Microbial and Animal Rhodopsins: Structures, Functions and Molecular Mechanisms

Oliver P. Ernst[1], David T. Lodowski[2], Marcus Elstner[3], Peter Hegemann[4], Leonid S. Brown[5], and Hideki Kandori[6]

[1]Departments of Biochemistry and Molecular Genetics, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada, [2]Center for Proteomics and Bioinformatics, Case Western Reserve University School of Medicine, 10900 Euclid Avenue, Cleveland, Ohio 44106, United States, [3]Institute for Physical Chemistry, Karlsruhe Institute of Technology, Kaiserstrasse 12, 76131 Karlsruhe, Germany, [4]Institute of Biology, Experimental Biophysics, Humboldt-Universität zu Berlin, Invalidenstrasse 42, 10115 Berlin, Germany, [5]Department of Physics and Biophysics Interdepartmental Group, University of Guelph, 50 Stone Road East, Guelph, Ontario N1G 2W1, Canada, [6]Department of Frontier Materials, Nagoya Institute of Technology, Showa-ku, Nagoya 466-8555, Japan

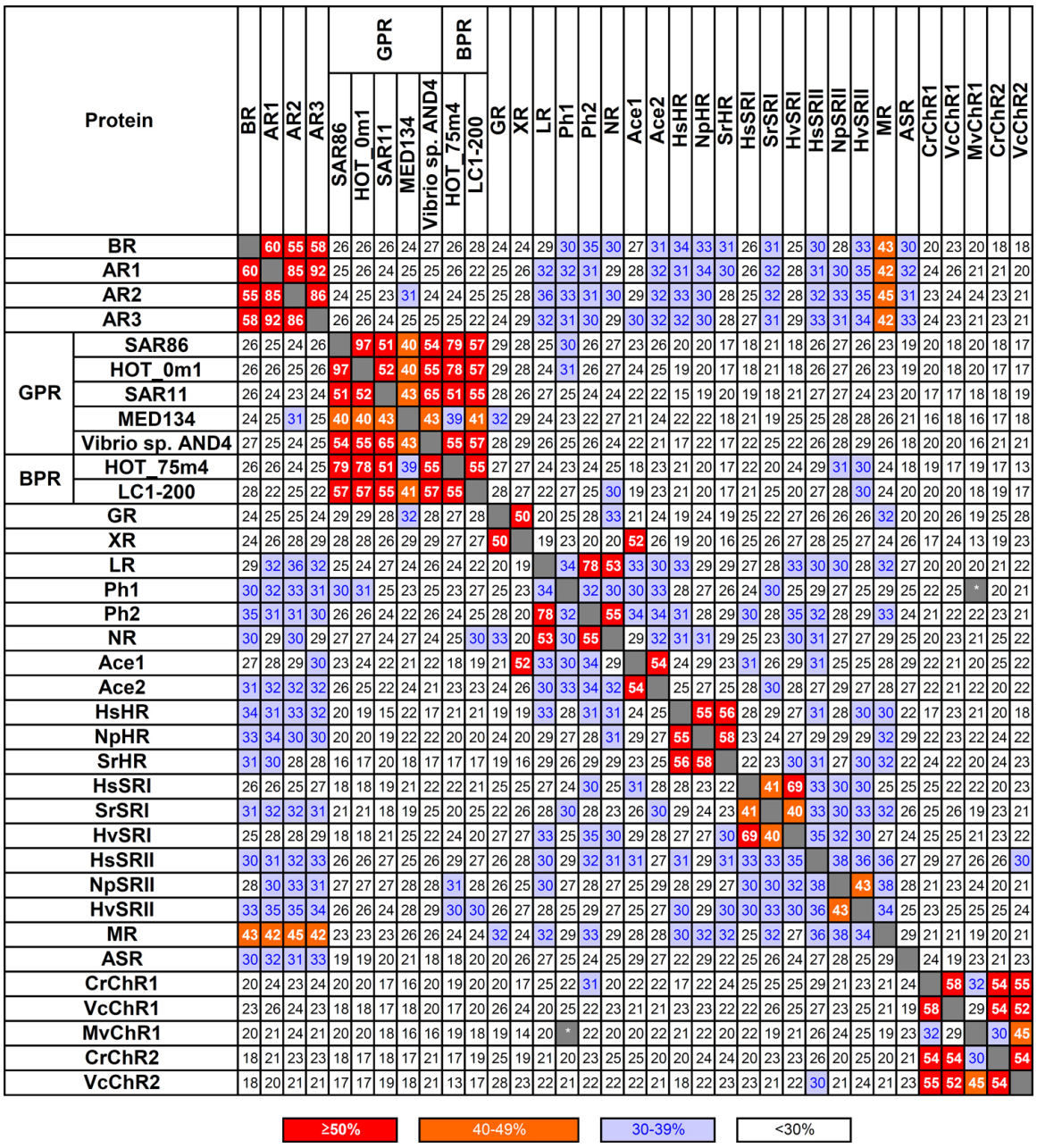
Helix		A	B	B	B	B	C	C	C	C	C	C	C	C	C	D	D	D	E	E	E	F	F	F	-	G	G	G	G	G	G	G
# in BR		20	49	53	56	57	82	83	85	86	89	90	91	93	96	115	118	122	141	145	150	182	185	186	189	194	204	208	212	215	216	219	220
BR	H^+ pump	M	V	A	M	Y	R	Y	D	W	T	T	P	L	D	D	M	G	S	M	Y	W	Y	P	W	E	E	F	D	A	K	F	G
AR1	H^+ pump	M	V	A	A	Y	R	Y	D	W	T	T	P	L	D	D	M	G	S	M	Y	W	Y	P	W	E	E	F	D	A	K	F	G
AR2	H^+ pump	M	V	A	A	Y	R	Y	D	W	T	T	P	L	D	D	M	G	S	F	Y	W	Y	P	W	E	E	F	D	A	K	F	G
AR3	H^+ pump	M	V	A	A	Y	R	Y	D	W	T	T	P	L	D	D	M	G	S	M	Y	W	Y	P	W	E	E	F	D	A	K	F	G
MR	?	M	I	A	A	Y	R	Y	D	W	T	T	P	L	D	D	M	G	S	F	Y	W	Y	P	W	G	E	Y	D	T	K	F	G
GPR[1]	H^+ pump	M	I	A	H	Y	R	Y	D	W	T	V	P	M	E	S	M	G	S	Y	Y	W	Y	P	Y	P	L	Y	D	N	K	F	G
GPR[2]	H^+ pump	L	V	A	H	Y	R	Y	D	W	T	V	P	L	E	T	M	G	G	W	Y	W	Y	P	Y	M	L	Y	D	N	K	F	G
GPR[3]	H^+ pump	L	V	A	H	Y	R	Y	D	W	T	V	P	L	E	T	M	G	G	W	Y	W	Y	P	Y	M	L	Y	D	N	K	F	G
BPR[4]	H^+ pump	M	V	A	H	Y	R	Y	D	W	T	V	P	Q	E	S	M	G	G	W	Y	W	Y	P	Y	F	L	Y	D	N	K	F	G
BPR[5]	H^+ pump	M	I	A	H	Y	R	Y	D	W	T	V	P	Q	E	S	M	G	G	W	Y	W	Y	P	Y	L	L	Y	D	N	K	F	G
GPR[6]	H^+ pump	L	V	A	H	Y	R	Y	D	W	T	V	P	L	E	S	M	G	G	W	Y	W	Y	P	Y	L	L	Y	D	N	K	F	G
GPR[7]	H^+ pump	L	V	A	H	Y	R	Y	D	W	T	V	P	L	E	S	M	G	G	W	Y	W	Y	P	Y	L	L	Y	D	N	K	F	G
GR	H^+ pump	M	V	A	H	Y	R	Y	D	W	T	V	P	L	E	S	M	G	S	F	Y	W	Y	P	Y	L	V	Y	D	A	K	F	G
XR	H^+ pump	M	V	A	H	Y	R	Y	D	W	T	V	P	L	E	A	M	G	S	F	Y	W	Y	P	Y	A	L	Y	D	A	K	Y	G
Ace1	H^+ pump	M	I	A	A	Y	R	Y	D	W	T	T	P	L	D	D	M	G	G	L	W	W	Y	P	W	G	E	M	D	A	K	Y	A
Ace2	H^+ pump	M	I	A	A	Y	R	Y	D	W	T	T	P	L	D	D	M	G	G	Q	Y	W	Y	P	W	G	E	M	D	A	K	F	G
LR	H^+ pump	M	I	A	S	Y	R	Y	D	W	T	T	P	L	D	D	M	G	A	Y	W	W	Y	P	W	G	E	Y	D	A	K	F	G
Ph1	H^+ pump	A	V	A	A	Y	R	Y	D	W	T	T	P	L	D	D	M	G	G	L	Y	W	Y	P	W	G	E	Y	D	A	K	F	G
Ph2	H^+ pump	M	I	A	S	Y	R	Y	D	W	T	T	P	L	D	D	M	G	A	Y	W	W	Y	P	W	G	E	Y	D	A	K	F	G
NR	?	M	I	A	S	Y	R	Y	D	W	T	T	P	L	E	D	M	G	G	Y	W	W	Y	P	W	G	E	Y	D	A	K	F	G
HsHR	Cl^- pump	A	I	S	S	Y	R	Y	T	W	S	T	P	I	A	D	M	G	S	F	S	W	Y	P	W	E	T	Y	D	A	K	F	A
NpHR	Cl^- pump	A	V	S	S	Y	R	Y	T	W	S	T	P	I	A	D	M	G	S	F	Y	W	Y	P	W	E	T	Y	D	A	K	F	A
SrHR	Cl^- pump	A	I	S	S	Y	R	Y	T	W	S	T	P	I	A	D	M	G	S	F	Y	W	Y	P	W	E	T	Y	D	A	K	F	T
HsSRI	Sensor	L	I	A	S	Y	R	Y	D	W	T	T	P	L	Y	D	M	G	S	H	A	W	Y	P	W	A	V	Y	D	A	K	Y	V
SrSRI	Sensor	M	I	A	M	Y	R	Y	D	W	T	T	P	L	Y	D	M	G	S	H	A	W	Y	P	W	-	I	I	D	A	K	F	V
HvSRI	Sensor	F	I	A	S	Y	R	Y	D	W	T	T	P	L	F	D	M	G	S	H	A	W	Y	P	W	E	V	Y	D	A	K	Y	V
HsSRII	Sensor	M	I	A	A	Y	R	Y	D	W	T	T	P	I	Y	D	V	G	G	Y	Y	W	Y	P	W	A	Y	V	D	S	K	F	V
NpSRII	Sensor	M	I	A	A	Y	R	Y	D	W	T	T	P	I	F	N	V	G	G	F	Y	W	Y	P	W	P	D	I	D	T	K	F	G
HvSRII	Sensor	E	I	A	A	Y	R	Y	D	W	T	T	P	N	F	Q	T	G	G	F	Y	W	Y	P	W	A	A	V	D	T	K	F	G
ASR	Sensor	M	I	S	A	Y	R	Y	D	W	T	T	P	L	S	Q	V	G	G	F	W	W	Y	P	W	S	D	F	P	S	K	F	S
CrChR1	Channel	S	V	E	K	F	R	Y	E	W	T	C	P	I	H	D	T	G	G	G	F	W	F	P	F	E	S	H	D	S	K	W	S
VcChR1	Channel	S	V	E	K	S	R	Y	E	W	T	C	P	L	H	D	C	G	S	G	F	W	F	P	F	E	S	H	D	A	K	W	G
MvChR1	Channel	S	V	E	I	Y	R	Y	E	W	T	C	P	I	A	D	C	G	G	G	W	W	F	P	F	E	S	H	D	S	K	W	G
CrChR2	Channel	S	V	E	K	V	R	Y	E	W	T	C	P	I	H	D	T	G	G	G	F	W	F	P	F	E	S	H	D	S	K	W	G
VcChR2	Channel	S	V	E	K	V	R	Y	E	W	T	C	P	I	H	D	T	G	G	G	F	W	F	P	F	E	S	H	D	S	K	W	G

1) MED134, 2) AND4, 3) SAR11, 4) LC1-200, 5) HOT 75m4, 6) SAR86, 7) HOT 0m1

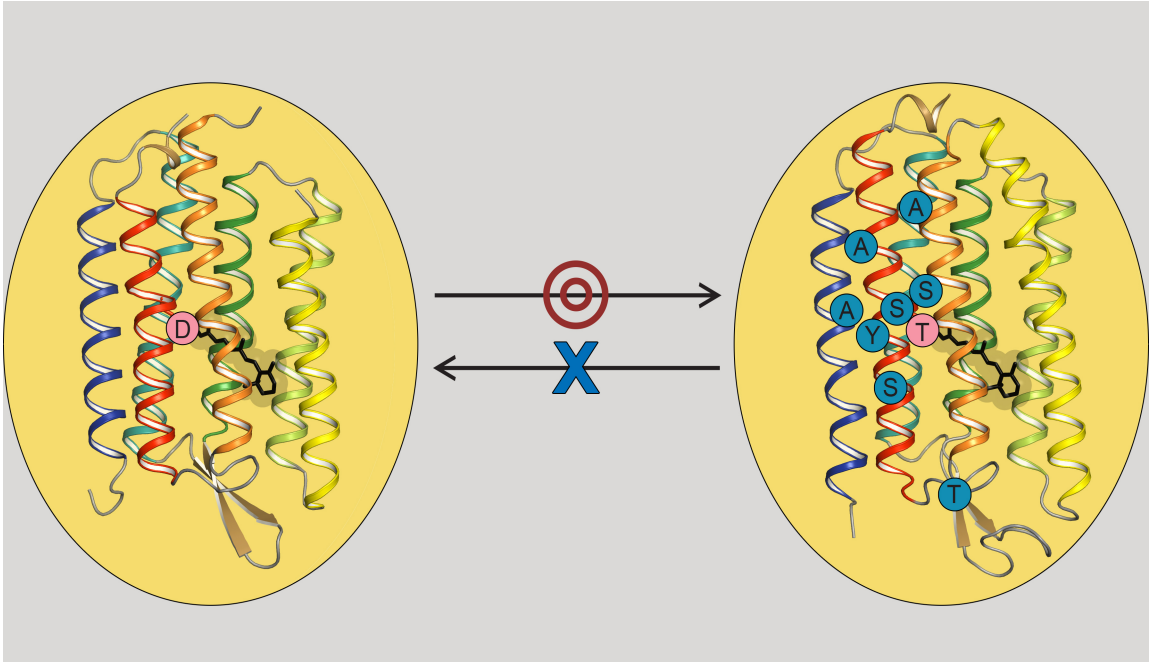
Supporting Information Figure 1. Conservation of the key amino acids for microbial rhodopsins shown in Figure 9. Despite the varying global sequence identities among microbial rhodopsins (Supporting Information Figure 2), a subset of residues is conserved in sequence and in position in the TM helices, and plays a role in activity and selectivity. The numbering scheme corresponds to the position in BR. Negatively charged residues are shown in red, positively charged in blue, aromatic in yellow, polar in green, non-polar in grey.

BR; Bacteriorhodopsin from Archaea (*Halobacterium salinarum*), AR1; Archaerhodopsin-1 from Archaea (*Halorubrum chaoviator*), AR2; Archaerhodopsin-2 from Archaea (*Halobacterium sp.* AUS-2), AR-3; Archaerhodopsin-3 from Archaea (*Halorubrum sadomense*), MR; Midrhodopsin from Eubacteria (*Haloquadratum walsbyi* DSM 16790), GPR(1); Green-absorbing Proteorhodopsin from Eubacteria (*Dokdonia donghaensis* MED134; EAQ40507), GPR(2); Green-absorbing Proteorhodopsin from

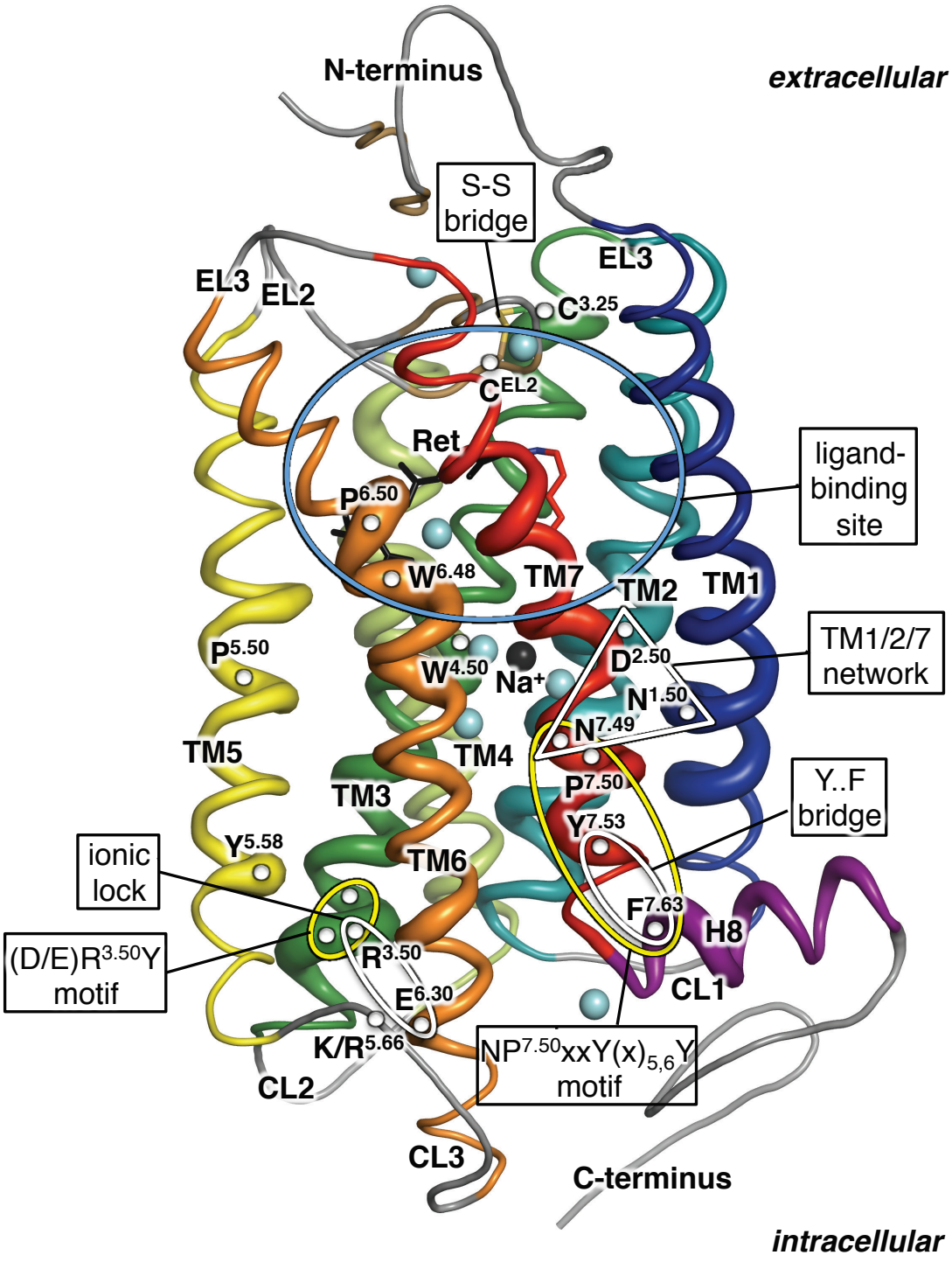
Eubacteria (*Vibrio sp*. AND4; ZP_02194911), GPR(3); Green-absorbing Proteorhodopsin from Eubacteria (*Candidatus Pelagibacter ubique* HTCC1062, "SAR11" group; Q4FMZ3), BPR(4); Blue-absorbing Proteorhodopsin from Eubacteria (*Photobacterium sp*. LC1-200; BAL68143), BPR(5); Blue-absorbing Proteorhodopsin from uncultured bacterium (HOT 75m4; AAK30179), GPR(6);Green-absorbing Proteorhodopsin from uncultured marine gamma proteobacterium ("SAR86" group; AAG10475), GPR(7); Green-absorbing Proteorhodopsin from uncultured bacterium (HOT 0m1; AAK30176), GR; Gloeobacter Rhodopsin from Eubacteria (*Gloeobacter violaceus*), XR; Xanthorhodopsin from Eubacteria (*Salinibacter ruber* DSM 13855), Ace1; Acetabularia Rhodopsin 1 from Eukaryote (*Acetabularia acetabulum*), Ace2; Acetabularia Rhodopsin 2 from Eukaryote (*Acetabularia acetabulum*), LR; Leptosphaeria Rhodopsin from Eukaryote (*Leptosphaeria maculans*), Ph1; Phaeosphaeria Rhodopsin 1 from Eucaryote (*Phaeosphaeria nodorum* SN15), Ph2; Phaeosphaeria Rhodopsin 2 from Eukaryote (*Phaeosphaeria nodorum* SN15), NR; Neurospora Rhodopsin from Eukaryote (*Neurospora crassa* OR74A), HsHR; Halorhodopsin from Archaea (*Halobacterium salinarum*), NpHR; Halorhodopsin from Archaea (*Natronomonas pharaonis*), SrHR; Halorhodopsin from Eubacteria (*Salinibacter ruber* DSM 13855), HsSRI; Sensory Rhodopsin I from Archaea (*Halobacterium salinarum*), SrSRI; Sensory Rhodopsin I from Eubacteria (Salinibacter ruber M8), HvSRI; Sensory Rhodopsin I from Archaea (*Haloarcula vallismortis* ATCC 29715), HsSRII; Sensory Rhodopsin II from Archaea (*Halobacterium salinarum*), NpSRII; Sensory Rhodopsin II from Archaea (*Natronomonas pharaonis*), HvSRII; Sensory Rhodopsin II from Archaea (*Haloarcula vallismortis* ATCC 29715), ASR; Anabaena Sensory Rhodopsin from Eubacteria (*Nostoc sp.* PCC 7120), CrChR1; Channelrhodopsin 1 from Eukaryote (*Chlamydomonas reinhardtii*), VcChR1; Channelrhodopsin 1 from Eukaryote (*Volvox carteri f. nagariensis*), MvChR1; Channelrhodopsin 1 from Eukaryote (*Mesostigma viride*), CrChR2; Channelrhodopsin 2 from Eukaryote (*Chlamydomonas reinhardtii*), VcChR2; Channelrhodopsin 2 from Eukaryote (*Volvox carteri f. nagariensis*).



Supporting Information Figure 2. Pairwise sequence identities for the selected microbial rhodopsins from Figure 9. Despite their similarity of topology and conservation of the key intramembrane residues, especially in the retinal binding pocket, global sequence homology amongst the microbial rhodopsins can be as low as ~15%.



Supporting Information Figure 3. Asymmetric functional conversion between light-driven proton and chloride pumps. Light-driven proton pump BR can be converted into a chloride pump by a single amino acid replacement (Asp-to-Thr mutation). In contrast, substituting the BR-like ten amino acids into HR creates a NpHR mutant which is not capable of pumping protons. FTIR spectroscopy revealed that the failure of functional conversion of HR is due to the lack of strongly hydrogen-bonded water(s), the functional determinant of proton pumping.[160,301]



Supporting Information Figure 4. Sequence and motif conservation and ordered water molecules in GPCRs mapped on bovine Rho (PDB ID:1U19) structure (cf. Figure 20). Greater "Tube" thickness indicates greater residue conservation at that position.[463] Ordered water molecules (shown in light blue) found by superposition of antagonist bound GPCR structures. A sodium ion (shown as a black sphere) has been found within

the A_{2A}-adenosine and PAR1 receptor structures (PDB ID: 4EIY, 3VW7) in a similar position occupied by water in Rho. TM helices are connected by cytoplasmic (CL1-CL3) and extracellular (EL1-EL3) loops. Asn2 and Asn15 are glycosylated (not shown). Cys322 and Cys323 at end of H8 are palmitoylated (not shown). Cys110 on TM3 and Cys187 in EL2 are linked by a disulfide bridge. Residues Asn55$^{1.50}$, Asp83$^{2.50}$, Arg135$^{3.50}$, Trp161$^{4.50}$, Pro215$^{5.50}$, Pro267$^{6.50}$ and Pro303$^{7.50}$ are most conserved residues in a TM helix. The superscript relates to Ballesteros–Weinstein numbering (cf. ref 521). Conserved motifs are the (D/E)RY motif in TM3 (Glu134-Arg135-Tyr136) and the NPxxY$(x)_{5,6}$F motif in TM7 and H8 (Asn302-Pro303-x-x-Tyr306-$x_{5,6}$-Phe313). The inverse agonist 11-*cis* retinal is attached to Lys296$^{7.43}$. Residues Tyr223 (Y$^{5.58}$) and Lys231 (K/R$^{5.66}$) of the Y$^{5.58}$ $(x)_7$K(R)$^{5.66}$ motif in TM5 are microswitches for stabilizing the active receptor state.[521]