Proton transfer pathways in photosynthetic reaction centers analyzed by profile Hidden Markov Models and Network calculations

Eva-Maria Krammer*[1], Mirco S. Till*[1], Pierre Sebban[2] and G. Matthias Ullmann[1]

*These authors contributed equally to this work.

[1] Structural Biology/Bioinformatics, University of Bayreuth, Universitätsstr. 30, BGI, Bayreuth, 95447, Germany.

[2] Laboratoire de Chimie Physique, MR 8000, University P. XI/CNRS, Bât. 350, Faculté d'Orsay, 91405 Orsay Cedex, France

Supporting Information

Species	GenInfo Identifier (bRC)	Identity	Species (PSII)	GenInfo Identifier	Identity
Rhodobacter sphaeroides	`gi|77463861`	100 %	*Rubrivivax gelatinosus*	`gi|7416804`	42 %
Rhodobacter capsulatus	`gi|132175`	64 %	*Methylobacterium chloromethanicum*	`gi|156452396`	39 %
Dinoroseobacter shibae	`gi|118735497`	60 %	*Rhodopseudomonas viridis*	`gi|3891313`	40 %
Roseobacter denitrificans	`gi|110677547`	58 %	*Halorhodospira halophila*	`gi|121998410`	40 %
Thiocapsa roseopersicina	`gi|31621269`	40 %	*Thermochromatium tepidum*	`gi|14549194`	41 %
Jannaschia sp. CCS1	`gi|89052646`	54 %	*Bradyrhizobium sp. ORS278*	`gi|146338712`	39 %
Rhodopseudomonas palustris	`gi|91978193`	45 %	*Rhodospirillum rubrum*	`gi|83591956`	39 %
Erythrobacter sp. NAP1	`gi|85710619`	38 %			

Table 1: Sequences used for the building of the pHMM for the H subunit. The idenity refers for the H subunit of *Rb. sphaeroides*.

Species	GenInfo Identifier
Rhodobacter sphaeroides 2.4.1	`gi|77463861`
unjcultured proteobacterium DelRiverFos13D03	`gi|61653198`
Rhodobacter capsulatus	`gi|132175`
Dinoroseobacter shibae DFL 12	`gi|118735497`
Roseobacter sp. AzwK-3b	`gi|149914231`
Roseobacter denitrificans OCh 114	`gi|110677547`
Roseobacter sp. CCS2	`gi|126734371`
Roseovarius sp. 217	`gi|85703508`
Roseovarius sp. TM1035	`gi|149202062`
Jannaschia sp. CCS1	`gi|89052646`
uncultured proteobacterium	`gi|21328669`
Loktanella vestfoldensis SKA53	`gi|84514954`
Rhodopseudomonas palustris BisB5	`gi|91978193`
Rhodopseudomonas palustris HaA2	`gi|86751083`
uncultured proteobacterium DelRiverFos06H03	`gi|61653231`
Thiocapsa roseopersicina	`gi|31621269`
Rubrivivax gelatinosus	`gi|7416804`
Bradyrhizobium sp. BTAi1	`gi|148257642`
Blastochloris viridis	`gi|132177`
Halorhodospira halophila SL1	`gi|121998410`
Methylobacterium extorquens PA1	`gi|153898039`
Thermochromatium tepidum	`gi|14549194`
Rhodopseudomonas palustris CGA009	`gi|39934618`
marine gamma proteobacterium HTCC2080	`gi|119504112`
Bradyrhizobium sp. ORS278	`gi|146338712`
uncultured proteobacterium	`gi|21328605`
Rhodopseudomonas palustris BisB18	`gi|90422830`
Rhodospirillum rubrum ATCC 11170	`gi|83591956`
Methylobacterium sp. 4-46	`gi|149119551`
Rhodopseudomonas palustris BisA53	`gi|115523374`
Erythrobacter sp. NAP1	`gi|85710619`
gamma proteobacterium KT 71	`gi|88706650`
Fulvimarina pelagi HTCC2506	`gi|114707550`

Table 2: Sequences of the H subunit used for the multiple sequence alignment

Species	GenBank identifier L subunit	GenBank identifier M subunit
Rhodobacter sphaeroides	`gi|443429`	`gi|22218909`
Rhodobacter azotoformans	`gi|27530687`	`gi|27530688`
Rhodobacter capsulatus	`gi|132180`	`gi|132186`
Roseovarius sp. 217	`gi|85703519`	`gi|85703518`
Rhodovulum euryhalinum	`gi|22758840`	`gi|22758841`
Roseobacter sp. AzwK-3b	`gi|149914242`	`gi|149914241`
Dinoroseobacter shibae DFL 12	`gi|159046063`	`gi|159046064`
Rhodovulum sulfidophilum	`gi|4630787`	`gi|4630788`
Bradyrhizobium sp. ORS278	`gi|146338692`	`gi|146338693`
Sphingomonas ursincola	`gi|7527323`	`gi|7527324`
Blastomonas natatoria	`gi|7527318`	`gi|7527319`
Rubrivivax gelatinosus	`gi|1710048`	`gi|29893490`
Jannaschia sp. CCS1	`gi|89052664`	`gi|89052663`
Rhodopseudomonas palustris BisB18	`gi|90422810`	`gi|90422811`
Loktanella vestfoldensis SKA53	`gi|84515016`	`gi|84515017`
Lamprocystis purpurea	`gi|30313694`	`gi|30313695`
Roseateles depolymerans	`gi|11907541`	`gi|11907542`
Rhodobacter blasticus	`gi|2521966`	`gi|2521967`
Phaeospirillum molischianum	`gi|2529234`	`gi|2529235`
Thermochromatium Tepidum	`gi|12084454`	`gi|157201197`
Roseospirillum parvum	`gi|37935706`	`gi|37935707`
Allochromatium vinosum	`gi|6226891`	`gi|6226892`
Halorhodospira halophila SL1	`gi|121998384`	`gi|121998383`
Roseobacter litoralis	`gi|3493355`	`gi|3493356`
Erythrobacter longus	`gi|8388769`	`gi|8388770`
Rhodobacter veldkampii	`gi|27530690`	`gi|27530691`
Thiocapsa roseopersicina	`gi|42821328`	`gi|42821329`
Rhodomicrobium vannielii	`gi|2521975`	`gi|3077785`
Rhodoferax fermentans	`gi|2521969`	`gi|2521970`
Acidiphilium acidophilum	`gi|3128132`	`gi|3128133`
Rhodocyclus tenuis	`gi|2521972`	`gi|2521973`
Blastochloris viridis	`gi|132183`	`gi|132189`
Rhodospirillum photometricum	`gi|1871499`	`gi|1710052`
Allochromatium vinosum	`gi|1230548`	`gi|2521961`
Rhodomicrobium vannielii	`gi|3077784`	`gi|2521976`
Roseococcus thiosulfatophilus	`gi|27496726`	`gi|27496727`
Acidiphilium organovorum	`gi|6093940`	`gi|3046846`
Porphyrobacter neustonensis	`gi|3097506`	`gi|3097507`
Rubritepida flocculans	`gi|27496608`	`gi|27496609`
Erythrobacter litoralis	`gi|3077787`	`gi|3077788`
Fulvimarina pelagi HTCC2506	`gi|114707541`	`gi|114707542`
Porphyrobacter sanguineus	`gi|3062896`	`gi|3062897`
Porphyrobacter tepidarius	`gi|3982480`	`gi|3982481`
Rhodoplanes elegans	`gi|5306110`	`gi|5306111`

Species	GenBank identifier L subunit	GenBank identifier M subunit
Erythrobacter sp. JL475	`gi|110798011`	`gi|110798010`
Blastochloris sulfoviridis	`gi|3551527`	`gi|3551528`
Citromicrobium sp. JL354	`gi|110798016`	`gi|110798015`
Chloroflexus aurantiacus J-10-fl	`gi|76259509`	`gi|76259508`
Chloroflexus aggregans DSM 9485	`gi|118046575`	`gi|118046574`
Rhodospirillum rubrum ATCC 11170	`gi|83594305`	`gi|83594304`

Table 3: Sequences used for the L213/M44 alignment.

Cluster number	Residue
1	HisL190, GluL212, SerL223, GluM234, N101, N102, O89, $\mathrm{Q_B}$
2	GluL205, ArgL207, ThrL208, AspL210, HisL211, GluM22, AspH124, HisH126, N103, O87, P93, X91, X92, X94, X95, X96
3	AspL213, ThrL214, ArgL217, AsnM44, N38, N39, N45, O88, O90, P118, P119, P120
4	AspL218, ArgM29, AsnM25, AsnM28, GlnM46, TyrM51, ArgM136, N14, N17, N46, O91, O92, X106
5	ThrL226, LysH130, AspH170, GluH173, O93, O94, P121, P125
6	ArgL231, TyrM3, AsnM5, SerM8, GlnM9, N6, N8, N27, N34, N35, N40, P123, P124, X8, X9, X11, X12
7	GlnM11, ArgM13, ThrM37, N9, N41, N42, P108, P122
8	SerH227, ArgH228, N96, P137, P164, X44, X46, X49
9	GluM232, ArgH177, GlnH194, N99, N100, P126, P136, X4, X5, X6, X10
10	ArgM233, GlnM237, ArgH117, GluH122, GluH230, AspH231, N98, P82, P83, P89, P90, P99, P157, P158, P159
11	GluM236, HisH68, ArgH70, N65, N104, N105, O86, P35, P91, P92, X72, X73, X97

Table 3: Water molecules and protein residues in the network spanning from the cytoplasm to $\mathrm{Q_B}$ in the bRC from *Rb. sphaeroides*. The water molecules with the chain id M, L, and H in the PDB file are named N, O, and P, respectively, or X if they were added in this study.

Cluster number	Residue
1	HisL190, GluL212, AsnL213, SerL223, GluM232, N575, O717, O729, P626, P656, P678, Q_B
2	ThrL208, GluL210, HisL211, GluM234, ThrM237, HisH72, LysH133, SerH176, N535, N539, O747, P620, P630, P650, P652, X131
3	ArgL217, TyrM34, GlnM45, HisH178, O759, O764, O791, O862, P654, P660, P661, P673, P685, X24
4	GlnL214, AspL218, SerM20, TrpM23, AspM25, ArgM28, SerM133, ArgM134, O757, O783, P647, X136
5	SerL228, ArgL231, ThrM8, AspM43, N621, N658, O731, O780, O781, O799, O818, X16, X18, X19, X20, X21, X25, X109
6	GlnM4, TyrH179, N518, O723, O727, O732, O741, O750
7	ArgM13, GlnM11, TyrM36, LysM40, AspH149, O756, O864
8	TyrM3, TyrM7, GlnM9, ArgM226, ThrH211, X48, X49, X52, X55, X57, X59, X60, X65, X101
9	AspM230, ArgH181, AspH174, GluH177, N511, N546, O737, O819, X17
10	ArgM231, GlnM235, ArgH120, AspH125, TyrH182, GluH235, AspH236, SerH239, N502, N506, N510, N512, N555, O702, O728, O751, O758, O785, O786, X10
11	AspM2, N513, N523, O726, X53 X58, X62

Table 3: Water molecules and protein residues in the network spanning from the cytoplasm to Q_B in the bRC from *B. viridis*. The water molecules with the chain id M, L, and H in the PDB file are named N, O, and P, respectively, or X if they were added in this study.