TR-WAXS reveals accelerated conformational changes in iodoretinal substituted proteorhodopsin

Erik Malmerberg, Ziad Omran, Jochen S. Hub, Xuewen Li, Gergely Katona , Sebastian Westenhoff, Linda C. Johansson, Magnus Andersson , Marco Cammarata, Michael Wulff, David van der Spoel, Jan Davidsson, Alexandre Specht, Richard Neutze

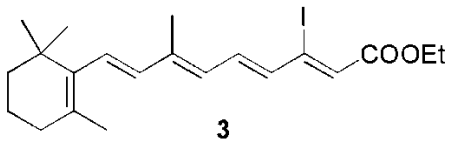
SUPPORTING MATERIAL

Supporting materials and methods:

Synthesis of 13-iodo-13-demethyl analogue of all-trans retinal:

13-Iodo-13-demethylretinal **4** (1) was obtained by the synthesis depicted in **Fig. S1**. Wittig reaction between phosphonium salt **1** (2) and aldehyde **2** (3) provides the ethyl ester **3** (4) with moderate yield. This ester **3** was then transformed to the aimed aldehyde **4** by DIBAL reduction followed by MnO_2 oxidation.

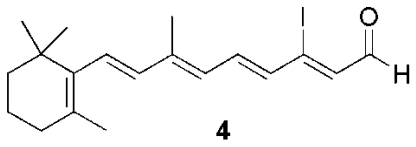
Ethyl (2*Z*,4*E*,6*E*,8*E*)-3-iodo-7-methyl-9-(2,6,6-trimethylcyclohex-1-en-1-yl)nona-2,4,6,8 tetraenoate **3**:



3

To a cooled (-78 °C) solution of β-ionylideneethyltriphenylphosphonium bromide **1** (1.08 g, 2 mmol) in THF (12 ml) was slowly added n-BuLi (2 ml, 1 M in THF, 2 mmol). After stirring for 30 minutes, a solution of aldehyde 2 (0.51 g, 2 mmol) in THF (8 mL) was added and the mixture was stirred at -78 °C for 1 h and at 25 °C for 3 h. 10% aqueous HCl was then added until neutral pH was reached, and the mixture was extracted with Et_2O (3 x 15 mL). The combined organic layers were washed with H_2O (3 x 10 mL) and brine (3 x 10 mL), dried (Na_2SO_4) and evaporated. Purification of the residue by flash chromatography using hexane as eluent afforded 0.22 g (25%) of 3. ^{1}H NMR (300 MHz, CDCl3) δ: 1.06 (s,6H), 1.28 (t, J = 7.1 Hz, 3H), 1.50 (m, 2H), 1.62 (m, 2H), 1.75 (s, 3H), 2.05 (s, 3H), 2.05 (m, 2H), 4.28 (q, J = 7.1 Hz, 2H), 6.15 (m, 2H), 6.36 (m, 2H), 6.42 (m, 1H), 7.29 (m, 1H). ^{13}C NMR (75 MHz, CDCl3) δ: 13.1, 14.7, 19.6, 22.2, 29.4, 33.5, 34.7, 40.0, 61.1, 106.4, 124.6, 126.9, 128.3, 129.9, 130.7, 133.9, 137.6, 138.0, 141.8, 165.0.

(2*Z*,4*E*,6*E*,8*E*)-3-Iodo-7-methyl-9-(2,6,6-trimethylcyclohex-1-en-1-yl)nona-2,4,6,8-tetraenal **4**:



4

To a cooled (0 °C) solution of ester **3** (0.22 g, 0.5 mmol) in THF (4ml) was slowly added DIBAL (1.10 ml, 1 M in hexane, 1.1 mmol). After stirring at 25 °C for 2 h, MeOH (2 mL) was added, followed by 5% aqueous HCl (2 mL), and stirring was continued for an additional 30 min. The layers were separated and the aqueous phase was extracted with Et_2O (3 x 20 mL). The combined organic layers were washed with H_2O (6 x 15 mL) and brine (4 x 15 mL), dried (Na_2SO_4), filtered and evaporated under reduced pressure to yield 0.187 g (94%) of the corresponding alcohol as yellow oil which was used in the next step without further purification. (0.1 g, 0.25 mmol) of the obtained alcohol was then dissolved in CH_2Cl_2 (10 ml) and (0.5 g, 5.7 mmol) of MnO_2 was added portionwise. After being stirred at 25 °C for 2 h, the mixture was filtered through a Celite pad, with thorough washing with CH_2Cl_2. The solvent was evaporated and the residue was purified by HPLC using an XTera C18 column and isocratic elution (0.1% TFA in water/ACN 3:7). 77 mg (81%) of the desired aldehyde 4 was thus obtained as yellow oil. Retention time: 64.6 min. ^{1}H NMR (300 MHz, CDCl3) δ: 1.06 (s,6H), 1.48 (m, 2H), 1.63 (m, 2H), 1.74 (s, 3H), 2.07 (m, 5H), 6.15 (d, J = 15.1 Hz, 1H), 6.21 (d, J = 15.1 Hz, 1H), 6.33 (m, 2H), 6.45 (d, J = 16.0 Hz, 1H), 7.49 (dd, J = 16.0, 11.8 Hz, 1H), 9.85 (d, J = 7.5, 1H). ^{13}C NMR (75 MHz, CDCl3) δ: 13.1, 19.6, 22.2, 29.3, 33.5, 34.7, 39.9, 111.3, 125.7, 130.7, 131.2, 133.1, 136.3, 136.4, 137.6, 138.0, 141.8, 190.6. MS (ESI): 397.1 $[M+H]^+$.

Protein expression, reconstitution and purification

SAR86 pR (GPR) with a C-terminal His_6 tag (engineered and used for Ni^{2+}-affinity column purification) was heterologously expressed in *E.coli* UT5600 strain using the pBAD-TOPO® plasmid.(5) The cells were grown in a 20L fermentor in TB-media at 37°C and 300 rpm. Expression of proteo-opsin (the apoprotein lacking the chromophore) was induced at OD_{600} 1.7-1.9 (~6 hours after 1:100 inoculation) with 1.25g/l L-arabinose (Merck) at 15°C and 200 rpm. Harvest followed ~17h after induction (16000 x g, 20 min, 4°C) and the resulting cell pellet was resuspended in minimal volume of 50mM MOPS pH 7.2. The cells were mixed with lysis buffer (50mM MOPS, pH 7.2, 1mM EDTA, 1 mg/ml Lysozyme, 12 µl benzonase (Merck), 3 tablets of Complete EDTA-free protease inhibitor (Roche)). Microliter quantities of either concentrated all-*trans* retinal (native pR) or 13-Iodo-13-demethylretinal (13-I-pR) dissolved in ethanol was added to the suspension to a final concentration of 12 µM. The solution was stirred at 4° C and ~15 min before the cells were disrupted by passing the suspension twice through a French Press (Thermo Fischer Scientific) at 20.000 Psi. Unlysed cells were spun down (12.000xg, 30 min) and the resulting supernatant was ultracentrifuged (150.000 x g, 2h) to recover the pelleted membrane fraction. The membranes were homogenized (50mM MOPS, pH 7.2, 1mM EDTA, 200mM NaCl) and then solubilized (50mM MOPS, pH 7.2, 1mM EDTA, 200mM NaCl, 6 % w/v β-OG) for 1.5h at 4°C. Unsolubilized debris was removed by ultracentrifugation (150.000 x g, 20 min). The supernatant was diluted 1:1 in 60mM MOPS, pH 7.2, 800mM NaCl and 12mM imidazole and applied onto Chelating Sepharose Fast Flow Gel (GE Healthcare) loaded with Ni^{2+} ions. The bound resin was washed (30mM MOPS, pH 7.2, 100mM NaCl, 40mM imidazole, 1% w/v β-OG) and subsequently eluted (same buffer solution containing 150mM imidazole) using an Äkta Explorer System (GE Healtcare). Desirable fractions were pooled, concentrated and diluted 1:50 (50mM KPi, pH 6.0, 1% w/v β-OG) and loaded onto a Mono S HR 10/10 (GE Healtcare) column for cation exchange chromatography. The resin was washed using the same buffer and the protein was eluted with a gradient

between pH 6.0 and pH 9.0 and the colored fractions were pooled and concentrated. Time-resolved WAXS data was collected on concentrated (15mg/ml protein, 25mM KPi, pH=9.0, 1% w/v β-OG) native pR or 13-I-pR. Under these conditions proteorhodopsins has a relative long half-degradation life, based on retinal ejection, of ~36h. This ensured that under the time course of each sample measurement (~2-3h) the sample suffered little degradation (<5%).

Singular value decomposition of the time-resolved difference WAXS data:

The number of signal-containing basis spectra in each of the datasets was determined based on inspection of the eigenvectors, the magnitude of the singular values and calculation of autocorrelation values of the basis spectra(6). The autocorrelation values were calculated according to;

$$C_i = \sum_{j=1}^{m-1} U_i(q_j)U_i(q_{j+1}) \quad \textbf{(1)}$$

Since the eigenvectors (U_i) are all normalized to unity, the vectors that display slow variations in q (i.e. "signal") will give values of C_i approaching 1 whereas vectors displaying fast variations in q (i.e. "noise") will result in autocorrelation values much less than 1, or possibly even negative values. The autocorrelations therefore provide a useful measure of the signal-to-noise ratio in the eigenvectors(6). The singular values and autocorrelation values in **Fig. S4** suggest that only the first singular vector, for each pR and 13-I-pR, contributes significantly to the signal in the experimental data. This is also evident from **Fig. S5**, illustrating the three most significant eigenvectors and amplitudes from SVD scaled according to their singular values. In an earlier study of native proteorhodopsin(7) two additional rapid states was observed to have formed and decayed on the ~10 μs time-scale. Due to a lower sampling frequency in this time regime in the current study such fast intermediates could not be distinguish using singular value decomposition.

In order to evaluate the similarity between the basis spectra for pR and 13-I-pR, the correlation coefficient between these two basis spectra was calculated according to,

$$R_{i,j} = \frac{cov(i,j)}{\sqrt{cov(i,i)cov(j,j)}} \quad \textbf{(2)}$$

where *i*, *j* represents the column vectors of the pR and 13-I-pR signal containing eigenvectors in the signal containing q-range (0.1-1.0 $Å^{-1}$).

Force field parameters and simulation details

In the absence of a crystallographic structure of pR a homology model of pR based on the bR structure was used (8). Interactions of the protein atoms were described by the AMBER99SB force field (9). Parameters for the detergent were taken from the general AMBER force field (GAFF) and generated through the Antechamber software (10-11). Partial charges for the detergent were obtained from electrostatic fitting of using the RESP procedure to wave functions calculated on the Hartree-Fock 6-31G* quantum level. Retinal parameters were taken from Hayashi *et al.* (12). Parameters of the 13-Iodo-13-demethylretinal were identical to the native retinal parameters, with the exception of the direct vicinity of the iodine. To detect the effect of the iodine on the partial charges, the charges for the iodo-retinal were derived on a B3LYP/gen quantum level, yielding the partial charge of the iodine to 0.044*e*, where *e* denotes the positive unit charge. In addition, partial charge of the carbon bound to the iodine (C13) was adapted from 0.0043*e* to 0.0919*e* to yield an overall charge of +1 for the 13-Iodo-13-demethylretinal. That procedure is justified because the quantum calculation of the 13-Iodo-13-demethylretinal yields very similar charges for the 13-Iodo-13-demethylretinal as compared to the topology by Hayashi *et al.(12)* The bond and angle parameters for the iodine were taken from the AMBER99SB force field (9), and dihedral parameters were taken from the native retinal.

For simulations in the 13-*cis* state of the retinal, the C13-C14 bond was stabilized by an additional Ryckhaert-Bellemans dihedral potential;

$$V_{rb}(\phi) = \sum_{n=0}^{5} C_n [\cos(\phi - 180)]^n \quad \textbf{(3)}$$

Here, ϕ is the dihedral definition according to the IUPAC/UIB convention. In addition, to avoid isomerization along the C15-N_ζ bond, that bond was stabilized in the *trans* state by a related additional dihedral potential. For the six parameters C_n to stabilize the *cis* and the *trans* angle we used (in kJ/mol) $C_{n,cis}$ = 50, 100, 50, 0, 0, 0 and $C_{n,trans}$ = 50, -100, 50, 0, 0, 0, respectively. These potentials have zero curvature in the energetic minimum in the *cis* and *trans* state, respectively, and are thus expected to have only a minor effect the dynamics in these dihedral states. Note that these dihedral potentials were only applied to stabilize a specific state and we do not aim to describe the energetics of the *trans* → *cis* isomerization.

All simulations were carried out using the Gromacs simulation software (13-14). Protons were added to the structure using the pdb2gmx software (14). Electrostatic interactions were calculated at every step with the particle-mesh Ewald method and a direct-space cut-off at 1 nm (15-16). Short-range repulsive and attractive dispersion interactions were described by a Lennard-Jones potential, which was cut off at 1.0 nm. The LINCS algorithm was used to constrain all bond lengths (17), allowing a time step of 2~fs. The simulation temperature was kept constant at 300 K through a stochastic dynamics integrator (τ=0.1ps) (18).

Free energy calculations

Free-energy differences were computed for transforming retinal into the iodo-retinal, with the retinal either in the all-*trans* or in the 13-*cis* state. Each of these transformations was conducted for the

respective retinal conformation inside pR or in vacuum. In unrestrained simulations of the homology model of pR, the protein was rather unstable on a time-scale of tens of nanosecond, as quantified from the RMSD of backbone atoms. Therefore, we performed all free-energy calculations with position restraints on the backbone atoms of pR, keeping the protein structure close to the original pR model. To exclude the possibility that a specific choice for the applied position restraining the force constant f_c would bias the qualitative results, we have conducted the free-energy calculations using a wide range of f_c values between 200 and 1000 kJ mol^{-1} nm^{-2}. Free-energy calculations with pR were conducted with the protein in vacuum, which can be justified by the required position restraints and by the fact that the retinal interacts only with atoms inside the protein matrix.

Free-energy differences were computed using thermodynamic integration (TI). Accordingly, the transformation from the native to the 13-iodoretinal was conducted along an alchemical reaction coordinate λ, where λ =0 corresponds to the native, and λ =1 to the 13-iodoretinal. The transformation was split into 11 equally spaced λ-values between 0 and 1. For each λ-value, the system was equilibrated for 200 ps and simulated for at least another 300 ps. The free-energy difference was subsequently computed by integrating $\langle \partial H/\partial \lambda \rangle$ from λ =0 to λ =1, where <…> denotes the average over the simulation. Statistical errors were computed by binning analysis and were found to be ≤ 1 kJ/mol.

Structural Refinement

The selection of rigid body movements was based on information from trapped intermediates of the homologous bacteriorhodopsin, as previously described (7). The cytoplasmic halves of helix E and F together (residues 157-201), and of the terminal portion of helix G (residues 215-240), were generated based on the observed movements between the bacteriorhodopsin D96G, F171C, F219L triple mutant structure (PDB ID 1FBK) and its corresponding resting state (PDB ID 1FBB) (19). In a similar fashion,

a movement of the extracellular portion of helix C (residues 91-101) was extracted from a trapped structure of the L-intermediate (PDB ID 1VJM) (20). Additional movements of the entire helix A (residues 29-49); the entire helix B (residues 55-82); the cytoplasmic half of helix C (residues 104-113); and all of helix D (residues 120-142) were also introduced into the structural refinement procedure. These movements were selected to avoid α-helices clashing into each other, and the points of helix inflexion were selected from the distribution of proline residues within the transmembrane portions of the proteorhodopsin homology model (extracellular helix C: Pro90, cytoplasmic helix C: Pro103), which typically mark the boundaries of flexible regions. For each rigid body, an atomic position difference vectors was calculated between the rigid body and the resting state. From this set of 7 difference vectors, a grid of PDB files were created by adding different combinations of these movements, with varying magnitudes, to the resting state. For each matrix element (candidate transient state conformation) in the grid, the WAXS difference scattering was predicted relative to the resting state using CRYSOL. These predicted spectra were then compared with the experimentally determined WAXS difference data and the transient state structure corresponding to the optimal fit was used as starting point for molecular dynamics refinement in the 125 β-OG micell system. It should be noted that control simulations of bacteriorhodopsin, without position restraints, in the same size and type of detergent micell (125 β-OG) using the same parameterization yielded a stable bR conformation (<1.5 Å rmsd deviation of backbone atoms) on the tens of nanosecond scale (data to be published elsewhere).

Supporting tables:

Sample	pR	13-I-pR
PR (500 nm) decay	66 ms	3.3 ms
O (610 nm) decay	73 ms	5.4 ms
M (410 nm) rise*	27μs	N/A
M (410 nm) decay*	32 μs	N/A
WAXS rise	2 ± 1 μs	N/A
WAXS decay	49 ± 11 ms	2.0 ± 0.4ms

***Exponential fitting values for the rise and decay of the positive absorbance change.**

Table S1. Spectroscopic and WAXS half-times of native pR and 13-I pR at pH 9.0 (23°C).

Supporting figures:

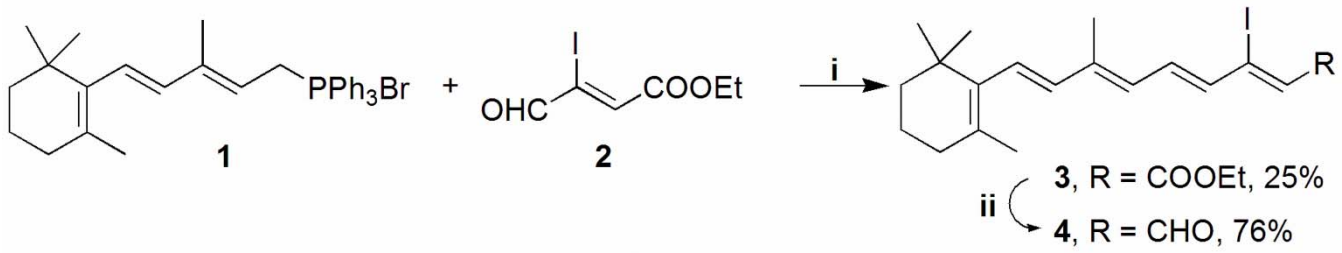


Figure S1. Chemical synthesis of 13-demethyl-13-iodoretinal. (**i**) *n*BuLi, THF, -20 °C to RT, 4h, (**ii**) a- DIBAL, CH_2Cl_2, -78 °C, 2h; b- MnO_2, Hexane, RT, 2h.

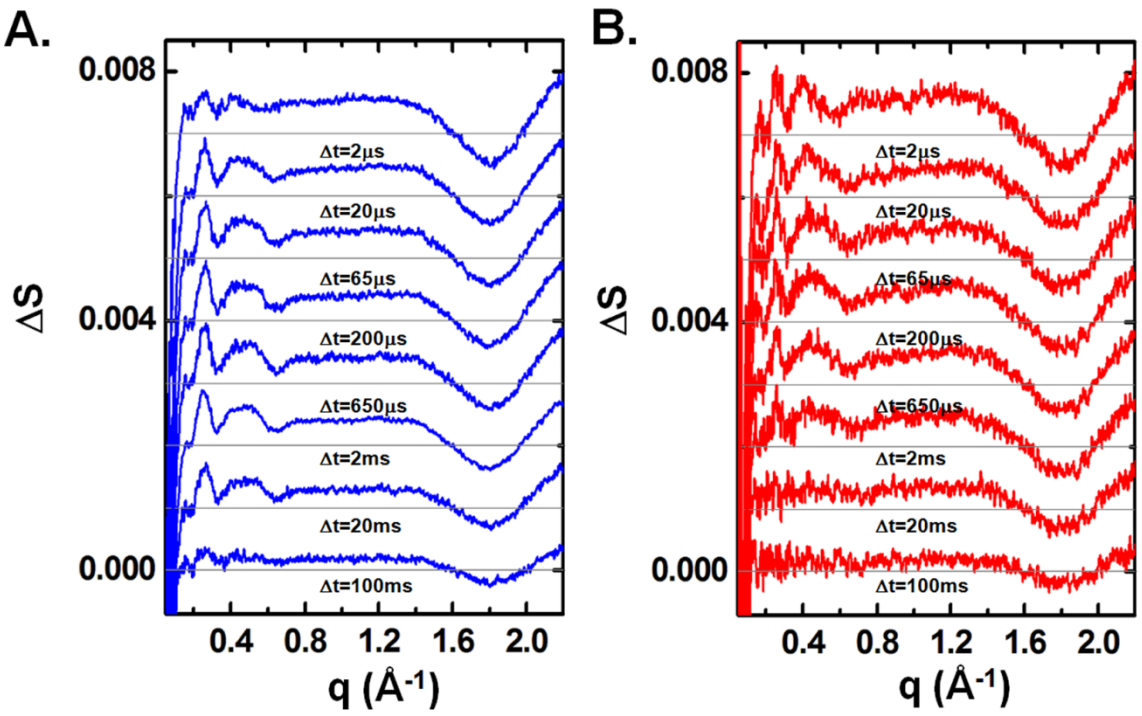


Figure S2. Time-resolved difference WAXS data collected at pH 9.0 (23 °C) before removal of the laser induced thermal signal. (A) native pR and (B) 13-I-pR for the indicated time delays following laser excitation. The gray line represents $\Delta S(q,\Delta t) = 0$.

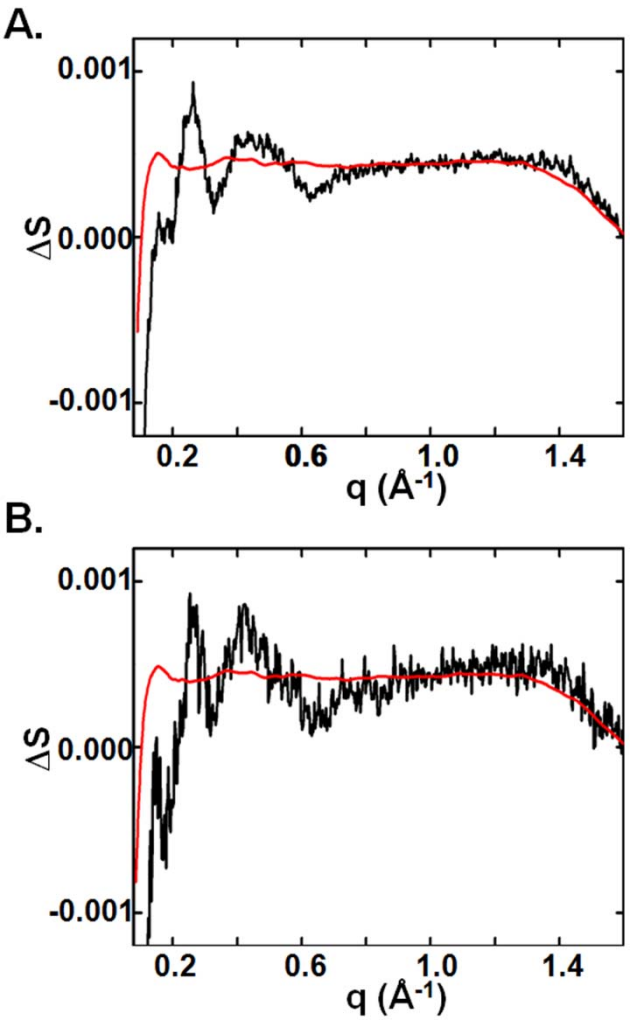


Figure S3. Correction of TR-WAXS difference data due to laser induced heating. Δt = 20μs (black) and

experimental heating signal (red) for native pR (A) and 13-I-pR (B). The thermal signal was filtered, to lower noise, and scaled to match the solvent induced heating ($q > 0.8$) in the time-resolved WAXS difference data and then subtracted to yield the thermal free signal. The same procedure was iterated for all time-delays.

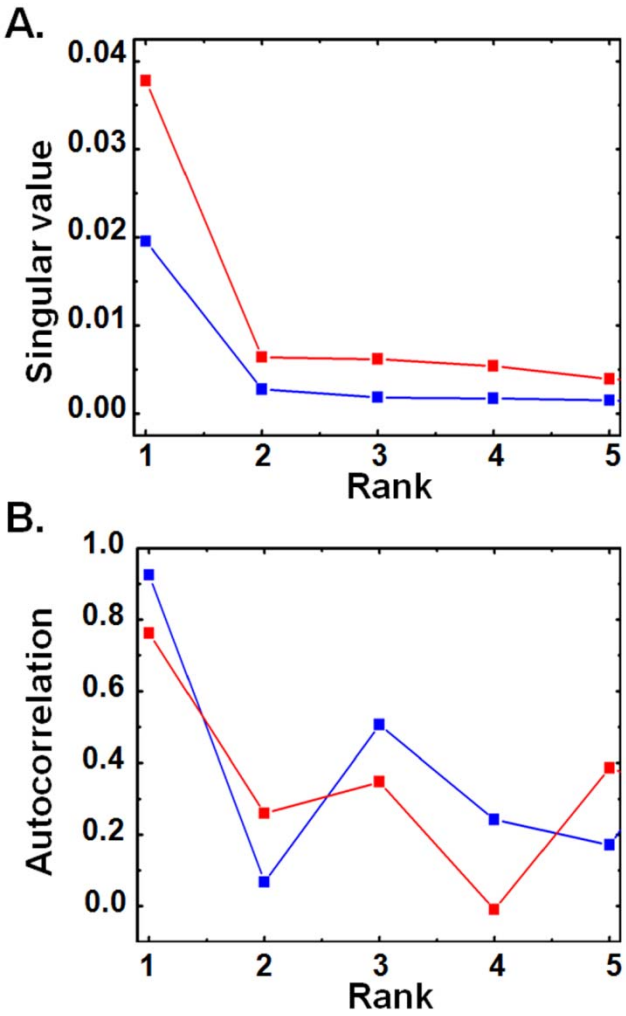


Figure S4. Singular values and eigenvector autocorrelation values from singular value decomposition of difference WAXS data. (A) Singular values for the first five components (s_1-s_5) from SVD. (B) Autocorrelation values (c_1-c_5) of the first five eigenvectors (U_1-U_5). pR (blue) and 13-I-pR (red).

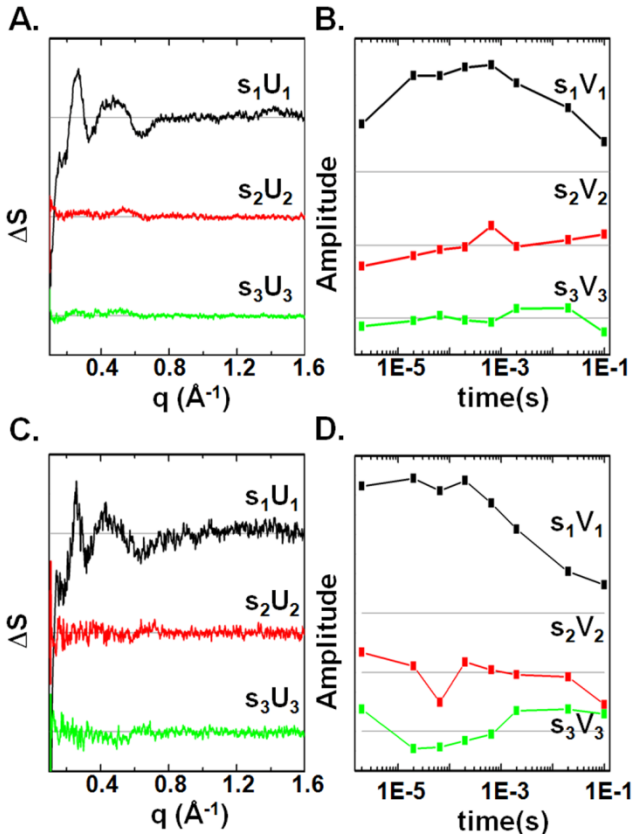


Figure S5. Eigenvectors and eigenvector amplitudes from SVD on difference WAXS data of WT-pR and 13-I-pR. (A/C) The first three eigenvectors scaled by their singular value for pR (A) and 13-I-pR (C). (B/D) The corresponding eigenvector amplitudes scaled by their singular value for pR (B) and 13-I-pR (B). The grey line indicates amplitude of zero.

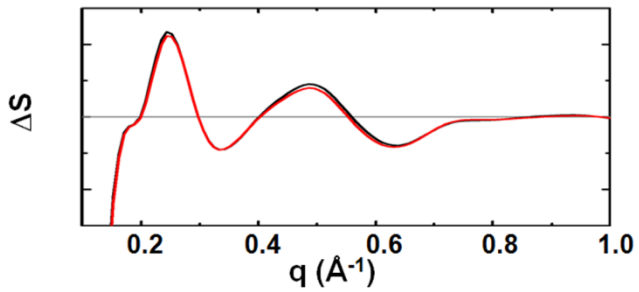


Figure S6. Theoretical effect of -CH_3 to I substitution on the WAXS difference scattering. Theoretical difference scattering between the resting state and the optimal transient conformational state for the native pR model (black line) and for the case were the retinal C20 methyl group was replaced for an iodine in the same models (red line).

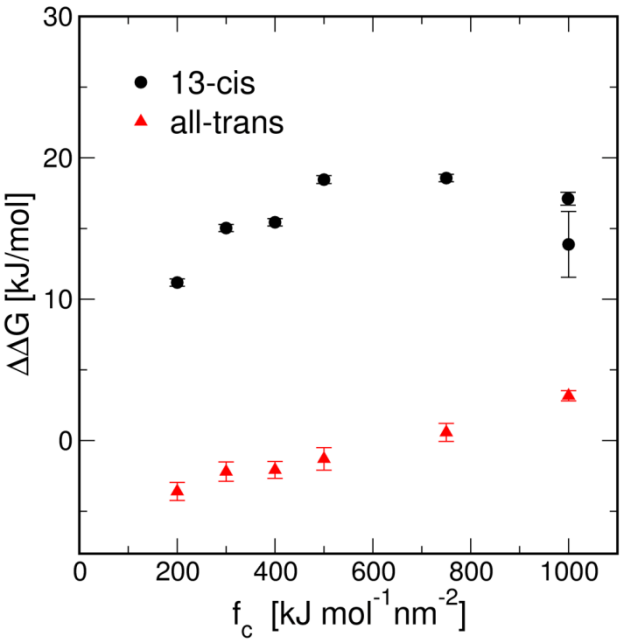


Figure S7. Destabilization of the protein in the resting state conformation by transforming the wild type retinal into 13-iodoretinal, either in the 13-*cis* state (black dots) or in the all-*trans* (red triangles). In all-*trans*, replacement of the C-20 methyl group by iodine has little influence on the free energy. In 13-*cis*, however, the iodine clashes into Leu105 and Trp197, leading to an unfavorable increased free energy of pR in that conformation. The destabilization was quantified from the free-energy difference $\Delta G^{(pR)}$ for the transformation inside proteorhodopsin (pR), minus the respective free-energy difference for the retinal in vacuum $\Delta G^{(vac)}$. That difference $\Delta\Delta G = \Delta G^{(pR)} - \Delta G^{(vac)}$ thus quantifies the impact of the iodine on the surrounding protein matrix. Here, $\Delta\Delta G$ is plotted versus the applied position restraint force constant f_c, confirming that the qualitative result of the free-energy calculations is unaffected by the specific choice for f_c. The calculation for f_c=1000 kJ mol^{-1}nm^{-2} was carried out twice and both results are included in the plot.

1. Hiraki, K., T. Hamanaka, X. G. Zheng, T. Shinada, J. M. Kim, K. Yoshihara, and Y. Kito. 2002. Bacteriorhodopsin analog regenerated with 13-desmethyl-13-iodoretinal. *Biophys. J.* 83:3460-3469.

2. Curley, R. W., Jr.; DeLuca, H.F. J. . 1984. Ethyl 13,14-dihydro-13,14-methyleneretinoates: analogs of all-trans- and 13-cis-retinoic acid. J. Org. Chem. 49:1941.

3. Shinada, T., and K. Yoshihara. 1996. Stereoselective functionalization at C-9 of retinae: Synthesis of 9-trans-19-nor-9-haloretinal analogs. Chem Pharm Bull 44:264-266.

4. Li, H. B., and C. Morin. 2004. Synthesis of iodinated analogues of all trans retinoic acid (ATRA) for SPECT imaging. Tetrahedron Lett 45:5673-5676.

5. Gourdon, P., A. Alfredsson, A. Pedersen, E. Malmerberg, M. Nyblom, M. Widell, R. Berntsson, J. Pinhassi, M. Braiman, O. Hansson, N. Bonander, G. Karlsson, and R. Neutze. 2008. Optimized in vitro and in vivo expression of proteorhodopsin: a seven-transmembrane proton pump. Protein Expr Purif 58:103-113.

6. Henry, E. R., and J. Hofrichter. 1992. Singular Value Decomposition: Application and Analysis of Experimental Data. Methods in Enzymologi 210:129-192.

7. Andersson, M., E. Malmerberg, S. Westenhoff, G. Katona, M. Cammarata, A. B. Wohri, L. C. Johansson, F. Ewald, M. Eklund, M. Wulff, J. Davidsson, and R. Neutze. 2009. Structural Dynamics of Light-Driven Proton Pumps. Structure 17:1265-1275.

8. Rangarajan, R., J. F. Galan, G. Whited, and R. R. Birge. 2007. Mechanism of spectral tuning in green-absorbing proteorhodopsin. Biochemistry-Us 46:12679-12686.

9. Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling. 2006. Proteins 65:712-725.

10. Wang, J., R. M. Wolf, J. M. Caldwell, P. A. Kollman, and D. A. Case. 2004. Development and

testing of a general amber force field. J Comput Chem 25.

11. Wang, J., W. Wang, P. A. Kollman, and D. A. Case. 2006. Automatic atom type and bond type perception in molecular mechanical calculations. Mol Graphics Modell 26:247260.

12. Hayashi, S., E. Tajkhorshid, and K. Schulten. 2002. Structural changes during the formation of early intermediates in the bacteriorhodopsin photocycle. Biophys J 83:1281-1297.

13. van der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen. 2005. GROMACS: Fast, flexible and free. J Comp Chem 26:701-1719.

14. Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl. 2008. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4:435-447.

15. Darden, T., D. York, and L. Pedersen. 1993. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. J Chem Phys 98:10089-10092.

16. Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh ewald potential. J Chem Phys 103:8577-8592.

17. Hess, B. 2008. P-LINCS: A parallel linear constraint solver for molecular simulation. J Chem Theory Comput 4:116-122.

18. van Gunsteren, W. F., and H. J. C. Berendsen. 1988. A leap-frog algorithm for stochastic dynamics. Mol Sim 1:173-185.

19. Subramaniam, S., and R. & Henderson. 2000. Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature* 406:653-657.

20. Edman, K., A. Royant, G. Larsson, F. Jacobson, T. Taylor, D. van der Spoel, E. M. Landau, E. Pebay-Peyroula, and R. & Neutze. 2004. Deformation of helix C in the low temperature L-intermediate of bacteriorhodopsin. *J. Biol. Chem.* 279:2147-2158.