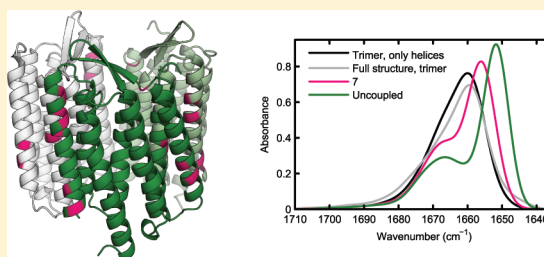


# Vibrational Coupling between Helices Influences the Amide I Infrared Absorption of Proteins: Application to Bacteriorhodopsin and Rhodopsin

Eeva-Liisa Karjalainen and Andreas Barth\*

Department of Biochemistry and Biophysics, Arrhenius Laboratories of Natural Sciences, Stockholm University, SE-106 91 Stockholm, Sweden

**ABSTRACT:** The amide I spectrum of multimers of helical protein segments was simulated using transition dipole coupling (TDC) for long-range interactions between individual amide oscillators and DFT data from dipeptides (la Cour Jansen et al. *J. Chem. Phys.* **2006**, *125*, 44312) for nearest neighbor interactions. Vibrational coupling between amide groups on different helices shift the helix absorption to higher wavenumbers. This effect is small for helix dimers ( $1\text{ cm}^{-1}$ ) at  $10\text{ Å}$  distance and only moderately affected by changes in the relative orientation between the helices. However, the effect becomes considerable when several helices are bundled in membrane proteins. Particular examples are the 7-helix membrane proteins bacteriorhodopsin (BR) and rhodopsin, where the upshift is  $4.3$  and  $5.3\text{ cm}^{-1}$ , respectively, due to interhelical coupling within a BR monomer. A further upshift of  $4.0\text{ cm}^{-1}$  occurs when BR monomers associate to trimers. We propose that interhelical vibrational coupling explains the experimentally observed unusually high wavenumber of the amide I band of BR.



## INTRODUCTION

Helices are the most abundant secondary structure in proteins. The  $\alpha$ -helix structure was defined by Pauling in 1951<sup>1</sup> and has since then been found to manifest great structural variation.<sup>2</sup> Helices play fundamental roles in protein functionality and specificity, e.g., by contributing to binding and catalytic sites, and by forming channels through membranes and thereby providing transport routes for ions.<sup>3–6</sup> Bundles of  $\alpha$ -helices constitute one out of two most important structural motifs in membrane proteins together with  $\beta$ -barrels.<sup>7,8</sup>

Determining the structures of peptides and proteins, particularly of membrane proteins, that are difficult to crystallize or too large for NMR is an important challenge for the life science community. Infrared (IR) spectroscopy is a classical spectroscopic method that has proven its value as a tool for protein investigation over the past decades including investigations of secondary structure content, catalytic reactions, and folding and misfolding processes among others.<sup>9–11</sup>

The strong amide I band of IR spectra is predominantly caused by absorption of the backbone carbonyl groups. It reflects the protein conformation as it is influenced by the coupling of transition dipole moments on different amide groups of the polypeptide backbone.<sup>12,13</sup> The interactions within a helix have been well characterized during the past decades.<sup>13–19</sup>

Several recent studies have instead focused on the IR spectral features of transition dipole coupling (TDC) between secondary structure elements, such as those between  $\beta$ -sheets<sup>20–22</sup> or  $\alpha$ -helices in a helix-dimer.<sup>23,24</sup> This coupling is important to consider as it influences the band positions associated with the secondary structure absorption and therefore the interpretation

of the IR spectra. Recently, the 1D and 2D IR spectra of the helix dimer pair have been investigated both experimentally and theoretically, and it was found that there is an exciton coupling interaction between the amide I transitions of the respective helices.<sup>23,24</sup> The ability to measure such coupling interactions could enable 3D structure determination by providing important constraints. Studies of the predominantly helical globular protein myoglobin have also concluded that it is not possible to accurately replicate the spectra of a many-helix protein without including long-range interactions between helices.<sup>25</sup>

To explore in further detail how helix–helix interactions manifest in IR spectra, we have performed a systematic study where the amide I band of different structural configurations of a helix pair as well as the resulting impact of helix bundling was simulated. The amide I band is here simulated using transition dipole coupling for long-range electrostatic effects and parametrized ab initio data for short-range interactions. The study focuses on the spectral changes that arise as a result of interactions between the individual helices and discusses in particular the IR spectrum of bacteriorhodopsin (BR) and rhodopsin.

Like many membrane proteins, these two proteins contain closely packed bundles of helices particularly in their transmembrane (TM) region. Bacteriorhodopsin (BR) is an integral ion pumping membrane protein that performs light-induced proton transport across the plasma membrane of *Halobacterium salinarum*.<sup>26–28</sup> It has been intensively studied over more than

Received: January 11, 2012

Revised: March 15, 2012

Published: March 21, 2012

40 years due to being an important prototype for  $\alpha$ -helical membrane transporters and particularly for seven transmembrane (TM) helix proteins. The seven helix bundle is a structural motif that is common for a large group of G-protein coupled receptors (GPCRs),<sup>29</sup> which in turn includes many highly important drug target receptors in humans.

BR has been extensively investigated using FTIR spectroscopy since long before atomic resolution structures were available, as reviewed in refs 30, 31. FTIR measurements show that BR exhibits an unexpectedly high IR absorption maximum.<sup>32,33</sup> Experimental measurements show that the maximum is around  $1664\text{ cm}^{-1}$  in  $\text{H}_2\text{O}$  and  $1662$  in  $\text{D}_2\text{O}$   $\text{cm}^{-1}$ <sup>32,34,35</sup> rather than in the typical range of helix absorption  $1645\text{--}1660\text{ cm}^{-1}$ .<sup>16,36</sup> This is, however, not the case for all proteins with similar structure. Rhodopsin,<sup>37,38</sup> a GPCR retinal protein that is structurally very similar but performs a different biological function, has an amide I absorption maximum around  $1657\text{ cm}^{-1}$ .<sup>35,39</sup> Several tentative explanations for this unusual IR spectral characteristic of BR have been proposed in the literature,<sup>33,40</sup> e.g., the occurrence of  $\alpha_{\text{II}}$ -helix structures that have other dihedral angles as well as longer hydrogen bonds.<sup>41</sup>

None of these hypotheses have, however, yet been conclusively confirmed. On the basis of the findings for model helix systems as well as simulations of the BR and rhodopsin structures presented in the Results and Discussion section, we propose that the high frequency amide I maximum of BR could, at least partially, be attributed to TDC between the TM helices and between subunits. This proposal is in line with experimental IR data from unfolding studies.<sup>40,42–44</sup>

## THEORETICAL METHODS

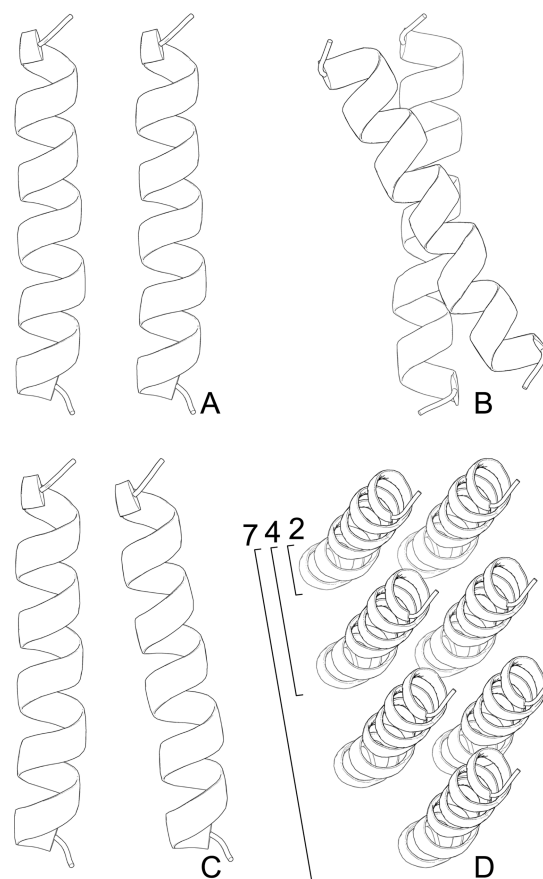
**Structural Information.** The helix data used for simulations are extracted from the Protein Data Bank (PDB)<sup>45</sup> files of real protein structures as specified in Table 1. The included

**Table 1. Structural Data for  $\alpha$ -Helices Used for Simulations**

PDB ID	name	residues
1FBB	bacteriorhodopsin	A104–126
1F88	rhodopsin	A149–169
1SU4	$\text{Ca}^{2+}$ -ATPase	A54–79
2ZXE	$\text{Na}^+/\text{K}^+$ -ATPase	A282–313
2ZXE	$\text{Na}^+/\text{K}^+$ -ATPase	A99–121
1VF5	cytochrome C	A115–137
2BG9	acetylcholine	A212–239
3NY8	adrenergic receptor	A146–171
3PBL	dopamine receptor	A34–57
1PV6	lactose permease	A165–186

helices are all more than 20 residue long TM helices that were selected to represent a variation of structural details present in naturally occurring helices.

In order to investigate coupling between  $\alpha$ -helices, the respective single-helix structures were duplicated, resulting in two identical helices. The two helices were subsequently translated and rotated with respect to each other. The  $z$ -axis is defined as parallel to the helix axis and was found by a fit of the  $\text{C}_\alpha$  positions of each helix. The fit was carried out using least-squares minimization of the perpendicular distance between the  $\text{C}_\alpha$  positions and a line intersecting the average position of all  $\text{C}_\alpha$  atoms. Together with the  $z$ -axis, the  $x$ - and  $y$ -axes form an orthonormal basis. The coordinate frame convention is defined



**Figure 1.** Helix systems used in simulations to investigate the effect of interhelix coupling on amide I spectra. All helices are identical and based on a helix structure from a PDB-file that was copied, translated, and rotated to create configurations with several helices. (A) Helices translated along the  $x$ -axis. (B) Helix rotated  $\xi_x$  degrees around the  $x$ -axis. (C) Rotation around the  $y$ -axis by  $\xi_y$  degrees. (D) Helix bundles with 2, 4, and 7 helices.

in Figure 1. Translation along the  $x$ -axis by  $8\text{--}20\text{ \AA}$ , rotation around the  $x$ -axis by  $\xi_x \in [0^\circ, 180^\circ]$ ,  $y$ -axis  $\xi_y \in [0^\circ, 14^\circ]$ , and  $\xi_z \in [0^\circ, 360^\circ]$  around the  $z$ -axis were simulated.

Helix bundle structures were generated by positioning two, four, and seven duplicates of a single helix equidistantly with a center-to-center distance of  $10\text{ \AA}$ , a typical distance in helix packing in proteins.<sup>46,47</sup> For BR and rhodopsin simulations, the PDB structures 1C3W<sup>48</sup> and 1HZS<sup>49</sup> were used as input, respectively. The helical residues as specified in the original papers were retained, whereas all other residues were deleted for the simulations of the structures indicated as “helix only”. For BR (rhodopsin), helical residues A104–127, A133–155 (A34–64, A71–100), were included in the simulation of the spectrum with two helices and with four helices also A81–101, A164–191 (A106–139, A286–309) were included. Finally, in the simulated seven helix spectrum A9–31, A36–63, A200–225 (A150–172, A200–225, A244–276) were included as well. The structures for the biological units were generated using the transformations specified in the PDB files. Control simulations were performed using PDB coordinates found in older structures, 1FBB<sup>50</sup> and 1F88<sup>51</sup> for BR and rhodopsin.

**Amide I Simulations.** To selectively investigate structure–spectra correlations of only the backbone conformation sensitive amide I mode, only a representation involving the amide units is required. A protein structure can then be reduced to

$N$  harmonic amide I oscillators where  $N$  is the number of peptide groups.

The Hamiltonian for a system of amide I oscillators has been found to be well described using exciton theory, a theory traditionally used for describing delocalized electron states in molecular aggregates.<sup>19,52,53</sup> Exciton theory is applicable to systems with similar, repeating units that can be considered uncoupled from other modes, thereby making it highly suitable for describing the interactions of the polypeptide backbone.<sup>19,54</sup> The form of the exciton coupling is furthermore well described using transition dipole coupling, which is the main effect determining the shape of the amide I band of polypeptides.<sup>13</sup>

The transition dipole coupling  $F_{jk}$  between amide I oscillators  $j$  and  $k$  is

$$F_{jk} = \frac{0.1}{\epsilon} \frac{\delta\mu_j \cdot \delta\mu_k - 3(\delta\mu_j \cdot \mathbf{n}_{jk})(\delta\mu_k \cdot \mathbf{n}_{jk})}{R_{jk}^3}$$

Here,  $\delta\mu_j$  and  $\delta\mu_k$  are the transition dipoles for oscillators  $j$  and  $k$  whereas  $R_{jk}$  and  $\mathbf{n}_{jk}$  are the distance and unit vector between the transition dipoles, respectively.  $\epsilon$  is the dielectric constant.

The interactions between helices presented in the Results and Discussion section have been modeled using transition dipole coupling to describe long-range interactions<sup>12,13</sup> together with previously published parametrized ab initio coupling maps for describing coupling between nearest neighbors in sequence.<sup>55</sup> Diagonalization of the Hamiltonian matrix yields the  $N$  normal modes of the protein, in a procedure analogous to that of normal-mode analysis.<sup>56,57</sup>

**Computational Details.** The amide I absorption spectra of the different helix geometries extracted from PDB-files were simulated using theoretical methods as described in the Amide I Simulations section. All simulations were performed using MATLAB (R2010a, The Mathworks, CA). The simulated spectra were generated by assigning a Gaussian function with a fwhm of  $8 \text{ cm}^{-1}$  to each normal mode. The simulations do not incorporate dynamical fluctuations, and all intrinsic frequencies were set to an identical value of  $1675 \text{ cm}^{-1}$ ; i.e., all residues were treated equally.

TDC requires parametrized values for the transition dipole magnitude, direction, and location. The simulations presented used a transition dipole moment of  $3.70 \text{ D/\AA}$  located on the C=O bond,  $0.868 \text{ \AA}$  away from the carbon atom and rotated by  $20^\circ$  toward the nitrogen atom.<sup>56</sup> The dielectric constant  $\epsilon$  was set to 1. All simulations were performed using two other different TDC parameter sets,<sup>58,59</sup> which yielded similar qualitative results and in all cases the same observed trends. Simulations using a different ab initio map<sup>60</sup> for nearest neighbor couplings also provided almost identical results.

The infrared intensity  $I_k$  of the respective modes is calculated from the eigenvector associated with that mode according to<sup>61</sup>

$$I_k = \sum_{i=1}^3 \left( \frac{\partial \mu_i}{\partial Q_k} \right)^2 = \sum_{i=1}^3 \left| \sum_{j=1}^N \left( \frac{\partial \mu_i}{\partial q_j} \right) \left( \frac{\partial q_j}{\partial Q_k} \right) \right|^2$$

Here,  $\partial q_j / \partial Q_k$  is the vibrational amplitude of a local amide I mode of group  $j$  to collective mode  $k$ ,  $\partial \mu_i / \partial q_j$  is the  $i$ th component of the transition dipole for peptide group  $j$ , and  $N$  is the number of peptide groups in the polypeptide.

All spectral intensities were divided by the number of peptide groups to facilitate comparison between systems of different size. For each helix system, spectra were calculated at several

different interhelix distances with a step length of  $1 \text{ \AA}$ . Rotations around the respective axes were performed at  $5^\circ$  ( $x$ ,  $z$ -axes) or  $2^\circ$  ( $y$ -axis) intervals. Band positions were calculated as the position of mean intensity in a frequency range defined by the bandwidth at 80% of maximum band intensity. Simulated absorption spectra are shown in the interval  $1640\text{--}1710 \text{ cm}^{-1}$ ; no significant absorption was present outside this spectral range.

## RESULTS AND DISCUSSION

We modeled the effect of interhelix coupling on the IR amide I spectrum using transition dipole coupling for long-range interactions and ab initio coupling maps for short-range couplings (see Theoretical Methods). TDC has been found inappropriate for describing short-range electrostatic effects,<sup>58,62</sup> but predicts correct qualitative trends for the long-range interactions that are at the focus of this study. More complex methods for simulating the amide I band including effects such as hydrogen bonding, solvation, and structural dynamics have been developed<sup>17,18</sup> but are not necessary to include in the simulation protocol for the present purpose. However, due to the simplicity of the modeling approach, it does not predict correct absolute band positions. Instead, the analyses focus on the relative differences in amide I maximum position as a function of long-range interactions between helices.

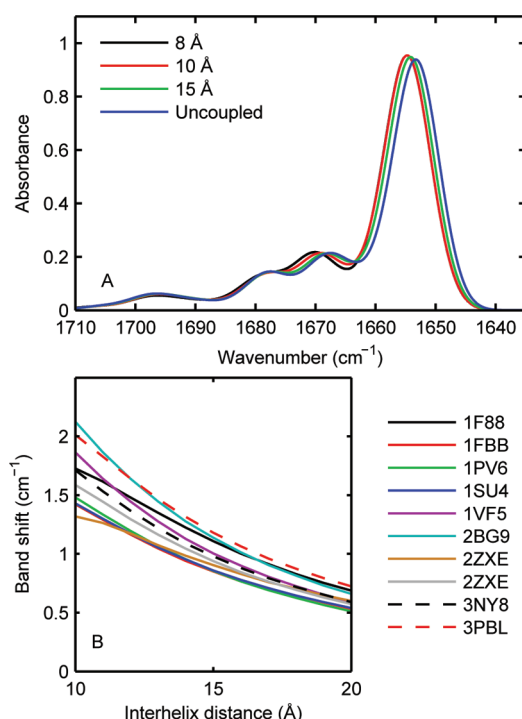
Section A presents the simulations with pairs of helices which depend on the interhelix distance and on their relative orientation. Within this section, spectra for a helix pair constructed from a 23 residue long helix from BR are shown first. Then simulation results are presented for 10 TM helices from different proteins to demonstrate that the trends and spectral impact found are generally valid and independent of the specific structural details.

In section B, the effects from interactions between an increasing number of helices, here grouped as bundles of four and seven identical helices, were simulated using the helices from different proteins as a starting point. Finally, section C discusses the amide I spectra of two real seven TM helix proteins, BR and rhodopsin, that were investigated and compared in order to explore whether helix–helix interactions provide an explanation for the unusually high wavenumber IR amide I maximum experimentally observed for BR.<sup>32</sup>

**A1. Helix–Helix Interaction Effects on Amide I Spectra.** *Amide I Band Position Depends on Interhelix Distance.* Simulated amide I absorption spectra of a pair of  $\alpha$ -helices at interhelix distances ranging from 8 to  $20 \text{ \AA}$  are shown in Figure 2A. The two-helix system structure was constructed using a helix from BR (1FBB) as specified in Table 1. IR amide I spectra were simulated for configurations with increasing separation distance along the  $x$ -axis. The thin black line spectrum shows the absorption spectrum of a single helix, with the main amide I maximum at  $1653 \text{ cm}^{-1}$  and two smaller bands at  $1668$  and  $1678 \text{ cm}^{-1}$ . The spectrum is typical of nonideal finite  $\alpha$ -helices with intensity almost over the full wavenumber range shown.

For the helix pair the amide I maximum shifts  $1.4 \text{ cm}^{-1}$  to higher wavenumbers upon decreasing the distance from effectively uncoupled helices to an interhelix distance of  $10 \text{ \AA}$  ( $0.9 \text{ cm}^{-1}$  at  $15 \text{ \AA}$ ). The amplitude of the bands is not significantly altered when the interhelix distance changes. Figure 2B shows the change in the amide I maximum band position with interhelix separation for all helices. The average band shift upon decreasing the interhelix distance from effectively uncoupled to  $10 \text{ \AA}$  is  $1.7 \text{ cm}^{-1}$  and  $1 \text{ cm}^{-1}$  when going from uncoupled to  $15 \text{ \AA}$ .





**Figure 2.** (A) Simulated amide I spectra for a two-helix system at different interhelix distances constructed from a helix from BR (1FBB). (B) Shift of the amide I maximum relative to the band position of a pair of uncoupled helices as a function of the interhelix distance for two-helix systems as specified in Table 1.

The trend in terms of shift of the main band position is the same for all helices independent of structural variation in the test set such as slight variation in length, curvature, or tightness of coil.

**A2. Effect of Interhelix Orientation.** The change in interhelix coupling as a result of changing the interhelix orientation at a fixed distance around all three symmetry axes was systematically investigated for all helices in the test set. The center-to-center interhelix distance was 10 Å in all simulations.

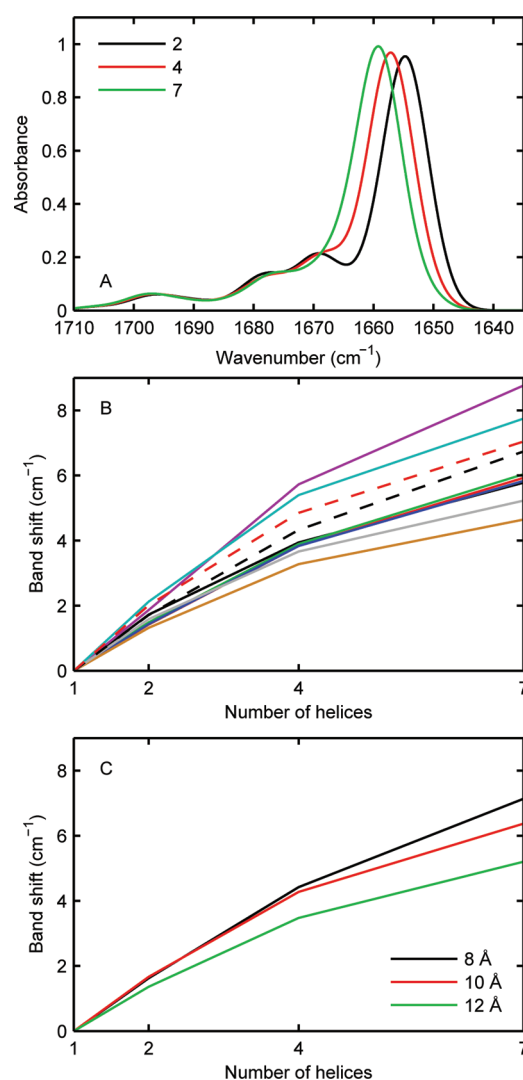
Rotation around the  $x$ -axis in the  $0$ – $180^\circ$  range yielded very small impact on the position of the amide I maximum. The shift with respect to the uncoupled case depends on interhelix orientation in a periodical manner, with the average shift for all helices of about  $0\text{ cm}^{-1}$  occurring at a rotation angle of  $90^\circ$ , indicating that perpendicular helices are essentially uncoupled. This is expected from the dipole arrangement as the distance between dipoles is maximal and several of the terms that contribute to the TDC coupling now approach zero.

Rotation around the  $y$ -axis, that is tilting of the helices toward each other (Figure 1C), is sterically hindered due to the two helices colliding. This allows only for investigation of a rotation angle  $\xi_y \in [0^\circ, 14^\circ]$ . Within this range, the effect of the rotation on the position of the amide I maximum was approximately zero. Finally, full rotation around the helical  $z$ -axis was simulated and was found to yield a shift of  $\pm 0.5\text{ cm}^{-1}$  with respect to the case where the duplicated helix was merely translated.

The main influence on the position of the amide I maximum is thus the distance between helices; the respective rotations play only a minor role in determining the band position at least under symmetrical conditions such as those discussed. The small changes due to rotation around the  $x$ - and  $y$ -axes at a 10 Å

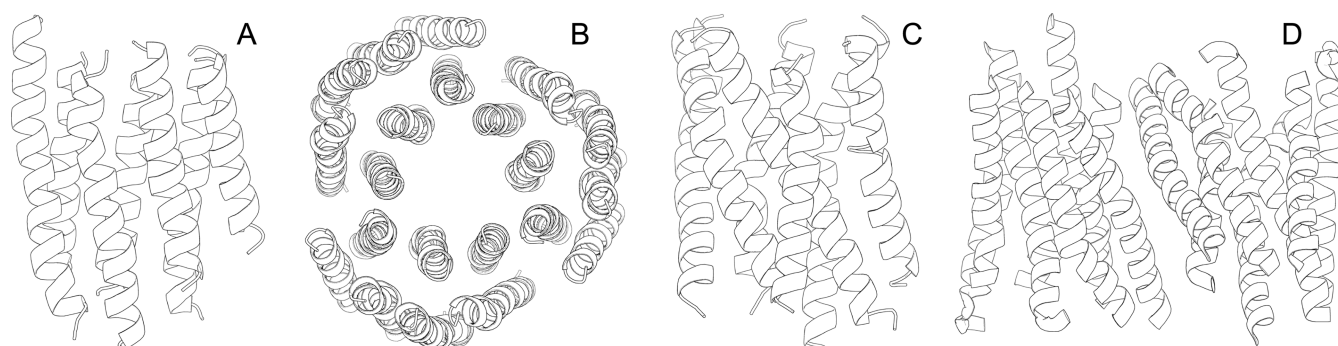
distance may not be experimentally detectable, but should still be considered when analyzing the amide I spectra of helix bundles with heterogeneous structure. No spectral changes in addition to the shift of the amide I maximum were observed as a result of interhelix coupling.

**B. Effect of Helix Bundling on IR Spectra.** In membrane proteins, helices can frequently occur in bundles of four or seven helices. Helix bundles with 2, 4, and 7 helices (as illustrated in Figure 1D) were created to investigate how the amide I band is influenced by an increasing number of helices that interact with each other. Figure 3A shows simulated amide I



**Figure 3.** (A) Simulated amide I spectra for bundles of 2, 4, and 7 helices based on a helix extracted from BR (1FBB). (B) Band shift of the amide I maximum relative to the band position of the one-helix spectrum for bundles of 2, 4, and 7 helices at an interhelix distance of 10 Å for all helices in the simulation set. Color code is the same as in Figure 2B. (C) Band shift of the amide I maximum as a function of the number of helices in a bundle relative to that of a single helix at different interhelix distances.

spectra for increasingly large helical bundles based on a helix from BR (see Table 1). The spectra show a shift of the amide I maximum to higher wavenumbers when the number of helices increases. The amide I maximum shifts  $2.4\text{ cm}^{-1}$  to higher wavenumbers when increasing the number of helices from two to four and  $4.5\text{ cm}^{-1}$  when considering the difference between



**Figure 4.** Structures of BR and rhodopsin displaying only helical residues for clarity. The (A) monomer and the (B) biological trimer structure of BR (1FBB). The (C) monomer and (D) dimer structures of rhodopsin (1F88).

the spectra of two and seven helices. Simulations were performed at an interhelix distance of 10 Å

Figure 3B demonstrates that the trend is the same for all structures in the test set: the amide I maximum shifts to higher wavenumbers with an increasing number of helices. The average band shift for all structures upon increasing the number of helices from two to four and from two to seven helices is 2.6 and 4.7  $\text{cm}^{-1}$ , respectively. As noted earlier in this section, the effect of interhelix coupling is distance dependent. The distance dependence of the shift of the amide I maximum as a function of the number of helices relative to that of a single helix is visualized in Figure 3C.

Interhelix coupling has previously been observed experimentally<sup>23</sup> in experiments with isotopically labeled amide groups at a specific site on each helix of a helix dimer. This further supports our simulation results presented here.

### C. Application to Bacteriorhodopsin and Rhodopsin.

Two structurally similar proteins built up from seven TM helices are BR and rhodopsin (illustrated in Figure 4). Here, we have simulated amide I spectra based on fragments of the protein structures to demonstrate to what extent the trend in amide I band shift is retained despite structural heterogeneity, as compared to the more ideal model bundles created through copy and translation of a single helix described in the previous section.

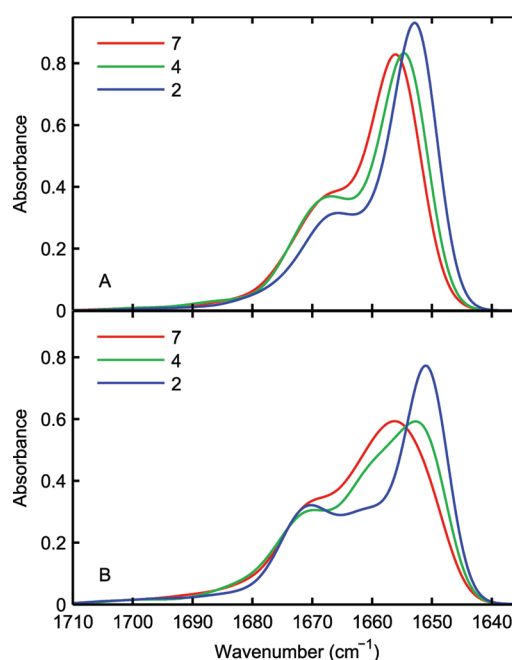
BR is a 248 residue integral membrane protein with a bundle of seven  $\alpha$  helices that traverse the membrane connected by short loops.<sup>48</sup> The  $\alpha$ -helical residues in the bundle, varying in length from 20 to 28 residues each, constitute 79% of the protein secondary structure content. Three of them are approximately perpendicular to the membrane whereas four of them are tilted on average 20° to the membrane normal.<sup>33</sup>

Bovine rhodopsin is also a seven TM helix protein, consisting of 348 residues distributed on seven helices that are 19–34 residues long (60% of all residues) as well as connecting loops.<sup>49</sup> The helices are irregular in terms of kinks and bends around Pro-Gly residues and are to a different extent tilted with respect to the membrane normal, with a tilt angle ranging from only a few degrees to 25–33°. In a comparison of the structures of BR and rhodopsin, it is evident that the rhodopsin structure has longer and more unstructured helices both in terms of tilt and regularity. The tilts and kinks result in shorter interhelix distances between parts of the helices and thereby coupling interactions which are somewhat different from those in BR. A superposition of the two protein structures reveals that there are substantial differences. For instance, helices I–III match

comparatively well, whereas helices IV and V cannot be overlaid.<sup>49</sup>

BR is known to form trimers in the membrane as visualized in Figure 4 that can subsequently associate into a hexagonal lattice.<sup>26</sup> Rhodopsin is generally considered to form dimers in a membrane environment.<sup>37</sup> Dimerization of GPCRs has been widely accepted,<sup>37,63</sup> but there is also evidence that the monomer can perform the biological function.<sup>64</sup>

Figure 5 shows amide I spectra of the helical components of (A) BR and (B) rhodopsin, analogous to the helix bundle

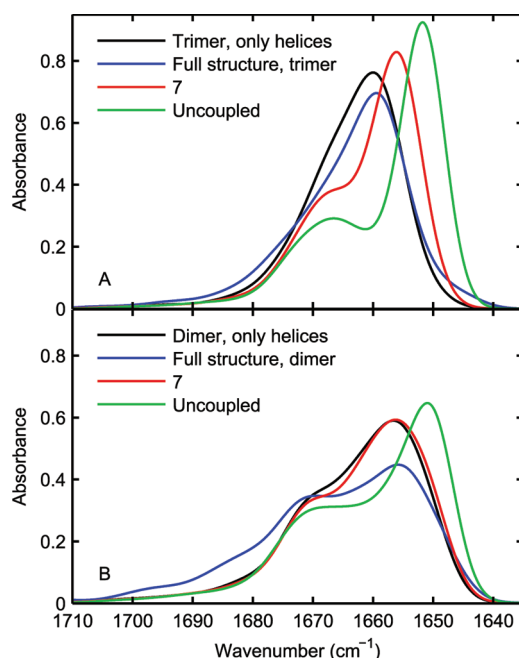


**Figure 5.** Simulated amide I spectra of bundles consisting of 2, 4, and 7 helices from (A) BR and (B) rhodopsin as specified in Structural Information.

spectra presented in the previous sections. In the case of BR, the amide I band maximum shifts 1.8  $\text{cm}^{-1}$  toward higher wavenumbers upon comparing the spectrum from four helices to that of only two and 3.2  $\text{cm}^{-1}$  when increasing the number of helices to seven. This is slightly less than for the model bundles and is explained by the extent of structural heterogeneity such as variations in interhelix distance, tilt, length, and curvature as compared to the model bundle where the helices are identical, equidistantly spaced and aligned along

the  $z$ -axis. For rhodopsin, the amide I maximum shift is more comparable to that of the model helix bundles and shows a shift of 1.7 and 5.2  $\text{cm}^{-1}$  toward higher wavenumbers between the spectra of the two helix systems and the four and seven helix systems, respectively. Different combinations of helix bundles gave similar results for both proteins.

**C1. Interhelix Coupling Is an Explanation for High Wavenumber Absorption of Bacteriorhodopsin.** Also, the amide I spectra of the biological units of BR and rhodopsin were investigated and compared to the subunit spectra as well as the sum of the absorption spectra of the uncoupled helix components of the respective proteins. Figure 6A shows the



**Figure 6.** Comparison between simulated spectra of BR and rhodopsin structures and the sums of the respective uncoupled constituent helices. (A) Simulated amide I spectra for the monomer unit of BR with only helical residues, the biological trimer unit with only helical residues, as well as the full trimer structure and the sum of the spectra of the individual uncoupled helices. (B) Simulated spectra of the rhodopsin monomer and dimer structures including only helical residues, the full dimer structure, and the sum of all individual uncoupled helices.

spectra of the sum of the individual uncoupled helices, the helix monomer and trimer spectra, as well as the full absorption spectrum of the trimer including nonhelical residues. The band shift between the uncoupled helices and the monomer with only helices is 4.3  $\text{cm}^{-1}$  whereas the corresponding difference for the biological trimer unit of BR is 8.3  $\text{cm}^{-1}$ .

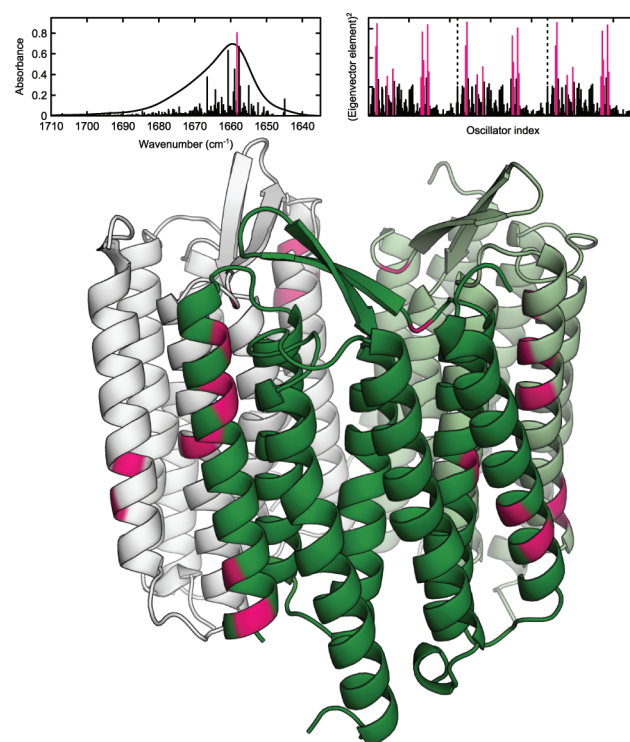
Figure 6B shows the corresponding simulated amide I spectra for rhodopsin and its component helices. The band shift for rhodopsin is 5.3  $\text{cm}^{-1}$  between amide I maxima of spectra of uncoupled helices and the structure with seven helices. There are two potential monomer structures, corresponding to the two chains in the PDB file, as well as two dimer structures that have been proposed as biological units for rhodopsin as specified in the PDB file. The amide I spectrum corresponding to the asymmetric unit (equivalent to one the biological units) was simulated and is shown in Figure 6. The amide I band shift between the dimer and the uncoupled helices is 5.64  $\text{cm}^{-1}$ , that is only another 0.35  $\text{cm}^{-1}$  as compared to the monomer unit.

This is attributed to a high degree of irregularity of the structure. The simulated amide I spectrum of the second dimer structure showed an amide I maximum very similar to that of the monomer; the band is shifted  $\sim 0.5 \text{ cm}^{-1}$  to lower wavenumbers as compared to the monomer spectrum.

Comparison of the amide I spectra of the full protein amide I spectra to spectra of structures with only helical residues retained shows that this reduction does not impact the band position significantly; the difference is less than 1  $\text{cm}^{-1}$  for both proteins. This supports the approach and results presented in earlier sections, where all but helical residues have been neglected. Simulations using older, less refined structures of BR and rhodopsin (1FBB and 1F88) yielded spectra that displayed the same trends in terms of band shifts, but showed fewer features and a slightly higher wavenumber position of rhodopsin.

## C2. Interhelix Coupling Is Present between Subunits.

Investigation of mode delocalization in the BR trimer revealed that the strongest absorbing modes are delocalized and couple over all helices of the trimer. The 21 residues that contribute most to the strongest mode (top left panel) are visualized in the bottom panel of Figure 7. The extent of contribution is further



**Figure 7.** Delocalization of the largest amide I mode in the BR trimer. Top left: Amide I absorption spectrum of the BR trimer. Top right: The largest modes and the extent of participation per oscillator, expressed as square of the eigenvector element per residue. Bottom: Visualization of the coupling between amide I oscillators in the strongest amide I mode in the BR trimer, with the 21 oscillators that participate most highlighted in pink. The monomer units of BR are colored in different shades for clarity.

detailed in the top right panel, which shows the square of the eigenvector elements for all residues in the BR trimer for the strongest mode. The most contributing residues are distributed symmetrically over the three subunits of BR. The extent of delocalization can also be quantified in terms of the inverse



participation ratio (IPR) value, which describes how many oscillators participate in a given mode.<sup>25,65</sup> For the strongest mode in the BR trimer, the IPR value is 197; i.e., 197 oscillators (corresponding to 37% of all oscillators) participate in this mode. For the BR monomer, the corresponding IPR value is 30 (17% of all oscillators).

The extent of localization is highly dependent on heterogeneity of the amide I oscillators, both in terms of intrinsic frequency as well as structure.<sup>25,66</sup> A fully homogeneous set of amide I oscillators in a symmetric environment allows in theory for a mode to be completely delocalized over all residues. Here, the heterogeneity in individual oscillator frequencies is neglected, as is dynamics, which could impose localization of the instantaneous normal modes due to structural fluctuations. The intrinsic frequencies of amide I oscillators in regular transmembrane helices such as those used in the present simulations can be expected to be rather constant because of the defined hydrogen bonding interactions and the absence of interaction with solvent. This mitigates the effect of assigning all oscillators the same intrinsic frequency, which of course is a crude approximation. Diagonal disorder of the intrinsic frequencies has previously been shown to result in a symmetrical broadening of the spectrum,<sup>25</sup> which could here in practice result in less prominent shifts than predicted. A more realistic system including diagonal disorder would generally also have more localized modes and thus lower IPR values. The above concerns do, however, not cloud the general conclusion from the analysis.

Our results convey that interhelix coupling plays a greater role in determining the position of the main band for BR than rhodopsin and that there is coupling between practically all helices within a BR subunit as well as between units in the trimer. The predicted higher wavenumber position of BR is due to the extensive coupling between the helices in the trimer. This provides an explanation for the unusually high experimental amide I band position of BR. The same proposal has been made before, in a conference abstract by Hunt in 1988<sup>67</sup> without further proof.

**C3. Unfolding Experiments Support Simulation Results.** The folding and unfolding properties of BR have been extensively investigated over the years,<sup>68–70</sup> including studies using IR spectroscopy where unfolding has been induced through heat-induced denaturation or through detergent solubilization.<sup>40,42–44,71,72</sup> It has been shown that there are two main temperature dependent transitions, one reversible premelting transition around  $\sim 78$  °C and an irreversible transition at  $\sim 97$  °C.<sup>73</sup> The premelting transition has been shown to be associated with gel to liquid transition of the hexagonal BR lattice<sup>74</sup> and a partial unfolding of the helices<sup>69</sup> using X-ray diffraction measurements.

Upon heating BR just above the premelting point, IR studies show that there is a shift of the amide I maximum from  $1665\text{ cm}^{-1}$  to a more typical  $\alpha$ -helix absorption of  $1652\text{ cm}^{-1}$ .<sup>42</sup> This spectral change was found to be accompanied by the appearance of a band around  $1623\text{ cm}^{-1}$ , characteristic of polypeptide aggregation.<sup>75</sup> Another IR study, where the effect of BR solubilization on IR spectra was investigated, yielded similar results and demonstrated that the exceptionally high wavenumber band is only present when the trimer is intact.<sup>43</sup> The observed band shift upon unfolding of the protein has been proposed to originate from an  $\alpha$  to  $\alpha_{\text{II}}$  helix conversion, as reviewed by Heyes.<sup>40</sup> There has, however, not been found any

evidence for the existence of  $\alpha_{\text{II}}$ -helices in the crystal structures.<sup>40,48</sup>

Our simulation results are consistent with the tertiary and quaternary structure and not a particular secondary structure being responsible for the high amide I maximum of BR. There is reasonable quantitative agreement between the experimental data by Torres et al.<sup>43</sup> and our simulations. Experimentally, the trimer to monomer transition leads to a  $7\text{ cm}^{-1}$  downshift; in our simulations, the downshift is  $4\text{ cm}^{-1}$ . Further loss of intramolecular helix interactions leads to a further downshift by  $3\text{ cm}^{-1}$  experimentally and by  $4.3\text{ cm}^{-1}$  computationally. Our simulations underestimate the effects of trimer formation, which might be due to vibrational coupling between different trimer units enabled by the close packing in the membrane. This would manifest in the experimental spectra but was not simulated by us.

**C4. Experimental Tests for Interhelix Coupling.** At present it is too early to conclude whether an amide I wavenumber that is larger than a certain threshold number (for example  $1660\text{ cm}^{-1}$ ) is indicative for interhelix coupling. Therefore, the following section discusses experimental tests to prove or disprove interhelix coupling. We consider two different cases: intermolecular coupling between different monomeric units of a protein complex and intramolecular coupling within a monomeric protein.

If intermolecular coupling between helices in different units is assumed, this hypothesis can be tested for example by (i) solubilization of membrane proteins, (ii) reconstitution of membrane proteins with high lipid/protein ratios, and (iii) mixing of unlabeled protein with uniformly labeled ( $^{13}\text{C}$  or  $^{13}\text{C}$  and  $^{18}\text{O}$ ) protein. The former two tests reduce the intermolecular coupling due to an increase in distance between the monomers and are similar to the experiments described above for BR. Option ii requires that the monomer–monomer interactions are rather weak and that the oligomers dissociate when the lipid/protein ratio is high. Test iii reduces the spectral effects of coupling due to the different vibrational frequencies in the labeled and unlabeled proteins. In all cases, our simulations predict that the measures lead to a lower amide I wavenumber. For the latter case, 2D IR spectroscopy provides an additional proof of the coupling, since cross peaks between the vibrations of the labeled and unlabeled proteins can be observed.

Intramolecular coupling between helices within a protein monomer is more difficult to prove. If a molten globule state of the protein in question can be generated, the coupling is expected to be reduced due to an increase in distance between the helices. In consequence, the amide I wavenumber will decrease. An unequivocal proof of interhelix coupling within monomeric units can be obtained by selective labeling of amide groups. If single residues in adjacent helices can be labeled, the coupling leads to a splitting of the absorption band of the labeled groups that can be detected by FTIR spectroscopy,<sup>23</sup> or to the appearance of cross peaks in 2D IR spectra when the two residues are differently labeled, e.g.,  $^{13}\text{C}=\text{O}$  and  $^{13}\text{C}=\text{O}$ .<sup>23</sup> While this approach is readily applicable for peptides, it is considerably more difficult for proteins. Here the labeling of entire helices by native chemical ligation seems to be a more straightforward option. In this case, interhelix coupling will manifest in cross peaks in the 2D IR spectrum between labeled and unlabeled helices.

## CONCLUSIONS

Our simulations showed that interhelix coupling influences the amide I absorption of proteins. The interhelix distance is the main parameter that determines the strength of coupling between helices, whereas the relative orientation only moderately influences the coupling strength. What was found to be significantly more relevant for the position of the amide I maximum was the number of helices bundled together, coupling to each other. The results showed that the amide I maximum was shifted to higher wavenumbers as the number of helices increased.

The fact that an increase in the number of bundled helices can be followed in an IR spectrum is interesting and might be useful for several reasons. It could for instance provide a marker for the oligomeric state of membrane proteins, or it could be used to monitor the folding of a protein, in particular when combined with isotope labeling for enhanced structural resolution.

This study also highlights that tertiary quaternary interactions between secondary structure elements influence the band positions that are traditionally used for as structure–spectra correlations for assigning secondary structure absorption in IR spectra. This fact has recently been recognized for the case of  $\beta$ -sheet structures.<sup>20–22</sup> The effect of coupling between secondary structure elements such as helices could be so large that the resulting band position falls outside the typical band assignment ranges. This could for instance be the case for BR, where a large number of helices in the biological trimer unit couple and shifts the band position to higher wavenumber. The experimentally observed high wavenumber IR amide I absorption has mystified spectroscopists for decades now, and several explanations have been provided as to its origin. This paper provides a foundation for a compellingly simple explanation for the unusually high wavenumber band position of BR.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: andreas.barth@dbb.su.se. Phone: +46 (0)8 162452. Fax: +46 (0)8 155597.

### Notes

The authors declare no competing financial interest.

## REFERENCES

- (1) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 729–740.
- (2) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601–19.
- (3) Wim, G.; Hol, J. *Prog. Biophys. Mol. Biol.* **1985**, *45*, 149–195.
- (4) Lemmon, M. A.; Engelman, D. M. *FEBS Lett.* **1994**, *346*, 17–20.
- (5) Popot, J. L.; Engelman, D. M. *Annu. Rev. Biochem.* **2000**, *69*, 881–922.
- (6) Sachs, J. N.; Engelman, D. M. *Annu. Rev. Biochem.* **2006**, *75*, 707–712.
- (7) von Heijne, G. *J. Intern. Med.* **2007**, *261*, 543–557.
- (8) Vinothkumar, K. R.; Henderson, R. *Q. Rev. Biophys.* **2010**, *43*, 65–158.
- (9) Barth, A. In *Biological and Biomedical Infrared Spectroscopy*, 1st ed.; Barth, A., Haris, P. I., Eds.; IOS Press: Amsterdam, 2009; pp 53–78.
- (10) Ganim, Z.; Chung, H. S.; Smith, A. W.; Deflores, L. P.; Jones, K. C.; Tokmakoff, A. *Acc. Chem. Res.* **2008**, *41*, 432–441.
- (11) Cho, M. *Chem. Rev.* **2008**, *108*, 1331–1418.
- (12) Krimm, S.; Abe, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 2788–2792.
- (13) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181–367.
- (14) Higgs, P. W. *Proc. R. Soc. London* **1953**, *A220*, 472–485.
- (15) Wang, J.; Hochstrasser, R. M. *Chem. Phys.* **2004**, *297*, 195–219.
- (16) Barth, A.; Zscherp, C. *Q. Rev. Biophys.* **2002**, *35*, 369–430.
- (17) Choi, J.-H.; Cho, M. In *Biological and Biomedical Infrared Spectroscopy*; Barth, A., Haris, P. I., Eds.; IOS Press: Amsterdam, 2009; pp 224–260.
- (18) Kubelka, J.; Bour, P.; Keiderling, T. A. In *Biological and Biomedical Infrared Spectroscopy*; Barth, A., Haris, P. I., Eds.; IOS Press: Amsterdam, 2009; pp 178–223.
- (19) Hamm, P.; Zanni, M. T. *Concepts and Methods of 2D Infrared Spectroscopy*, 1st ed.; Cambridge University Press: Cambridge, 2011.
- (20) Strasfeld, D. B.; Ling, Y. L.; Gupta, R.; Raleigh, D. P.; Zanni, M. T. *J. Phys. Chem. B* **2009**, *113*, 15679–15691.
- (21) Measey, T. J.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2011**, *133*, 1066–1076.
- (22) Karjalainen, E.-L.; Ravi, H. K.; Barth, A. *J. Phys. Chem. B* **2011**, *115*, 749–57.
- (23) Fang, C.; Senes, A.; Cristian, L.; DeGrado, W. F.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16740–16745.
- (24) Remorino, A.; Korendovych, I. V.; Wu, Y.; DeGrado, W. F.; Hochstrasser, R. M. *Science* **2011**, *332*, 1206–1209.
- (25) Choi, J.-H.; Ham, S.; Cho, M. *J. Chem. Phys.* **2002**, *117*, 6821–6832.
- (26) Haupts, U.; Tittor, J.; Oesterhelt, D. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 367–399.
- (27) Lanyi, J. K. *Annu. Rev. Physiol.* **2004**, *66*, 665–688.
- (28) Kühlbrandt, W. *Nature* **2000**, *406*, 569–570.
- (29) Rosenbaum, D. M.; Rasmussen, S. R. G. F.; Kobilka, B. K. *Nature* **2009**, *459*, 356–363.
- (30) Rothschild, K. J. *J. Bioenerg. Biomembr.* **1992**, *24*, 147–167.
- (31) Siebert, F.; Hildebrandt, P. *Vibrational Spectroscopy in Life Science*, 1st ed.; Wiley-VCH GmbH & Co.: Weinheim, 2008.
- (32) Rothschild, K. J.; Clark, N. A. *Science* **1979**, *204*, 311–312.
- (33) Tamm, L. K.; Tatulian, S. A. *Q. Rev. Biophys.* **1997**, *30*, 365–429.
- (34) Rothschild, K. J.; Clark, N. A. *Biophys. J.* **1979**, *25*, 473–87.
- (35) Haris, P. I.; Coke, M.; Chapman, D. *Biochim. Biophys. Acta* **1989**, *995*, 160–167.
- (36) Byler, D. M.; Susi, H. *Biopolymers* **1986**, *25*, 469–487.
- (37) Palczewski, K. *Annu. Rev. Biochem.* **2006**, *75*, 743–767.
- (38) Smith, A. W.; Lessing, J.; Ganim, Z.; Peng, C. S.; Tokmakoff, A.; Roy, S.; la Cour Jansen, T.; Knoester, J. *J. Phys. Chem. B* **2010**, *114*, 10913–10924.
- (39) Rothschild, K. J.; Sanches, R.; Hsiao, T. L.; Clark, N. A. *Biophys. J.* **1980**, *31*, 53–64.
- (40) Heyes, C. D.; El-sayed, M. A. *J. Phys. Chem. B* **2003**, *107*, 12045–12053.
- (41) Krimm, S.; Dwivedi, A. M. *Science* **1982**, *216*, 407–408.
- (42) Taneva, S. G.; Caaveiro, J. M. M.; Muga, A.; Gofii, F. M. *FEBS Lett.* **1995**, *367*, 297–300.
- (43) Torres, J.; Sepulcre, F.; Padrs, E. *Biochemistry* **1995**, *34*, 16320–16326.
- (44) Wang, J.; El-sayed, M. A. *Biophys. J.* **1999**, *76*, 2777–2783.
- (45) RCSB Protein Data Bank (PDB): <http://www.pdb.org>.
- (46) Chothia, C.; Levitt, M.; Richardson, D. J. *Mol. Biol.* **1981**, *145*, 215–250.
- (47) Sheridan, R. P.; Levy, R. M.; Salemme, F. R. *Biochemistry* **1982**, *21*, 4545–4549.
- (48) Luecke, H.; Schobert, B.; Richter, H.-t.; Cartailler, J.-p.; Lanyi, J. K. *J. Mol. Biol.* **1999**, *291*, 899–911.
- (49) Teller, D. C.; Okada, T.; Behnke, C. A.; Palczewski, K.; Stenkamp, R. E. *Biochemistry* **2001**, *40*, 7761–7771.
- (50) Subramaniam, S.; Henderson, R. *Nature* **2000**, *406*, 653–657.
- (51) Palczewski, K.; Kumasaka, T.; Hori, T.; Trong, I. L.; Teller, D. C.; Okada, T. *Science* **2000**, *289*, 739–745.
- (52) Hexter, R. M. *J. Chem. Phys.* **1960**, *33*, 1833–1841.
- (53) Frenkel, J. *Phys. Rev.* **1931**, *37*, 17–44.



- (54) Hamm, P.; Lim, M.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123–6138.
- (55) la Cour Jansen, T.; Dijkstra, A. G.; Watson, T. M.; Hirst, J. D.; Knoester, J. *J. Chem. Phys.* **2006**, *125*, 44312.
- (56) Torii, H.; Tasumi, M. *J. Chem. Phys.* **1992**, *96*, 3379–3387.
- (57) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations. The Theory of Infrared and Raman Vibrational Spectra*; Dover Publications, Inc.: Mineola, NY, 1955.
- (58) Torii, H.; Tasumi, M. *J. Raman Spectrosc.* **1998**, *29*, 81–86.
- (59) Lee, C.; Cho, M. *J. Phys. Chem. B* **2004**, *108*, 20397–20407.
- (60) Gorbunov, R. D.; Kosov, D. S.; Stock, G. *J. Chem. Phys.* **2005**, *122*, 224904.
- (61) Torii, H.; Tasumi, M. In *Infrared Spectroscopy of Biomolecules*; Mantsch, H. H., Chapman, D., Eds.; Wiley Liss: New York, 1996; pp 1–17.
- (62) Hamm, P.; Woutersen, S. *Bull. Chem. Soc. Jpn.* **2002**, *75*, 985–988.
- (63) Fotiadis, D.; Jastrzebska, B.; Philippsen, A.; Müller, D. J.; Palczewski, K.; Engel, A. *Curr. Opin. Struct. Biol.* **2006**, *16*, 252–259.
- (64) Chabre, M.; le Maire, M. *Biochemistry* **2005**, *44*, 9395–9403.
- (65) Thouless, D. *Phys. Rep.* **1974**, *13*, 93–142.
- (66) Demirdöven, N.; Cheatum, C. M.; Chung, H. S.; Khalil, M.; Knoester, J.; Tokmakoff, A. *J. Am. Chem. Soc.* **2004**, *126*, 7981–7990.
- (67) Hunt, J. F. *Biophys. J.* **1988**, *53*, 97a.
- (68) Booth, P. J. *Biochim. Biophys. Acta* **2000**, *1460*, 4–14.
- (69) Müller, J.; Münster, C.; Salditt, T. *Biophys. J.* **2000**, *78*, 3208–3217.
- (70) Curnow, P.; Booth, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18970–18975.
- (71) Wang, J.; El-sayed, M. A. *Biophys. J.* **2000**, *78*, 2031–2036.
- (72) Wang, J.; Heyes, C. D.; El-sayed, M. A. *J. Phys. Chem. B* **2002**, *106*, 723–729.
- (73) Jackson, M. B.; Sturtevant, J. M. *Biochemistry* **1978**, *17*, 911–915.
- (74) Hiraki, K.; Hamanaka, T.; Mitsui, T.; Kito, Y. *Biochim. Biophys. Acta* **1981**, *647*, 18–28.
- (75) Zandomeneghi, G.; Krebs, M. R. H.; Mccammon, M. G.; Fändrich, M. *Protein Sci.* **2004**, *13*, 3314–3321.