Supplementary Information (SI) for Physical Chemistry Chemical Physics.

Retinal Chromophore Environment in an Inward Light-Driven Proton Pump Studied by Solid-state NMR and Hydrogen-Bond Network Analysis

Marie Pinto,[a] Maryam Saliminasab,[a] Andrew Harris,[a] Michalis Lazaratos,[b] Ana-Nicoleta Bondar,[c] Vladimir Ladizhansky[a] and Leonid S. Brown*[a]*

[a.] *Department of Physics and Biophysics Interdepartmental Group, University of Guelph, Guelph, Ontario N1G 2W1, Canada.*
[b.] *Freie Universität Berlin, Physics Department, Theoretical Molecular Biophysics Group, D-14195 Berlin, Germany*
[c.] *University of Bucharest, Faculty of Physics, Măgurele 077125, Romania; Forschungszentrum Jülich, Institute for Computational Biomedicine (IAS-5/INM-9), 52428 Jülich, Germany*

SUPPLEMENTARY INFORMATION

SUPPLEMENTARY FIGURES

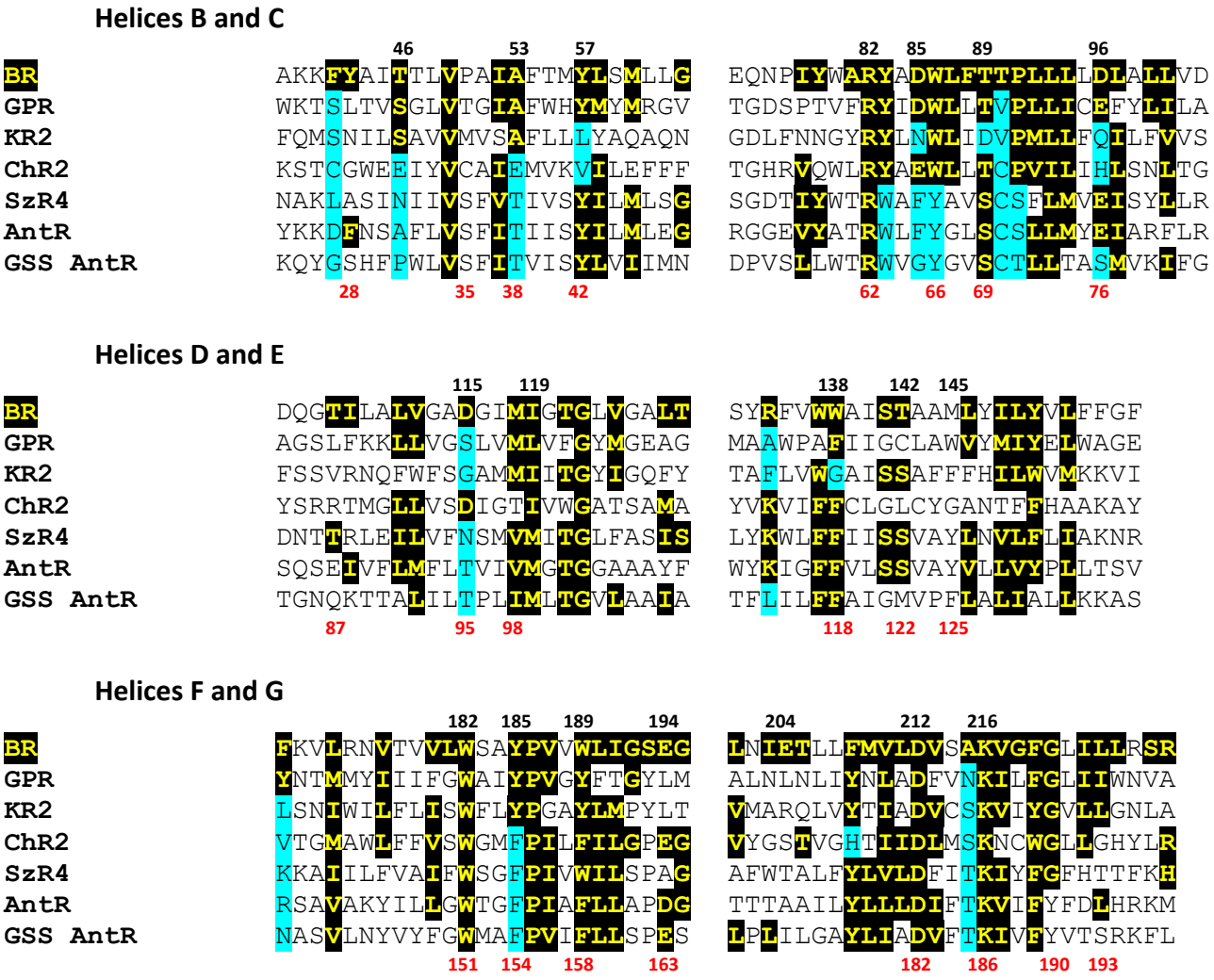

Figure S1. Sequence alignment (produced using CLUSTALO)[1] for helices B-G for some microbial rhodopsins discussed in the main text. BR numbering of selected residues is shown on top, GSS AntR numbering at the bottom in red. Residues conserved in BR sub-family are shown in yellow on black, important differences from BR are highlighted in cyan. Abbreviations: BR – *Halobacterium salinarum* bacteriorhodopsin, GPR – green-absorbing proteorhodopsin EBAC31A08, KR2 – *Krokinobacter eikastus* sodium pumping rhodopsin, ChR2 – Channelrhodopsin-2 from *Chlamydomonas reinhardtii*, SzR4 – Schizorhodopsin-4 AM_5_00977, AntR – Antarctic SzR Ga0105045_102227662, GSS AntR – Antarctic SzR Ga0307935_10093322. The sequences were taken from public databases: NCBI Proteins and DOE JGI IMG/M.[2]

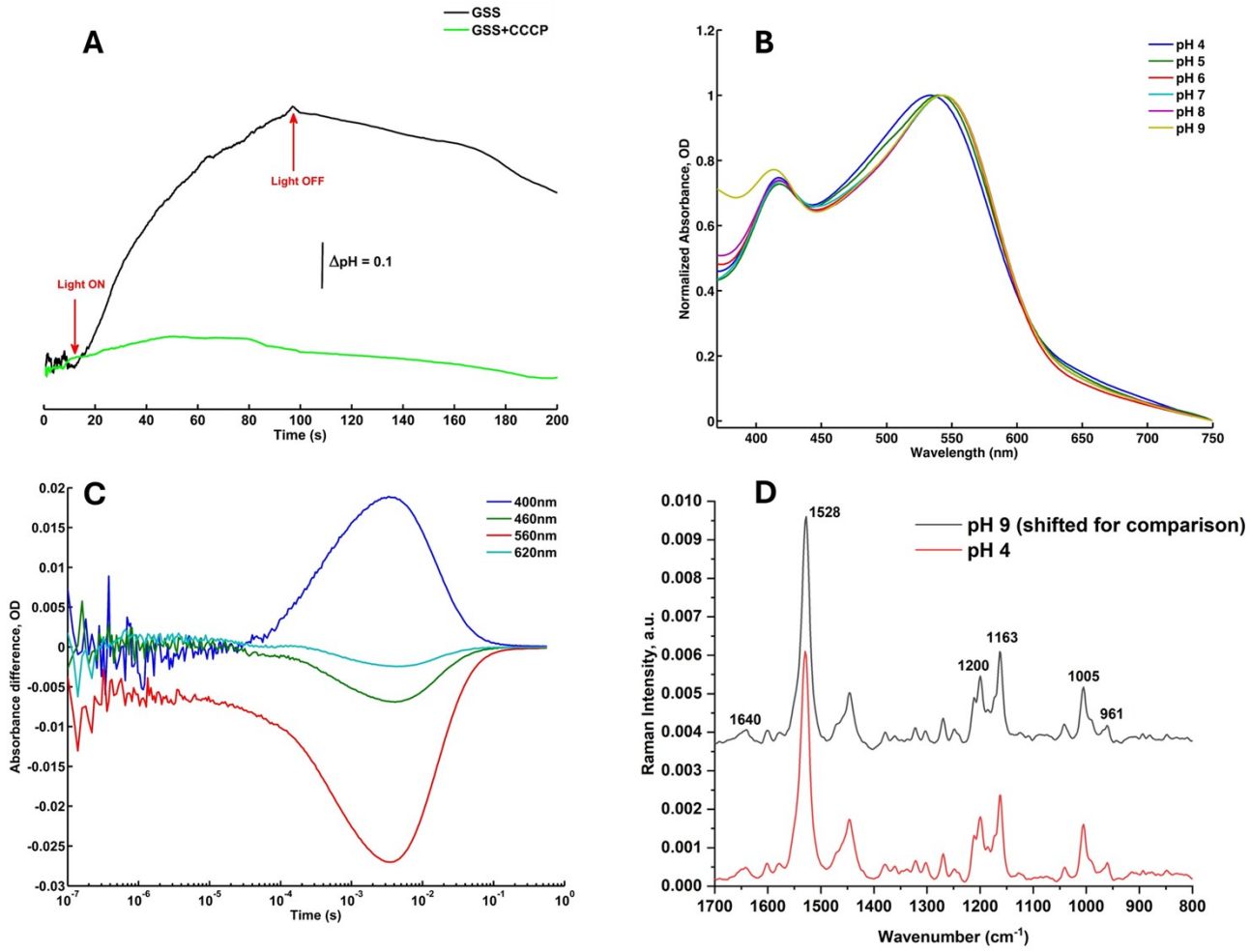

Figure S2. Initial characterization of GSS AntR. (**A**) Light-induced pH changes in the unbuffered suspension of *E. coli* cells expressing GSS AntR demonstrating the active inward proton transport (black and green, without and with the proton uncoupler CCCP). (**B**) Normalized visible absorption spectra of gel-encased *E. coli* membrane fragments with dark-adapted GSS AntR at varying pH, equilibrated with 50 mM NaCl and 50 mM potassium acetate, KH_2PO_4, MES, CHES, and Tris. (**C**) Photocycle kinetics of GSS AntR shown at representative wavelengths, measured using gel-encased *E. coli* membrane fragments with GSS AntR equilibrated with 50 mM NaCl and 25 mM CHES at pH 9. (**D**) Normalized FT-Raman spectra of lipid-reconstituted dark-adapted purified GSS AntR at pH 9 (10 mM NaCl, 25 mM CHES) and 4 (5 mM NaCl and 10 mM potassium acetate). See the **Experimental** section in the main text for the samples and measurements details.

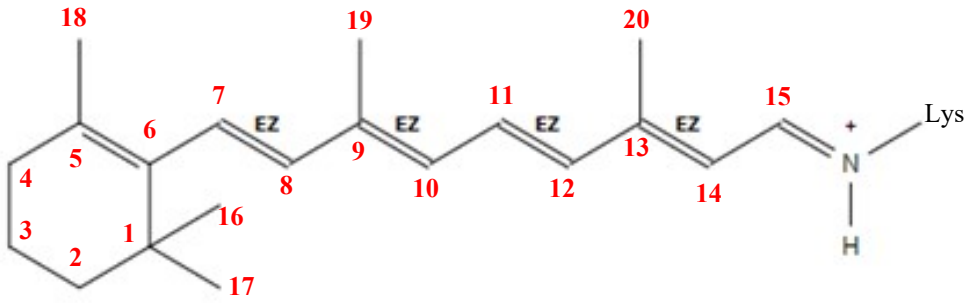

Figure S3. Structure of protonated all-*trans*-6-s-*trans*-retinal Schiff base with carbon atoms numbering (produced by CAS Draw).

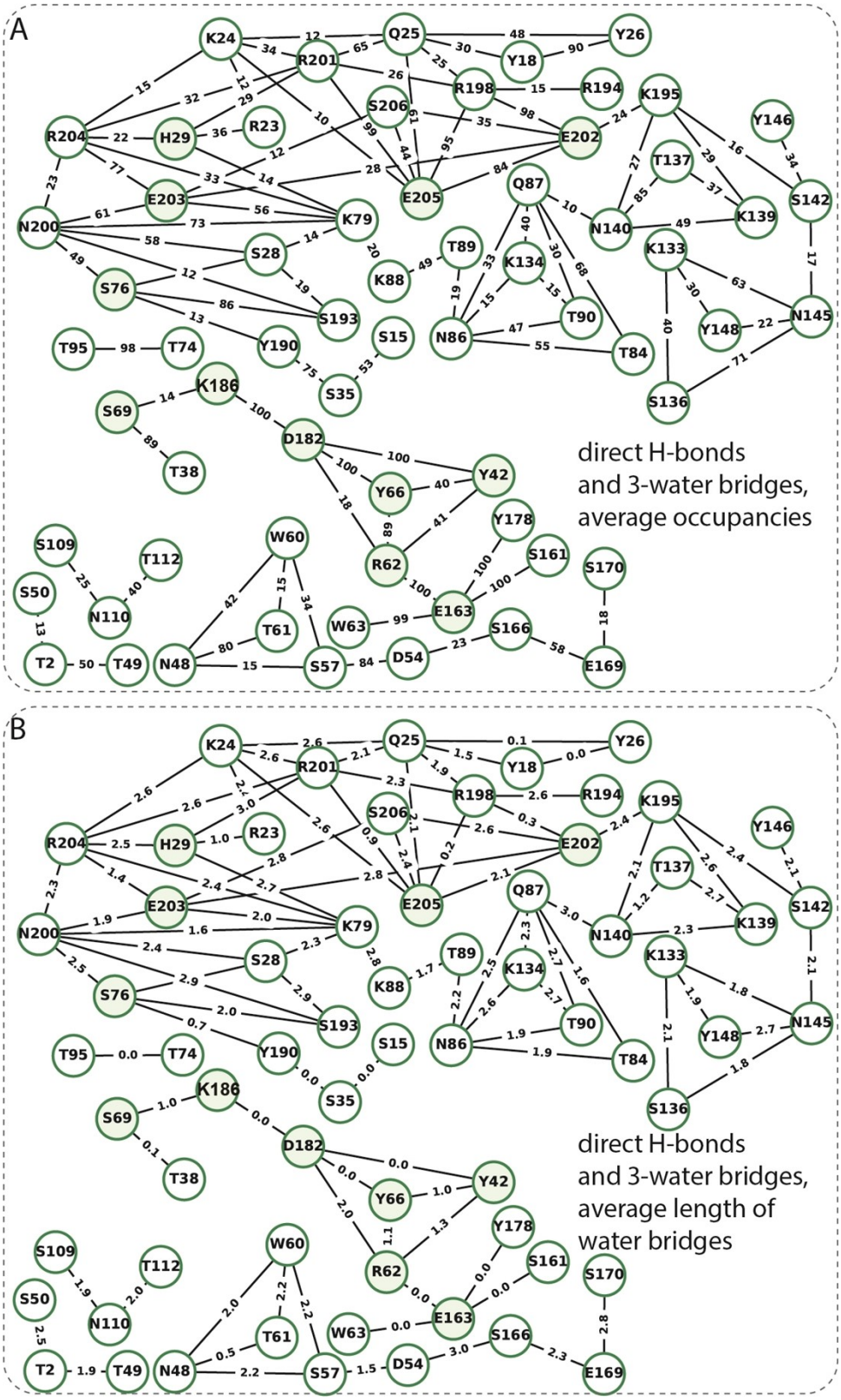

Figure S4. Protein-water H-bond network computed from the main GSS AntR simulation. The minimum H-bond occupancy shown is 10%. Each edge represents a direct or water-mediated bridge with up to three water molecules in bridge. Selected nodes discussed in main text are shown as filled circles. (A, B) H-bond graph with edges showing the average occupancy (panel A) or the average number of water molecules in bridge (panel B).

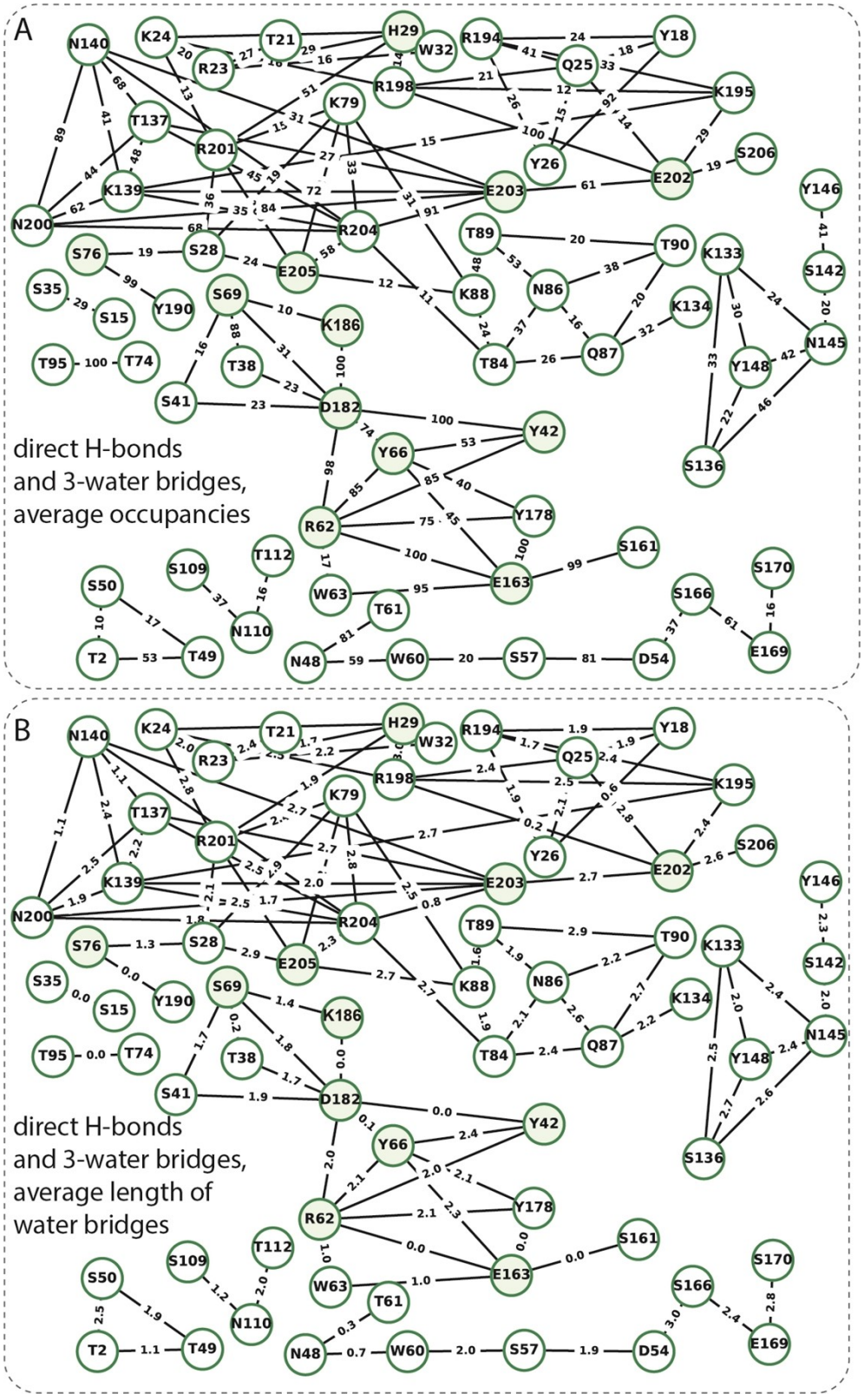

Figure S5. Protein-water H-bond network computed from repeat #1 (ColabFold-generated structural model of GSS AntR based on a manual choice of multiple templates). The minimum H-bond occupancy is 10%, color coding is the same as in **Fig. S4**. (A, B) H-bond graphs with edges showing average H-bond occupancies (panel A) or average numbers of water molecules in bridge (panel B).

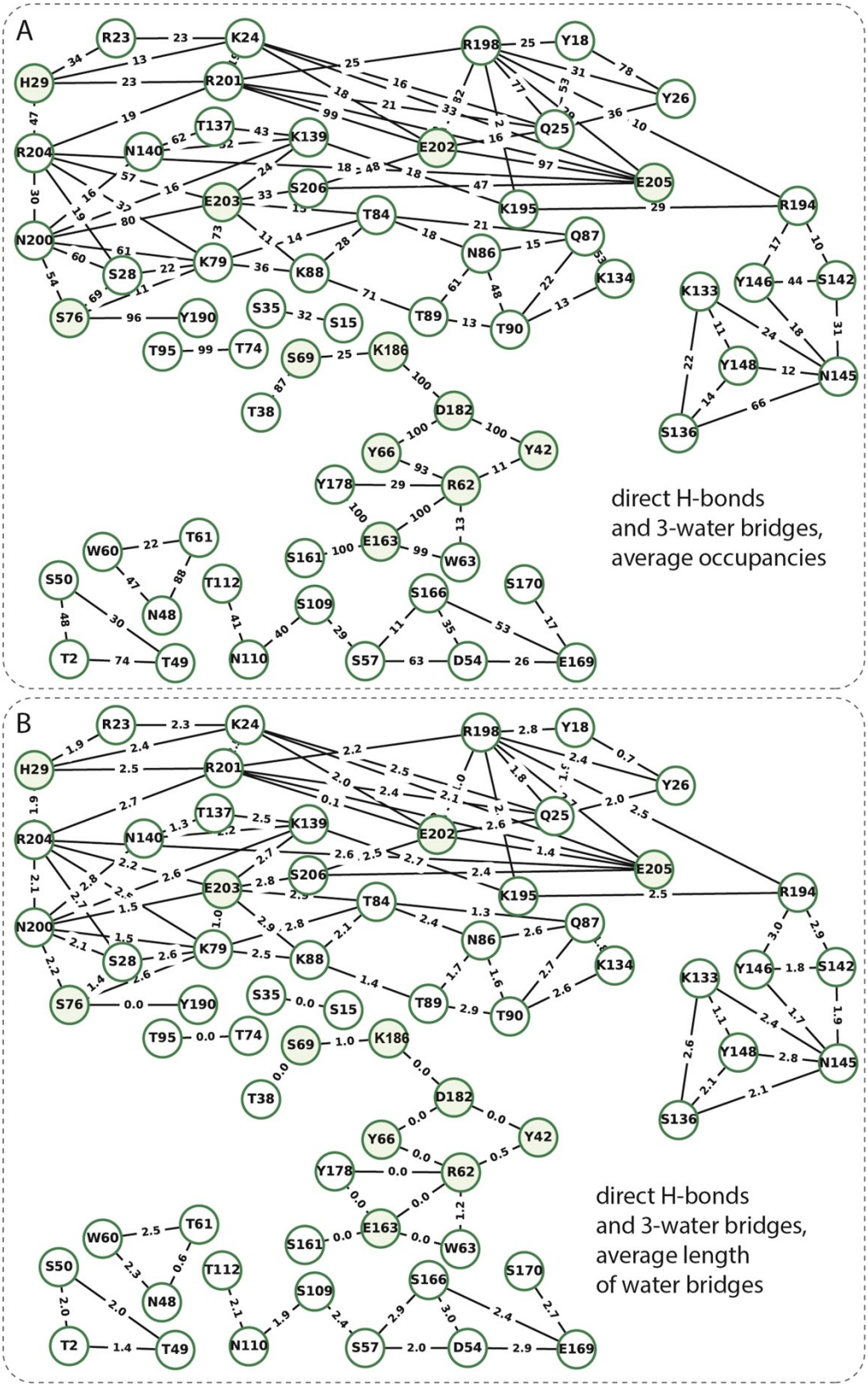

Figure S6. Protein-water H-bond network computed from repeat #2 (ColabFold-generated structural model of GSS AntR based on the structure of SzR4 as a template). The minimum H-bond occupancy is 10%, color coding is the same as in **Fig. S4**. (A, B) H-bond graphs with edges showing average H-bond occupancies (panel A) or average numbers of water molecules in bridge (panel B).

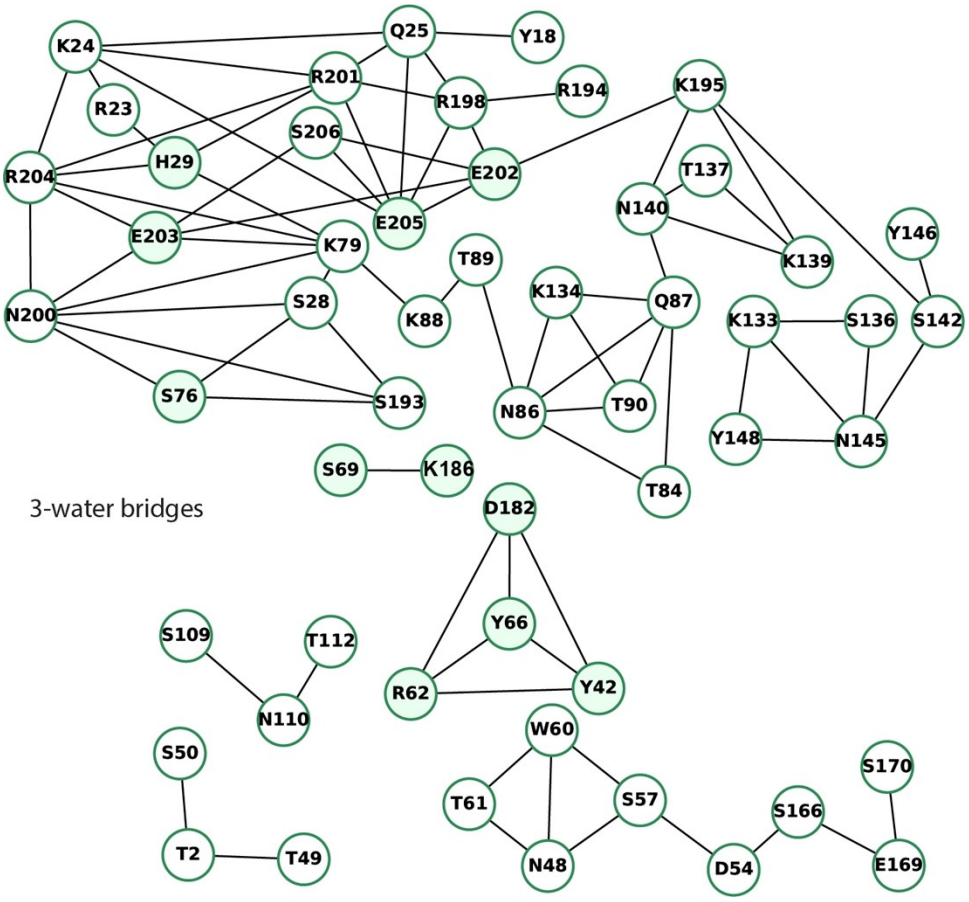

Figure S7. Protein-water H-bond network computed from the main GSS AntR simulation without direct H-bonds between the sidechains. Each edge represents a water-mediated bridge with up to three water molecules in bridge. The minimum H-bond occupancy is 10%, color coding is the same as in **Fig. S4**.

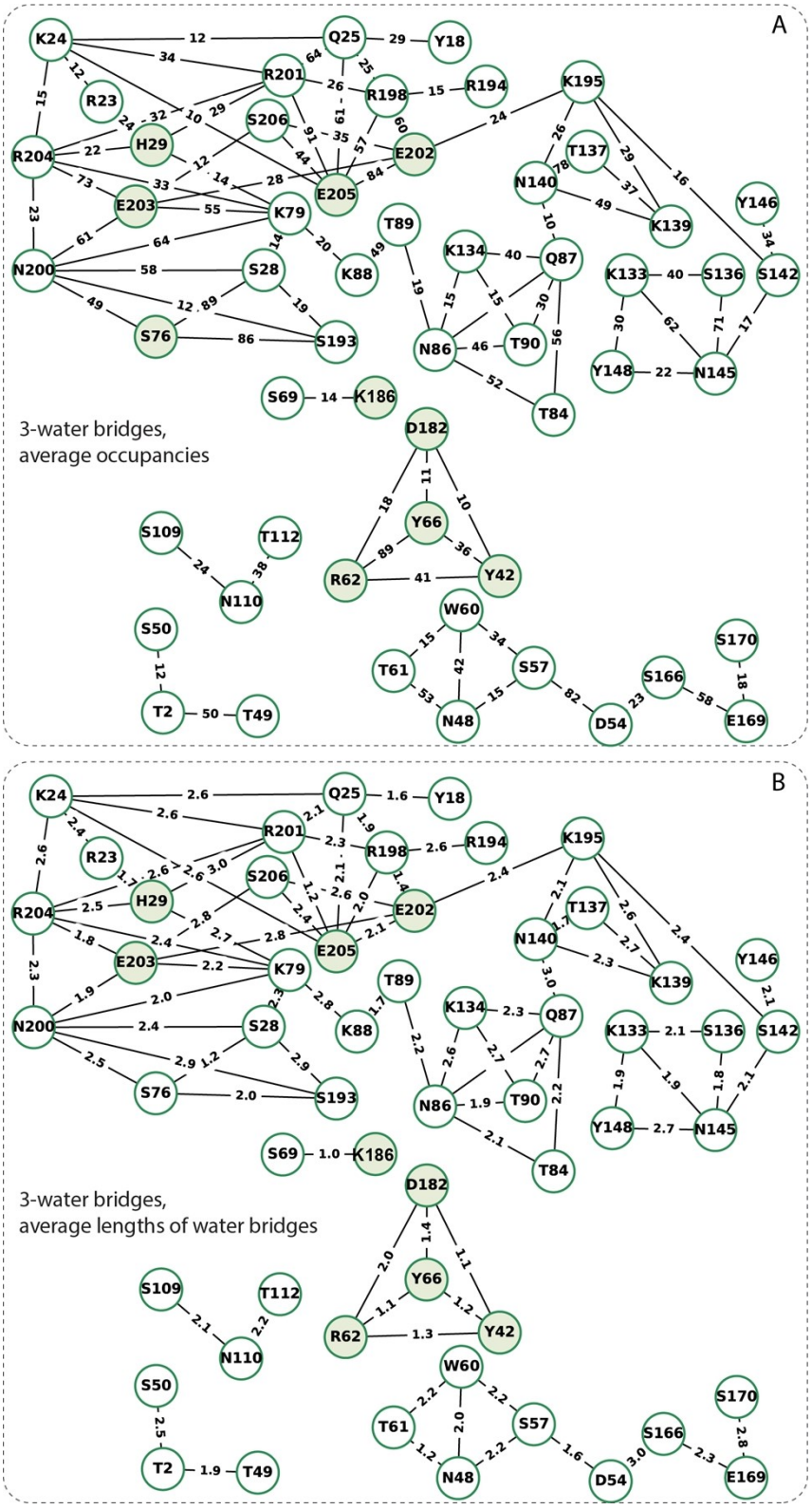

Figure S8. Occupancies and length of bridges in the protein-water H-bond network computed from the main GSS AntR simulation without direct H-bonds between the sidechains (**Fig. S7**). Each edge represents a water-mediated bridge with up to three water molecules in bridge. The minimum H-bond occupancy is 10%, color coding is the same as in **Fig. S4**. (A, B) H-bond graphs with edges showing average H-bond occupancies (panel A) or average numbers of water molecules in bridge (panel B).

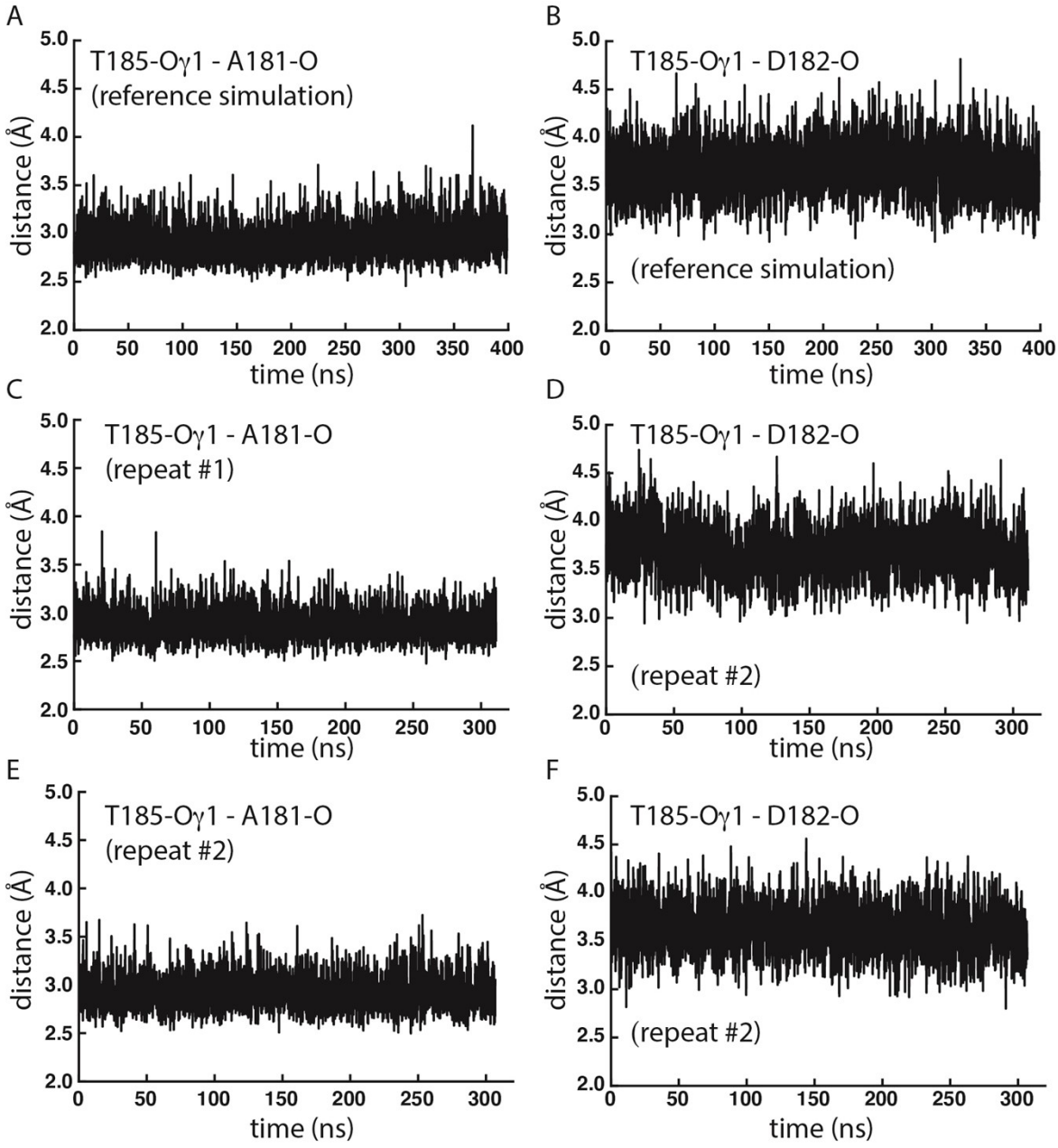

Figure S9. Time series of the distances between Thr185-Oγ1 and the backbone carbonyl oxygen atoms of Ala181 and Asp182 in the three independent simulations performed. All distances are reported in Å. For clarity, coordinates were read with a step of 100 ps. Distances were monitored from the reference simulations (panels A, B), repeat #1 (panels C, D), and repeat #2 (panels E, F).

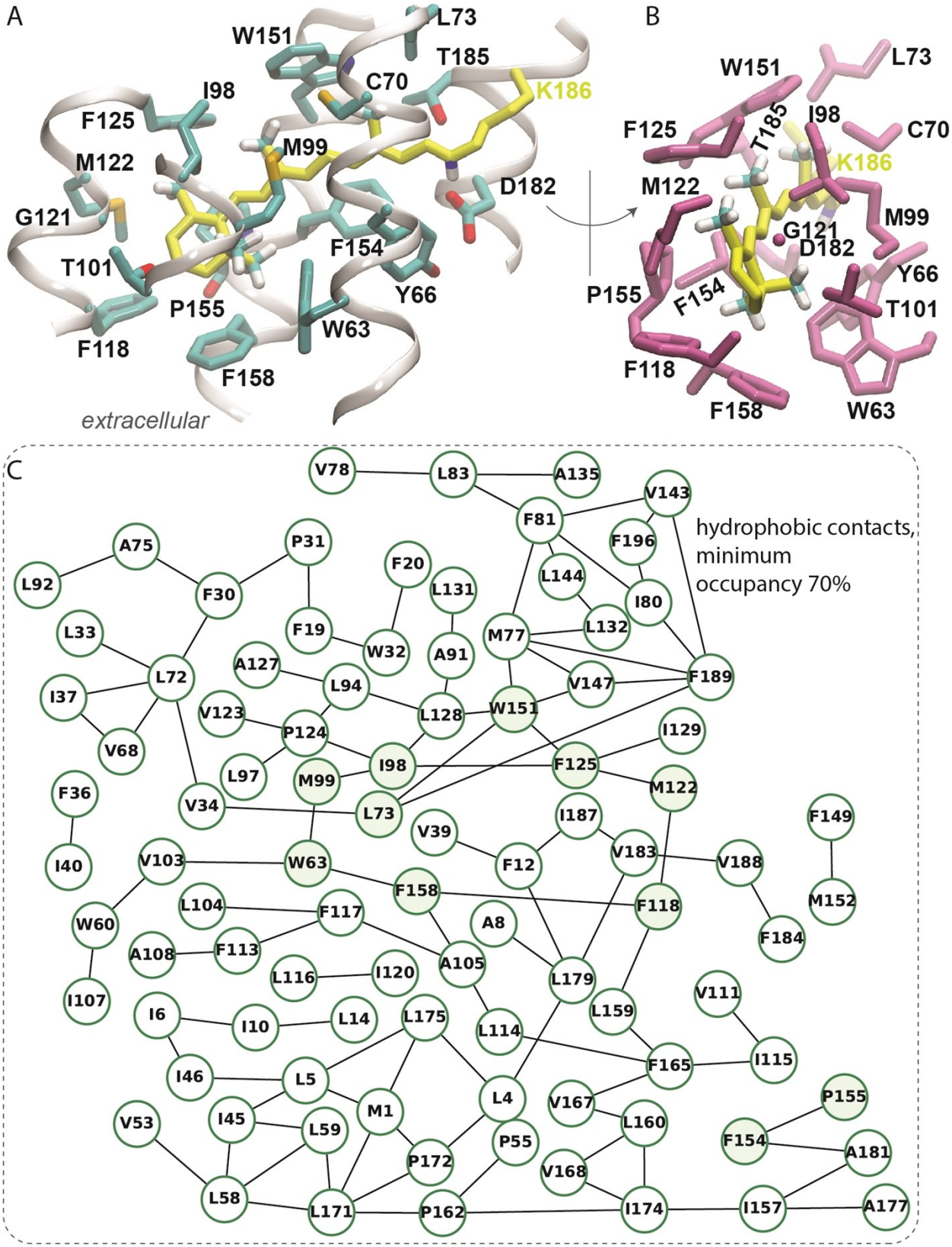

Figure S10. Close contacts between the retinal and protein sidechains in the main simulation of GSS AntR. (A, B) Close view of the retinal and protein sidechains with at least one hetero-atom within 4.5 Å of any of the retinal carbon atoms. (C) Hydrophobic contacts map shown at a minimum occupancy of 70%. The nodes filled with green color are within hydrophobic contact with retinal carbon atoms at the end of the simulation.

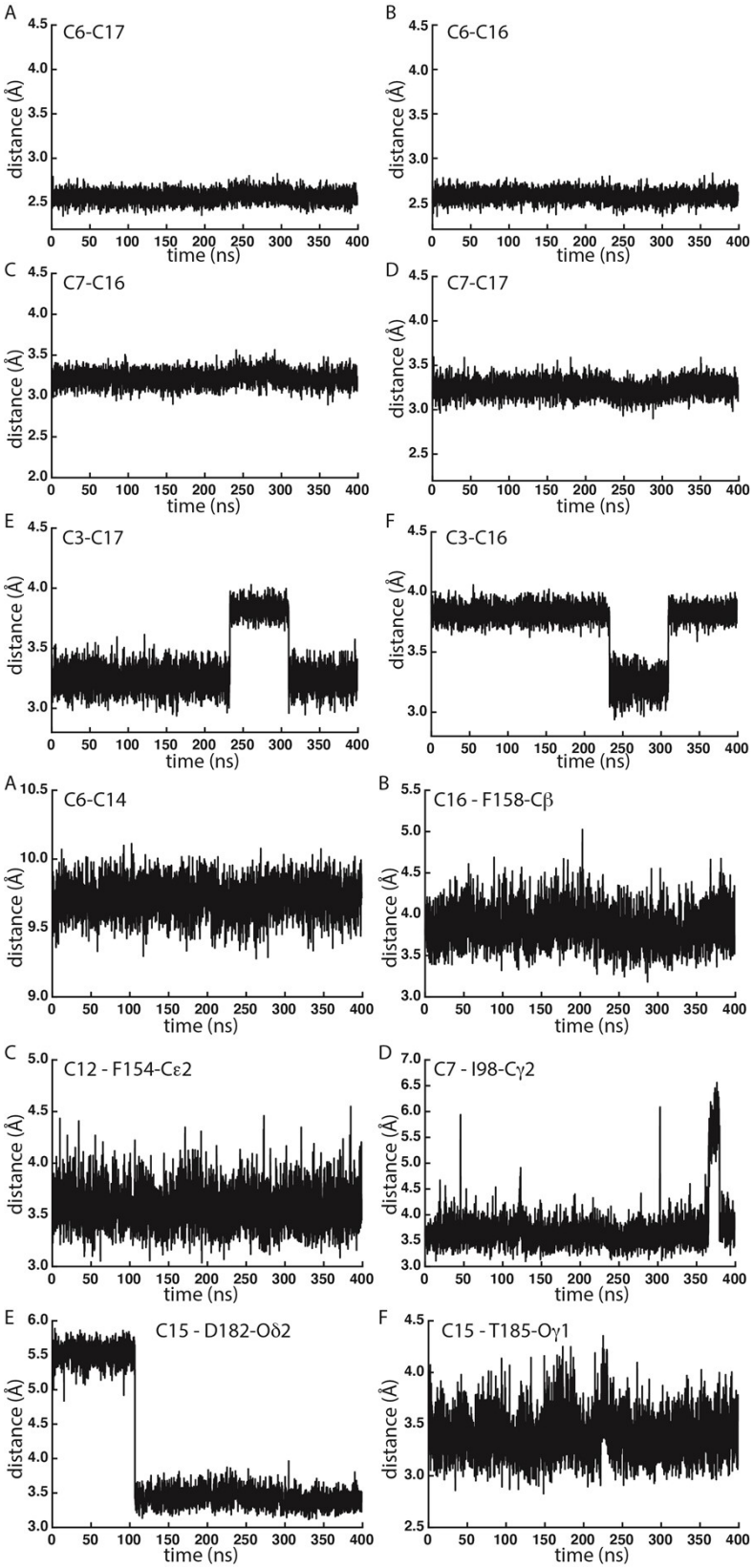

Figure S11. Time series of selected intra-retinal and retinal-sidechain distancesd monitored along the reference simulation. All distances are reported in Å. For clarity, coordinates were read with a step of 100 ps. The C6-C14 distance serves as a control of retinal conformation stability. Note the occasional β-ionone ring inversion and sidechain flips of Asp182 and Ile98.

SUPPLEMENTARY TABLES

Table S1. Templates used for structural modeling of GSS AntR using ColabFold. For each set of structural modeling computations, we indicate the range of pLDDT scores. The structural model with the highest-ranking score was selected from each set.

Protein	**Organism**	**PDB ID**	**Res. (Å)**	**Reference**
***GSS_1**, default templates selection, pLDDT scores 92.9-93.8*				
Proton-pumping mutant	*M. repens*	6wp8	2.5	[3]
Chloride importer		6k6j	2.5	-
Chloride importer N63A/P118A		6k6k	2.2	-
Deltarhodopsin-3	*Haloterrigena thermotolerans*	4fbz	2.7	[4]
Archaerhodopsin-1	*H. ezzemoulense*	1uaz	3.4	[5]
Archaerhodopsin-2	*Halobacterium sp*. AUS-2	2z55	2.5	-
Archaerhodopsin-3	*Halorubrum sodomense*	6guz	1.9	-
BR early M intermediate	*H. salinarum*	1kg8	2.0	[6]
BR M56A		1pxs	2.2	[7]
BR K41P		1tn5	2.2	[8]
BR T24S		1s51	2.0	[9]
BR T24A		1s54	2.2	[9]
BR L111A		3hap	1.6	[10]
Proton pump	*Leptosphaeria maculans*	7bmh	2.2	[11]
SpaR	*Sphingomonas paucimobilis*	8anq	2.8	[12]
Proton pump	*Coccomyxa subellipsoidea*	6gyh	2.0	[13]
Schizorhodopsin-4, SzR4	Asgard archaea	7e4g	2.1	[14]
***GSS_2**, custom templates selection, pLDDT scores 93.8-94.6*				
Proton-pumping mutant	*M. repens*	6wp8	2.5	[3]
Deltarhodopsin-3	*H. thermotolerans*	4fbz	2.7	[4]
Archaerhodopsin-1	*H. ezzemoulense*	1uaz	3.4	[5]

Archaerhodopsin-2	*Halobacterium sp*. AUS-2	2z55	2.5	-
BR	*H. salinarum*	7z09	1.05	[15]
Proton pump	*Leptosphaeria maculans*	7bmh	2.2	[11]
SpaR	*Sphingomonas paucimobilis*	8anq	2.8	[12]
SzR4	Asgard archaea	7e4g	2.1	[14]
***GSS_3**, single template SzR4, pLDDT scores 93.8-94.6*				
SzR4	Asgard archaea	7e4g	2.1	[14]

Table S2. Carbon chemical shifts determined for retinal of GSS AntR compared with the literature values for other microbial rhodopsins. In the cases where referencing was done for TMS the chemical shift values were adjusted to correspond to the DSS referencing by adding 1.8 ppm.[16]

Retinal Carbon #	**BR all-*trans*[17]**	**GSS AntR**	**BR 13-*cis*-15-*syn*[17]**	**GPR[18]**	**GPR[19]**	**GPR L105Q[19]**	**KR2[20]**
1	36.3	**35.3**	36.3	37.1			
2	44.5	**44.1**	44.5				
3	20.4	**20**	20.4				
4	36.4	**36.8**	36.4				
5	146.6	**146.2**	146.6				
6	137.2	**138.3**	136.7	139.9			
7	131.3	**133.8**	132.5	130.5			
8	134.5	**134**	133.4	134.7			
9	148.2	**148.6**	150.2	147			
10	134.8	**135.7**	131.5	132.4	132.7	133.2	134.2
11	140.9	**137.6**	137.2	139.8	140.1	138.8	139.4
12	136.1	**139.3**	126	132.5	132.8	133.8	133.1
13	166.6	**169.1**	170.5	166.7	166.4	166.6	170.5
14	123.8	**125.2**	112.3	121.8	122.4	122.8	122.7
15	161.8	**166.1**	165	163.7	164	167.8	168.7
16	30.7	**32.4**	30.7	32.1	32.6	33.2	31.5
17	30.7	**27.8**	30.7	32.1	27.9	28.7	30.3
18	23.8	**23.3**	23.8		23.4	23.3	23.7
19	13.1	**14.8**	13.1	15.2			
20	15.1	**15.3**	23.8	15.7	16.2	15.7	

Table S3. Signal-to-noise ratios for C16 and C17 cross-peaks.

Cross-peak	Signal-to-Noise Ratio		
	RFDR	10 ms DARR	30 ms DARR
C16/C2	--	--	8.4
C17/C2	--	--	8.7
C2/C16	--	5.2	6.3
C2/C17	--	5.2	8.8
C16/C7(8)	--	--	6.9
C17/C7(8)	--	--	5.2
C3/C17	--	--	6.2
C6/C16	6.8	--	--
C6/C18	5.8	--	--
C7(8)/C16	--	--	6.8
C7(8)/C17	--	--	5.1

SUPPLEMENTARY REFERENCES

1. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J. D. Thompson and D. G. Higgins, *Mol Syst Biol*, 2011, **7**, 539.
2. I. A. Chen, K. Chu, K. Palaniappan, M. Pillay, A. Ratner, J. Huang, M. Huntemann, N. Varghese, J. R. White, R. Seshadri, T. Smirnova, E. Kirton, S. P. Jungbluth, T. Woyke, E. A. Eloe-Fadrosh, N. N. Ivanova and N. C. Kyrpides, *Nucleic Acids Res*, 2019, **47**, D666-D677.
3. J. E. Besaw, W.-L. Ou, T. Morizumi, B. T. Eger, J. D. Sanchez Vasquez, J. H. Y. Chi, A. Harris, L. S. Brown, R. J. D. Miller and O. P. Ernst, *J. Biol. Chem.*, 2020, **295**, 14793-14804.
4. J. Zhang, K. Mizuno, Y. Murata, H. Koide, M. Murakami, K. Ihara and T. Kouyama, *Proteins*, 2013, **000**, 1-8.
5. N. Enami, H. Okumua and T. Kouyama, *J. Photosci.*, 2003, **9**, 320-322.
6. M. T. Facciotti, S. Rouhani, F. T. Burkard, F. M. Betancourt, K. H. Downing, R. B. Rose, G. McDermott and R. M. Glaesser, *Biophys. J.*, 2001, **81**, 3442-3445.
7. S. Faham, D. Yang, E. Bare, S. Yohannan, J. P. Whitelegge and J. U. Bowie, *J. Mol. Biol.*, 2004, **335**, 297-305.
8. S. Yohannan, D. Yang, S. Faham, G. Boulting, J. Whitelegge and J. U. Bowie, *J. Mol. Biol.*, 2004, **341**, 1-6.
9. S. Yohannan, S. Faham, D. Yang, D. Grosfeld, A. K. Chamberlain and J. U. Bowie, *J. Am. Chem. Soc.*, 2004, **126**, 2284-2285.
10. N. H. Joh, A. Oberai, D. Yang, J. P. Whitelegge and J. U. Bowie, *J. Am. Chem. Soc.*, 2009, **131**, 10846-10847.
11. D. Zabelskii, N. Dmitrieva, O. Volkov, V. Shevchenko, K. Kovalev, T. Balandin, D. Soloviov, R. Astashkin, E. Zinovev, A. Alekseev, E. Round, V. Polovinkin, I. Chizhov, A. Rogachev, I. Okhrimenko, V. Borschchevskiy, V. Chupin, G. Büldt, E. Bamberg, E. Koonin and V. Gordeliy, *Nature Comm. Biol.*, 2021, **4**, 821.
12. I. S. Okhrimenko, K. Kovalev, L. E. Petrovskaya, N. Ilyinsky, A. A. Alekseev, E. Marin, T. I. Rokitskaya, Y. N. Antonenko, S. A. Siletsky, P. A. Popov, Y. A. Zagryadskaya, D. V. Soloviov, I. V. Chizov, D. V. Zabelskii, Y. L. TRyzhykau, A. V. Vlasov, A. I. Kuklin, A. O. Bogorodskiy, A. E. Mikhailov, D. V. Sidorov, S. Bukhalovich, F. Tsybrov, S. Bukhdruker, A. D. Vlasova, V. I. Borschchevskiy, D. A. Dolkigh, M. P. Kirpichnikov, E. Bamberg and V. Gordeliy, *Nature Comm. Chemistry*, 2023, **6**, 88.

13. R. Fudim, M. Szczepek, J. Vierock, A. Vogt, A. Schmidt, G. Kleinau, P. Fischer, F. Bartl, P. Scheerer and P. Hegemann, *Sci. Signal.*, 2019, **12**.
14. A. Higuchi, W. Shihoya, M. Konno, T. Ikuta, H. Kandori, K. Inoue and O. Nureki, *PNAS*, 2021, **118**, e2016328118.
15. V. Borshchevskiy, K. Kovalev, E. Round, R. Efremov, R. Astashkin, G. Bourenkov, T. Baladin, I. Chiznov, C. Baeken, I. Guschkin, A. Kuzmin, A. Alekseev, A. Rogachev, D. Willbold, M. Engelhard, E. Bamberg, G. Büldt and V. Gordeliy, *Nature Struct. Mol. Biol.*, 2022, **29**, 440-450.
16. C. R. Morcombe and K. W. Zilm, *J Magn Reson*, 2003, **162**, 479-486.
17. S. O. Smith, H. J. de Groot, R. Gebhard, J. M. Courtin, J. Lugtenburg, J. Herzfeld and R. G. Griffin, *Biochemistry-Us*, 1989, **28**, 8897-8904.
18. R. A. Munro, J. de Vlugt, M. E. Ward, S. Y. Kim, K. A. Lee, K. H. Jung, V. Ladizhansky and L. S. Brown, *J Biomol Nmr*, 2019, **73**, 49-58.
19. J. Mao, X. Jin, M. Shi, D. Heidenreich, L. J. Brown, R. C. D. Brown, M. Lelli, X. He and C. Glaubitz, *Sci Adv*, 2024, **10**, eadj0384.
20. J. Kaur, C. N. Kriebel, P. Eberhardt, O. Jakdetchai, A. J. Leeder, I. Weber, L. J. Brown, R. C. D. Brown, J. Becker-Baldus, C. Bamann, J. Wachtveitl and C. Glaubitz, *J Struct Biol*, 2019, **206**, 55-65.