RESEARCH ARTICLES

Conservation of Distantly Related Membrane Proteins: Photosynthetic Reaction Centers Share a Common Structural Core

Sumedha Sadekar, Jason Raymond,† and Robert E. Blankenship‡*

*Computational Biosciences Program, Arizona State University; †Microbial Systems Division, Biosciences Directorate, Lawrence Livermore National Laboratory, Livermore, California; and ‡Departments of Biology and Chemistry, Washington University

Photosynthesis was established on Earth more than 3 billion years ago. All available evidences suggest that the earliest photosynthetic organisms were anoxygenic and that oxygen-evolving photosynthesis is a more recent development. The reaction center complexes that form the heart of the energy storage process are integral membrane pigment proteins that span the membrane in vectorial fashion to carry out electron transfer. The origin and extent of distribution of these proteins has been perplexing from a phylogenetic point of view mostly because of extreme sequence divergence. A series of integral membrane proteins of known structure and varying degrees of sequence identity have been compared using combinatorial extension–Monte Carlo methods. The proteins include photosynthetic reaction centers from proteobacteria and cyanobacterial photosystems I and II, as well as cytochrome oxidase, bacteriorhodopsin, and cytochrome *b*. The reaction center complexes show a remarkable conservation of the core structure of 5 transmembrane helices, strongly implying common ancestry, even though the residual sequence identity is less than 10%, whereas the other proteins have structures that are unrelated. A relationship of sequence with structure was derived from the reaction center structures; with characteristic decay length of 1.6 Å. Phylogenetic trees derived from the structural alignments give insights into the earliest photosynthetic reaction center, strongly suggesting that it was a homodimeric complex that did not evolve oxygen.

Introduction

Proteins that share a recent common evolutionary origin have high primary sequence identity and similar 3-dimensional structures. As the evolutionary distance between proteins increases, the sequence identity decreases and the structural similarity diminishes. Below about 25% sequence identity, it is usually not possible to reliably infer common ancestry from sequence comparisons alone, a situation often described as the "twilight zone" of molecular evolution studies (Doolittle 1986; Rost 1999). However, similar structures can persist well into the twilight zone, and structural comparisons can sometimes be used to infer common ancestry of more distantly related proteins. Several examples of this situation are known, and in some cases, it is clear that the proteins have a common ancestor despite a low sequence identity (Valencia et al. 1991; Bork et al. 1994; Murzin et al. 1995; Al-Lazikani et al. 2001; Torrents et al. 2002).

Photosynthetic reaction center complexes are multisubunit integral membrane protein complexes (Blankenship 2002). They sensitize the light-driven electron transfer processes of photosynthetic energy storage that form the basis of all primary productivity and are at the base of all food chains. The reaction center complexes are all divided into 2 main classes, known as Type I and Type II, based on the identity of the early electron acceptors. It has long been apparent from biophysical analysis and sequence comparisons that reaction centers within each of the 2 classes are structurally and functionally similar and probably are descended from a single common ancestor (Williams et al. 1984; Youvan et al. 1984; Michel and Deisenhofer 1988). However, it had not been realized until recent structural data became available that each of the 2 broad classes of reaction centers are probably themselves descended from a very distant common ancestor (Fromme et al. 1994; Schubert et al. 1998) as the residual sequence identity between the 2 classes is less than 10%, putting them well into the twilight zone. Xiong and Bauer (2002) have proposed that the evolutionary ancestor of Type II photosynthetic reaction centers was a *b*-type cytochrome, based on putative conserved key residues that coordinate cofactors in both systems.

We have carried out detailed structural comparisons of the core portions of all available high-resolution reaction center complexes from both Type I and Type II complexes and use the results to build evolutionary trees of reaction centers based solely on structural conservation. In addition, we derive sequence alignments and thereby phylogenetic trees based on structure, compare these results with more traditional sequence-based trees, and use the results to make inferences about the nature of the earliest photosynthetic reaction centers.

Materials and Methods

Structural Alignment

The protein structures were retrieved from the Protein Data Bank (PDB) (Berman et al. 2000). The PDB entry file names, PDB IDs, the protein chains studied, and their resolution are α-proteobacteria: *Rhodobacter sphaeroides* (1AIJ), L, M chains, 2.20 Å; *Blastochloris* (*Rhodopseudomonas*) *viridis* (1DXR), L, M chains, 2.00 Å; γ-proteobacteria: *Thermochromatium tepidum* (1EYS), L, M chains, 2.20 Å; Cyanobacteria: *Thermosynechococcus elongatus* (1S5L), D1, D2 chains, 3.50 Å; *Synechococcus elongatus* (1JB0), A1, A2 chains, 2.50 Å.

Multiple structural alignments, based on conventional methods like combinatorial extension (CE) (Shindyalov and Bourne 1998) or DALI (Holm and Sander 1993), use "master–slave" pairwise alignments, which are not done by all-to-all comparison of proteins but by "pile-up" of structural neighbors, whereas methods like HOMSTRAD (Mizuguchi et al. 1998) and CAMPASS

Key words: photosynthesis, evolution, reaction center, protein structure.

E-mail: blankenship@wustl.edu.

Mol. Biol. Evol. 23(11):2001–2007. 2006
doi:10.1093/molbev/msl079
Advance Access publication August 3, 2006

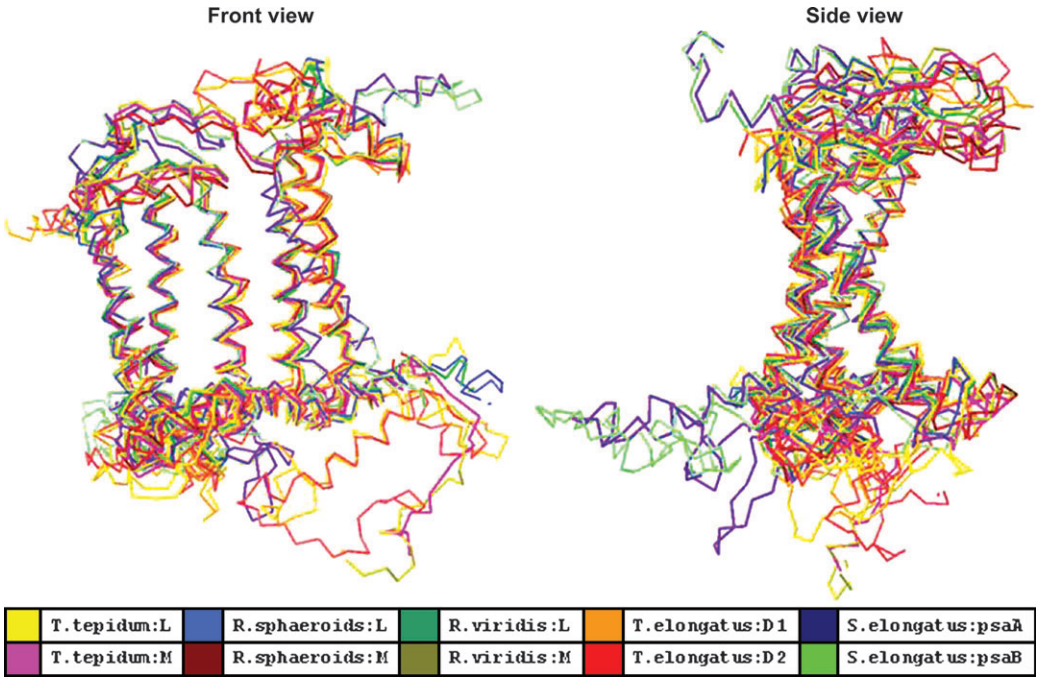

FIG. 1.—Structural alignments of all photosynthetic reaction center proteins: α-proteobacteria: *Rhodobacter sphaeroides* (1AIJ), L, M chains; *Rhodopseudomonas viridis* (1DXR), L, M chains; *Thermochromatium tepidum* (1EYS), L, M chains; Cyanobacteria: *Thermosynechococcus elongatus* (1S5L), D1, D2 chains of photosystem II; and *Synechococcus elongatus* (1JB0), A1, A2 chains of photosystem I. The 6 N-terminal helices that constitute the antenna domain of the photosystem I complex are not shown but were included in the data set used for alignment. The unaligned thread-like portions on the top and the bottom are the loops outside the membranes, joining the transmembrane helices. The left figure shows a front view of the 10 overlaid structures, whereas the right figure shows the side view of the same complexes rotated by 90°.

(Sowdhamini et al. 1998) provide multiple alignments only for predefined protein families. Therefore, the CE–Monte Carlo (MC) server (http://bioinformatics.albany.edu/~cemc/) was used to generate the overlays. The alignments were generated using the CE algorithm and iteratively optimized using MC simulations (Guda et al. 2004). This algorithm does not require a master–slave alignment and allows input structures in PDB format. All 11 helices from PsaA and PsaB were included in the structural alignment. The CE–MC software intelligently aligned the last 5 C-terminal helices due to high similarity with the other proteins, and the first 6 N-terminal helices remained unaligned, which is as expected because they code for an antenna domain that is not present in the other complexes. The 6 N-terminal helices were manually trimmed after obtaining the results to facilitate visualization. After aligning the structures, root mean square distances (RMSDs) between the α-carbon atoms of the backbone chains of all possible protein pairs were calculated (table 1). From the CE–MC alignment results, the variation of identity versus RMSD values was plotted for all possible protein pairs, as shown in figure 2. The identity values were calculated from *p* distances evaluated using MEGA2 (Nei and Kumar 2000). The sequences for the structurally aligned proteins are shown in the Supplementary Material online. We also performed alignments of unrelated transmembrane proteins such as subunit 1 of cytochrome *c* oxidase from *Paracoccus denitrificans* (supplementary fig. 1, Supplementary Material online), the cytochrome *b* subunit of the cytochrome bc_1 complex from *Bos taurus* (supplementary fig. 2, Supplementary Material online), and bacteriorhodopsin from *Halobacterium salinarium* (supplementary fig. 3, Supplementary Material online). Additional details of sequence selection alignment and tree building are given in the Supplementary Material online.

Results

The results of the structural overlay of the photosynthetic reaction centers are shown in figure 1. This comparison includes proteobacterial reaction centers and cyanobacterial photosystems I and II. All known reaction centers have a dimeric core of proteins, in most cases heterodimers that consist of 2 similar but distinct subunits (Schubert et al. 1998; Blankenship 2002). Examples of these heterodimers are the "Type II" reaction centers consisting of the L and M subunits of the proteobacterial reaction centers and the D1 and D2 proteins of the cyanobacterial photosystem II, as well as the "Type I" reaction centers exemplified by the PsaA and PsaB subunits of the cyanobacterial photosystem I. For the comparisons shown in figure 1, all 10 subunits were superimposed from 5 independent structures. The structures of these integral membrane complexes are remarkably well conserved in the 5 transmembrane helical domains but are less conserved in the loops linking the transmembrane domains. The 2 halves of the heterodimeric reaction center complexes in each case are very similar to each other, indicating that they have arisen from gene duplication and divergence events (see below). The RMSD between corresponding amino acid residues in the photosynthetic reaction centers are given in table 1, along with sequence identities resulting from multiple sequence alignments derived

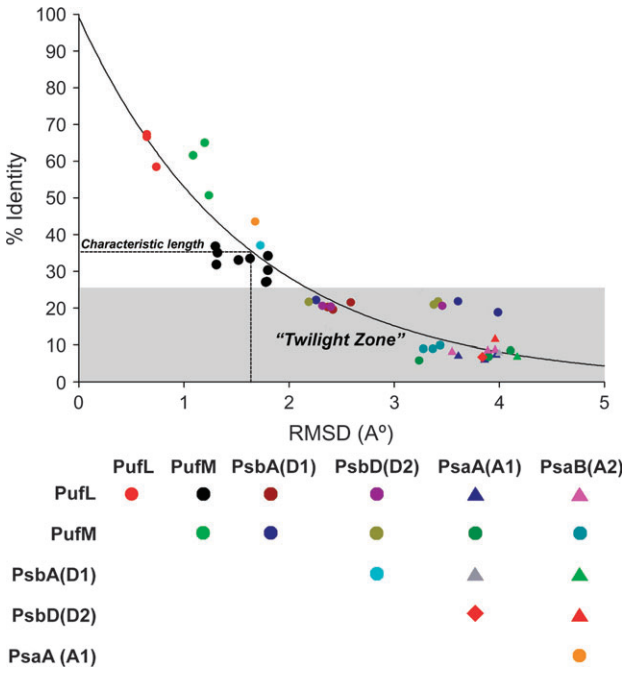

FIG. 2.—Plot of variation of sequence identity between protein pairs versus structural similarity represented by RMSD values, using the data from table 1 given above. The solid line is an exponential fit to the data with characteristic decay length 1.6 Å.

from the structural overlays. The sequence identities range from 60% to 70% for the same subunits from different species of proteobacteria to less than 10% for comparisons between the Type I and Type II complexes. Figure 2 shows a comparison of the percentage sequence identity in the multiple alignments to the RMSDs determined from the structural overlays. This figure illustrates the gradual decay of structural similarity as sequence identity decreases. The curve is fit to an exponential function with a characteristic decay length of 1.6 Å.

The overlaid structures of the other membrane proteins compared (cytochrome oxidase, bacteriorhodopsin, and the cytochrome *b* subunit of the cytochrome bc_1 complex) were compared with those of the reaction centers. These proteins are structurally distinctly different both from each other and from the reaction centers (supplementary figs. 1–3, Supplementary Material online). These results, taken with figure 1, suggest that the reaction centers form a coherent class of proteins that are evolutionarily related, albeit very distantly, whereas the other integral membrane proteins chosen either are completely independent evolutionary innovations from the reaction centers or have diverged so long ago so that not only has any hint of sequence conservation disappeared but also the structures have lost any discernable relationship.

Unrooted phylogenetic trees were constructed using the data derived from the structural comparisons. Figure 3*A* shows a tree constructed directly from the RMSD, which were used as proxies for evolutionary distances. Sequence alignments derived from the structural alignments were also used to construct an evolutionary tree using standard methods of phylogenetic analysis, shown in figure 3*B*. This tree is identical in topology to the tree based directly on the RMSD. These trees are in turn topologically identical for the taxa present to a sequence-based tree shown in figure 4, which was produced from a larger group of reaction center sequences representing all known classes of photosynthetic organisms. Sequence alignments used to construct the trees are given in the Supplementary Material online as well as detailed subtrees. Three putative ancient gene duplication events are indicated by stars in figures 3 and 4. These represent gene duplications and subsequent divergences that led to a heterodimeric reaction center core structure. The ancestral state is inferred to be a homodimeric structure. Figure 4 includes 2 groups of photosynthetic reaction centers for which structural information is not yet available, the green sulfur bacteria and the heliobacteria. These are both Type I reaction centers generally similar in biochemical and biophysical properties to photosystem I, although in both cases they are known to be homodimeric complexes (Buttner et al. 1992; Liebl et al. 1993). The statistical bootstrap values for figure 4 are given in the Supplementary Material online.

Discussion

It is a well-known fact in structural biochemistry and protein evolution that proteins with very similar sequences have similar structures and that as the sequence gradually diverges through divergent evolution the structures become less similar. The most extensive compilation of this effect is that given for soluble proteins by Chothia and Lesk (1986). Although they calculated the relationship of structure and sequence somewhat differently than we do here, the characteristic distance that describes their data (estimated at 1.5 Å from Chothia and Lesk 1986, Figure 2) are remarkably similar to what we found for these membrane proteins. It thus appears that a sequence/structure relationship of this sort describes a wide range of proteins, both soluble and membrane.

The tree that was derived only from RMSD derived from structural comparisons does not rely on the standard assumptions that underlie tree-building routines based only on sequence comparisons, such as substitution matrices or corrections for uneven rates of evolution. Therefore, it is not expected to be subject to the pervasive problem of long-branch attraction, which can dramatically distort evolutionary trees where very distantly related proteins are compared. Interestingly, the sequence alignment derived from the structural comparisons is not identical to the alignment derived from sequence comparisons, although they do give rise to evolutionary trees with the same topology. This suggests that over a long period of time, a protein can drift away from an original sequence while maintaining a structure that is conserved by functional constraints. It will be interesting to see other examples of this effect, which should serve as a cautionary note to avoid overinterpretation of sequence data.

The fact that the structure-based tree has an identical topology to trees based on sequence alignments suggests that they both represent the true evolutionary relationship of photosynthetic reaction centers, including 3 independent gene duplication events that gave rise to heterodimeric complexes. There has previously been discussion as to whether these apparent multiple gene duplication events,

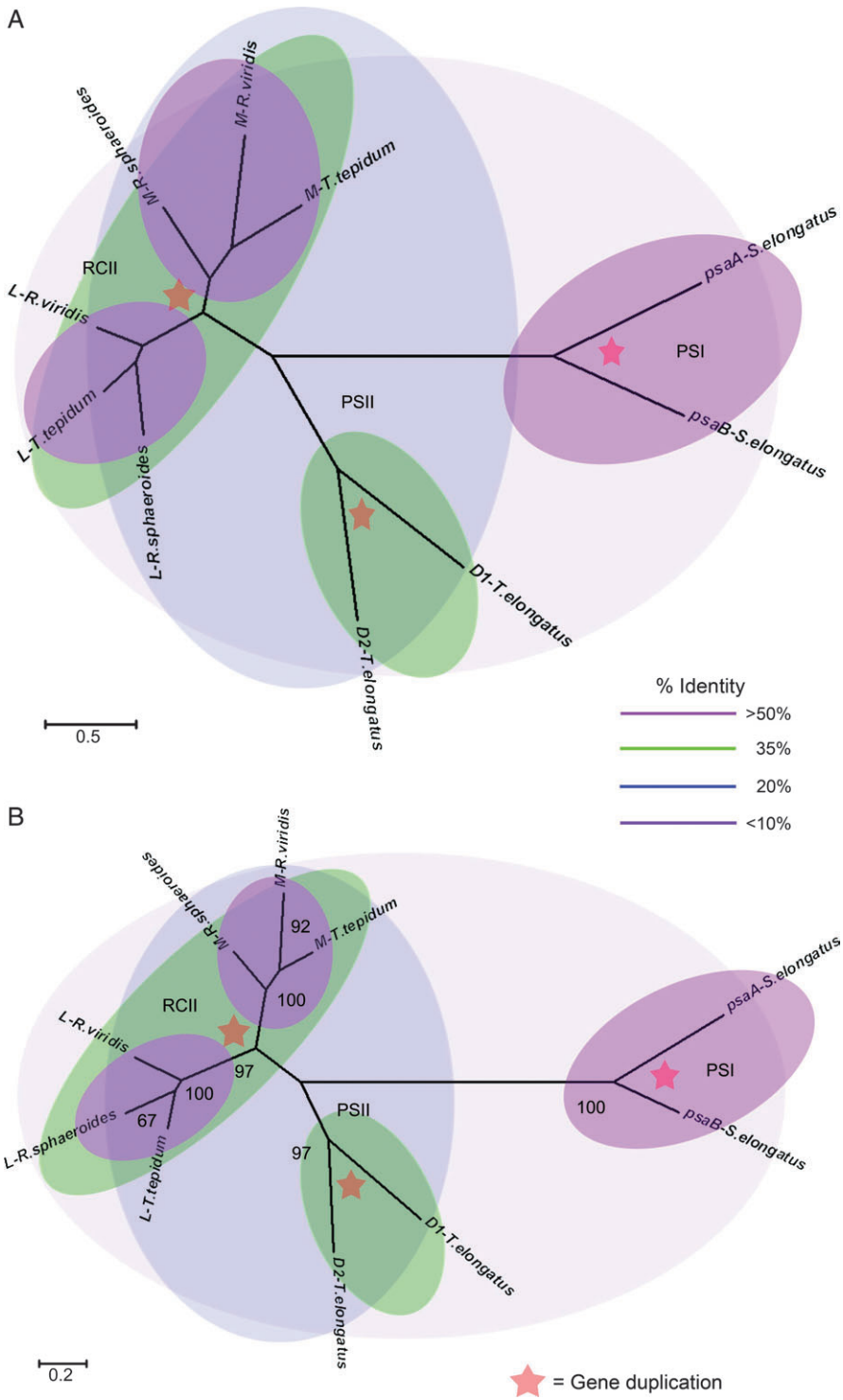

FIG. 3.—(*A*) Unrooted phylogenetic tree constructed using RMSDs derived from structural alignments (table 1) directly as proxies for evolutionary distances. (*B*) Unrooted Neighbor-Joining phylogenetic tree of photosynthetic reaction centers based on a sequence alignment derived from the structural alignments shown in figure 1. The red stars represent inferred gene duplication events. The colored boundaries enclose proteins sharing a particular percentage of similarity.

especially with respect to the Type II reaction centers, represented a correct topology or have been distorted by other processes such as gene conversion (Blankenship 1994; Lockhart et al. 1996; Blankenship 2002). The results are most consistent with multiple gene duplication events. However, our results do not appear to lend support to the proposal from Xiong and Bauer (2002) that the evolutionary ancestor of all Type II photosynthetic reaction centers was a *b*-type cytochrome. The comparison of the structure of the *b* cytochrome subunit from the cytochrome bc_1 complex from *B. taurus* does not show any apparent structural relationship to the reaction center structure

Table 1
Sequence and Structural Comparisons of Photosynthetic Reaction Centers. Upper Section: RMSD between Any Pair of Proteins in Angstrom Units, Calculated between the α-Carbon Atoms Forming the Protein Backbone. Lower Section: Sequence Identities (%) Derived from the Structural Overlays

	Rhodobacter sphaeroides, L	*Rhodopseudomonas viridis*, L	*Thermochromatium tepidum*, L	*Rhodobacter sphaeroides*, M	*Rhodopseudomonas viridis*, M	*Thermochromatium tepidum*, M	*Thermosynechococcus elongatus*, D1	*Thermosynechococcus elongatus*, D2	*Synechococcus elongatus*, A	*Synechococcus elongatus*, B
Rhodobacter sphaeroides, L	—	0.74	0.65	1.8	1.8	1.52	2.59	3.46	3.86	3.55
Rhodopseudomonas viridis, L	58.36	—	0.65	1.3	1.78	1.63	2.37	2.32	3.61	3.96
Thermochromatium tepidum, L	67.28	66.54	—	1.32	1.79	1.31	2.42	2.4	3.97	3.89
Rhodobacter sphaeroides, M	34.2	36.8	35.02	—	1.24	1.2	2.26	2.19	3.9	3.44
Rhodopseudomonas viridis, M	30.24	27.02	27.17	50.6	—	1.09	3.99	3.38	3.24	3.37
Thermochromatium tepidum, M	33.05	33.48	31.8	64.94	61.51	—	3.61	3.42	4.11	3.28
Thermosynechococcus elongatus, D1	21.46	20.17	19.5	22.08	18.72	21.74	—	1.73	4	4.17
Thermosynechococcus elongatus, D2	20.5	20.5	20.24	21.61	20.83	21.7	36.98	—	3.84	3.96
Synechococcus elongatus, A	05.91	06.99	07.29	06.77	05.76	08.38	07.73	06.74	—	1.71
Synechococcus elongatus, B	08.06	08.6	08.38	09.9	08.9	08.9	06.74	11.46	42.42	—

(supplementary fig. 1, Supplementary Material online), with average RMSD value of 6.1 Å, compared with an average of 2.6 Å for the reaction center comparisons (table 1). This does not definitively rule out the possibility that cytochrome *b* was the ancestor of the reaction centers but also does not provide any positive support for this proposal.

The remarkable structural conservation of the reaction center complexes and the overall topology of figure 4 can be used to draw some tentative conclusions about the nature of the earliest reaction center complexes. These complexes were almost certainly homodimeric, indicated by the colored homodimeric zone in figure 4. Because this is an unrooted tree, it is not possible to determine whether the Type I or Type II reaction centers are the ancestral state. However, if the tree is rooted at the midpoint of the long edge connecting the Type I and Type II complexes (which implicitly assumes equal rates of evolution on the 2 main branches), then the ancestral reaction center may well have been intermediate in structure between the 2 types and not easily categorized as either Type I or Type II. Recent biophysical evidence (M. F. Hohmann-Marriott and R. E. Blankenship, unpublished data) suggests that the reaction centers in the green sulfur bacteria, usually categorized as Type I, may indeed have functional aspects that are intermediate between those exhibited by the heterodimeric Type I (photosystem I) reaction centers and the type II complexes. This is consistent with a more primitive nature for these complexes as indicated by both their homodimeric composition and their relative position in figure 4. A similar proposal has been made by Allen (2005).

The data also suggest that the 3 gene duplications that gave rise to the 3 groups of heterodimeric reaction center complexes were independent events. The possible functional advantages of a heterodimeric reaction center have long been discussed (Blankenship 1992, 2002; Buttner et al. 1992; Liebl et al. 1993; Schubert et al. 1998; Baymann et al. 2001; Allen 2005) but are still not clear especially because 2 groups of extant organisms utilize homodimeric complexes. However, the fact that 3 independent duplications and divergences of reaction center genes have almost certainly taken place and the majority of photosynthetic organisms contain heterodimeric complexes suggests that there is strong selection pressure for this to take place and it must have an important functional role.

Was the ancestral photosynthetic organism anoxygenic (non–oxygen evolving) or oxygenic (oxygen evolving)? Previous work (Xiong et al. 2000) has suggested that it was anoxygenic, based primarily on the much greater simplicity of the subunit composition of the reaction centers found in anoxygenic photosynthetic organisms (Raymond and Blankenship 2004a) and the very difficult chemistry involved in the oxidation of water (Blankenship and Hartman 1998; Dismukes et al. 2001), which argues against this being a metabolic capability of a primitive organism. Our results strongly support this view. Recent biogeochemical data from 3.4 billion–year-old cherts derived from microbial mats have been interpreted to indicate that the photosynthetic organisms that built these mats were anoxygenic, which is also consistent with the view that the earliest photosynthetic reaction centers were not capable of oxygen evolution (Tice and Lowe 2006).

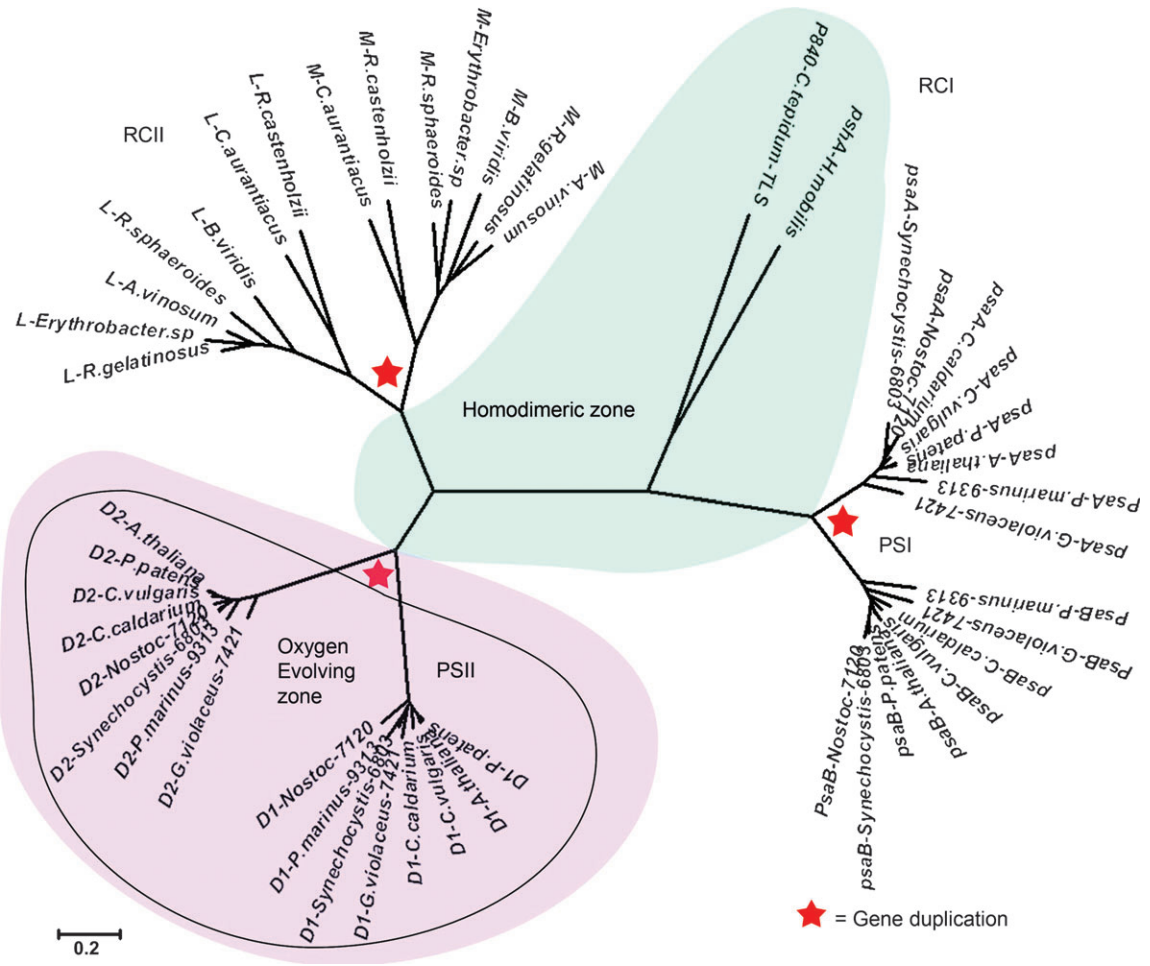

FIG. 4.—Unrooted Neighbor-Joining phylogenetic tree based on an extensive set of sequences of photosynthetic reaction centers. Only the C-terminal electron transfer domains of the Type I reaction centers were used in the analysis. The red stars represent inferred gene duplication events. The blue colored region represents sequence space of reaction centers either known or inferred to have a homodimeric core protein structure, whereas all others have a heterodimeric structure. The red colored region represents sequence space of those reaction centers that evolve oxygen. The dashed line indicates a development of oxygen evolution capability that is well after the gene duplication event that led to a heterodimeric reaction center, whereas the solid line indicates an earlier development of oxygen evolution capability.

The most deeply branching reaction centers are the homodimeric Type I reaction centers found in the anoxygenic (and strictly anaerobic) green sulfur bacteria and heliobacteria (Mix et al. 2005). The only group of oxygenic photosynthetic prokaryotes is the cyanobacteria, which contain both the oxygen-evolving photosystem II and the non–oxygen-evolving photosystem I. Only photosystem II is capable of oxygen evolution, so that further restricts the oxygen evolution phenotype to only one branch of the tree, strongly suggesting that it is a derived trait.

It is not yet possible to place reliable dates either on the appearance of the major groups of bacteria or on the appearance of the oxygen evolution phenotype in phototrophs. Even if one could date the appearance of the cyanobacteria as has been claimed (Summons et al. 1999; Battistuzzi et al. 2004), it would not be possible to determine with certainty if the earliest cyanobacteria were oxygenic organisms or developed this capability at a later time. Biogeochemical evidence strongly suggests that the ability to evolve oxygen appeared by at least 2.2–2.4 billion years ago as the level of free atmospheric oxygen began to increase at about that time (reviewed in Knoll 1999). It could have appeared significantly earlier, but constraints on this date are controversial (Raymond and Blankenship 2004b).

In summary, our results indicate that all photosynthetic reaction centers have derived from a single homodimeric anoxygenic primitive reaction center. Despite extensive sequence divergence, the structures of the transmembrane portions of core reaction center complexes have remained remarkably conserved.

Supplementary Material

Supplementary figures 1–3 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

Supported by a grant from the National Aeronautics and Space Administration Exobiology (NNG04GK59G) program. J. R. acknowledges support through a Lawrence Fellowship.

Literature Cited

Al-Lazikani B, Sheinerman FB, Honig B. 2001. Combining multiple structure and sequence alignments to improve sequence detection and alignment: application to the SH2 domains of Janus kinases. Proc Natl Acad Sci USA 98:14796–801.

Allen JF. 2005. A redox switch hypothesis for the origin of two light reactions in photosynthesis. FEBS Lett 579:963–8.

Battistuzzi FU, Feijao A, Hedges SB. 2004. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. BMC Evol Biol 4:44.

Baymann F, Brugna M, Muhlenhoff U, Nitschke W. 2001. Daddy, where did PS-1 come from? Biochim Biophys Acta 1507: 291–310.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. Nucleic Acids Res 28:235–42.

Blankenship RE. 1992. Origin and early evolution of photosynthesis. Photosynth Res 33:91–111.

Blankenship RE. 1994. Protein structure, electron transfer and evolution of prokaryotic photosynthetic reaction centers. Antonie Leeuwenhoek 65:311–29.

Blankenship RE. 2002. Molecular mechanisms of photosynthesis. Oxford: Blackwell Science.

Blankenship RE, Hartman H. 1998. The origin and evolution of oxygenic photosynthesis. Trends Biochem Sci 23:94–7.

Bork P, Holm L, Sander C. 1994. The immunoglobulin fold. Structural classification, sequence patterns and common core. J Mol Biol 242:309–20.

Buttner M, Xie DL, Nelson H, Pinther W, Hauska G, Nelson N. 1992. Photosynthetic reaction center genes in green sulfur bacteria and in photosystem 1 are related. Proc Natl Acad Sci USA 89:8135–9.

Chothia C, Lesk A. 1986. The relation between the divergence of sequence and structure in proteins. EMBO J 5:823–6.

Dismukes GC, Klimov VV, Baranov SV, Kozlov YN, DasGupta J, Tyryshkin A. 2001. The origin of atmospheric oxygen on Earth: the innovation of oxygenic photosynthesis. Proc Natl Acad Sci USA 98:2170–5.

Doolittle RF. 1986. Of urfs and orfs: a primer on how to analyze derived amino acid sequences. Mill Valley, CA: University Science Books.

Fromme P, Schubert WD, Krauss N. 1994. Structure of photosystem-I. Suggestions on the docking sites for plastocyanin, ferredoxin and the coordination of P700. Biochim Biophys Acta 1187:99–105.

Guda C, Lu SF, Scheeff ED, Bourne PE, Shindyalov IN. 2004. CE-MC: a multiple protein structure alignment server. Nucleic Acids Res 32:W100–3.

Holm L, Sander C. 1993. Protein-structure comparison by alignment of distance matrices. J Mol Biol 233:123–38.

Knoll AH. 1999. A new molecular window on early life. Science 285:1025–6.

Liebl U, Mockensturm-Wilson M, Trost JT, Brune DC, Blankenship RE, Vermaas W. 1993. Single core polypeptide in the reaction center of the photosynthetic bacterium *Heliobacillus mobilis*: structural implications and relations to other photosystems. Proc Natl Acad Sci USA 90: 7124–8.

Lockhart P, Steel M, Larkum A. 1996. Gene duplication and the evolution of photosynthetic reaction center proteins. FEBS Lett 385:193–96, 390:242.

Michel H, Deisenhofer J. 1988. Relevance of the photosynthetic reaction center from purple bacteria to the structure of photosystem II. Biochemistry 27:1–7.

Mix LJ, Haig D, Cavanaugh CM. 2005. Phylogenetic analyses of the core antenna domain: investigating the origin of photosystem I. J Mol Evol 60:153–63.

Mizuguchi K, Deane CM, Blundell TL, Overington JP. 1998. HOMSTRAD: a database of protein structure alignments for homologous families. Protein Sci 7:2469–71.

Murzin AG, Brenner SE, Hubbard T, Chothia C. 1995. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247:536–40.

Nei M, Kumar S. 2000. Molecular evolution and phylogenetics. Oxford: Oxford University Press.

Raymond J, Blankenship RE. 2004a. The evolutionary development of the protein complement of photosystem 2. Biochim Biophys Acta 1655:133–9.

Raymond, J, Blankenship RE. 2004b. Biosynthetic pathways, gene replacement, and the antiquity of life. Geobiology 2: 199–201.

Rost B. 1999. Twilight zone of protein sequence alignments. Protein Eng 12:85–94.

Schubert WD, Klukas O, Saenger W, Witt HT, Fromme P, Krauss N. 1998. A common ancestor for oxygenic and anoxygenic photosynthetic systems: a comparison based on the structural model of photosystem I. J Mol Biol 280:297–314.

Shindyalov IN, Bourne PE. 1998. Protein structure alignment by incremental combinatorial extension CE of the optimal path. Protein Eng 11:739–47.

Sowdhamini R, Burke DF, Huang JF, Mizuguchi K, Nagarajaram HA, Srinivasan N, Steward RE, Blundell TL. 1998. CAMPASS: a database of structurally aligned protein superfamilies. Structure 6:1087–94.

Summons RE, Jahnke LL, Hope JM, Logan GA. 1999. 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. Nature 400:554–7.

Tice M, Lowe D. 2006. Hydrogen-based carbon fixation in the earliest known photosynthetic organisms. Geology 34:37–40.

Torrents E, Aloy P, Gibert I, Rodriguez-Trelles F. 2002. Ribonucleotide reductases: divergent evolution of an ancient enzyme. J Mol Evol 55:138–52.

Valencia A, Kjeldgaard M, Pai EF, Sander C. 1991. GTPase domains of ras p21 oncogene protein and elongation factor Tu: analysis of three-dimensional structures, sequence families, and functional sites. Proc Natl Acad Sci USA 88:5443–7.

Williams JC, Steiner LA, Feher G, Simon MI. 1984. Primary structure of the L subunit of the reaction center from *Rhodopseudomonas sphaeroides*. Proc Natl Acad Sci USA 81:7303–7.

Xiong J, Bauer CE. 2002. A cytochrome *b* origin of photosynthetic reaction centers: an evolutionary link between respiration and photosynthesis. J Mol Biol 322:1025–37.

Xiong J, Fischer W, Inoue K, Nakahara M, Bauer CE. 2000. Molecular evidence for the early evolution of photosynthesis. Science 289:1724–30.

Youvan DC, Bylina EJ, Alberti M, Begusch H, Hearst JE. 1984. Nucleotide and deduced polypeptide sequences of the photosynthetic reaction-center, B870 antenna, and flanking polypeptides from *R. capsulata*. Cell 37:949–57.

Michele Vendruscolo, Associate Editor

Accepted July 12, 2006