LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-405754

Spectral Signatures of the Pentagonal Water Cluster in Bacteriorhodopsin

M. Baer, G. Mathias, I-F. W. Kuo, D. J. Tobias, C. J. Mundy, D. Marx

July 29, 2008

ChemPhysChem

Disclaimer

Spectral Signatures of the Pentagonal Water Cluster in Bacteriorhodopsin

Marcel Baer,[a]*, Gerald Mathias,[a] I-Feng W. Kuo,[b] Douglas J. Tobias,[c], Christopher J. Mundy[d] and Dominik Marx[a]

[a] Lehrstuhl für Theoretische Chemie, Ruhr–Universität Bochum, 44780 Bochum, Germany

[b] Lawrence Livermore National Laboratory 7000 East Avenue, Livermore, CA 94550

[c] University of California, 4118 NS1, Irvine, CA 92697

[d] Pacific Northwest National Laboratory, PO Box 999, Richland, WA 99352

July 25, 2008

The exchange of protons between basic and acidic groups within proteins often involves transient protonation of amino acids and water molecules embedded in the protein matrix.[1;2] One of the best studied proteins in this respect is Bacteriorohodopsin (BR), which works in the membrane of Halobacterium salinarium as a light–driven proton pump.[3] The pumping process is triggered in the initial bR state by a photon absorption of an *all-trans* retinylidene chromophore, which is linked via a protonated Schiff base (pRSB) to the sidechain of Lys216. The subsequent photocycle comprises a series of intermediate states J, K, L, M, N and O, which are characterized by conformational and absorbance changes of the chromophore accompanying several elementary proton transfer processes.[4] Upon completion of the photocycle one net proton has been transferred from the cyctoplasmic to the extracellular side against the proton gradient across the membrane. These proton exchange reactions can be mon-

*To whom correspondence should be addressed. E-mail: marcel.baer@theochem.rub.de

itored by time resolved infrared (IR) spectroscopy of the BR wild type and site specific mutants, which allow the localization of absorbance changes within the protein.[5;6] Furthermore, these measurements have revealed the fundamental importance of internal water molecules in these processes[6–9] as supported by recent large-scale QM/MM molecular dynamics studies of anharmonic IR spectra.[10–12].

Within BR the most prominent ensemble of water molecules is located beneath the Schiff base, depicted in Figure 1. In the bR initial state this region has

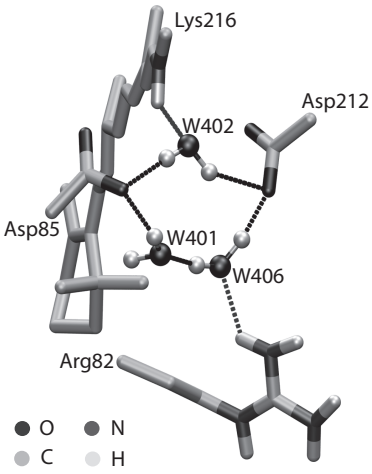

Figure 1: Structure of the Schiff Base region in BR generated from PDB entry 1C3W[13]. The oxygen atoms of three distinct water molecules (W401, W402, W406) form a roughly pentagonal cluster together with an oxygen from Asp85 and an oxygen from Asp212. Aliphatic hydrogens are omitted for clarity.

an electrostatic quadrupole structure with positive charges located at pRSB and Arg82, which are counterbalanced by negative charges at Asp85 and Asp212. Within this quadrupolar arrangement of charges, three internal water molecules, namely W401, W402, and W406, are resolved in the crystal structure.[13] Their hydrogen-bonded network spans a roughly planar pentagonal hydrogen-bonded cluster, composed of the three water oxygens and one of the carboxy oxygens of both Asp85 and Asp212.

Experimentally two specific features of the IR spectrum are unambiguously assigned to the pentagonal water cluster. The first one is a so called dangling bond at $3644\,\mathrm{cm}^{-1}$ that is attributed to W401 (see Refs.[9;14–16]), which has no hydrogen-bonding partner in the most likely arrangement of hydrogens as shown in Figure 1. Note, however, that hydrogen atoms are not resolved by x-ray crystallography. The second feature is an unusually broad band between $2400\,\mathrm{cm}^{-1}$ and $3000\,\mathrm{cm}^{-1}$, which is present in the bR state and bleaches in the K and L intermediates.[9] A similar band has been observed in low-temperature spectra of the K intermediate and it has been proposed that this band consists predominantly of absorptions of strong hydrogen-bonded waters superimposed by the N–H stretch of the strongly hydrogen-bonded Schiff Base.[16] Because W402 is hydrogen-bonded to the positively charged Schiff base and to the two negative carboxylates Asp85 and Asp212 it is the most likely origin of this broad absorbance.[6;8;9;16–19]

Computer simulations can relate such spectra to structural and functional models via combined quantum mechanical / molecular mechanical (QM/MM) calculations,[10–12;20;21] which treat the region of interest with quantum mechanical electronic structure methods, whereas the surrounding environment is described by a classical mechanical force field. In pioneering static QM/MM studies Hayashi *et al.* determined harmonic frequencies for deuterated water molecules W401 and W402 and the N–D stretch of the Schiff base,[20;21] which were found to be in good agreement with experiment.[6;16;17] Therein, the QM part was treated within the Hartree–Fock approximation which typically overbinds hydrogen bonds. Furthermore, the underlying normal mode analysis neglects anharmonic effects and mode couplings, which might be important in hydrogen–bonded networks. As a remedy the frequencies where scaled down by 10 % in order to match the right frequency range, as is usually done in such calculations.[22] Most importantly, however, this methodology always yields *sharp* IR resonances and, by construction, is unable to explain the origin of the extremely broad band observed experimentally below $3000\,\mathrm{cm}^{-1}$.

In this work we address the effects of finite temperatures in conjunction with

the dynamical interaction with the protein environment by employing three model systems of systematically increased size within a QM/MM framework. Model A, the minimal model, includes only the three water molecules W401, W402 and W406 treated using QM. Model B adds the sidechains of the two deprotonated aspartic acids Asp85 and Asp212 to the QM part. Model C comprises, furthermore, the protonated retinal Schiff base including the Lys216 sidechain and the positive Arg82, the last missing charged sidechain in this region. Hence, the whole pRSB, the complex counterion and the pentagonal water cluster are treated on the QM level in model C. Together with the two smaller models it allows to dissect explicitly the effects of different residues and the Schiff base on the measured IR spectra.

Within a pure force field description the water molecules beneath the Schiff base are highly mobile and the pRSB tends to from a salt bridge with Asp85,[23;24] which was attributed to a lack of polarizability in these models. In all three QM/MM models considered here, we find the pRSB mostly hydrogen bonded to W402 and, at the same time, the water molecules are far less mobile than in the aforementioned pure force field simulations. For models A and B, the pentagonal water cluster turns out to be conformationally stable throughout the whole simulation, whereas in model C it is seen to undergo interesting local changes of the hydrogen-bonding topology. In particular, a salt bridge forms between Arg82 and Asp212 within a few picoseconds and breaks the hydrogen bond between Arg82 and W406. After the rearrangement of the hydrogen-bonded network W401 becomes the donor in the hydrogen bond with W406.

Figure 2.a) shows the total IR absorption in the O–H stretch region of the three water molecules resulting from models A, B, and C of increasing complexity. All models exhibit absorption in the range from $2800\,\mathrm{cm}^{-1}$ to $3650\,\mathrm{cm}^{-1}$, which is typical for water molecules depending on their coupling and hydrogen-bonding strength.[25] An important observation is that only the largest model, C, features a significant intensity below $3000\,\mathrm{cm}^{-1}$ with a flat maximum at about $2800\,\mathrm{cm}^{-1}$ in accordance with the experimentally observed broad absorbance feature.

For a more detailed understanding and assignment we now turn to the vibrational spectra of the *individual* water molecules. Table 1 lists the O–H stretch frequencies of the water molecules obtained by a Fourier transform of the corresponding bond length autocorrelation function. The resulting frequency distributions of model A and B are characterized by dominant single peaks. For

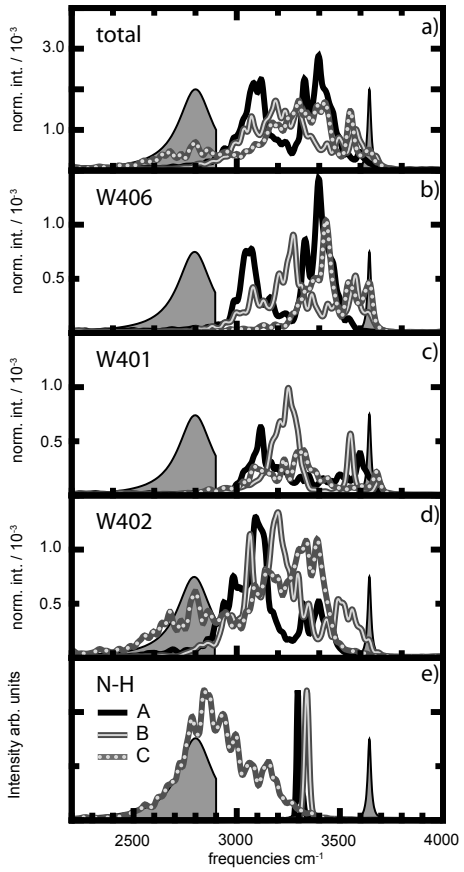

Figure 2: Calculated IR–spectra obtained from models A, B, and C (labeled according to inset): a) for all water molecules, b) water W406, c) W401 and d) W402. e) power spectrum of the Schiff base N–H bond length. Experimental frequencies (taken from Ref.[9], and assigned to the water cluster therein) are marked by grey areas. Their intensities are chosen arbitrarily as they result from difference spectra and are not readily comparable to computed intensities.

	A / cm^{-1}	B / cm^{-1}	C / cm^{-1}
W401	3590	3677	3570/3690
	3120	3325	3200/3270
W402	3086	3550	3420
	2980	3180	3190
W406	3394	3570	3649
	3060	3270	3340/3450

Table 1: Frequencies of the individual O–H stretches for water W401, W402 and W406 within models A, B and C obtained by a Fourier transform of the bond length.

model C we find also bimodal distributions which correspond to different conformational states of the hydrogen bonded network discussed above. All stretch frequencies are separated by about $300\,cm^{-1}$, except for W401 in model A. Here, the splitting is smaller because the hydrogen bonding is symmetric and the force field parameterization for both aspartic acids is identical. The O–H stretches of W402 experience also the largest blue shift upon increasing the QM region, i.e. when switching to models B and finally to C.

The individual IR absorptions of each water molecule are disentangled in Figures 2.b) to d). W406 has three major absorption peaks for all models: in model A at 3050, 3320, and $3400\,cm^{-1}$, in model B at 3070, 3275, and $3570\,cm^{-1}$ and in model C at 3420, 3530, and $3640\,cm^{-1}$. For each model two peaks correspond to the bare O–H stretching frequencies given in Table 1. The remaining peaks have to be assigned to coupling terms between the water molecules and with the environment. Because of the structural changes in model C discussed above, W406 has one hydrogen that is not involved in a hydrogen bond, which gives rise to the high frequency band at $3640\,cm^{-1}$, the spectral position of the dangling bond.

The O–H stretching frequencies of water W401 (Figure 2.c) are distributed

in the range between 3000 cm^{-1} and 3700 cm^{-1} for all models. For all models we find peaks corresponding to the positions of the bond stretching frequencies given in Table 1. Compared to W406, Figure 2.b), the spectral density is separated into two regions, below and above 3400 cm^{-1}. This reflects the hydrogen bonding arrangement in which W401 donates a hydrogen bond to the negatively charged Asp85 and accepts (model A, B) one from W406 resulting in a redshift of the O–H stretching. This effect is not as pronounced for W406 in model C, as the carboxylic group of Asp212 is also involved in a salt bridge to Arg82. Thus, we find consistently the signatures of the dangling bond in all three models. For the minimal model A it shows up slightly red shifted at W401. In model B the signal appears both for W401 and W406 at even lower frequencies. This is probably due to the imbalanced description of the QM aspartic acids and the MM argenine. Finally, Model C comes closest to the experimental peak position in addition to featuring the conformational change. These findings emphasize the necessity to choose the QM region carefully and in a balanced fashion.

The peculiar spectral properties stemming from W402 in Figure 2.d) reveal the most striking differences between the models. The main spectral features for models A and B are in line with the peaks of the power spectra of the O–H bonds given in Table 1. In contrast in model C we find in addition a very broad band peaked around 2800 cm^{-1} with a shoulder ranging down to 2500 cm^{-1}, which cannot be solely attributed to the O–H stretches of this water molecule. To examine the previously proposed coupling to the pRSB N–H stretch, we provide the power spectrum of this particular mode in Figure 2.e). For models A and B we find a single sharp peak at 3300 cm^{-1} resulting from the force field parameterization. Within the full QM treatment in model C this mode is extremely broad spanning the range between 2500 cm^{-1} and 3300 cm^{-1} with a maximum at 2850 cm^{-1}. In particular, the wing to the red coincides perfectly with the band we observe in water W402 and, most importantly, with the continuum band in the experimental spectrum, which is shown as grey shaded area (note that no frequency scaling whatsoever was applied). We therefore attribute this continuum band to a polarization coupling between the pRSB N–H stretch

and water W402, which amplifies the IR absorption greatly. This effect can only be seen if both moieties are treated on the QM level, i.e. in the large model C. Furthermore, reference calculation on model A, in which we reparametrized the N–H stretch to 2800 cm^{-1}, showed no significant change of absorption below 3000 cm^{-1}. Thus, the origin of this particular continuum band is the mutual polarization between the Schiff base proton and water W402 in the field of the two aspartic acids.

In conclusion, these extensive finite-temperature QM/MM molecular dynamics simulations of the pentagonal water cluster in the initial bR state yield detail insight into vibrational spectra, which allow a most direct interpretation of experimental spectra without any adjustment. In particular, we reveal in detail how different topologies of the hydrogen-bonded network do affect the IR absorption and its interpretation. Furthermore, intermolecular couplings in highly charged environments are subject to complex polarization effects. The resulting broad continuum bands are only faithfully accessible via a complete quantum mechanical treatment of the "hot spot" region at finite temperatures.

Acknowledgments:

Helpful discussions with Florian Garczarek, Erik Freier, and Klaus Gerwert are gratefully acknowledged. We thank Deutsche Forschungsgemeinschaft (DFG FOR 436), CCMS at LLNL, Fonds der Chemischen Industrie (FCI) and the National Science Foundation (grant CHE-0750175) for partial financial support. The simulations were carried out at NIC Jülich (DEISA/DECI grant), BOVILAB@RUB (Bochum), and Ressourcenverbund Nordrhein–Westfalen (RV–NRW). Part of this work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. CJM is supported by the Department of Energy, Office of Basic Energy Sciences Chemical Sciences Division. Pacific Northwest National Laboratory is operated by Battelle for the Department of Energy.

Computational Methods:

A detailed description how the system was set up and classically equilibrated can be found also in this issue.[24] The system was equilibrated for 8 ns using

Gromacs-3.2.1 only restraining the heavy atoms of model C to the crystal structure. All QM/MM calculations were performed using the software suite `CP2k` (http://cp2k.berlios.de). The interaction energy for the QM region was computed via the `QuickStep`[26;27] module within `CP2k`. The `QuickStep` module performs Kohn-Sham density functional calculations using a dual basis set method (Gaussian plus plane wave basis functions) in conjunction with Goedecker-Teter-Hutter (GTH) pseudopotentials[28;29] to describe the core electrons. A triple-zeta Gaussian basis set augmented with two sets of d-type or p-type polarization functions (TZV2P) was used and a 280 Ry plane wave cutoff for the electron density. Exchange and correlation energy was computed within the GGA approximation using the BLYP functional.[30;31] For every time step, the electronic structure was explicitly quenched to a tolerance of 10^{-7} Hartree using a time step of 0.5 fs. The temperature was controlled using one Nosé-Hoover chain thermostat[32] for the whole system. The electrostatic coupling between the QM and MM regions was computed via a real space multigrid procedure.[33] For models B and C covalent bonds between the QM and MM regions (i.e. C_α and C_β) were treated within the IMMOM link atom approach.[34] The IR spectra were calculated as the Fourier transforms of the dipole autocorrelation functions to which harmonic quantum correction factors were applied.[35] The necessary molecular dipoles were obtained from atomic core charges together with nearest Wannier function centers sampled every 1 fs.

References

[1] a) T. Decoursey, *Physiol. Rev.* **2003**, *83*, 475–579; b) T. DeCoursey, *Physiol. Rev.* **2003**, *83*, 1067–1067; c) T. DeCoursey, *Physiol. Rev.* **2004**, *84*, 1479–1479.

[2] S. Cukierman, *Biochim. Biophys. Acta* **2006**.

[3] U. Haupts, J. Tittor, D. Oesterhelt, *Ann. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 367–399.

[4] J. K. Lanyi, *Biochim. Biophys. Acta* **2006**, *1757*, 1012–1018.

[5] F. Siebert, *Methods in Enzymology* **1995**, *246*, 501–526.

[6] H. Kandori, *Biochim. Biophys. Acta-Bioenerg.* **2000**, *1460*, 177–191.

[7] R. Rammelsberg, G. Huhn, M. Lübben, K. Gerwert, *Biochemistry* **1998**, *37*, 5001–5009.

[8] F. Garczarek, L. Brown, J. Lanyi, K. Gerwert, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 3633–3638.

[9] F. Garczarek, K. Gerwert, *Nature* **2006**, *439*, 109–112.

[10] R. Rousseau, V. Kleinschmidt, U. Schmitt, D. Marx, *Phys. Chem. Chem. Phys.* **2004**, *6*, 1848–1859.

[11] R. Rousseau, V. Kleinschmidt, U. Schmitt, D. Marx, *Angew. Chem.-Int. Edit.* **2004**, *43*, 4804–4807.

[12] G. Mathias, D. Marx, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 6980–6985.

[13] H. Luecke, B. Schobert, H. Richter, J. Cartailler, J. Lanyi, *J. Mol. Biol.* **1999**, *291*, 899–911.

[14] A. Maeda, J. Sasaki, Y. Yamazaki, R. Needleman, J. Lanyi, *Biochemistry* **1994**, *33*, 1713–1717.

[15] H. Kandori, Y. Yamazaki, J. Sasaki, R. Needleman, J. Lanyi, A. Maeda, *J. Am. Chem. Soc.* **1995**, *117*, 2118–2119.

[16] N. Shibata, H. Kandori, *Biochemistry* **2005**, *44*, 7406–7413.

[17] H. Kandori, N. Kinoshita, Y. Shichida, A. Maeda, *J. Phys. Chem. B* **1998**, *102*, 7899–7905.

[18] T. Tanimoto, Y. Furutani, H. Kandori, *Biochemistry* **2003**, *42*, 2300–2306.

[19] H. Kandori, N. Kinoshita, Y. Yamazaki, A. Maeda, Y. Shichida, R. Needleman, J. Lanyi, M. Bizounok, J. Herzfeld, J. Raap, J. Lugtenburg, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 4643–4648.

[20] S. Hayashi, I. Ohmine, *J. Phys. Chem. B* **2000**, *104*, 10678–10691.

[21] S. Hayashi, E. Tajkhorshid, H. Kandori, K. Schulten, *J. Am. Chem. Soc.* **2004**, *126*, 10516–10517.

[22] J. Neugebauer, B. A. Hess, *J. Chem. Phys.* **2003**, *118,*, 7215–7225.

[23] C. Kandt, J. Schlitter, K. Gerwert, *Biophys. J.* **2004**, *86*, 705–717.

[24] A. Chaumont, M. Baer, G. Mathias, D. Marx, *ChemPhysChem* **2008**.

[25] H. Kandori, Y. Shichida, *J. Am. Chem. Soc.* **2000**, *122*, 11745–11746.

[26] G. Lippert, J. Hutter, M. Parrinello, *Mol. Phys.* **1997**, *92*, 477–487.

[27] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, *Comput. Phys. Commun.* **2005**, *167*, 103–128.

[28] S. Goedecker, M. Teter, J. Hutter, *Phys. Rev. B* **1996**, *54*, 1703–1710.

[29] C. Hartwigsen, S. Goedecker, J. Hutter, *Phys. Rev. B* **1998**, *58*, 3641–3662.

[30] A. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.

[31] C. Lee, W. Yang, R. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.

[32] G. Martyna, M. Klein, M. Tuckerman, *J. Chem. Phys.* **1992**, *97*, 2635–2643.

[33] T. Laino, F. Mohamed, A. Laio, M. Parrinello, *J. Chem. Theory Comput.* **2005**, *1*, 1176–1184.

[34] F. Maseras, K. Morokuma, *J. Comput. Chem.* **1995**, *16*, 1170–1179.

[35] R. Ramírez, T. López-Ciudad, P. P. Kumar, D. Marx, *J. Chem. Phys.* **2004**, *121*, 3973–3981.