparison, Figure 14a also shows a BR-M FTIR difference spectrum and a RR spectrum of BR_{568}. These spectra have been scanned from the theses quoted in the caption. Quite obviously, the comparisons of the depicted spectra are hampered by three facts:

(i) The RR and IR intensities follow from different mechanisms. Nevertheless, for BR, the spectra look quite similar with the notable exception of the third fingerprint band at 1255 cm^{-1} which is intense in IR and weak in RR. (ii) The FTIR spectrum (a) is a difference spectrum between the initial state BR_{568} of the photocycle and the long-lived intermediate M. However, because the chromophore is deprotonated in M, the IR intensities of the M chromophore are small. Correspondingly small are the negative bands contributed by the M chromophore to the difference spectrum (a). (iii) In contrast to RR, FTIR contains additional contributions from the protein. Particularly in the spectral range between 1500 and 1700 cm^{-1}, the FTIR difference spectrum shows signatures of strong amide bands. These difference bands signify that the protein backbone undergoes conformational changes in the BR to M transition. Thus, RR spectra are necessary to identify the chromophore bands in FTIR difference spectra.

Independent of these difficulties, the main spectral features in the regions of the C—C single and double bond stretching vibrations are clearly visible in the FTIR difference and in the calculated IR spectra: Three fingerprint bands are seen near 1200 cm^{-1}, and above 1500 cm^{-1}, one immediately recognizes the strong ethylenic and C═N stretching vibrations. The visual inspection additionally suggests that the calculated band widths are quite close to the observations.

However, one also sees the remaining shortcomings of the description: Whereas FTIR determines for the three fingerprint bands nearly equal intensities, the calculation assigns too small intensities to the two companions of the central and very intense $[C-C]_+$ band. Furthermore, the calculated fingerprint bands are not as widely spread as the experimental ones and their center is shifted by about 13 cm^{-1} to the blue.

In the discussion of Figure 8, we have suggested that these shortcomings are most likely caused by a delocalization of positive charge which is smaller in the chromophore of BR_{568} than in its DFT/MM model. For the model, a reduction of the delocalization requires that the localizing electric field becomes strengthened, e.g., by moving some negative charge density a little closer to the Schiff base group.[36] Therefore, the remaining differences seem to indicate that our electrostatics model is—despite the use of a polarized force field for the apoprotein—not yet good enough.

The Effect of Polarized Water Molecules. As explained in part I of this work (DOI 10.1021/jp902428x),[1] we restricted the polarization of the MM fragment to the protein material in the chromophore binding pocket of BR and chose for the three crystallographic water molecules in this pocket the standard nonpolarized TIP3P water model.[50] However, after completion of the MM−MD simulations and of the computation of the associated room-temperature spectra, we found by DFT/MM calculations that the dipole moments of these three water molecules are by about 10% larger than those of water molecules in the bulk liquid.[1]

To estimate the influence of more polar water models on the chromophore spectra, we repeated for the structure 1C3W/MD the DFT/MM calculation "L" of the chromophore's cryogenic IR spectrum (cf. Figure 10) with such water models. As demonstrated by Figures 23 and 24 in the SI, the larger water dipoles left the C═N stretch frequency invariant and shifted the C═C stretch frequencies on average by 2 cm^{-1} to the blue and the frequencies of the C—C fingerprint bands by about 1.5 cm^{-1} to the red. Therefore, the deviations of the calculated stretch frequencies from the experimental data reduced from 7.8 to 7.2 cm^{-1} for the double bonds and from 15.1 to 13.8 cm^{-1} for the single bonds (see the SI for further details).

As a result, the use of polarized water molecules should not only further stabilize the native structure in MD simulations[1] but should also further improve the match between the experimental data and the INMA spectra. Transferring, for instance, the 2 cm^{-1} blue shift of the C═C stretches, which is caused by the larger water dipoles (Figure 23 in the SI), to the room-temperature conformational mixture reduces the frequency deviation from 6.5 to 5.6 cm^{-1} for the double bonds.

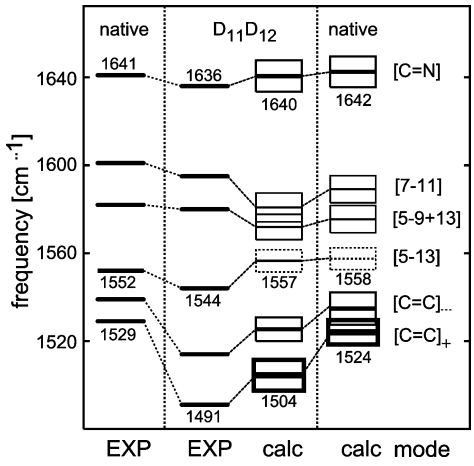

Figure 15. Effect of deuteration at C_{11} and C_{12} on the C═C and C═N stretch frequencies of the BR chromophore. RR frequencies[35] labeled as EXP are compared with DFT/MM results (calc) obtained for our 300 K conformational mixture (0.58 C_1 + 0.42 C_2). The assignments are indicated by the dashed lines and are explained in the text (cf. additionally the caption to Figure 12).

Accuracy Exemplified for an Isotope Effect. The remaining deviations between the calculated spectra and the experimental data represent important guidelines for further improvements. Nevertheless, the thus achieved description is not too bad either. Quite certainly, it is much better than the combination of semiempirical quantum chemistry with one-counterion models applied earlier[36] to the vibrational analysis of the BR chromophore or even the purely empirical modeling in ref 35. However, a precise quality assessment requires the consideration of isotope effects. In that respect, the chromophore of BR is a particularly fortunate case because RR spectra on a plethora of isotopically labeled chromophores are available.[35] At the same time, this wealth of data makes the presentation of corresponding results quite cumbersome, because one has to analyze many more details.

Therefore, we decided to focus this paper on the analysis of how the protein binding pocket and the conformational heterogeneity affect the chromophore spectra and to what extent an increasingly more accurate and complex modeling can improve the descriptions. Correspondingly, here we present solely one carefully selected isotope effect. The SI discusses three further isotopic substitutions. One is the spectral effect of deuterating the Schiff base, one the effect of inserting a $^{13}C_{14}$,$^{13}C_{15}$ double label into the chromophores carbon skeleton, and one is the effect of ^{15}N substitution.

Among all isotopically labeled chromophores investigated by Smith et al.,[35] there is only one species which shows six nonoverlapping and clearly visible RR bands in the high-frequency region of the double bond stretches. This species is the 11,12, dideutero $RSBH^+$. Because the chromophore contains six double bonds, the RR spectrum of this compound uniquely defines the frequencies of six double bond stretching modes. All other studied isotopomers[35] show a considerable overlap of bands and, therefore, the determination of all six C═C frequencies requires error prone deconvolution techniques.

The central columns of Figure 15 contain the RR and DFT/MM results for the double bond stretch frequencies of the 11D,12D-chromophore. Here, the assignment of the calculated to the experimental frequencies was exclusively based on the respective spectral sequences. The modes calculated for the labeled chromophores could be easily assigned to those of native chromophore by visual comparison of mode compositions. The resulting assignment of the labeled to the unlabeled RR frequencies and to the associated modes generally agrees with that by Smith et al.[35]

An exception is the 11D,12D-band at 1544 cm^{-1} which our DFT/MM calculation assigns to an almost pure C_5═C_6 stretch calculated at 1557 cm^{-1}. Despite its considerable intensity, Smith et al.[35] had assigned this band, like the much weaker RR band at 1552 cm^{-1} in the native chromophore, to a combination mode of unknown composition. This judgement was caused by the results of their empirical normal-mode analysis which consistently predicted the C_5═C_6 stretch near 1600 cm^{-1}. Now, the pattern of isotopic shifts shown in Figure 15 clearly demonstrates that the RR bands at 1552 cm^{-1} (native) and 1544 cm^{-1} (11D,12D) both belong to the mode [5-13]. Note that Grossjean et al.[36] arrived at the same conclusion though with much weaker arguments.

The above discussion is a typical example for the kind of reasoning required in the vibrational analysis of complex spectra. It shows that the assignment of bands to normal modes is greatly simplified if a calculation predicts the various normal modes in the correct spectral sequence and with reasonable compositions. Then, the inevitable differences between the observed frequencies of anharmonic fundamentals and the calculated harmonic frequencies do not necessarily preclude correct assignments.

Can One Distinguish Conformations? As outlined in the Introduction, this work was motivated by the question whether current DFT/MM hybrid methods can describe the vibrational spectra of the BR_{568} chromophore *in situ* with an accuracy that could suffice to decide such intricate questions like that of the chromophore structure in the L intermediate.

A partial answer to this question has been given by the comparison in Figure 13. The two dominantly populated lysine conformations C_1 and C_2 do not induce significant spectral shifts and, therefore, cannot be distinguished by IR. Here, the lack of effects of the conformational states on the IR spectrum is due to strong similarities of the chromophore structures and of the electrostatics within the binding pocket. In particular, the hydrogen bonded network stabilizing the binding pocket [cf. Figure 2 in part I (DOI 10.1021/jp902428x)] is preserved for both conformations. However, for the purpose of identifying the chromophore structure in L, the DFT/MM descriptions must solely be accurate enough to distinguish strongly different chromophore geometries and hydrogen bonding patterns in the chromophore binding pocket. The latter issue will now be checked.

For this purpose, recall that in part I (DOI 10.1021/jp902428x) we encountered a non-native conformation C_{nn} of the hydrogen bonded network within the binding pocket. C_{nn} was determined as the artificial attractor which hampers all MD simulations modeling the apoprotein by a nonpolarizable force field like C22 [cf. part I (DOI 10.1021/jp902428x)]. C_{nn} is characterized by an erroneous hydrogen bond between the $RSBH^+$ and Asp85. To analyze the spectral effect of this altered conformation, we have selected five snapshots from the C22-MD trajectory documented in Figure 4 of part I (DOI 10.1021/jp902428x) and have calculated an INMA-IR spectrum by DFT/MM. As demonstrated by Figure 27 in the SI, all selected snapshots belong to the non-native conformation C_{nn}.

Figure 16 shows that all double bond stretching frequencies obtained from the C22-INMA treatment are redshifted with respect to the PBR-INMA frequencies (the latter data were shown already in Figure 15). Due to this redshift, which on average amounts to 6 cm^{-1}, the root-mean-square deviation between the calculated and experimental double bond stretching

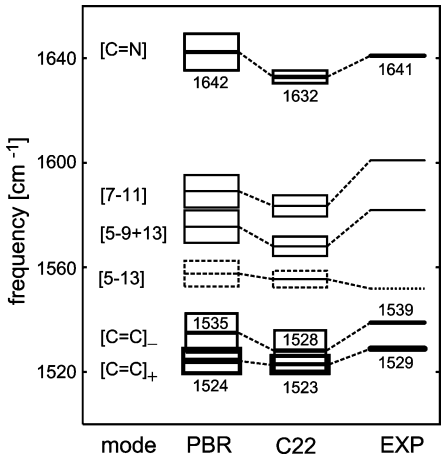

Figure 16. Comparison of the INMA double bond frequencies obtained for the PBR force field (43 snapshots: 25 C_1, 18 C_2) and for the C22 force field (5 snapshots, conformation C_{nn}). Experimental data are given as a reference.

frequencies increases from 6.5 cm^{-1} (PBR) to 11 cm^{-1} (C22). Here, the C═N double bond stretch shows the largest redshift of about 10 cm^{-1}.

The just identified redshift is exclusively due to the differences between the conformational states, because a blue shift is expected if one solely replaces the PBR by the C22 force field while leaving the conformational ensemble unchanged. The latter expectation derives from Figures 7 and Figure 26 (in the SI) which both show that the double bond frequencies are shifted on average by 6 cm^{-1} to the blue, if the description of the apoprotein is switched from PBR to C22 while the underlying structure (1C3W/MD) remains unchanged. The only difference between Figures 7 and 26 (in the SI) is the choice of the DFT/MM partitioning (Figure 7, M; Figures 26 and 16, L).

In summary, the IR spectrum calculated for the C_{nn} conformation with its non-native hydrogen bonding pattern sizably differs from the spectrum calculated for the native pattern (Figure 28 in the SI compares these spectra within the range 1100−1700 cm^{-1}). Furthermore, the C_{nn} spectrum deviates much more strongly from the experimental spectrum than the one obtained for our mixture model.

The SI analyzes the causes of these spectral shifts in some more detail. Here, Figure 29 shows the structural changes induced in the chromophore by the transition to the C_{nn} conformation. Although the charge localization and bond alternation are seen to increase in this transition, these changes do not entail the correspondingly expected blue shift of the double bond stretches, because a modified pattern of twists around C−C bonds within the Schiff base region of the $RSBH^+$ apparently leads to an overcompensating redshift.

These results show that not all conformational changes in the binding pocket of the BR chromophore are spectroscopically silent. Instead, all those changes, which are connected with distinct variations of the local electrostatics and with associated changes of the chromophore geometry, should generate clear signatures in the calculated IR spectra. The results furthermore demonstrate that substantial deviations of the calculated structural ensemble from the experimental one are associated with strongly increased deviations between calculated and experimental frequencies. Consequently, room-temperature ensembles calculated by MD with a conventional nonpolarizable force field, which therefore feature a non-native hydrogen bonded network, are badly suited as starting points for DFT/MM calculations of the chromophore's IR spectra. Note that the latter conclusion represents an *a posteriori* justification for our tedious approach of first constructing a PBR force field for MD simulations.

The fact that each improvement of the physical modeling (increasing the size of the DFT fragment, accounting for effects of room-temperature fluctuations and of conformational flexibility, including the polarization also of the water molecules, and choosing a proper conformational ensemble through the use of PBR instead of C22) brought the DFT/MM descriptions closer to the experimental findings suggests that, conversely, strong deviations between calculated and experimental spectra can be taken as clues for rating suggested conformations as less likely.

Summary and Outlook

The above results have made clear that the aim to compute the room-temperature vibrational spectra of the retinal chromophore in BR_{568} by DFT/MM hybrid techniques as accurately as possible requires utter technical care.

As a prerequisite, one must generate by MM−MD simulations a quite precise and physically realistic model of the room-temperature conformational ensemble in the chromophore region of the protein. As we have shown in part I of this work (DOI 10.1021/jp902428x),[1] MD simulations that are based on standard nonpolarizable MM force fields like C22 produce a non-native conformation C_{nn}, which exhibits a strongly altered electrostatics in the surroundings of the $RSBH^+$ and is associated with a changed chromophore geometry. Above, we have shown that such deviations from the native conformational ensemble preclude accurate computations of the chromophore's vibrational spectra, thus proving one of our working hypotheses. Because we attributed the apparent instability of the native conformation found in previous MD descriptions to the lacking inclusion of the electronic polarizability and because well-established polarizable force fields for proteins were not yet available, we had to generate the polarized force field PBR by DFT/MM calculations.[1]

The PBR-MD simulations then reproduced the crystallographic conformation C_1 as the most stable state at cryogenic and ambient temperatures. For 300 K, the simulations revealed the existence of a second conformational state C_2 at sizable amounts (42%).[1] As we have shown above, the structural differences of C_1 and C_2 (which are localized in the backbone and in the aliphatic chain linking the $RSBH^+$ to the backbone) hardly affect the vibrational spectra of the chromophore.

With a room-temperature conformational ensemble of the chromophore region in BR_{568} at hand, the DFT/MM computation of the chromophore's room-temperature IR spectra by applying the INMA procedure had become possible. For this purpose, a suitable partitioning of the system into DFT and MM fragments had to be chosen. Here, the largest DFT fragment L (cf. Figure 2) turned out to be the best choice and the correspondingly enhanced computational effort had to be spent. Similarly, also all further improvements of the modeling, in which initially simplified assumptions were given up in favor of physically more realistic ones, led to an increasingly better description of the well-known chromophore spectra. In these improvements, e.g., the rigid cryogenic X-ray structure was replaced by the 300 K conformational ensemble resulting from our PBR-MD simulations or the standard TIP3P water models were replaced by more polar ones, thus accounting for the corresponding DFT/MM results.[1]

Concerning the structure of the chromophore in BR_{568}, the resulting DFT/MM description is much more accurate than the available experimental structures (cf. Figure 5 and the associated

discussion). Concerning the chromophore's vibrational spectra, however, this description still leaves room for further improvements. Here, particularly the restriction to the most simple density functionals, which currently is imposed by the use of the program CPMD[39] as the DFT code in our hybrid program, is a major obstacle. Thus, an extension of our QM/MM interface to a more versatile DFT package (e.g., ref 51) is required for more accurate descriptions.

Nevertheless, the current description is quite clearly good enough to explain the experimental spectra. Evidence for this claim has been given by considering several selected effects of isotope substitutions (in the text and in the SI). Furthermore, such a DFT/MM description is good enough to distinguish non-native from native conformations by comparing calculated spectra with observed ones, if the electrostatics in the chromophore binding pocket and the chromophore geometry are significantly different for these conformations. Also, this second claim has been substantiated by providing an example. Here, it has been shown that replacing the native conformational mixture by the non-native conformation C_{nn} strongly deteriorates the description of the chromophore's IR spectrum in BR_{568}.

These results nourish the hope that DFT/MM analyses of chromophore's vibrational spectra in the early photocycle intermediates K and L can help to identify the associated chromophore structures. As we have suggested in an initial section serving to motivate our work, particularly the analysis of L at room temperature should be capable of shedding light on the yet unknown mechanism of light-driven proton pumping in BR. As long as a highly resolved, room-temperature X-ray structure of L remains elusive, one will have to choose a MD-based modeling approach instead. Here, one will have to start with various assumptions on the primary photoisomerization and will have to carry out subsequent MD simulations extending over microsecond time spans to compute the altered distribution of water molecules characterizing the L state of the protein. Quite clearly, a standard nonpolarizable force field will not be good enough for such MD simulations because these simulations aim at the reliable prediction of the protein conformation in L. One will have to use a polarizable or polarized force field instead. Subsequently, the quality of the conformation computed for the room-temperature L intermediate can be scrutinized by comparing the vibrational chromophore spectra observed at room temperature by time-resolved spectroscopy with DFT/MM predictions. A match of comparable quality as the one achieved in this work for BR_{568} would then indicate that the chromophore structure in L has been finally identified. With this structure at hand, the mechanism of proton pumping in BR should be accessible.

The DFT/MM technology used by us in this work is computationally not efficient enough for the purpose sketched above or for standard applications to other systems like, e.g., the tetrapyrrole chains in phytochrome. First, for accurately computing the chromophore spectra, one will need a more efficient and more accurate DFT module.[51] Second, for the speedy and repeated computation of the polarization in an extended neighborhood of the chromophore, one may resort to an even faster though less accurate approach such as the self-consistent charge density functional tight-binding method.[52] Of course, instead of such a polarized force field, the use of a reliable polarizable force field for proteins would be desirable. However, such force fields are not yet established. Nevertheless, within a MM−MD scenario involving polarized force fields, a prediction of the L intermediate in the BR photocycle through microsecond simulations seems to be technically feasible, as is demonstrated by the recent 10 μs simulation of a small protein in solution.[53] As a result, there is hope that the mechanism converting BR light energy into proton pumping can be revealed.

Acknowledgment. The authors gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (SFB 533, projects C1 and C3). The authors thank B. Schropp for technical support in the DFT/MM calculations.

Supporting Information Available: On 14 pages, the SI provides 11 figures and associated text explaining and discussing the additional material. The SI visualizes the C−C stretching modes of an $RSBH^+$ (Figures 17, 18), illustrates the effects of the DFT fragment size, of thermal fluctuations, and of conformational transitions on the chromophore's structure and IR spectra (Figures 19−22), discusses the spectral effects of water polarization (Figures 23, 24) and of various isotope substitutions (Figure 25), and explains the structural and spectral consequences of altering the hydrogen bonds in the chromophore binding pocket (Figures 26−29). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Babitzki, G.; Denschlag, R.; Tavan, P. *J. Phys. Chem. B*. DOI: 10.1021/jp902428x.
(2) Lanyi, J. K. *Annu. Rev. Physiol.* **2004**, *66*, 665–688.
(3) Kühlbrandt, W. *Nature* **2000**, *406*, 569–570.
(4) Schulten, K.; Tavan, P. *Nature* **1978**, *272*, 85–86.
(5) Schulten, K.; Schulten, Z.; Tavan, P. An Isomerization Model for the Pump Cycle of Bacteriorhodopsin. In *Information and Energy Transduction in Biological Membranes*; Bolis, L., Helmreich, E. J. M., Passow, H., Eds.; Allan R. Liss, Inc.: New York, 1984.
(6) Smith, S. O.; Myers, A.; Pardoen, J.; Winkel, C.; Mulder, P.; Lugtenburg, J.; Mathies, R. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 2055–2059.
(7) Alshuth, T. Kinetische und strukturelle Untersuchungen am Chromophor von Bacteriorhodopsin mit Hilfe zeitaufgelöster Resonanz-Raman-Spektroskopie. Doktorarbeit, Goerg-August Universität Göttingen, 1985.
(8) Gerwert, K. Transduktion der Lichtenergie in Protonen-Transfer-Reaktionen beim Bacteriorhodopsin: Eine Untersuchung mit Hilfe der zeitaufgelösten IR- und statischen FTIR-Differenzspektroskopie. Doktorarbeit, Albert-Ludwigs-Universität Freiburg im Breisgau, 1985.
(9) Gerwert, K.; Siebert, F. *EMBO J.* **1986**, *5*, 805–811.
(10) Fodor, S. P. A.; Pollard, W. T.; Gebhard, R.; van den Berg, E. M. M.; Lugtenburg, J.; Mathies, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2156–2160.
(11) Groβjean, M. F.; Tavan, P.; Schulten, K. *Eur. Biophys. J.* **1989**, *16*, 341–349.
(12) Fahmy, K.; Groβjean, M. F.; Siebert, F.; Tavan, P. *J. Mol. Struct.* **1989**, *214*, 257–288.
(13) Hage, W.; Kim, M.; Frei, H.; Mathies, R. A. *J. Phys. Chem.* **1996**, *100*, 16026–16033.
(14) Rödig, C.; Chizhov, I.; Weidlich, O.; Siebert, F. *Biophys. J.* **1999**, *76*, 2687–2701.
(15) Lanyi, J. K. *Biochim. Biophys. Acta* **2004**, *1658*, 14–22.
(16) Lanyi, J. K.; Schobert, B. *J. Mol. Biol.* **2007**, *365*, 1379–1392.
(17) Braiman, M.; Mathies, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 403–407.
(18) Bondar, A.-N.; Suhai, S.; Fischer, S.; Smith, J. C.; Elstner, M. *J. Struct. Biol.* **2007**, *157*, 454–469.
(19) Royant, A.; Edman, K.; Ursby, T.; Pebay-Peyroula, E.; Landau, E. M.; Neutze, R. *Nature* **2000**, *406*, 569–570.
(20) Lanyi, J. K.; Schobert, B. *J. Mol. Biol.* **2003**, *328*, 439–450.
(21) Edman, K.; Royant, A.; Larsson, G.; Jacobson, F.; Taylor, T.; Van Der, S. D.; Landau, E. M.; Pebay-Peyroula, E.; Neutze, R. *J. Biol. Chem.* **2004**, *297*, 2147–2158.
(22) Kouyama, T.; Nishikawa, T.; Tokuhisa, T.; Okumura, H. *J. Mol. Biol.* **2004**, *335*, 531–546.
(23) Mak-Jurkauskas, M. L.; Bajaj, V. S.; Hornstein, M. K.; Belenky, M.; Griffin, R. G.; Herzfeld, J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 883–888.
(24) Fahmy, K.; Siebert, F.; Tavan, P. *Biophys. J.* **1991**, *60*, 989–1001.
(25) Lórenz-Fonfría, V. A.; Furutani, Y.; Kandori, H. *Biochemistry* **2008**, *47*, 4071–4081.
(26) Dioumaev, A. K.; Lanyi, J. K. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9621–9626.

(27) For physiological temperatures, the kinetic competition of the $L \rightarrow BR_{568}$ shunt reaction with the kinetically preferred deprotonation of the $RSBH^+$ marking the $L \rightarrow M_1$ transition has been first discussed in ref 5.

(28) Bondar, A.-N.; Elstner, M.; Suhai, S.; Smith, J. C.; Fischer, S. *Structure* **2004**, *12*, 1281–1288.

(29) Braun-Sand, S.; Sharma, P. K.; Chu, Z. T.; Pisliakov, A. V.; Warshel, A. *Biochim. Biophys. Acta* **2008**, *1777*, 441–452.

(30) Diller, R.; Stockburger, M. *Biochemistry* **1988**, *27*, 7641–7651.

(31) Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1999**, *110*, 10452–10467.

(32) The case of an unprotonated RSB featuring a polyene-like pattern of alternating single and double bonds is much simpler. Here, even an empirical modeling of the force field may be feasible. See, e.g., ref 54.

(33) Curry, B.; Broek, A.; Lugtenburg, J.; Mathies, R. A. *J. Am. Chem. Soc.* **1982**, *104*, 5274–5286.

(34) Smith, S. O.; Myers, A.; Mathies, R.; Pardoen, J.; Winkel, C.; van den Berg, E.; Lugtenburg, J. *Biophys. J.* **1985**, *47*, 653–664.

(35) Smith, S. O.; Braiman, M. S.; Myers, A. B.; Pardoen, J. A.; Courtin, J. M. L.; Winkel, C.; Lugtenburg, J.; Mathies, R. A. *J. Am. Chem. Soc.* **1987**, *109*, 3108–3125.

(36) Groβjean, M. F.; Tavan, P.; Schulten, K. *J. Phys. Chem.* **1990**, *94*, 8059–8069.

(37) Schmitz, M.; Tavan, P. On the art of computing the IR spectra of molecules in condensed phase. In *Modern methods for theoretical physical chemistry of biopolymers*; Starikov, E. B., Tanaka, S., Lewis, J., Eds.; Elsevier: Amsterdam, The Netherlands, 2006.

(38) Luecke, H.; Schobert, B.; Richter, H.-T.; Cartailler, J.-P.; Lanyi, J. K. *J. Mol. Biol.* **1999**, *291*, 899–911.

(39) Hutter, J.; Alavi, A.; Deutsch, T.; Bernasconi, M.; Goedecker, S.; Marx, D.; Tuckerman, M.; Parinello, M. "CPMD V3.9", Copyright IBM Corp and MPI für Festkörperforschung Stuttgart, 2004; see www.cpmd.org.

(40) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(41) Perdew, J.; Yue, W. *Phys. Rev. B* **1986**, *33*, 8800–8802.

(42) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993–2005.

(43) Nonella, M.; Mathias, G.; Tavan, P. *J. Phys. Chem. A* **2003**, *107*, 8638–8647.

(44) Nonella, M.; Mathias, G.; Eichinger, M.; Tavan, P. *J. Phys. Chem. B* **2003**, *107*, 316–322.

(45) Groβjean, M. F.; Tavan, P. *J. Chem. Phys.* **1988**, *88*, 4884–4896.

(46) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–145.

(47) MacKerell, A.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(48) Lansing, J. C.; Hohwy, M.; Jaroniec, C. P.; Creemers, A. F. L.; Lugtenburg, J.; Herzfeld, J.; Griffin, R. G. *Biochemistry* **2002**, *41*, 431–438.

(49) Alshuth, T.; Stockburger, M. *Photochem. Photobiol.* **1986**, *43*, 55–66.

(50) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(51) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103–128.

(52) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.

(53) Freddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K. *Biophys. J.* **2008**, *94*, L75–L77.

(54) Ames, J. B.; Fodor, S. P. A.; Gebhard, R.; Raap, J.; van den Berg, E. M. M.; Lugtenburg, J.; Mathies, R. A. *Biochemistry* **1989**, *28*, 3681–3687.

JP902432E