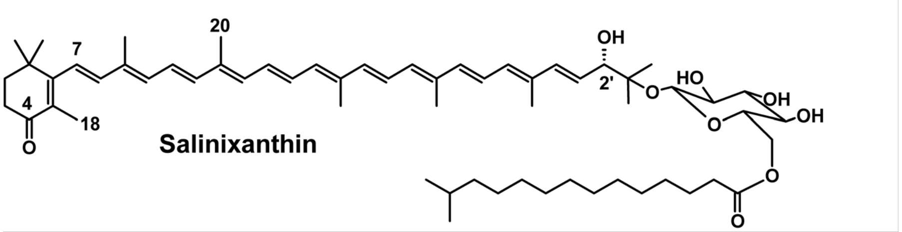
SUPPORTING INFORMATION

Section I. Distances between salinixanthin (*1*) and side chain atoms of residues within 4 Å in xanthorhodopsin based on the 3DLL coordinate set (*2*). Homologous residues in gloeobacter rhodopsin and bacteriorhodopsin are shown in the right column (residues identical to xanthorhodopsin are in bold).



Distance (Å)	Residue in XR	Carotenoid part	Residue in GR / BR
2.58	Arg184, NH1	Glycoside, O3*	Ala206 / Glu166
3.08	Asn191, ND2	Glycoside, O1*	**Asn213** / Val173
3.11	Thr160, OG1	Chain, C7	**Thr182 / Thr142**
3.18	Met211, CE	Keto group, O21	**Met233** / S193
3.25	Tyr207, CE2	Keto group, O21	**Tyr229** / Trp189
3.29	Leu194, CD2	Chain, C8′	Trp216 / Asp176
3.54	Leu148, CD2	Ring, C3	Asp170 / Val130
3.56	Gly156, CA	Keto group, O21	**Gly178** / W138
3.64	Met208, SD	Ring, C18	Leu230 / Leu190
3.73	Pro204, CB, CG	Chain, C9, C10, C11	**Pro226 / Pro186**
3.73	Gly201, CA	Chain, C13	**Gly223** / Ser183
3.76	Phe163, CE2	Chain, C20	**Phe185** / Met154
3.77	Leu188, CD2	Glycoside, O3*	**Leu210** / Thr170
3.95	Ile205, CD	Chain, C14	**Ile227** / Val187
3.97	Ala198, CB	Chain, C12′	Leu220 / Val180
4.09	Leu197, CD1	Chain, C20	**Leu219** / Val179

Section II. Xanthorhodopsin clade: genes of retinal proteins with high homology to xanthorhodopsin.

The xanthorhodopsin sequence (*3, 4*) contains 20.5% residues identical to those in bacteriorhodopsin (*3*) and 26.1% residues, identical with those in green light absorbing proteorhodopsin (*2*). There are organisms that carry retinal-protein genes with substantially higher homology, forming a clade with xanthorhodopsin (*4-6*). These are genes in *Gloeobacter violaceous*, *Pyrocystis lunula*, numerous members of Actinobacteria, Proteobacteria and other species (see below and ref. (*5*)). Most of the sequences suggest that these proteins can bind retinal and act as light-driven proton pumps since they contain homologues of the key residues in proton transport, Asp96 serving as the proton acceptor from the Schiff base and Glu108 as the internal proton donor (numbers from xanthorhodopsin sequence). An interesting exception is the retinal protein of *Fulvimarina pelagi*, in which these residues are replaced with asparagine and glutamine, respectively. These replacements suggest that it is a sensor rather than a proton pump. As in xanthorhodopsin and the proteorhodopsins, most of these sequences contain a histidine homologous to His62 in xanthorhodopsin which makes a strong hydrogen bond with Asp96 in xanthorhodopsin.

Some sequences (like the highly homologous protein from *Methylophilales bacterium* HTCC2181 (*6*)) contain tryptophan in the retinal ring binding site (homologous to Trp138 in bacteriorhodopsin), however in a large number of sequences it is replaced with the smaller glycine as in xanthorhodopsin of *S. ruber*, which makes these proteins potential candidates for binding of carotenoid antenna with a ring similar to that in salinixanthin.

Below, we present incomplete list of these organisms and their retinal protein sequence accession number followed by corresponding sequences of the E and F helix obtained from NCBI using BLASTP. Marked with red are residues that interact with carotenoid in xanthorhodopsin and conserved in other proteins. To the right the percentages of identities to xanthorhodopsin are shown, as well as the four residues homologous to His56, Asp96, Glu108 and G178 of xanthorhodopsin. The sequence of bacteriorhodopsin, BR, (*7, 8*) from *Halobacterium salinarum* is given for comparison.

1) Salinibacter ruber DSM 13855 (**YP_445623**)	100%	H D E G
2) Gloeobacter violaceus PCC 7421 (**NP_923144**)	53%	H D E G
3) Thermus aquaticus Y51MC23 ctg62 (**ZP_03495873**)	53%	H D E G
4) Roseiflexus sp. RS-1 (**YP_001277280**)	52%	H D E G
5) Methylophilales bacterium HTCC2181 (**ZP_01551538**)	49%	H D E W
6) Alpha proteobacterium BAL199 (**EDP63929**)	48%	H D E G
7) Octadecabacter antarcticus 238 (**ZP_05063020**)	47%	H D E G
8) Actinobacterium MWH-EgelM2-3.D6 (**ACN42852**)	46%	H D E G
9) Fulvimarina pelagi HTCC2506 (**ZP_01440547**)	34%	S N Q G
10) Geodermatophilus odscurus DSM 43160 (**ZP_03889903**)	32%	A D E G
11) Kineococcus radiotolerans SRS30216 (**EAM73404**)	32%	S D E G
12) Exiguobacterium sp. AT1b (**YP_002885111**)	31%	H D K G
13) Polaribacter sp. MED152 (**ZP_05108337**)	32%	H D E G
14) Dokdonia donghaensis MED134 (**EAQ40507**)	31%	H D E G
15) Polaribacter irgensii 23-P (**ZP_01117885**)	31%	H D E G
16) Flavobacteria bacterium BAL38 (**ZP_01734914**)	29%	H D E G
17) Psychroflexus torquis ATCC 700755 (**ZP_01253360**)	29%	H D E G

BR	**YRFVWWAISTAAMLYILYVLFFGFTSKAE--SMRPEVASTFKVLRNVTVVLWSAYPVVWLIGS**
1)	**TRGLWGFLSTIPFVWILYILFTQLGDTIQ--RQSSRVSTLLGNARLLLLATWGFYPIAYMIPM**
2)	**TRIIWGTVSTIPFAYILYVLWVELSRSLV--RQPAAVQTLVRNMRWLLLLSWGVYPIAYLLPM**
3)	**PRTLWGALSSVPFFYILYVLWVELGQAIREAKFGPRVLELLGATRLVLLMSWGFYPIAYALGT**
4)	**TRLVWGTLSSIPFLYILYVLWRDLGAVME--RESAEVRVLLRNLRLLLLASWGVYPIAYLLPL**
5)	**TRWMWWILAMIPFLYIVKDLVSGLGASIS--KQPKAVQGLVSNARTLTILSWTFYPIVFILPM**
6)	**TRLIWGALSMIPFLYIVMTLYSGLSKSIDQ--QPENVRGLVSMARSLVVVSWFFYPIVYFLPF**
7)	**VRWIFGLLSMIPFLWIVFQLYVGLGKAIES--QPENARGLVRTARNVTVGSWCFYPVVY----**
8)	**TKILWGVLSTIPFLYILYVLFVELGKSLER--QPAGVAATVGRLRWLLLATWGVYPISYLFPL**
9)	**RLNVWGVISTIFFVWLIIEVRGVISRAI-AMG-PAELAAWPKNIWWFFLAFWGLYPIAYALPQ**
10)	**ALVVWGLISTAFFAYLYVALIGAVRRSLPTMG--PEAAVSLRNATIVLLSSFGIYPLVYAVPV**
11)	**ALIAWGVVSSAFYAALHVVLIRVLRRSRGAMG--AQAYTSYRNATILLLSVWGWYPIAYALQV**
12)	**QLGLWGYLVGCIAWIYIIYLLFTNVTKAAETSP-PPIRKALLNMRLFILIGWAIYPIGYAVTL**
13)	**QAWFWGLVSGIAYFVIVYDIWLGKAKKLAVA-AGGSVLKAHKTLCWFVLVGWAIYPIGYMAG-**
14)	**NAWLWGLISGAAYFVIVYDIWLGSAKKLAV-RAGGAVLSAHKTLCWFVLVGWAIYPIGYMAG-**
15)	**NAQIWGLASGIAYFVIVYDIWLGTAKKLAVA-AGGEVLKAHKALCWFVLVGWAIYPIGYMAG-**
16)	**NAWIWGLVSGIAYFAIAYEIWLGEASKLAK-AAGGNVLEAHKILCWFVLVGWAIYPLGYMLG-**
17)	**FSWQWGLASGIAYFVIVYIVWFGEAKKLAVS-AGGEVLKAHNILCWFVLVGWAIYPLGYILG-**

Some marked residues (T and P) are conserved in bacteriorhodopsin, and hence should not be viewed as specifically involved in carotenoid binding in xanthorhodopsin and other proteins.

Section III. Alignment of xanthorhodopsin and gloeobacter rhodopsin.

Helical segments in 3DLL structure of xanthorhodopsin (XR) are underlined, helices are labeled by letters A, B, C, D, F and G. Identical residues are labeled with red. Upper numbers correspond to gloeobacter rhodopsin sequence from *G. violaceous* genome (NC_005125.1) (*9*), lower numbers, to xanthorhodopsin (*3*), genome sequence NC_007677.1 (*4*) and crystal structure (*2*). Residues of xanthorhodopsin not resolved in the crystal structure are shown with grey font.

```
GR MLMTVFSSAPELALLGSTFAQVDPSNLSVSDSLT
XR MLQELPT.........................LT

   35                           64      72             87        97
GR YGQFNLVYNAFSFAIAAMFASALFFFSAQALVGQRYRLALLVSAIVVSIAGYHYFRIFNSWDAAYVL
XR PGQYSLVFNMFSFTVATMTASFVFFVLARNNVAPKYRISMMVSALVVFIAGYHYFRITSSWEAAYAL
   10           A               39      47          B  62        72

                116  121         132        144                     167
GR ENGVYSLTSEKFNDAYRYVDWLLTVPLLLVETVAVLTLPAKEARPLLIKLTVASVLMIATGYPGEI
XR QNGMYQPTGELFNDAYRYVDWLLTVPLLTVELVLVMGLPKNERGPLAAKLGFLAALMIVLGYPGEV
                91   96    C    107          119        D          142

           173 178            193         205                       231
GR SDDIT---TRIIWGTVSTIPFAYILYVLWVELSRSLVRQPAAVQTLVRNMRWLLLLSWGVYPIAYLL
XR SENAALFGTRGLWGFLSTIPFVWILYILFTQLGDTIQRQSSRVSTLLGNARLLLLATWGFYPIAYMI
           151 156     E      171         183          F            209

  232          239                               274
GR PM-----LGVSGTSAAVGVQVGYTIADVLAKPVFGLLVFAIALVKTKADQESSEPHAAIGAAANKSG
XR PMAFPEAFPSNTPGTIVALQVGYTIADVLAKAGYGVLIYNIAKAKSEEEGFNVSEMVEPATASA
  210         222              G                 257             273

      298
GR GSLIS
XR
```

REFERENCES

1. Lutnaes, B. F., Oren, A., and Liaaen-Jensen, S. (2002) New C_{40}-carotenoid acyl glycoside as principal carotenoid in *Salinibacter ruber*, an extremely halophilic eubacterium, *J. Nat. Prod. 65*, 1340-134.

2. Luecke, H., Schobert, B., Stagno, J., Imasheva, E. S., Wang, J. M., Balashov, S. P., and Lanyi, J. K. (2008) Crystallographic structure of xanthorhodopsin, the light-driven proton pump with a dual chromophore, *Proc. Natl. Acad. Sci. U.S.A. 105*, 16561-16565.

3. Balashov, S. P., Imasheva, E. S., Boichenko, V. A., Antón, J., Wang, J. M., and Lanyi, J. K. (2005) Xanthorhodopsin: A proton pump with a light-harvesting carotenoid antenna, *Science 309*, 2061-2064.

4. Mongodin, E. F., Nelson, K. E., Daugherty, S., DeBoy, R. T., Wister, J., Khouri, H., Weidman, J., Walsh, D. A., Papke, R. T., Sanchez Perez, G., Sharma, A. K., Nesbø, C. L., MacLeod, D., Bapteste, E., Doolittle, W. F., Charlebois, R. L., Legault, B., and Rodriguez-Valera, F. (2005) The genome of *Salinibacter ruber*: Convergence and gene exchange among hyperhalophilic bacteria and archaea, *Proc. Natl. Acad. Sci. U.S.A. 102*, 18147-18152.

5. Sharma, A. K., Zhaxybayeva, O., Papke, R. T., and Doolitle, W. F. (2008) Actinorhodopsins: proteorhodopsin-like gene sequences found predominantly in non-marine environments, *Environental Microbiology 10*, 1039-1056.

6. Giovannoni, S. J., Hayakawa, D. H., Tripp, H. J., Stingl, U., Givan, S. A., Cho, J.-C., Oh, H.-M., Kitner, J. B., Vergin, K. L., and Rappé, M. S. (2008) The small genome of an abundant coastal ocean methylotroph, *Environ. Microbiol. 10*, 1771–1782.

7. Ovchinnikov, Y. A., Abdulaev, N. G., Feigina, M. Y., Kiselev, A. V., and Lobanov, N. A. (1979) The structural basis of the functioning of bacteriorhodopsin: An overview, *FEBS Lett. 100*, 219-224.

8. Khorana, H. G., Gerber, G. E., Herlihy, W. C., Gray, C. P., Anderegg, R. J., Nihei, K., and Biemann, K. (1979) Amino acid sequence of bacteriorhodopsin, *Proc. Natl. Acad. Sci. U.S.A. 76*, 5046-5050.

9. Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H., Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N., Shimpo, S., Takeuchi, C., Yamada, M., and Tabata, S. (2003) Complete genome structure of Gloeobacter violaceus PCC 7421, a cyanobacterium that lacks thylakoids, *DNA Research 10*, 137-145.