Profile Hidden Markov Models for Analyzing Similiarities and Dissimiliarities in the Bacterial Reaction Center and Photosystem II.

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

[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

a) Example 1

Rhodobacter sphaeroides	L	207	RTPDHEDIHFRDLVGYSIGTLGIHRLGLLL	237
Blastochloris viridis	L	207	KTAEHENQYFRDVVGYSIGALSIHRLGLFL	237
Rhodobacter capsulatus	L	207	RTPDHEDIYFRDLMGYSVGTLGIHRLGLLL	237
Thermochromatium tepidum	L	215	KTSEHENIHFRDIVGYSIGALAIHRLGLFL	245
Rhodospirillum rubrum	L	208	KGPEHENIYFQDTIGYSVGTLGIHRVGLIL	238
Thermosynechococcus elongatus	D1	246	YNIVAAHGYFGRLIFQYASFNNSRSLHFFL	276
Prunus serrulata	D1	246	YNIVAAHGYFGRLVFQYASFNNSRSLHFFL	276
Pyrocystis lunula	D1	142	YSISAAHGYFGRLIFQYASFNNSRSLHFFL	172
Rhodosorus marinus	D1	233	YNIVAAHGYFGRLIFQYASFNNSRSLHFFL	263
Adenoides eludens	D1	246	YSISAAHGYFGRLIFQYASFNNSRSLHFFL	276

b) Example 2

Rhodobacter sphaeroides	L	207	RTPDHEDIHFRDLVGYSIGTLGIHRLGLLL	237
Rhodobacter capsulatus	L	207	RTPDHEDIYFRDLMGYSVGTLGIHRLGLLL	237
Rhodospirillum rubrum	L	207	KGPEHENIYFQDTIGYSVGTLGIHRVGLIL	237
Allochromatium vinosum	L	213	KTSEHENIHFRDIVGYSIGALAIHRLGLFL	243
Acidiphilium acidophilum	L	207	KFTEHEDIHFRDFIGYSIGTLGIHRLGLFL	237
Spinacia oleracea	D1	247	RFGQEEETYNIVAAHGYFGRLIFQYASFNN	277
Chara vulgaris	D1	238	KFGQEEETYNIVAAHGYFGRLIFQYASFNN	268
Adiantum capillus-veneris	D1	238	KFGQEEETYNIVAAHGYFGRLIFQYASFNN	268
Phaeodactylum tricornutum	D1	238	KFGQEEETYNIVAAHGYFGRLIFQYASFNN	268
Karlodinium micrum	D1	172	KFGQEEETYNIVAAHGYFGRLIFQYASFNN	202

Figure 1: Examples of misalgined sequences using ClustalW. (a) The MSA made with ClustalW resulted in an accurate alignment of the marker residues L212 and L213 (black) but not of L223 (green); Ten of fourty-four sequences shown as an example. (b) The MSA made with ClustalW using another set of sequences resulted in a misalignment of the marker residuesL212/213 and L223. Ten of thirty sequences are shown as an example. The sequence word around every marker residue (L212, L213 and L223) is framed in black. The alignments were done using the Gonnet250 matrix. Using other scoring matrixes did not improve the alignments. The full-length MSAs are given in the Supporting Information.

Species (bRC)	GenInfo Identifier	Identity	Species (PSII)	GenInfo Identifier	Identity
Rhodobacter sphaeroides	gi|146278067	96 %	*Thermosynechococcus elongatus*	gi|22299386	100 %
Rhodobacter azotoformans	gi|27530687	95 %	*Thermosynechococcus vulcanus*	gi|1709826	99 %
Rhodobacter capsulatus	gi|132180	78 %	*Nodularia spumigena*	gi|119512256	89 %
Rhodovulum euryhalinum	gi|22758840	72 %	*Nostoc punctiforme*	gi|23126457	90 %
Dinoroseobacter shibae	gi|118735480	73 %	*Anabaena variabilis*	gi|75908357	89 %
Roseobacter denitrificans	gi|110677518	70 %	*Trichodesmium erythraeum*	gi|113474085	89 %
Rhodovulum sulfidophilum	gi|4630787	70 %	*Chara vulgaris*	gi|108773236	88 %
Rhodopseudomonas palustris	gi|39934599	71 %	*Conocephalum conicum*	gi|5688982	88 %
Sphingomonas ursincola	gi|7527323	71 %	*Marchantia polymorpha*	gi|11466695	88 %
Blastomonas natatoria	gi|7527318	70 %	*Andreaea wilsonii*	gi|38043790	87 %
Bradyrhizobium sp. BTAi1	gi|148257662	67 %	*Bryum cellulare*	gi|27228025	88 %
Rubrivivax gelationosus	gi|1710048	69 %	*Amaranthus powellii*	gi|16119014	88 %
Rhodopseudomonas palustris	gi|91978213	71 %	*Pinus thunbergii*	gi|7524753	87 %
Rhodospirillum rubrum	gi|83594305	70 %	*Lepidium virginicum*	gi|139388890	87 %
Jannaschia sp. CCS1	gi|89052664	70 %	*Nicotiana tabacum*	gi|4589848	87 %
Loktanella vestfoldensis	gi|84515016	72 %	*Actinidia deliciosa*	gi|32493034	84 %
Halorhodospira halophila	gi|121998384	64 %	*Euglena stellata*	gi|77744807	85 %
Erythrobacter longus	gi|8388769	64 %	*Prunus glandulosa*	gi|15529723	83 %
Rhodospirillum centenum	gi|18654138	64 %	*Emiliania huxleyi*	gi|71842227	84 %
Rhodoferax fermentans	gi|2521969	64 %	*Prunus persica*	gi|15529727	82 %
Acidiphilium acidophilum	gi|3128132	65 %	*Dunaliella salina*	gi|68004437	85 %
Rhodopseudonomas viridis	gi|230260	58 %	*Cycas taitungensis*	gi|150251436	83 %
Thermochromatium tepidum	gi|12084454	65 %	*Zea mays*	gi|11467171	87 %
			Rhodomonas salina	gi|149072085	86 %
			Panax ginseng	gi|52220790	87 %
			Prochlorococcus marinus	gi|33239704	84 %
			Mastigocladus laminosus	gi|5759152	87 %

Table 1: Sequences used for the building of the L/D1 HMM model. The column Identity refers for the L subunit of *Rb. sphaeroides* and for the D1 subunit to *T. elongatus*.

Species (bRC)	GenInfo Identifier	Identity	Species (PSII)	GenInfo Identifier	Identity
Rhodobacter sphaeroides	gi|443430	100 %	*Thermosynechococcus elongatus*	gi|22297998	100 %
Rhodobacter capsulatus	gi|132186	76 %	*Anabaena variabilis*	gi|75907465	93 %
Rhodovulum sulfidophilum	gi|4630788	69 %	*Nostoc punctiforme*	gi|23129896	92 %
Rhodobacter veldkampii	gi|27530691	74 %	*Ostreococcus tauri*	gi|113170413	90 %
Dinoroseobacter shibae	gi|118735481	66 %	*Nephroselmis olivacea*	gi|11467794	90 %
Jannaschia sp. CCS1	gi|89052663	66 %	*Oltmannsiellopsis viridis*	gi|108773343	90 %
Rubrixivax gelatinosus	gi|1710051	66 %	*Chlorella vulgaris*	gi|7524799	90 %
Rhodopseudomonas palustris	gi|91978212	64 %	*Pseudendoclonium akinetum*	gi|108796979	89 %
Acidiphilium rubrum	gi|3046828	63 %	*Chara vulgaris*	gi|108773238	90 %
Roseobacter denitrificans	gi|110677517	62 %	*Cyanophora paradoxa*	gi|11467390	89 %
Methylobacterium extorquens	gi|153897616	64 %	*Marchantia polymorpha*	gi|11466698	90 %
Roseateles depolymerans	gi|11907542	63 %	*Lactuca sativa*	gi|78675163	89 %
Thermochromatium tepidum	gi|12084455	62 %	*Anthoceros formosae*	gi|28202167	90 %
Halorhodospira halophila	gi|121998383	60 %	*Odontella sinensis*	gi|11467543	89 %
Roseospirillum parvum	gi|37935707	61 %	*Solanum lycopersicum*	gi|89280630	89 %
Rhodospirillum rubrum	gi|83594304	64 %	*Pinus thunbergii*	gi|7524716	89 %
Fulvimarina pelag	gi|114707542	59 %	*Drimys granadensis*	gi|115604929	89 %
Psychroflexus torquis	gi|91222803	66 %	*Triticum aestivum*	gi|14017557	88 %
Rhodospeudomonas viridis	gi|230261	49 %	*Nicotiana tabacum*	gi|11465949	89 %
Roseobacter sp. CCS2	gi|126734150	72 %	*Trichodesmium erythraeum*	gi|113474993	86 %
Loktanella vestfoldensis	gi|4630788	71 %	*Pisum sativum*	gi|473493	88 %
Phaeospirillum molischianum	gi|2529235	69 %	*Panax ginseng*	gi|38455760	88 %
			Stigeoclonium helveticum	gi|115349988	88 %
			Emiliania huxleyi	gi|71842284	86 %
			Zea mays	gi|11467176	87 %
			Rhodomonas salina	gi|149072099	86 %
			Guillardia theta	gi|11467626	85 %
			Prochlorococcus marinus	gi|124022552	84 %

Table 2: Sequences used for the building of the M/D2 HMM model. The column Identity refers for the M subunit of *Rb. sphaeroides* and for the D2 subunit to *T. elongatus*.

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|146278067	*Rb. sphaeroides ATCC 17025*	gi|6226891	*Allochromatium vinosum*
gi|27530687	*Rb. azotoformans*	gi|21328595	*up*
gi|132180	*Rb. capsulatus*	gi|6518472	*ap MBIC3951*
gi|85703519	*Roseovarius sp. 217*	gi|119504121	*mgp HTCC2080*
gi|149202073	*Roseovarius sp. TM1035*	gi|50593423	*up*
gi|21328651	*up*	gi|50593421	*up*
gi|61653209	*up DelRiverFos13D03*	gi|121998384	*Halorhodospira halophila SL1*
gi|22758840	*Rhodovulum euryhalinum*	gi|88706660	*gp KT 71*
gi|149914242	*Roseobacter sp. AzwK-3b*	gi|3493355	*Roseobacter litoralis*
gi|118735480	*Dinoroseobacter shibae DFL 12*	gi|5019582	*Erythrobacter sp. MBIC3960*
gi|50952773	*up eBACred25D05*	gi|8388769	*Erythrobacter longus*
gi|110677518	*Roseobacter denitrificans*	gi|3493352	*Roseobacter denitrificans*
gi|4630787	*Rhodovulum sulfidophilum*	gi|18654138	*Rhodospirillum centenum*
gi|146338692	*Bradyrhizobium sp. ORS278*	gi|27530690	*Rb. veldkamgpii*
gi|126734151	*Roseobacter sp. CCS2*	gi|85710602	*Erythrobacter sp. NAP1*
gi|39934599	*Rps. palustris CGA009*	gi|4650839	*Blastomonas sp. NT12*
gi|7527323	*Sphingomonas ursincola*	gi|42821328	*Thiocapsa roseopersicina*
gi|7527318	*Blastomonas natatoria*	gi|3046827	*Acidiphilium rubrum*
gi|148257662	*Bradyrhizobium sp. BTAi1*	gi|2521975	*Rhodomicrobium vannielii*
gi|1710048	*Rubrivivax gelatinosus*	gi|57231451	*Sphingomonas sp. PB180*
gi|91978213	*Rps. palustris BisB5*	gi|2521969	*Rhodoferax fermentans*
gi|86751103	*Rps. palustris HaA2*	gi|3128132	*Acidiphilium acidophilum*
gi|83594305	*Rhodospirillum rubrum ATCC 11170*	gi|2521972	*Rhodocyclus tenuis*
gi|89052664	*Jannaschia sp. CCS1*	gi|132183	*Rhodopseudomonas viridis*
gi|90422810	*Rps. palustris BisB18*	gi|1710049	*Rhodospirillum photometricum*
gi|84515016	*Loktanella vestfoldensis SKA53*	gi|124579201	*mgp HTCC2246*
gi|30313694	*Lamgprocystis purpurea*	gi|27496726	*Roseococcus thiosulfatophilus*
gi|11907541	*Roseateles depolymerans*	gi|6093939	*Acidiphilium multivorum*
gi|115523354	*Rps. palustris BisA53*	gi|6093940	*Acidiphilium organovorum*
gi|2521966	*Rb. blasticus*	gi|3097506	*Porphyrobacter neustonensis*
gi|2529234	*Phaeospirillum molischianum*	gi|2521978	*Thiocystis gelatinosa*
gi|2338775	*Ectothiorhodospira shaposhnikovii*	gi|57231454	*Sphingomonas sp. PB229*

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|12084454	*T. tepidum*	gi|6093938	*Acidiphilium cryptum*
gi|37935706	*Roseospirillum parvum*	gi|27496608	*Rubritepida flocculans*
gi|153897615	*Methylobacterium extorquens PA1*	gi|124579174	*mgp HTCC2148*
gi|61653248	*up DelRiverFos06H03*	gi|3097503	*Erythromicrobium ramosum*
gi|3551492	*Rps. palustris*	gi|15809648	*up*
gi|3077787	*Erythrobacter litoralis*	gi|15809641	*up*
gi|3551483	*Erythrobacter sp.*	gi|15809635	*up*
gi|114707541	*Fulvimarina pelagi HTCC2506*	gi|15809656	*up*
gi|3062896	*Porphyrobacter sanguineus*	gi|15809639	*up*
gi|3982480	*Porphyrobacter tepidarius*	gi|19569282	*Rhodoplanes roseus*
gi|5306110	*Rhodoplanes elegans*	gi|15809652	*up*
gi|26986419	*Roseovarius tolerans*	gi|51317890	*Rb. sp. HTCC515*
gi|19682868	*up*	gi|15809654	*up*
gi|110798011	*Erythrobacter sp. JL475*	gi|15809637	*up*
gi|3551527	*Blastochloris sulfoviridis*	gi|26986413	*Roseovarius mucosus*
gi|110798016	*Citromicrobium sp. JL354*	gi|26986409	*Roseibium sp. DFL-11*
gi|51317892	*bp HTCC528*	gi|26985517	*Ahrensia sp. DFL-13*
gi|149124078	*Methylobacterium sp. 4-46*	gi|26986411	*Sphingomonas sp. HEL-42*

Table 4: All used sequences of the L subunit for the profile alignment using the build pHMM files. Used Abbrebations: up, ap, bp, mgp, and ub, for uncultured proteobacterium, for alpha proteobacterium, beta proteobacterium, for marine gamma proteobacterium and, for uncultured bacterium, respectively.

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|146278066	*Rb. sphaeroides ATCC 17025*	gi|6518473	*ap MBIC3951*
gi|27530688	*Rb. azotoformans*	gi|7527324	*Sphingomonas ursincola*
gi|132186	*Rhodobacter capsulatus*	gi|7527319	*Blastomonas natatoria*
gi|21328652	*up*	gi|3982481	*Porphyrobacter tepidarius*
gi|126734150	*Roseobacter sp. CCS2*	gi|32765600	*Erythrobacter sp. BA13*
gi|61653208	*up*	gi|3128133	*Acidiphilium acidophilum*
gi|84515017	*Loktanella vestfoldensis SKA53*	gi|3097504	*Erythromicrobium ramosum*
gi|85710603	*Erythrobacter sp. NAP1*	gi|4650840	*Blastomonas sp. NT12*
gi|149202072	*Roseovarius sp. TM1035*	gi|37379369	*Rhodospirillum rubrum*
gi|68138148	*ub*	gi|3077788	*Erythrobacter litoralis*
gi|4630788	*Rhodovulum sulfidophilum*	gi|2521973	*Rhodocyclus tenuis*
gi|85703518	*Roseovarius sp. 217*	gi|15081346	*ap R2A163*
gi|149914241	*Roseobacter sp. AzwK-3b*	gi|2521979	*Thiocystis gelatinosa*
gi|27530691	*Rb. veldkampii*	gi|2521976	*Rhodomicrobium vannielii*
gi|118735481	*Dinoroseobacter shibae DFL 12*	gi|132189	*Blastochloris viridis*
gi|146338693	*Bradyrhizobium sp. ORS278*	gi|3097507	*Porphyrobacter neustonensis*
gi|88706659	*gamma proteobacterium KT 71*	gi|3046846	*Acidiphilium organovorum*
gi|89052663	*Jannaschia sp. CCS1*	gi|110798015	*Citromicrobium sp. JL354*
gi|1710051	*Rubrivivax gelatinosus*	gi|32765594	*Erythrobacter sp. MG3*
gi|149124077	*Methylobacterium sp. 4-46*	gi|32765592	*Erythrobacter sp. COL13*
gi|91978212	*Rps. palustris BisB5*	gi|3046843	*Acidiphilium multivorum*
gi|3046828	*Acidiphilium rubrum*	gi|57231448	*Sphingomonas sp. PB56*
gi|132185	*Roseobacter denitrificans OCh 114*	gi|15667219	*ub*
gi|156449437	*Methylobacterium chloromethanicum*	gi|32765596	*Erythrobacter sp. MG22*
gi|39934600	*Rps. palustris CGA009*	gi|3493353	*Roseobacter denitrificans*
gi|2521967	*Rb. blasticus*	gi|2521961	*Allochromatium vinosum*
gi|86751102	*Rps. palustris HaA2*	gi|15081344	*ap R2A130*
gi|90422811	*Rps. palustris BisB18*	gi|146424629	*Rhodopseudomanslichen*
gi|11907542	*Roseateles depolymerans*	gi|15667215	*ub*
gi|6226892	*Allochromatium vinosum*	gi|50593194	*Roseobacter sp. BS90*
gi|61653247	*up DelRiverFos06H03*	gi|42821329	*Thiocapsa roseopersicina*
gi|121998383	*Halorhodospira halophila SL1*	gi|3493356	*Roseobacter litoralis*

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|37935707	*Roseospirillum parvum*	gi|22758841	*Rhodovulum euryhalinum*
gi|21328596	*up*	gi|32765598	*Erythrobacter sp. NJ3Y*
gi|50952772	*up eBACred25D05*	gi|50593210	*up*
gi|115523355	*Rps. palustris BisA53*	gi|2521970	*Rhodoferax fermentans*
gi|119504120	*mgp HTCC2080*	gi|50593200	*marine bacterium BS110*
gi|83594304	*Rhodospirillum rubrum ATCC 11170*	gi|15667217	*ub*
gi|132187	*Rhodospirillum rubrum*	gi|50593208	*up*
gi|114707542	*Fulvimarina pelagi HTCC2506*	gi|50593212	*up*
gi|91222803	*Psychroflexus torquis ATCC 700755*	gi|5306111	*Rhodoplanes elegans*
gi|2338776	*Ectothiorhodospira shaposhnikovii*	gi|3551493	*Rps. palustris*
gi|68138150	*Citromicrobium sp. CV44*	gi|27496727	*Roseococcus thiosulfatophilus*
gi|3046840	*Acidiphilium cryptum*	gi|27496609	*Rubritepida flocculans*
gi|30313695	*Lamprocystis purpurea*	gi|50593198	*Roseobacter sp. SYOP2*
gi|50593196	*Roseobacter sp. S03*	gi|15667213	*ub*
gi|2529235	*Phaeospirillum molischianum*	gi|15667221	*ub*
gi|5019583	*Erythrobacter sp. MBIC3960*	gi|15081328	*ub*
gi|32765590	*Erythrobacter sp. NAP1*	gi|15081334	*ub*
gi|50593204	*up*	gi|3551528	*Blastochloris sulfoviridis*
gi|110798010	*Erythrobacter sp. JL475*	gi|1710052	*Rhodospirillum photometricum*
gi|50593221	*up*	gi|15667225	*ub*
gi|2521964	*Erythrobacter longus*	gi|15081364	*ap R2A62*
gi|3551484	*Erythrobacter sp.*	gi|51317894	*bp HTCC528*
gi|50593219	*up*	gi|124579202	*mgp HTCC2246*
gi|3062897	*Porphyrobacter sanguineus*	gi|124579175	*mgp HTCC2148*
gi|50593202	*up*		

Table 5: All used sequences of the M subunit for the profile alignment using the build pHMM files. Used Abbrebations: up, ap, bp, mgp, and ub, for uncultured proteobacterium, for alpha proteobacterium, beta proteobacterium, for marine gamma proteobacterium and, for uncultured bacterium, respectively.

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|22299386	*Thermosynechococcus elongatus BP-1*	gi|15529715	*Prunus serrulata var. pubescens*
gi|119512256	*Nodularia spumigena CCY9414*	gi|87301309	*Synechococcus sp. WH 5701*
gi|75908357	*Anabaena variabilis ATCC 29413*	gi|113170441	*Ostreococcus tauri*
gi|119483344	*Lyngbya sp. PCC 8106*	gi|116075458	*Synechococcus sp. RS9916*
gi|56750175	*Synechococcus elongatus PCC 6301*	gi|148242540	*Synechococcus sp. RCC307*
gi|113474085	*Trichodesmium erythraeum IMS101*	gi|15529751	*Prunus takesimensis*
gi|16329178	*Synechocystis sp. PCC 6803*	gi|32493032	*Actinidia deliciosa*
gi|5759152	*Mastigocladus laminosus*	gi|111378709	*Paulinella chromatophora*
gi|108773236	*Chara vulgaris*	gi|32493020	*Actinidia polygama*
gi|86605956	*Synechococcus sp. JA-3-3Ab*	gi|21449452	*Thrixspermum formosanum*
gi|5688982	*Conocephalum conicum*	gi|15529719	*Prunus serrulata var. tomentella*
gi|11466695	*Marchantia polymorpha*	gi|32493022	*Actinidia melanandra*
gi|22711934	*Chaetosphaeridium globosum*	gi|12963407	*Prunus x yedoensis*
gi|86607868	*Synechococcus sp. JA-2-3B'a(2-13)*	gi|116073752	*Synechococcus sp. RS9916*
gi|28202164	*Anthoceros formosae*	gi|88807295	*Synechococcus sp. WH 7805*
gi|18860290	*Psilotum nudum*	gi|15529749	*Prunus spinulosa*
gi|67924669	*Crocosphaera watsonii WH 8501*	gi|15529713	*Prunus padus*
gi|34501422	*Physcomitrella patens subsp. patens*	gi|15529709	*Prunus leveilleana*
gi|126657141	*Cyanothece sp. CCY0110*	gi|15529717	*Prunus serrulata var. spontanea*
gi|124112079	*Chlorokybus atmophyticus*	gi|60117202	*Huperzia lucidula*
gi|69217624	*Ginkgo biloba*	gi|76262849	*Tetraselmis suecica*
gi|159161234	*Cuscuta exaltata*	gi|15983604	*Prunus leveilleana*
gi|153012231	*Medicago truncatula*	gi|116070816	*Synechococcus sp. BL107*
gi|27228057	*Plagiobryum zieri*	gi|30352077	*Adiantum capillus-veneris*
gi|118430282	*Agrostis stolonifera*	gi|115443567	*Bigelowiella natans*
gi|48478755	*Saccharum hybrid cultivar SP-80-3280*	gi|124302882	*Angiopteris evecta*
gi|14017552	*Triticum aestivum*	gi|32493026	*Actinidia eriantha*
gi|118430368	*Hordeum vulgare subsp. vulgare*	gi|148242539	*Synechococcus sp. RCC307*
gi|159161140	*Ceratophyllum demersum*	gi|52220790	*Panax ginseng*
gi|11467171	*Zea mays*	gi|114329948	*Nandina domestica*
gi|139387434	*Phaseolus vulgaris*	gi|91983972	*Vitis vinifera*
gi|122893970	*Ranunculus macranthus*	gi|113954773	*Synechococcus sp. CC9311*

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|69217632	*Yucca schidigera*	gi|41179021	*Chlamydomonas reinhardtii*
gi|50346762	*Nymphaea alba*	gi|11467381	*Cyanophora paradoxa*
gi|27228025	*Bryum cellulare*	gi|33866451	*Synechococcus sp. WH 8102*
gi|75755639	*Acorus calamus*	gi|32493024	*Actinidia macrosperma*
gi|69217630	*Typha latifolia*	gi|62199492	*Dunaliella tertiolecta*
gi|11466764	*Oryza sativa*	gi|15529731	*Prunus salicina*
gi|157325510	*Ipomoea purpurea*	gi|159902787	*Prochlorococcus marinus str. MIT 9211*
gi|91214123	*Glycine max*	gi|78211859	*Synechococcus sp. CC9605*
gi|7525013	*Arabidopsis thaliana*	gi|90994403	*Porphyra yezoensis*
gi|11465935	*Nicotiana tabacum*	gi|78185380	*Synechococcus sp. CC9902*
gi|108802623	*Eucalyptus globulus subsp. globulus*	gi|51209867	*Gracilaria tenuistipitata var. liui*
gi|56549784	*Pellia appalachiana*	gi|32493030	*Actinidia chinensis*
gi|78103231	*Phalaenopsis aphrodite subsp. formosana*	gi|37520348	*Gloeobacter violaceus PCC 7421*
gi|114107026	*Liriodendron tulipifera*	gi|11465678	*Porphyra purpurea*
gi|114329636	*Citrus sinensis*	gi|115349985	*Stigeoclonium helveticum*
gi|149390335	*Dioscorea elephantipes*	gi|118411049	*Phaeodactylum tricornutum*
gi|149390520	*Buxus microphylla*	gi|11467479	*Odontella sinensis*
gi|156618807	*Cuscuta reflexa*	gi|15529707	*Prunus japonica var. nakaii*
gi|7524753	*Pinus thunbergii*	gi|15529735	*Prunus choreiana*
gi|32480823	*Calycanthus floridus var. glaucus*	gi|77744782	*Euglena anabaena*
gi|149390422	*Illicium oligandrum*	gi|118411180	*Thalassiosira pseudonana*
gi|7524765	*Chlorella vulgaris*	gi|77744792	*Euglena mutabilis*
gi|116617088	*Coffea arabica*	gi|77744785	*Euglena granulata*
gi|108773358	*Oltmannsiellopsis viridis*	gi|15529729	*Prunus persica*
gi|2764565	*Picea abies*	gi|32493034	*Actinidia deliciosa*
gi|114199077	*Amaranthus retroflexus*	gi|77744807	*Euglena stellata*
gi|16904362	*Amaranthus powellii*	gi|15529723	*Prunus glandulosa*
gi|149390248	*Chloranthus spicatus*	gi|71842227	*Emiliania huxleyi*
gi|159106782	*Cuscuta obtusiflora*	gi|68004437	*Dunaliella salina*
gi|121720593	*Nuphar advena*	gi|34980272	*Euglena myxocylindracea*
gi|108773067	*Scenedesmus obliquus*	gi|150251436	*Cycas taitungensis*
gi|29565719	*Pinus koraiensis*	gi|126660350	*Cyanothece sp. CCY0110*
gi|11465541	*Cyanidium caldarium*	gi|77744799	*Lepocinclis spirogyroides*

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|68164783	*Cucumis sativus*	gi|15529753	*Prunus grayana*
gi|30468083	*Cyanidioschyzon merolae strain 10D*	gi|15529733	*Prunus tomentosa*
gi|158335322	*Acaryochloris marina MBIC11017*	gi|12620388	*Prunus serrulata var. quelpaertensis*
gi|67968296	*Selaginella uncinata*	gi|12620386	*Prunus sargentii*
gi|32493018	*Actinidia kolomikta*	gi|16224269	*Prunus leveilleana var. pendula*
gi|108796909	*Pseudendoclonium akinetum*	gi|83596386	*Fucus vesiculosus*
gi|32493016	*Actinidia arguta*	gi|15529705	*Prunus buergeriana*
gi|149072085	*Rhodomonas salina*	gi|15529743	*Prunus apetela*
gi|11466347	*Mesostigma viride*	gi|21913583	*Flintiella sanguinaria*
gi|108796812	*Staurastrum punctulatum*	gi|15529725	*Prunus mume*
gi|153805612	*Leptosira terrestris*	gi|15529739	*Prunus serrulata var. lannesiana*
gi|148239120	*Synechococcus sp. WH 7803*	gi|21913575	*Bangiopsis subsimplex*
gi|11467647	*Guillardia theta*	gi|12620390	*Prunus subhirtella var. ascendens*
gi|87124059	*Synechococcus sp. RS9917*	gi|15529721	*Prunus armeniaca var. ansu*
gi|126022810	*Amborella trichopoda*	gi|15529741	*Prunus incisa*
gi|15529711	*Prunus maximowiczii*	gi|21913585	*Porphyridium aerugineum*
gi|11467759	*Nephroselmis olivacea*	gi|21913601	*Chroomonas sp. SAG 980-1*
gi|33866004	*Synechococcus sp. WH 8102*		

Table 6: All used sequences of the D1 subunit for the profile alignment using the build pHMM files.

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|22297998	*Thermosynechococcus elongatus BP-1*	gi|33328123	*Widdringtonia cedarbergensis*
gi|86608336	*Synechococcus sp. JA-2-3B'a2-13*	gi|45386417	*Xyris jupicai*
gi|86606984	*Synechococcus sp. JA-3-3Ab*	gi|27435898	*Sagittaria latifolia*
gi|126656960	*Cyanothece sp. CCY0110*	gi|33328099	*Piper betle*
gi|148241358	*Synechococcus sp. RCC307*	gi|5738988	*Coleochaete nitellarum*
gi|56752457	*Synechococcus elongatus PCC 6301*	gi|45386450	*Philesia magellanica*
gi|17231782	*Nostoc sp. PCC 7120*	gi|11022870	*Cabomba caroliniana*
gi|11467794	*Nephroselmis olivacea*	gi|138278115	*Phyllocladus alpinus*
gi|33865211	*Synechococcus sp. WH 8102*	gi|45386453	*Prosartes trachycarpa*
gi|108773343	*Oltmannsiellopsis viridis*	gi|33328072	*Euonymus alatus*
gi|87124873	*Synechococcus sp. RS9917*	gi|33317750	*Cedrus deodara*
gi|7524799	*Chlorella vulgaris*	gi|130750346	*Aphelia brizula*
gi|131298	*Synechococcus sp. PCC 7002*	gi|45386414	*Strelitzia reginae*
gi|1709835	*Prochlorothrix hollandica*	gi|45386393	*Helmholtzia glaberrima*
gi|78213509	*Synechococcus sp. CC9605*	gi|37721077	*Roystonea princeps*
gi|108796979	*Pseudendoclonium akinetum*	gi|45386396	*Cyperus papyrus*
gi|78184245	*Synechococcus sp. CC9902*	gi|9081958	*Nymphaea odorata*
gi|159161155	*Ceratophyllum demersum*	gi|11022879	*Cercidiphyllum japonicum*
gi|108773238	*Chara vulgaris*	gi|45386462	*Tricyrtis sp. Waterway s.n.*
gi|11467390	*Cyanophora paradoxa*	gi|45386423	*Doryanthes palmeri*
gi|159026300	*Microcystis aeruginosa PCC 7806*	gi|37721050	*Anticlea elegans*
gi|11466698	*Marchantia polymorpha*	gi|37721047	*Stemona tuberosa*
gi|78675163	*Lactuca sativa*	gi|37721143	*Aphyllanthes monspeliensis*
gi|28202167	*Anthoceros formosae*	gi|37721068	*Mayaca fluviatilis*
gi|11465589	*Cyanidium caldarium*	gi|11022902	*Liriodendron tulipifera*
gi|32480838	*Calycanthus floridus var. glaucus*	gi|37721074	*Philydrum lanuginosum*
gi|109156584	*Oryza sativa indica cultivar-group*	gi|37721065	*Hydrothrix gardneri*
gi|16330392	*Synechocystis sp. PCC 6803*	gi|33317756	*Cycas revoluta*
gi|34500909	*Amborella trichopoda*	gi|56236596	*Abies lasiocarpa*
gi|69217531	*Ginkgo biloba*	gi|37721137	*Xeronema callistemon*
gi|87301087	*Synechococcus sp. WH 5701*	gi|37721155	*Smilacina racemosa*
gi|18860308	*Psilotum nudum*	gi|27435873	*Ephedra sinica*

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|114329651	*Citrus sinensis*	gi|37721152	*Lomandra longifolia*
gi|121720608	*Nuphar advena*	gi|11022864	*Acorus calamus*
gi|118430287	*Agrostis stolonifera*	gi|33317747	*Bowenia serrulata*
gi|114329741	*Platanus occidentalis*	gi|37721167	*Yucca glauca*
gi|113953456	*Synechococcus sp. CC9311*	gi|11022896	*Illicium parviflorum*
gi|114329963	*Nandina domestica*	gi|33317753	*Ceratozamia miqueliana*
gi|131296	*Secale cereale*	gi|37721029	*Butomus umbellatus*
gi|69217537	*Typha latifolia*	gi|11022867	*Asarum canadense*
gi|118410982	*Phaeodactylum tricornutum*	gi|45386459	*Smilax rotundifolia*
gi|78103248	*Phalaenopsis aphrodite*	gi|37721035	*Tieldia glutinosa*
gi|48478759	*Saccharum hybrid cultivar SP-80-3280*	gi|400884	*Euglena gracilis*
gi|149390349	*Dioscorea elephantipes*	gi|138278112	*Cunninghamia lanceolata*
gi|159106849	*Lolium perenne*	gi|11022893	*Gnetum gnemon*
gi|69217539	*Yucca schidigera*	gi|37721149	*Chlorophytum comosum*
gi|34501420	*Physcomitrella patens subsp. patens*	gi|8132548	*Zamia furfuracea*
gi|110227074	*Populus alba*	gi|37721041	*Narthecium ossifragum*
gi|50346777	*Nymphaea alba*	gi|37544959	*Trimenia moorei*
gi|11466770	*Oryza sativa japonica cultivar-group*	gi|89363015	*Euglena gracilis*
gi|119486382	*Lyngbya sp. PCC 8106*	gi|37721116	*Iris missouriensis*
gi|149390263	*Chloranthus spicatus*	gi|11022882	*Dioscorea bulbifera*
gi|124112082	*Chlorokybus atmophyticus*	gi|37521892	*Gloeobacter violaceus PCC 7421*
gi|126507951	*Lactuca sativa*	gi|37721062	*Ensete ventricosum*
gi|115605017	*Piper cenocladum*	gi|11022885	*Drimys winteri*
gi|225115	*Chlamydomonas reinhardtii*	gi|45386408	*Kingia australis*
gi|11467543	*Odontella sinensis*	gi|45386420	*Agapanthus africanus*
gi|114107128	*Daucus carota*	gi|37721119	*Ixiolirion tataricum*
gi|51209878	*Gracilaria tenuistipitata var. liui*	gi|45386435	*Calochortus apiculatus*
gi|89280630	*Solanum lycopersicum*	gi|27435879	*Hydrastis canadensis*
gi|108796815	*Staurastrum punctulatum*	gi|68004454	*Dunaliella salina*
gi|139389259	*Aethionema grandiflorum*	gi|45386447	*Tripladenia cunninghamii*
gi|122893985	*Ranunculus macranthus*	gi|37721053	*Ananas comosus*
gi|68164798	*Cucumis sativus*	gi|116009001	*Phyllostachys pubescens*
gi|13518433	*Lotus japonicus*	gi|82698228	*Gonyaulax polyedra*

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|150251458	*Cycas taitungensis*	gi|45386438	*Campynema lineare*
gi|7524716	*Pinus thunbergii*	gi|45386399	*Ecdeiocolea monostachya*
gi|115604929	*Drimys granadensis*	gi|56683016	*Microcycas calocoma*
gi|139388905	*Lepidium virginicum*	gi|154466712	*Olea europaea*
gi|14017557	*Triticum aestivum*	gi|45386402	*Elegia fenestrata*
gi|114804260	*Morus indica*	gi|130750349	*Trithuria submersa*
gi|131291	*Galdieria sulphuraria*	gi|83584433	*Heterocapsa triquetra*
gi|88696765	*Lemna minor*	gi|90994548	*Porphyra yezoensis*
gi|11465949	*Nicotiana tabacum*	gi|41581189	*Amphidinium operculatum*
gi|115531914	*Pelargonium x hortorum*	gi|13374964	*Amphidinium carterae*
gi|119512824	*Nodularia spumigena CCY9414*	gi|55585028	*Adenoides eludens*
gi|22711937	*Chaetosphaeridium globosum*	gi|62432771	*Porphyridium aerugineum*
gi|118411148	*Thalassiosira pseudonana*	gi|62432791	*Heterosigma akashiwo*
gi|153805577	*Leptosira terrestris*	gi|5739000	*Anthoceros punctatus*
gi|91214138	*Glycine max*	gi|62432775	*Palmaria palmata*
gi|124302903	*Angiopteris evecta*	gi|45386444	*Medeola virginiana*
gi|28261712	*Atropa belladonna*	gi|62432807	*Glaucocystis nostochinearum*
gi|108773048	*Scenedesmus obliquus*	gi|62432757	*Bangia atropurpurea*
gi|7525028	*Arabidopsis thaliana*	gi|62432793	*Pylaiella littoralis*
gi|29337172	*Pisum sativum*	gi|62432787	*Pavlova gyrans*
gi|115443541	*Bigelowiella natans*	gi|62432789	*Pavlova lutheri*
gi|139389637	*Arabis hirsuta*	gi|83638135	*Rhodella violacea*
gi|52220805	*Panax ginseng*	gi|62432761	*Cyanidium sp. Sybil cave*
gi|29565677	*Pinus koraiensis*	gi|62432767	*Galdieria sulphuraria*
gi|30468115	*Cyanidioschyzon merolae strain 10D*	gi|83638137	*Flintiella sanguinaria*
gi|13518320	*Oenothera elata subsp. hookeri*	gi|62432763	*Galdieria maxima*
gi|139390147	*Lobularia maritima*	gi|62432795	*Skeletonema costatum*
gi|139388089	*Barbarea verna*	gi|62432765	*Galdieria sulphuraria*
gi|11497520	*Spinacia oleracea*	gi|157678835	*Chattonella ovata*
gi|11466345	*Mesostigma viride*	gi|62432769	*Rhodosorus marinus*
gi|669098	*Antirrhinum majus*	gi|157678839	*Haramonas dimorpha*
gi|11465823	*Porphyra purpurea*	gi|62432781	*Rhodomonas abbreviata*
gi|156618821	*Cuscuta reflexa*	gi|5738976	*Physcomitrella patens*

GenInfo Identifier	Species	GenInfo Identifier	Species
gi|157325525	*Ipomoea purpurea*	gi|157678837	*Cricosphaera roscfensis*
gi|30352031	*Adiantum capillus-veneris*	gi|62432759	*Compsopogon coeruleus*
gi|159161248	*Cuscuta exaltata*	gi|157678841	*Proteomonas sulcata*
gi|18203605	*Populus deltoides*	gi|62432779	*Pyrenomonas helgolandii*
gi|159106792	*Cuscuta obtusiflora*	gi|62432799	*Karenia brevis*
gi|113474993	*Trichodesmium erythraeum IMS101*	gi|5739010	*Sphagnum fallax*
gi|473493	*Pisum sativum*	gi|67968347	*Selaginella uncinata*
gi|156618902	*Cuscuta gronovii*	gi|62432797	*Peridinium foliaceum*
gi|158334264	*Acaryochloris marina MBIC11017*	gi|62432773	*Chondrus crispus*
gi|125964870	*Porphyra yezoensis*	gi|5739020	*Haplomitrium mnioides*
gi|115349988	*Stigeoclonium helveticum*	gi|62432803	*Akashiwo sanguinea*
gi|71842284	*Emiliania huxleyi*	gi|62432805	*Heterocapsa niei*
gi|108796713	*Zygnema circumcarinatum*	gi|62432777	*Chroomonas sp. SAG 980-1*
gi|11467176	*Zea mays*	gi|62432801	*Karenia mikimotoi*
gi|139387449	*Phaseolus vulgaris*	gi|28630971	*Citrus limon*
gi|149072099	*Rhodomonas salina*	gi|27435916	*Welwitschia mirabilis*
gi|11467626	*Guillardia theta*	gi|31745675	*Pyrus communis*
gi|124022552	*Prochlorococcus marinus str. MIT 9303*	gi|33317759	*Dioon purpusii*
gi|60117200	*Huperzia lucidula*	gi|95115474	*Synechococcus sp. WH 8020*
gi|27435893	*Rheum x cultorum*	gi|29468202	*Ceratium horridum*
gi|71063182	*Xerophyta humilis*		

Table 7: All used sequences of the D2 subunit for the profile alignment using the build pHMM files.

Subunit	Stricly Conserved Residue
L	IleL64, ProL68, ProL69, GlyL74, LeuL75, AlaL78, ProL79, GlyL83, GlyL84, GlnL87, PheL97, TrpL100, ArgL103, GluL106, IleL107, ArgL109, LysL110, LeuL111, HisL116, ProL118, AlaL124, AlaL127, LeuL131, ArgL135, ProL136, GlyL140, TrpL142, PheL146, ProL147, TyrL148, GlyL149, HisL153, LeuL154, TrpL156, ValL157, GlyL161, TyrL162, PheL164, TyrL166, ProL170, AsnL171 HisL173, MetL174, PheL181, HisL190, AsnL199
D1	TrpD1-14, PheD1-17, ThrD1-24, AsnD1-26, ArgD1-27, TyrD1-29, GlyD1-31, TrpD1-32, PheD1-33, GlyD1-34, MetD1-37, ProD1-39, LeuD1-41, LeuD1-42, AlaD1-44, PheD1-52, AlaD1-54, AlaD1-55, ProD1-56, ValD1-58, AspD1-59, IleD1-60, AspD1-61, GlyD1-62, IleD1-63, ArgD1-64, GluD1-65, ProD1-66, ThrD1-67, GlyD1-69, SerD1-70, LeuD1-71, TyrD1-73, GlyD1-74, AsnD1-75, AsnD1-76, AlaD1-81, ProD1-84, AlaD1-88, GlyD1-90, HisD1-92, PheD1-93, TyrD1-94, ProD1-95, IleD1-96, TrpD1-97, GluD1-98, AlaD1-99, TrpD1-105, LeuD1-106, TyrD1-107, AsnD1-108, GlyD1-109, GlyD1-110, ProD1-111, PheD1-119, GlyD1-122, ArgD1-129, PheD1-132, LeuD1-133, SerD1-134, LeuD1-137 GlyD1-138, MetD1-139, ArgD1-140, ProD1-141, TrpD1-142, IleD1-143, ValD1-145, AlaD1-145, AlaD1-148, ProD1-149, AlaD1-154, PheD1-158, TyrD1-161, ProD1-162, GlyD1-164, GlnD1-165, GlyD1-166, SerD1-167, GlyD1-168, SerD1-169, GluD1-170, MetD1-172, ProD1-173, LeuD1-174, GlyD1-175, IleD1-176, AlaD1-177, GlyD1-178, PheD1-180, AsnD1-181, MetD1-183, ValD1-185, PheD1-186, GlnD1-187, AlaD1-188, HisD1-190, AsnD1-191, IleD1-192, LeuD1-193, MetD1-194, HisD1-195, ProD1-196, PheD1-197, HisD1-198, GlyD1-201, ThrD1-202, AlaD1-203, GlyD1-204, ValD1-205, PheD1-206, GlyD1-207, GlyD1-208, LeuD1-210, PheD1-211, SerD1-212, AlaD1-213, MetD1-214, HisD1-215, GlyD1-216, SerD1-217, LeuD1-218, ValD1-219, ThrD1-220, SerD1-221, SerD1-222, SerD1-223, ArgD1-224, GluD1-225, SerD1-232, AsnD1-234, GlnD1-241, GluD1-242, GluD1-243, GluD1-244, TyrD1-246, AsnD1-247, ValD1-249, AlaD1-250, HisD1-252, GlyD1-253, TyrD1-254, PheD1-255, GlyD1-256, ArgD1-257, LeuD1-258, PheD1-260, GlnD1-261, TyrD1-262, AlaD1-263, SerD1-264, PheD1-265, AsnD1-267, SerD1-268, ArgD1-269, LeuD1-271, HisD1-272, PheD1-273, PheD1-274, LeuD1-275, TrpD1-278, ProD1-279, ValD1-280, IleD1-283, TrpD1-284, ThrD1-286, ThrD1-292, MetD1-293, AlaD1-294, PheD1-295, AsnD1-296, LeuD1-297, AsnD1-298, GlyD1-299, PheD1-300, AsnD1-301, PheD1-302, AsnD1-303, SerD1-305, AspD1-308, GlyD1-311, ValD1-313, ThrD1-316, TrpD1-317, AlaD1-318, AspD1-319, AsnD1-322, ArgD1-323, AlaD1-324, LeuD1-326, GlyD1-327, GluD1-329, ValD1-330, MetD1-331, HisD1-332
M	LeuM94, ProM96, ProM97, GlyM102, ProM108, GlyM112, GlyM113, TrpM129, ArgM132, PheM150, GlyM161, ProM165, GlyM169, ProM176, GlyM178, HisM182, LeuM183, ProM200, IleM206, GlyM211, HisM219, GlyM220, ThrM232

D2	GlnD2-72, SerD2-79, HisD2-87, GlyD2-94, GlyD2-108, GlyD2-109, LeuD2-110, TrpD2-111, HisD2-117, GlyD2-118, GlyD2-124, PheD2-125, ArgD2-128, PheD2-130, GluD2-131, AlaD2-133, ArgD2-139, ProD2-140, TyrD2-141, AsnD2-142, AlaD2-143, GlyD2-146, SerD2-147, ProD2-149, IleD2-150, ValD2-156, PheD2-157, TyrD2-160, ProD2-161, LeuD2-162, GlyD2-163, GlnD2-164, ValD2-167, PheD2-168, PheD2-169, AlaD2-170, ProD2-171, SerD2-172, GlyD2-174, AlaD2-176, IleD2-178, PheD2-179, ArgD2-180, PheD2-181, LeuD2-183, PheD2-184, GlnD2-186, GlyD2-187, PheD2-188, HisD2-189, AsnD2-190, SerD2-192, LeuD2-193, AsnD2-194, ProD2-195, PheD2-196, HisD2-197, MetD2-198, MetD2-199, GlyD2-200, ValD2-201, AlaD2-202, GlyD2-203, LeuD2-205, GlyD2-206, AlaD2-208, LeuD2-209, LeuD2-210, IleD2-213, HisD2-214, GlyD2-215, AlaD2-216, ThrD2-217, ValD2-218, AsnD2-220, ThrD2-221, AspD2-225, GlyD2-226, ThrD2-231, PheD2-232, AlaD2-234, PheD2-235, AsnD2-237, GlnD2-239, GluD2-241, GluD2-242, ThrD2-243, TyrD2-244, SerD2-245, MetD2-246, ValD2-247, ThrD2-248, AlaD2-249, AsnD2-250, ArgD2-251, PheD2-252, TrpD2-253, SerD2-254, GlnD2-255, PheD2-257, GlyD2-258, AlaD2-260, PheD2-261, SerD2-262, AsnD2-263, LysD2-264, ArgD2-265, TrpD2-266, LeuD2-267, HisD2-268, PheD2-269, PheD2-270, LeuD2-272, PheD2-273, ValD2-274, ProD2-275, GlyD2-278, TrpD2-281, SerD2-282, GlyD2-285, GlyD2-288, LeuD2-289, AlaD2-290, AsnD2-292, LeuD2-293, ArgD2-294, AlaD2-295, TyrD2-296, AspD2-297, PheD2-298, SerD2-300, GlnD2-301, GluD2-302, AlaD2-305, GluD2-307, AspD2-308, ProD2-309, GluD2-310, PheD2-311, GluD2-312, ThrD2-313, PheD2-314, TyrD2-315, ThrD2-316

Table 8: Strictly conserved residues in the subunits L, D1, M, and D2. A residue is considered strictly conserved if the same aminoacid is found in at least 98% of all considered sequences.

Subunits	bRC	PSII
	ProL106	ProD1-84
	GlyL112	GlyD1-90
	GlyL131	GlyD1-109
	GlyL132	GlyD1-110
L/D1	ArgL151	ArgD1-129
	LeuL159	LeuD1-137
	ProL184	ProD1-162
	ProL195	ProD1-173
	GlyL197	GlyD1-175
	PheL229	PheD1-206
	HisL238	HisD1-215
	GlyM142	GlyD2-108
	GlyM143	GlyD2-109
	ArgM162	ArgD2-128
	ProM195	ProD2-129
M/D2	GlyM208	GlyD2-174
	ProM230	ProD2-195
	GlyM241	GlyD2-206
	HisM249	HisD2-214
	GlyM250	GlyD2-215
	ThrM252	ThrD2-217

Table 8: Strictly conserved residues in the RC (in L/D1 or in M/D2). A residue is considered strictly conserved if the same aminoacid is found in at least 98% of all considered sequences.