Supporting Information

Color Tuning in Rhodopsins: The Origin of the Spectral Shift between the Chloride-bound and Anion-free Forms of Halorhodopsin

Mikhail N. Ryazantsev,[1] Ahmet Altun,[2] Keiji Morokuma[1,3]*

[1] Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322

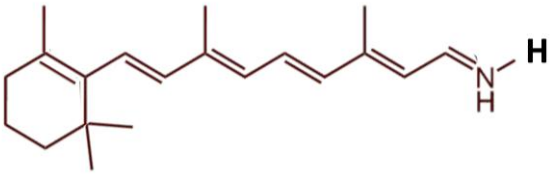
[2] Department of Physics, Fatih University, 34500 B. Cekmece, Istanbul, Turkey

[3] Fukui Institute for Fundamental Chemistry, Kyoto University, 34-4 Takano Nishihiraki-cho, Kyoto 606-8103, Japan

Table of Contents

I. Computational details. (p. 2)
II. The impact of the polar residues situated in the 5.0 Å region from the chromophore on λ_{max} (p. 3)
III. Geometric parameters of the retinal for the models described in the main text. (p. 4)
IV. Cartesian coordinates of the retinal for the models described in the main text (in Å) (p. 5-9)
IV. Geometrical parameters for the $^{+}$N-H region of PSBT (p. 10)
V. H-bond network in the $^{+}$N-H region of PSBT (p.11)
VI. References (p.12)

I. Computational details.



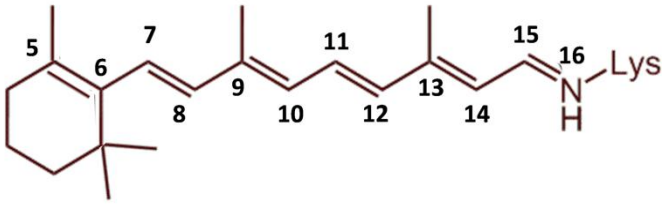
To construct the models we started from the 2.0 Å X-ray structure (PDB code 3A7K)[1] for the chloride-bound form (phR-Cl) and the 1.80 Å structure (3QBG) for the anion-free form (phR-af).[2] PROPKA[3] and PDB2PQR[4] programs in conjunction with visual inspection are used to assign protonation states of the titratable residues (pH=7) and add hydrogen atoms. Based on this information standard protonation states for titrable residues were assumed. The buried charged residues within 15 Å from the chromophore, that affect the spectral properties considerably, are Asp252(-1), Asp156(-1), Glu234(-1), Arg123(+1), Arg176(+1). The resulting structure is then optimized using two-layer ONIOM (QM:MM-EE) scheme[5] (QM=B3LYP/6-31G*; MM=AMBER96 for aminoacids and Cl^- and TIP3P for water, EE=electronic embedding)[6] implemented in the Gaussian09 package.[7] The QM part comprises the full all-trans-retinal chromophore along with covalently bound NH moiety of Lys256 for SBT linkage, resulting in total of 51 atoms including the link atom (see Figure above). The rest of the protein is treated at the MM level. To calculate the spectral properties of the chromophore in the presence and absence of the protein environment (described as AMBER96 point charges), we employ SORCI+Q/6-31G(d) level of the theory implemented in ORCA6.0 package.[8] We used equally-weighted 3 root complete active space self-consistent field (CASSCF) wavefunction as the reference wavefunction for the subsequent spectroscopy oriented configuration interaction (SORCI) calculations with the Davidson correction (+Q) for excitations higher than doubles (SORCI+Q). CASSCF active space includes 6 π-orbitals and 6-electrons. Computational efficiency of SORCI+Q calculations was enhanced by setting T_{pre}, T_{nat} and T_{sel} thresholds to 10-4, 10-6 and 10-6 *E*h, respectively. The basis set 6-31G* was used in all calculations. For a more detailed description and justifications of the employed methodology see our previous studies.[9-11]

II. The impact of the polar residues situated in the 5.0 Å region from the chromophore on λ_{max}

	*$\Delta\lambda_{max}$ (nm)	
	phR-Cl	**phR-af**
SER78	+6	0
TYR124	-2	-2
THR126	+3	-4
THR127	+6	+5
SER130	+6	-4
THR131	+1	+3
CYS160	+9	+5
TYR180	+8	+9
SER183	+5	+8
CYS184	+3	+1
THR222	+8	+3
TYR225	-4	-2
TYR257	+4	0
TYR82	+8	+6

*the influence of a reside was estimated by setting the charges of this residue to zero

III. Geometric parameters of the retinal for the models described in the main text.



.

Bond (in Å)	phR-Cl	phR-af	*gas phase	**phR-Cl-me	**phR-af-me
C5=C6	1.371	1.373	1.381	1.380	1.380
C6-C7	1.462	1.455	1.444	1.447	1.443
C7=C8	1.366	1.365	1.376	1.376	1.374
C8-C9	1.443	1.437	1.428	1.427	1.422
C9=C10	1.384	1.384	1.398	1.399	1.397
C10-C11	1.416	1.408	1.401	1.399	1.396
C11=C12	1.382	1.382	1.395	1.397	1.393
C12-C13	1.417	1.407	1.404	1.401	1.398
C13=C14	1.401	1.403	1.415	1.417	1.413
C14-C15	1.388	1.381	1.380	1.377	1.377
C15=N16	1.333	1.336	1.340	1.345	1.341
#BLA	**0.064**	**0.045**	**0.027**	**0.024**	**0.026**

* **gas phase** model – geometry optimized at B3LYP/6-31G* level with no external charges

***phR-Cl-me and phR-af-me** – models with mechanical embedding scheme (see the main text for the explanation)

#BLA is defined as the difference between the average of the single bonds and the average of the double bonds.

Dihedral	phR-Cl	phR-af	Gas-phase
C5=C6-C7=C8	167.4	168.7.	170.7
C6-C7=C8-C9	178.4	176.5	-179.2
C7=C8-C9=C10	-176.5	-176.0	179.8
C8-C9=C10-C11	177.8	-178.4	179.0
C9=C10-C11=C12	-179.6	-173.81	-179.8
C10-C11=C12-C13	-179.8	-174.2	-179.7
C11=C12-C13=C14	176.3	-177.0	-179.8
C12-C13=C14-C15	-172.3	-161.0	179.8
C13=C14-C15=N	176.5	-173.2	-179.9

IV. Optimized cartesian coordinates of the retinal for the models described in the main text (in Å)

phR-Cl

N	5.718257	0.856296	22.978758
H	6.128327	-0.047587	22.735692
C	10.055235	-2.964991	33.511976
C	10.151096	-4.329482	32.782663
C	11.237864	-4.193870	31.688187
H	12.186413	-3.887104	32.139182
H	11.371053	-5.150922	31.174466
H	10.990367	-3.435631	30.942118
C	8.815719	-4.803616	32.167942
H	8.970673	-5.786821	31.708192
H	8.056399	-4.932145	32.938530
H	8.402158	-4.134569	31.415625
C	10.614928	-2.801213	34.753376
C	10.720668	-1.495664	35.501407
H	11.776626	-1.196390	35.527701
H	10.152094	-0.661704	35.092268
H	10.408832	-1.640326	36.546673
C	10.559399	-5.478657	33.742231
H	9.691614	-5.757569	34.357996
H	10.803140	-6.358277	33.134520
C	11.705250	-5.103302	34.666029
H	11.992969	-5.946296	35.299055
H	12.585443	-4.834180	34.076284
C	11.265589	-3.928082	35.530191
H	10.556658	-4.292712	36.292381
H	12.100155	-3.522036	36.109763
C	9.469060	-1.805952	32.840290
H	9.334768	-0.929533	33.464878
C	9.072002	-1.662728	31.540752
H	9.203558	-2.478260	30.842516
C	8.468181	-0.488918	30.955640
C	8.275492	0.755843	31.782365
H	9.235426	1.106444	32.176804
H	7.835821	1.573254	31.208636
H	7.627798	0.554480	32.643098
C	8.076445	-0.595770	29.632458
H	8.271871	-1.548369	29.144200
C	7.427026	0.370498	28.827252
H	7.191175	1.340809	29.255069
C	7.092993	0.103852	27.512929
H	7.361451	-0.874376	27.125493
C	6.448286	0.979637	26.603934
C	5.953812	2.322413	27.069128
H	5.840034	3.042159	26.256490
H	4.974688	2.202343	27.546494
H	6.630704	2.758732	27.803549
C	6.273577	0.533500	25.287870
H	6.572825	-0.482215	25.055634
C	5.805163	1.301792	24.231666
H	5.490586	2.328299	24.392054
H	5.653285	1.459183	22.147599

phR-af

N	5.422380	0.716952	23.891999

H	4.850328	-0.125140	24.038371
C	10.274511	-3.231933	33.984839
C	10.319603	-4.550691	33.170681
C	11.405657	-4.394299	32.078551
H	12.380081	-4.172709	32.522872
H	11.484284	-5.328148	31.513035
H	11.185928	-3.589642	31.372851
C	8.964385	-4.936040	32.529993
H	9.054324	-5.931505	32.081941
H	8.177164	-4.996703	33.286112
H	8.619893	-4.261489	31.747190
C	10.834578	-3.156035	35.235612
C	10.966170	-1.892151	36.046938
H	10.699253	-2.095635	37.091747
H	12.017650	-1.578500	36.041471
H	10.373163	-1.048317	35.698721
C	10.693533	-5.761866	34.064851
H	9.814562	-6.057472	34.654918
H	10.919151	-6.615844	33.415422
C	11.844986	-5.473585	35.015402
H	12.109143	-6.369665	35.581953
H	12.738228	-5.185106	34.453301
C	11.426366	-4.347771	35.954159
H	10.674404	-4.730253	36.664258
H	12.249148	-4.013119	36.592876
C	9.705179	-2.027166	33.401524
H	9.583737	-1.188144	34.077201
C	9.291359	-1.805785	32.119492
H	9.417836	-2.576659	31.371791
C	8.670871	-0.608371	31.624750
C	8.425503	0.569509	32.528163
H	7.811659	0.269881	33.382991
H	9.357867	0.968781	32.933872
H	7.902358	1.377704	32.015457
C	8.259409	-0.648132	30.303837
H	8.406628	-1.589716	29.777085
C	7.655877	0.372635	29.543994
H	7.566519	1.373546	29.957473
C	7.193724	0.110276	28.267737
H	7.250737	-0.919959	27.926495
C	6.675975	1.043489	27.350574
C	6.561702	2.496992	27.700499
H	5.647977	2.640163	28.291170
H	7.404916	2.844069	28.298657
H	6.465822	3.136372	26.819177
C	6.203866	0.544286	26.127216
H	6.003414	-0.521196	26.074000
C	6.002732	1.264557	24.963269
H	6.376830	2.279575	24.852913
H	5.415402	1.133924	22.956049

gas phase

N	44.683745	10.247022	25.688924
H	44.955420	10.450362	26.642134
H	45.383736	9.831312	25.092057
C	33.929292	14.765482	29.107770
C	34.437265	14.834626	30.569772
C	35.606867	15.842298	30.698909
H	35.989409	15.838774	31.726097
H	35.283775	16.861678	30.465030
H	36.442706	15.610842	30.033662
C	34.882416	13.436100	31.068021

H	34.091629	12.695038	30.906665
H	35.079653	13.483978	32.145069
H	35.790109	13.063068	30.589647
C	32.697081	15.271391	28.743398
C	32.086759	15.159717	27.362286
H	32.292456	14.213766	26.855584
H	32.436285	15.973782	26.712267
H	30.999247	15.261166	27.431478
C	33.298496	15.276448	31.525751
H	32.617684	14.428884	31.688496
H	33.740263	15.509974	32.501606
C	32.489460	16.455091	30.995851
H	31.735534	16.768424	31.726020
H	33.136166	17.323912	30.822906
C	31.814429	16.040476	29.692532
H	30.933640	15.410454	29.904086
H	31.410126	16.911109	29.156097
C	34.730743	14.191746	28.052192
H	34.226480	14.114820	27.095934
C	36.040388	13.769679	28.053018
H	36.634863	13.842328	28.954800
C	36.714654	13.228580	26.916511
C	35.987090	13.076874	25.603124
H	35.124447	12.410041	25.713695
H	36.616225	12.671806	24.810667
H	35.608013	14.045274	25.257800
C	38.049195	12.848756	27.084071
H	38.477223	13.005515	28.073198
C	38.896271	12.282729	26.122734
H	38.503280	12.105612	25.127041
C	40.219306	11.933945	26.392686
H	40.577323	12.126650	27.403069
C	41.141050	11.358894	25.504126
C	40.787751	11.032672	24.071802
H	41.453774	11.557976	23.377939
H	39.766600	11.310051	23.813177
H	40.892833	9.958296	23.881493
C	42.440677	11.080907	25.989937
H	42.655202	11.329920	27.027726
C	43.450904	10.515657	25.238705
H	43.287742	10.248761	24.198817

phR-Cl-me

N	5.627370	0.860734	22.990689
H	5.989328	-0.061170	22.748025
C	10.076534	-2.955737	33.507923
C	10.159883	-4.313623	32.764378
C	11.237516	-4.174052	31.663128
H	12.196614	-3.879950	32.098517
H	11.371577	-5.125366	31.140901
H	10.989389	-3.417036	30.916204
C	8.806371	-4.765351	32.172687
H	8.917764	-5.766877	31.742678
H	8.053156	-4.839045	32.952696
H	8.401794	-4.112907	31.401576
C	10.626298	-2.800407	34.764609
C	10.696267	-1.515781	35.545901
H	11.742246	-1.184226	35.589967
H	10.103665	-0.688782	35.158320
H	10.389981	-1.695103	36.583280
C	10.566992	-5.470045	33.716509
H	9.697911	-5.756892	34.323824

H	10.816180	-6.343860	33.102701
C	11.714558	-5.107824	34.643159
H	12.004857	-5.957728	35.264839
H	12.600059	-4.846157	34.056740
C	11.271953	-3.940739	35.519720
H	10.541480	-4.305913	36.260494
H	12.093435	-3.541853	36.124080
C	9.501437	-1.795290	32.861698
H	9.381014	-0.920859	33.490848
C	9.088212	-1.645542	31.557339
H	9.215249	-2.461770	30.860333
C	8.481656	-0.486157	30.988249
C	8.277654	0.768427	31.796538
H	9.241728	1.239177	32.020501
H	7.662681	1.505378	31.279171
H	7.798069	0.547053	32.754624
C	8.094398	-0.592811	29.648192
H	8.300068	-1.547067	29.167278
C	7.451218	0.357491	28.847404
H	7.216258	1.331077	29.265422
C	7.110198	0.088884	27.519749
H	7.382712	-0.893207	27.140511
C	6.465434	0.940466	26.612897
C	5.982820	2.308156	27.016761
H	6.510762	3.096434	26.470576
H	4.914952	2.416423	26.803407
H	6.125208	2.487670	28.079923
C	6.267791	0.484128	25.285863
H	6.548365	-0.537558	25.047043
C	5.787664	1.269934	24.261535
H	5.503103	2.298347	24.454653
H	5.589513	1.480113	22.171023

phR-af-me

N	5.425528	0.774867	23.875210
H	4.920368	-0.108583	23.967820
C	10.275906	-3.235536	33.992855
C	10.319726	-4.549270	33.170893
C	11.402486	-4.395175	32.075871
H	12.384562	-4.199937	32.516114
H	11.468204	-5.317816	31.494831
H	11.198232	-3.576080	31.382245
C	8.955290	-4.925683	32.543212
H	9.039625	-5.916325	32.083956
H	8.178015	-4.982601	33.308937
H	8.601839	-4.245605	31.769159
C	10.816227	-3.164165	35.260773
C	10.909799	-1.918349	36.101281
H	10.654147	-2.150061	37.140403
H	11.947941	-1.561392	36.108606
H	10.280607	-1.090062	35.782708
C	10.690141	-5.765900	34.060447
H	9.807607	-6.070157	34.640108
H	10.933229	-6.606878	33.400785
C	11.833864	-5.484760	35.021683
H	12.098860	-6.381845	35.585213
H	12.734068	-5.202990	34.466987
C	11.403997	-4.364344	35.963686
H	10.632788	-4.745050	36.654038
H	12.214591	-4.032866	36.620856
C	9.729894	-2.026876	33.423032

H	9.630370	-1.190340	34.102937
C	9.310349	-1.800649	32.134215
H	9.422021	-2.585194	31.399851
C	8.711012	-0.618171	31.619165
C	8.489976	0.630079	32.434490
H	8.721178	0.497819	33.491664
H	9.111478	1.450196	32.059540
H	7.450211	0.958629	32.358278
C	8.290395	-0.684550	30.289149
H	8.419284	-1.639537	29.783094
C	7.694961	0.329550	29.536596
H	7.601994	1.326240	29.960348
C	7.220029	0.085959	28.249732
H	7.276584	-0.942109	27.898056
C	6.679851	1.018298	27.359280
C	6.538461	2.464153	27.735722
H	5.595455	2.595045	28.281585
H	7.353143	2.795914	28.380568
H	6.491759	3.127685	26.870003
C	6.193705	0.537289	26.122531
H	6.019631	-0.533429	26.046381
C	5.974185	1.285178	24.987234
H	6.296623	2.320234	24.926561
H	5.410166	1.217940	22.951513

IV. Geometrical parameters for the $^{+}$N-H region

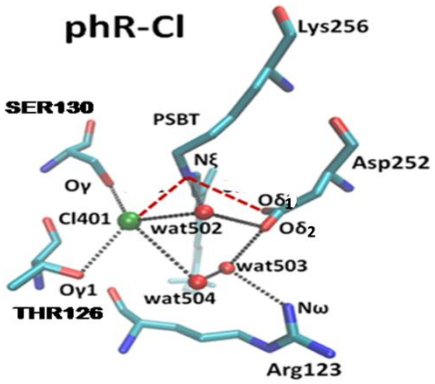
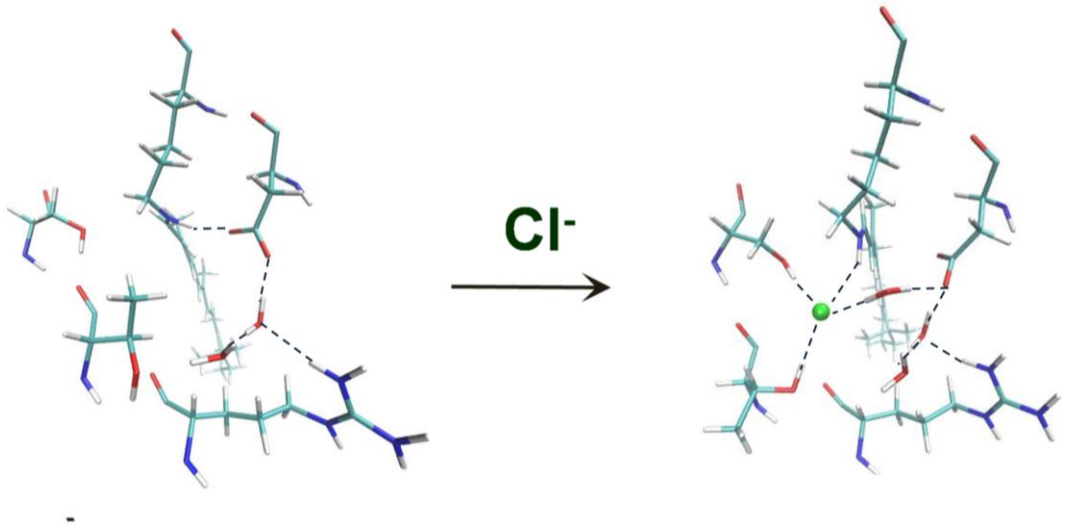


Table 1. Selected Geometrical Parameters of the Retinal Binding Site for QM/MM Optimized Structure (the present study) and X-ray structure[a]

	phR-Cl			phR-Cl	
	present work	X-ray structure[b] (chain A, B, D)		present work	X-ray [b] (chain A, B, D)
N_{ξ}-$O_{\delta 1}$	3.43	3.62, 3.71, 3.15	Cl401-O_{γ}	3.21	3.08, 3.09, 3.10
N_{ξ}-Cl401	3.37	3.46, 3.38, 3.41	Cl401-$O_{\gamma 1}$	3.23	3.71, 3.07, 2.92
$O_{\delta 2}$-wat503	2.68	2.91, 2.83, 2.89	Cl401-wat504	3.85	4.37, 4.50, 4.22
wat503-wat504	2.70	2.71, 2.79, 2.75	Cl401-wat502	3.13	3.30, 2.94, 2.97
wat503-N_{ω}	2.88	3.14, 2.71, 3.12			

[a]Distances in Å, residues and atoms numbers specified in Figure 1. [b]PDB code 3A7K.

V. H-bond network in the $^{+}$N-H region of PSBT



VI. References

(1) Kouyama, T.; Kanada, S.; Takeguchi, Y.; Narusawa, A.; Murakami, M.; Ihara, K. *J. Mol. Biol.* **2010**, *396*, 564.

(2) Kanada, S.; Takeguchi, Y.; Murakami, M.; Ihara, K.; Kouyama, T. *J. Mol. Biol.* **2011**, *413*, 162.

(3) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704.

(4) Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. A. *Nucl. Ac. Res.* **2007**, *35*, W522.

(5) Vreven, T.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760.

(6) 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. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(7) Gaussian 03, R. C., M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.

(8) Neese, F. *J. Chem. Phys.* **2003**, *119*, 9428.

(9) Altun, A.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* **2008**, *112*, 16883.

(10) Altun, A.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* **2008**, *112*, 6814.

(11) Altun, A.; Yokoyama, S.; Morokuma, K. *Photochemistry and Photobiology* **2008**, *84*, 845.