

# Supporting Information

## 1 Structural Models

There are currently four high-resolution structures of bovine visual rhodopsin reported in the literature.<sup>1-4</sup> The structure with PDB identifier 1F88,<sup>1</sup> solved at 2.8 Å resolution, is the first high-resolution structure of a GPCR and reveals the major features of the protein previously inferred from a variety of experimental studies, including cryomicroscopy.<sup>5</sup> A refined model (PDB identifier 1HZX)<sup>2</sup> adds some amino acid residues not identified in structure 1F88. A more recent refinement (PDB identifier 1L9H) improves the resolution to 2.6 Å and resolves essential structural components of the chromophore binding site, including two bound-water molecules next to the retinyl chromophore.<sup>3</sup> Finally, the most recent structure (PDB identifier 1U19) completely resolves the polypeptide chain at 2.2 Å resolution and provides further conformational details of the retinyl chromophore.<sup>4</sup>

The energy storage during the primary photochemical event involves a mechanism that depends on localized structural rearrangements within the retinyl chromophore and its close neighboring residues. The protein surrounding the chromophore is well characterized in all of the crystal structures mentioned above. QM/MM modeling of the energy storage mechanism has shown that bound water molecules observed in the 1L9H structure, but not resolved at lower resolution, are critical for the underlying conformational changes. It is therefore expected that as long as these water molecules are present in higher resolution structures, the reported calculations should be robust.

The computational models developed by recent DFT QM/MM studies<sup>6,7</sup> are based on the refinement of crystal structure 1F88.<sup>1</sup> The rhodopsin cavity is set neutral, consistently with FTIR experiments.<sup>8</sup> The pSB linkage between Lys-296 and the chromophore bears a net-positive charge NH(+), forming a salt-bridge with the negative counter-ion Glu-113.<sup>9,10</sup> Amino-acid residues Glu-122, Asp-83 and Glu-181, all within the protein core, are assumed to be neutral as indicated by FTIR experiments<sup>8</sup> and UV-vis spectroscopic measurements of site-directed mutants.<sup>11</sup> Further, model systems based on the crystal structure 1L9H, monomer A, solved at 2.6 Å resolution<sup>3</sup> have also been prepared. A crucial difference between the 1F88 and 1L9H crystal structures is the presence of two water molecules next to the retinyl chromophore in the 1L9H structure, not observed at

lower resolution. These two water molecules are included in QM/MM models prepared from the 1F88 structure since they play a crucial structural role in the stabilization of the chromophore as revealed by the DFT QM/MM studies.<sup>6</sup> Specifically, bound-water molecules contribute to the stability of an extended hydrogen-bonding network involving highly conserved amino-acid residues in the EII loop, including Glu-181, Cys-187, Ser-186 and Glu-113 (see Fig. 1). Such a structural feature is consistent with the counter-ion switch mechanism recently proposed for a subsequent step in the rhodopsin photobleaching sequence.<sup>12,13</sup>

DFT QM/MM computational models of rhodopsin have been obtained at the ONIOM Electronic Embedding (B3LYP/ 6-31G\*: Amber) level of theory.<sup>6,7</sup> Equilibrium configurations have been computed by geometry relaxation with respect to the conformation of the retinyl chromophore, bound-water molecules and amino-acid residues within a 20 Å radius from the retinyl chromophore. The protein structure beyond the 20 Å radius has been subject to harmonic constraints in order to preserve the natural shape of the protein.

## 2 QM/MM method

The *ONIOM Electronic-Embedding* (ONIOM-EE) method,<sup>14–20</sup> depicted in Fig. 1, involves a two-layer link-atom scheme where the prosthetic group of interest is modeled within a molecular layer described at the Quantum Mechanical (QM) level of theory, while the otherwise classical protein environment is described according to a Molecular Mechanics (MM) force field.

The energy evaluation involves the combination of three independent calculations as follows:

$$E = E^{\text{MM,full}} + E^{\text{QM,red}} - E^{\text{MM,red}}, \quad (1)$$

where  $E^{\text{MM,full}}$  is the energy of the full-system computed at the MM level of theory and  $E^{\text{QM,red}}$  is the energy of the reduced-system computed at the QM level of theory, including the electrostatic interactions between the QM electronic density and the MM layer. Finally,  $E^{\text{MM,red}}$  is the energy of the reduced-system computed at the MM level of theory, including the classical electrostatic interactions between the distribution of atomic charges in the reduced system and the MM layer. Note that since electrostatic interactions between the two layers are included in the calculation of both terms  $E^{\text{QM,red}}$  and  $E^{\text{MM,red}}$ , electrostatic

contributions included at the MM level in  $E^{\text{MM,red}}$  and  $E^{\text{MM,full}}$  cancel. Therefore, the resulting QM/MM evaluation of the total energy involves a rigorous QM description of the polarization of the prosthetic group induced by the surrounding protein environment.

### ONIOM-EE Energy Evaluation:

QM = DFT B3LYP/6-31G\*

MM = Amber Force Field

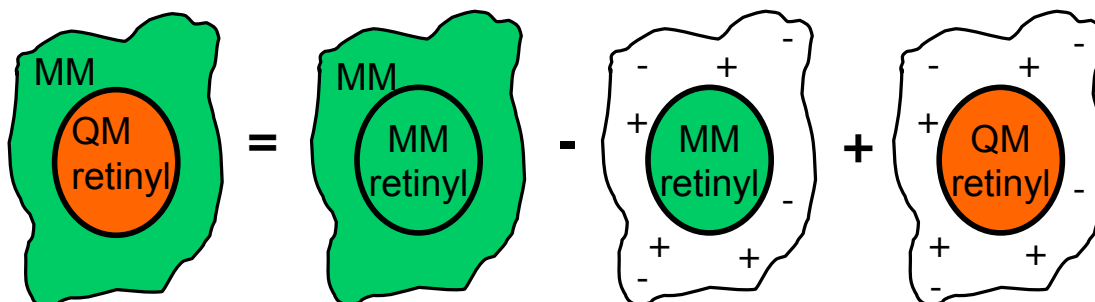


Figure 1: DFT QM/MM approach implemented in the description of molecular rearrangements due to 11-*cis*/all-*trans* isomerization in visual rhodopsin.<sup>6,7</sup> The methodology involves the ONIOM Electronic Embedding hybrid approach,<sup>14-20</sup> where the retinyl chromophore and nearby amino-acid residues are modeled within the Quantum Mechanics (QM) layer, including the underlying induced polarization due to the rest of the protein that is modeled according to a classical Molecular Mechanics (MM) level of theory.

The DFT QM/MM studies of energy storage and molecular rearrangements due to 11-*cis*/all-*trans* isomerization in visual rhodopsin have applied the ONIOM-EE method, as implemented in Gaussian03.<sup>21</sup> In these calculations, the rhodopsin binding site is partitioned into QM and MM layers with a frontier at the C<sub>δ</sub>-C<sub>ε</sub> bond of the Lys-296 side chain (*i.e.*, two bonds beyond the C-NH(+) linkage).<sup>6</sup> The QM layer includes 48 atoms of the retinyl chromophore, five atoms of Lys-296 (NH<sup>+</sup>, CH<sub>2</sub>) and a link hydrogen atom that saturates the extra valence of the terminal -C-H<sub>2</sub> at the boundary. We have recently investigated minimum energy structures obtained with extended QM layers including Glu-113, Glu-181, Ser-186, Cys-187, Lys-296, wat2a, and wat2b. We considered these residues since they play a dominant role in the underlying energy storage mechanism. Even for this

extended QM-MM setup with extended QM layer we have found no significant structural and energetic differences, demonstrating convergence relative to the size of the QM layer. In addition, a higher level of theory has been tested by re-evaluating the energy storage through geometry optimization of the *cis* and *trans* configurations at the ONIOM(mp2/6-31g\*:Amber), finding a difference of less than 1 kcal/mol with respect to the DFT calculation. Computational studies of the NMR spectroscopy, however, require modeling a more extended QM layer in order to accurately describe magnetic shielding effects.<sup>7</sup> Single point calculations of chemical-shifts explicitly consider extended QM layers, including amino-acid residues with significant steric interactions with the chromophore as well as nearby residues with aromatic functional groups (e.g., Trp-265, Tyr-268, Ser-186, Cys-187 and Gly-188).<sup>7</sup> The explicit inclusion of these residues results in a QM layer with 80 additional atoms, including the 6 link-hydrogen atoms placed at the corresponding frontiers. All ONIOM-EE studies<sup>6,7</sup> reported here describe the rest of the protein according to the Amber MM force field.<sup>22</sup>

### 3 Finite Temperature calculations

Helmholtz's free energies calculations of the *cis/trans* isomerization were based on hybrid Monte Carlo sampling of thermal configurations and thermodynamic integration at the MM level, followed by QM/MM calculations of statistical weighting factors associated with the complete thermodynamic cycle.<sup>23,24</sup> This approach allows for DFT QM/MM free energy calculations, by-passing the need of sampling thermal configurations at a high level of theory.

The QM/MM isomerization free energy change, between states 11-*cis* and 11-*trans*, is calculated as the sum of three terms: (1) the isomerization free energy difference at the classical MM level; (2) the negative free energy term between *cis* at the classical level and *cis* at the QM/MM level; (3) the free energy difference between *trans* at the classical level and *trans* at the QM/MM level. Both (2) and (3) are evaluated as  $\Delta A = -kT \ln \langle \exp(-\Delta V/kT) \rangle_{\text{MM}}$ , where  $\Delta V = E_{\text{QM/MM}} - E_{\text{MM}}$  is the energy difference between QM/MM and MM levels of theory, and MM denotes a molecular mechanics ensemble average. The

MM free energy difference is evaluated as

$$\Delta A = -kT \ln \frac{\langle \exp(-E_{trans}/kT) \rangle_{MM}}{\langle \exp(-E_{cis}/kT) \rangle_{MM}}. \quad (2)$$

The ensemble averages were performed at 300 K, by sampling 2000 configurations during 60 ps trajectories of molecular dynamics generated with the AMBER MM force field.<sup>22</sup> Due to the intrinsic limitations of the MM force field, MD simulations were carried out with the chromophore harmonically constrained to the *cis*, or *trans*, conformation obtained at the DFT QM/MM level. Residues with  $\alpha$  carbons beyond a 20 Å radius were also subject to harmonic constrains.

The resulting DFT QM/MM free energy storage is 30 kcal/mol, 4 kcal/mol smaller than the 34 kcal/mol computed at 0 K, in very good agreement with the experimental values of 32-35 kcal/mol. The individual free energy components of the cycle, described above, are: (1) -12 kcal/mol, (2) + 21 kcal/mol, and (3) 21 kcal/mol.

## References

- [1] Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289*, 739-745.
- [2] Teller, D.; Okada, T.; Behnke, C.; Palczewski, K.; Stenkamp, R. Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry* **2001**, *40*, 7761-7772.
- [3] Okada, T.; Yoshinori, Y.; Silow, M.; Navarro, J.; Landau, J.; Schichida, Y. Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc. Natl. Acad. Sci.* **2002**, *99*, 5982-5987.
- [4] Okada, T.; Sugihara, M.; Bondar, A.; Elstner, M.; Entel, P.; Buss, V. The Retinal Conformation and its Environment in Rhodopsin in Light of a New 2.2 Å Crystal Structure. *J. Mol. Biol.* **2004**, *342*, 571-583.
- [5] Schertler, G.; Hargrave, P. Preparation and analysis of two-dimensional crystals of rhodopsin. *Methods Enzymol.* **2000**, *315*, 91-107.
- [6] Gascón, J. A.; Batista, V. S. QM/MM Study of Energy Storage and Molecular Rearrangements due to the Primary Event in Vision. *Biophys. J.* **2004**, *87*, 2931-2941.
- [7] Gascón, J. A.; Sproviero, E. M.; Batista, V. S. QM/MM Study of the NMR Spectroscopy of the Retinyl Chromophore in Visual Rhodopsin. *J. Chem. Theory Comp.* **2005**, *1*, 674-685.
- [8] Fahmy, K.; Jäger, F.; Beck, M.; Zvyaga, T. A.; Sakmar, T. P.; Siebert, F. Protonation states of membrane-embedded carboxylic-acid groups in rhodopsin and metarhodopsin II - A Fourier-transform infrared-Spectroscopy study of site-directed mutants. *Proc. Natl. Acad. Sci.* **1993**, *90*, 10206-10210.
- [9] Sakmar, T. P.; Franke, R. R.; Khorana, H. G. Glutamic acid-113 serves as the retinylidene schiff-base counterion in bovine rhodopsin. *Proc. Natl. Acad. Sci.* **1989**, *86*, 8309-8313.
- [10] Zhukovsky, E. A.; Oprian, D. D. Effect of carboxylic-acid side-chains on the absorption maximum of visual pigments. *Science* **1989**, *246*, 928-930.
- [11] Yan, E. C. Y.; Kazmi, M. A.; De, S.; Chang, B. S. W.; Seibert, C.; Marin, E. P.; Mathies, R. A.; Sakmar, T. P. Function of extracellular loop 2 in rhodopsin: Glutamic acid 181 modulates stability and absorption wavelength of metarhodopsin II. *Biochemistry* **2002**, *41*, 3620-3627.
- [12] Birge, R. R.; Knox, B. E. Perspectives on the counterion switch-induced photoactivation of the G protein-coupled receptor rhodopsin. *Proc. Natl. Acad. Sci.* **2003**, *100*, 9105-9107.

- [13] Yan, E. C. Y.; Kazmi, M. A.; Ganim, Z.; Hou, J.; Pan, D.; Chang, B. S. W.; Sakmar, T. P.; Mathies, R. A. Retinal counterion switch in the photoactivation of the G protein-coupled receptor rhodopsin. *Proc. Natl. Acad. Sci.* **2003**, *100*, 9262-9267.
- [14] Maseras, M.; Morokuma, K. IMOMM - A new integrated ab-initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition-states. *J. Comp. Chem.* **1995**, *16*, 1170-1179.
- [15] Svensson, M.; Humbel, S.; Froese, R.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A multilayered integrated MO+MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and Pt(P(t-Bu)(3))(2)+H-2 oxidative addition. *J. Phys. Chem.* **1996**, *100*, 19357-19363.
- [16] Humbel, S.; Sieber, S.; Morokuma, K. The IMOMO method: Integration of different levels of molecular orbital approximations for geometry optimization of large systems: Test for n-butane conformation and S<sub>N</sub>2 reaction: RCl+Cl<sup>-</sup>. *J. Chem. Phys.* **1996**, *105*, 1959-1967.
- [17] Dapprich, S.; Komaromi, K.; Byun, K.; Morokuma, K.; Frisch, M. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Mol. Str. (Theochem)* **1999**, *461*, 1-21.
- [18] Vreven, T.; Morokuma, K. On the application of the IMOMO (integrated molecular orbital plus molecular orbital) method. *J. Comp. Chem.* **2000**, *16*, 1419-1432.
- [19] Vreven, T.; Mennucci, B.; daSilva, C.; Morokuma, K.; Tomasi, J. The ONIOM-PCM method: Combining the hybrid molecular orbital method and the polarizable continuum model for solvation. Application to the geometry and properties of a merocyanine in solution. *J. Chem. Phys.* **2001**, *115*, 62-72.
- [20] Vreven, T.; Morokuma, K. Investigation of the S<sub>0</sub> → S<sub>1</sub> excitation in bacteriorhodopsin with the ONIOM(MO:MM) hybrid method. *Theo. Chem. Acc.* **2003**, *109*, 125-132.
- [21] Frisch, M. J. *et al.* . *Gaussian 03, Revision A.1* **2003**, Gaussian, Inc., Pittsburgh PA.
- [22] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
- [23] Rod, T. H.; Ryde, U. Quantum Mechanical Free Energy Barrier for an Enzymatic Reaction. *Phys. Rev. Lett.* **2005**, *94*, 138302.
- [24] Rod, T. H.; Ryde, U. Accurate QM/MM Free Energy Calculations of Enzyme Reactions: Methylation by Catechol O-methyltransferase. *J. Chem. Theor. Comput.* **2005**, *94*, 138302.