S1-003

The remarkable chlorosome

N-U Frigaard[1], EV Vassilieva, H Li[1], KJ Milks[1], J Zhao[2], DA Bryant[1]

[1]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA. nxf10@psu.edu, dab14@psu.edu
[2]College of Life Sciences, Peking University, Beijing 100871, People's Republic of China. jzhao@pku.edu.cn

Keywords: chlorosome, Chlorobium tepidum*,* Chloroflexus aurantiacus*, light-harvesting antenna, interposon mutagenesis*

Introduction

Nature has devised several types of photosynthetic light-harvesting antennae. Chlorosomes are the major antennae in green sulfur bacteria (*Chlorobiaceae*) and green filamentous bacteria (*Chloroflexaceae*) (Blankenship et al. 1995, Oelze and Golecki 1995, Olson 1998, Vassilieva et al. 2000). These large structures resemble other photosynthetic antennae by being rather adaptable and flexible structures. For example, their physical size varies with growth conditions; there seems to be no fixed stoichiometry between BChl *c* and the chlorosome proteins; and nine of the ten known chlorosome proteins in *Chlorobium tepidum* are not needed for viability and growth, chlorosome formation or apparent light-harvesting functionality. We are primarily working with *Cb. tepidum* for which we have established reliable techniques for gene inactivation (Frigaard and Bryant 2001). This report summarizes our current understanding of the chlorosome and some of our recent findings derived from biochemical and molecular biological analyses.

Structure and function of chlorosomes

The chlorosome is the most efficient light-harvesting antenna known and enables cells that form these structures to grow at extremely low light intensities. The organization of the major chlorophyll in chlorosomes (typically BChl *c*) is based upon pigment-pigment interactions and not upon pigment-protein interactions as in other photosynthetic antenna systems (van Rossum et al. 2001). Chlorosomes are also unusual in that the protein-to-pigment ratio is very low (about 0.25 w/w) when compared to other antenna complexes. Chlorosomes essentially seem to be sacks formed from a lipid and protein monolayer envelope and filled with bacteriochlorophyll. About ten major proteins have been shown to be present in chlorosomes from several green sulfur bacteria. Our research is focused on the structure and function of these proteins.

The green sulfur bacteria are strictly anaerobic, but the complete genome sequence of *Cb. tepidum* reveals that it encodes several defense mechanisms against oxidative damage (Eisen et al. 2001). One of these mechanisms resides in the chlorosomes, which exhibit a unique, redox-regulated quenching of the energy transfer under oxidizing conditions. This mechanism eliminates photosynthetic generation of low potential reductants under oxic conditions which otherwise might generate toxic oxygen radicals (Frigaard and Matsuura 1999).

The size of the chlorosome antenna varies with growth conditions. In fact, the content of BChl *c* in *Chloroflexus aurantiacus* may vary ~35-fold on a cellular dry weight basis depending on the light intensity (Oelze and Golecki 1995), and BChl *c* is totally absent from cells grown aerobically. In *Cb. tepidum*, the size of the chlorosome antenna can be reduced about 10-fold by inhibition of BChl *c* synthesis with acetylene without any significant change in growth rate under light-saturating conditions (Ormerod et al. 1990, Frigaard and Ormerod 1995). Interestingly, little or no change in the amounts of chlorosome proteins are seen when BChl *c* synthesis is severely reduced in either *Cf. aurantiacus* (Foidl et al. 1998) or *Cb. tepidum* (Vassilieva et al. 2001b). These observations are consistent with the notion that chlorosome envelopes are assembled and then filled with BChl *c* as required. Both *Cf. aurantiacus* and *Cb. tepidum* have additional small BChl *a*-based antennae located in or associated with the cytoplasmic membrane. It is therefore possible that chlorosomes are not required in these cells if sufficient light is available. The chlorosome antenna might be a late acquisition in the evolutionary development of one or both groups of green bacteria. Interposon mutagenesis is currently being used in an attempt to eliminate BChl *c* from *Cb. tepidum* to test this hypothesis.

The chlorosome proteins

Analyses of chlorosomes isolated from *Cb. tepidum* have shown the presence of ten proteins, all of which are exclusively located in the chlorosome envelope (Vassilieva et al. 2000, 2001c). These proteins are denoted CsmA (6.2 kDa), CsmB (7.5 kDa), CsmC (14.3 kDa), CsmD (11.1 kDa), CsmE (6.7 kDa), CsmF (7.6 kDa), CsmH (21.6 kDa), CsmI (25.9 kDa), CsmJ (21.8 kDa), and CsmX (24.0 kDa). These ten proteins co-purify to an identical extent during chlorosome isolation (Vassilieva et al. 2001c). CsmA is the smallest and most abundant of these proteins. Together with CsmE, CsmA is produced as a precursor with a 20-amino acid extension at its C-terminus that is proteolytically processed during chlorosome assembly and maturation. The second most abundant chlorosome protein in *Cb. tepidum* is CsmB (Chung and Bryant 1996a). Using antibodies against the proteins of *Cb. tepidum*, it was found that the chlorosomes of *Cb. vibrioforme* (BChl *d*) and *Cb. phaeobacteroides* (BChl *e*) had very similar, if not identical protein compositions, although the relative amounts and masses of some component proteins seemed to differ (see Fig. 1). Very little is known about the functions of these proteins. CsmA has been proposed to be a BChl *a*-binding protein (Sakuragi et al. 1999), but there is no direct evidence for this. The genes encoding the chlorosome proteins are generally scattered throughout the genome, although three dicistronic operons are found: *csmCA*, *csmED*, and *csmXJ* (Chung et al. 1994, Chung and Bryant 1996b, Eisen et al. 2001).

Sequence analyses of the chlorosome proteins have revealed that some proteins are related to other chlorosome proteins and that only four basic sequence groups occur (Vassilieva et al. 2000, 2001c). Six of the proteins contain only a single sequence motif whereas the remaining four are fusions of two different motifs. Three of the sequence motifs (CsmA/CsmE, CsmB/CsmF, and CsmC/CsmD) share no obvious relationship with any other proteins in the databases, but the fourth motif occurring in the N-terminal regions of CsmI, CsmJ, CsmX is a ferredoxin/adrenodoxin-like sequence that ligates a [2Fe-2S] cluster (Vassilieva et al. 2001a). The C-terminal

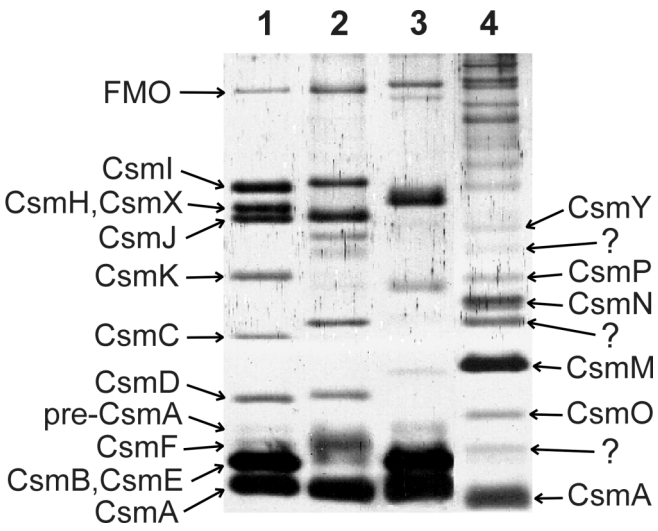

Fig. 1. Silver-stained SDS-PAGE gel of isolated chlorosomes from 1) *Cb. tepidum*, 2) *Cb. phaeobacteroides* 1549, 3) *Cb. vibrioforme* 8327, 4) *Cf. aurantiacus* J10-fl.

regions of CsmI, CsmJ, and CsmX resemble the CsmA/CsmE sequence motif. No other prosthetic groups have been detected in the chlorosome proteins. The CsmI, CsmJ, and CsmX proteins may be involved in the redox-regulated quenching mechanism or some other electron transfer process(es). CsmH is particularly interesting because in contrast to other chlorosome proteins, it can be overproduced in *Escherichia coli* as a water-soluble protein and because it is a fusion of the CsmB/CsmF and CsmC/CsmD motifs (Vassilieva et al. 2001c). Structural studies on CsmH could thus provide information on five chlorosome proteins simultaneously. The CsmH protein is currently being overproduced in *Escherichia coli* in preparation for structural investigations by NMR and/or X-ray crystallography.

The significance of the structural degeneracy of chlorosome proteins is not clear. It could have its origins in the specific interactions between or among the proteins, and it is possible that there is a fixed stoichiometry between some of the chlorosome proteins, (or even between BChl *a* and a chlorosome protein), but this has not yet been demonstrated. It is also interesting to note that the proteins consisting of only single sequence motifs occur in pairs (CsmA/CsmE, CsmB/CsmF, and CsmC/CsmD). Specific interactions are currently being investigated by using chemical cross-linking of the proteins in intact chlorosomes followed by immunochemical detection of the cross-linked products. It will also be useful to obtain quantitative information on the chlorosome proteins. Preliminary experiments show that CsmA can be cross-linked to form multimers including species with the apparent mass of dimers, trimers, tetramers, pentamers and hexamers. Possible interactions between CsmA and CsmF have tentatively been identified in this manner as well.

Production and characterization of mutants lacking chlorosome proteins

Interposon mutagenesis has been employed in attempts to inactivate all ten genes that encode chlorosome proteins in *Cb. tepidum*. At present, only *csmA* has not yet been inactivated. Single-locus mutants carrying insertions in or deletions of all other *csm* genes have been obtained and verified by PCR or Southern blotting or both. The absence of the targeted protein(s) from chlorosomes has also been verified by immunoblotting.

Characterization of the mutants is still in progress, but surprisingly it appears that the loss of any single protein (except CsmA) has no significant effect on the absorption and fluorescence properties of the cells; no significant effect on chlorosome formation, stability, and protein composition; and little noticeable effect on growth rate under saturating light conditions. One phenotypic effect that has been observed so far is an apparent increase in the rate of oxygen-induced BChl *c* fluorescence decay in mutants lacking CsmI, CsmJ, or CsmX, which may indicate that these proteins may have a role in the regulating the redox-dependent quenching. Chlorosomes of all mutants will be investigated by electron microscopy and analyzed with respect to pigment content (BChl *c*, BChl *a*, carotenoids, and isoprenoid quinones). Finally, the growth rate of each mutant will be determined under limiting light intensity. Previous work indicated that *csmC* mutants in *Cb. vibrioforme* had a slower growth rate than the wild-type strain (Chung et al. 1998), and preliminary observations suggest that the same may be true for the *csmC* mutant of *Cb. tepidum*. Mutants deficient in more than one chlorosome protein are now being constructed (currently three selection markers are available in *Cb. tepidum* (Frigaard and Bryant 2001)). The following double mutants have already been constructed: *csmD csmE*, *csmI csmJ*, and *csmI csmX*. Interestingly, an initial attempt to construct a double mutant lacking both CsmJ and CsmX failed. This result may suggest that functional chlorosomes from both groups of green bacteria require the presence of at least one Fe-S protein for some specific but thus far unknown function.

The preliminary observations with these mutants are very puzzling and evoke some interesting questions. If the chlorosome proteins are not needed for growth or chlorosome formation, what is their role and why are they so highly conserved? Why are ten different proteins present, although they only represent four basic sequence/structural motifs? The chlorosome protein compositions in the three strains of green sulfur bacteria that we have investigated (*Cb. tepidum*, *Cb. vibrioforme*, and *Cb. phaeobacteroides*) are apparently identical. The four basic protein families have persisted during evolution and speciation of the green sulfur bacteria and are also observed in chlorosomes of the distantly related *Cf. aurantiacus*.

Using genomics to identify new chlorosome proteins

The determination of the genome sequences of *Cb. tepidum* (Eisen et al. 2001) and *Cf. aurantiacus* (Joint Genome Institute, 2001) allows exciting new approaches to the study of these organisms including the genetics of chlorosomes. For example, *csmX* was first identified in the genome sequence of *Cb. tepidum* as a homolog of *csmI* and *csmJ*. Subsequent immunoblotting of isolated chlorosomes using polyclonal antibodies raised against recombinant CsmX revealed that CsmX indeed is truly a chlorosome protein (Vassilieva et al. 2001c).

Until recently, only three chlorosome proteins had been identified in *Cf. aurantiacus* (CsmA, CsmM, CsmN) (Blankenship et al. 1995). CsmA is obviously related to CsmA in the green sulfur bacteria, is synthesized as a precursor protein, and is similarly the most abundant chlorosome protein in *Cf. aurantiacus* chlorosomes. CsmM and CsmN are distantly related to the CsmC/CsmD sequence motif family (Vassilieva et al. 2000, 2001c). As noted above, chlorosomes from *Cb. tepidum* contain three Fe-S proteins (Vassilieva et al. 2001a, 2001c), and a related protein denoted CsmY was recently identified in the genome sequence of *Cf. aurantiacus*. CsmY has strong sequence similarity to the proteins of the CsmI/CsmJ/CsmX family,

and this protein has now been detected by N-terminal sequence analysis in chlorosome fractions from *Cf. aurantiacus*. We have also identified the *csmO* gene in the *Cf. aurantiacus* genome sequence. CsmO is a member of the CsmB/CsmF family, and we have confirmed its presence in isolated chlorosomes by N-terminal sequence analysis. A 17.2 kDa protein, denoted CsmP, has also been detected in chlorosome fractions from *Cf. aurantiacus* (Fig. 1). The *csmP* gene is possibly transcribed as part of an operon that includes the *csmM* and *csmN* genes. Interestingly, the *csmP* gene has no homolog in the genome of *Cb. tepidum*, but related proteins are found in the cyanobacteria *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002 and occur as the "bacterio-opsin linked product" in *Halobacterium* sp. Finally, we have identified a gene, tentatively denoted *csmQ,* in *Cf. aurantiacus* for which the N-terminal sequence of the deduced protein is 38% identical to CsmM. CsmQ is predicted to be ~30 kDa and could be a novel chlorosome protein. In support of this hypothesis, a polypeptide of this size is observed in the gel shown in Fig. 1.

At present, only ten proteins are known to be truly localized in the chlorosome envelopes in *Cb. tepidum* (Vassilieva et al. 2000, 2001c). However, it is reasonable to expect that several other proteins are involved in chlorosome formation and that such proteins may also be co-localized in the chlorosomes as well as in the cytoplasm or the cytoplasmic membrane. The genomic sequences of *Cb. tepidum* and *Cf. aurantiacus* may provide some clues to identify such proteins. Genes with related function are often located in operons in procaryotes and the organization of such genes is often conserved among different species (even after horizontal gene transfer). In other words, if the same gene organization is found in two otherwise distantly related organisms, it is likely that these genes have a related function and/or may been co-transferred by horizontal gene transfer. Such a pattern is observed with *csmA* in *Cb. tepidum* and *Cf. aurantiacus* (Fig. 2).

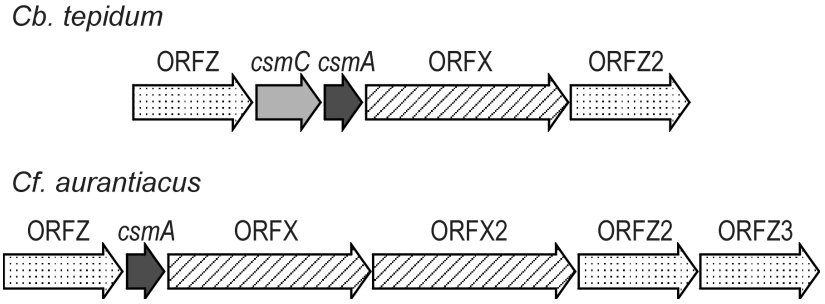

Fig. 2. Gene organization around *csmA* in *Cb. tepidum* and *Cf. aurantiacus*.

The *csmCA* operon in *Cb. tepidum* and *csmA* in *Cf. aurantiacus* are both flanked upstream by an *arsA* homolog which tentatively has been named ORFZ (Chung et al. 1994). ArsA is an extrinsic component of a membrane-bound, ATP-dependent arsenite exporter (Zhou et al. 2000). As shown in Fig. 2, one or two additional *arsA* homologs (ORFZ2, ORFZ3) are located downstream of *csmA*. Related genes are also located immediately downstream of *csmA* in both species (ORFX) which in *Cf. aurantiacus* has been duplicated (ORFX2). The ArsA protein family has specifically expanded in the ancestors of *Cb. tepidum* and *Cf. aurantiacus,* which contain other *arsA* homologs elsewhere in their genomes (Eisen et al. 2001). The functions of ORFZ, ORFZ2 and ORFX in *Cb. tepidum* are being studied by targeted gene

inactivation. The availability of the genomic sequences and the ability to manipulate *Cb. tepidum* genetically should provide many new insights into chlorosome structure, function and assembly.

Acknowledgments

The research in this article was supported by grant DE-FG02-94ER20137 from the U.S. Department of Energy, Division of Energy Biosciences, to DAB. N-U Frigaard was additionally supported by a postdoctoral fellowship from The Danish Natural Science Research Council.

References

Blankenship RE, Olson JM, Miller M (1995) In: Blankenship RE, Madigan MT, Bauer CE (eds.) *Anoxygenic Photosynthetic Bacteria*, Kluwer Academic Publishers, Dordrecht, pp. 399-435.

Chung S, Bryant DA (1996a) *Archives of Microbiology* **166**, 234-244.

Chung S, Bryant DA (1996b) *Photosynthesis Research* **50**, 41-59.

Chung S, Frank G, Zuber H, Bryant DA (1994) *Photosynthesis Research* **41**, 261-275.

Chung S, Shen G, Ormerod JG, Bryant DA (1998) *FEMS Microbiology Letters* **164**, 353-361.

Eisen JA, Nelson, KE, Paulsen, IT, et al. (2001) Submitted to *Nature.*

Foidl M, Golecki JR, Oelze J (1998) *Photosynthesis Research* **55**, 109-114.

Frigaard N-U, Bryant DA (2001) *Applied and Environmental Microbiology* **67**, 2538-2544.

Frigaard N-U, Matsuura K (1999) *Biochimica et Biophysica Acta* **1412**, 108-117.

Frigaard N-U, Ormerod J (1995) In: Mathis P (ed.) *Photosynthesis: From Light to Biosphere*, Kluwer Academic Publishers, Dordrecht, vol. I, pp. 163-166.

Joint Genome Institute (2001) http://www.jgi.doe.gov

Oelze J, Golecki JR (1995) In: Blankenship RE, Madigan MT, Bauer CE (eds.) *Anoxygenic Photosynthetic Bacteria*, Kluwer Academic Publishers, Dordrecht, pp. 259-278.

Olson JM (1998) *Photochemistry and Photobiology* **67**, 61-75.

Ormerod JG, Nesbakken T, Beale SI (1990) *Journal of Bacteriology* **172**, 1352-1360.

Sakuragi Y, Frigaard N-U, Shimada K, Matsuura K (1999) *Biochimica et Biophysica Acta* **1413**, 172-180.

van Rossum B-J, Steensgaard DB, Mulder FM, Boender GJ, Schaffner K, Holzwarth AR, de Groot, HJM (2001) *Biochemistry* **40**, 1587-1595.

Vassilieva EV, Frigaard N-U, Bryant DA (2000) *The Spectrum* **13**, 7-13.

Vassilieva EV, Antonkine ML, Zybailov BL, Yang F, Jakobs CU, Golbeck JH, Bryant DA (2001a) *Biochemistry* **40**, 464-473.

Vassilieva EV, Ormerod JG, Bryant, DA (2001b) *Photosynthesis Research*, in press.

Vassilieva EV, Stirewalt VL, Jakobs CU, Frigaard N-U, Baker MA, Sotak A, Bryant DA. (2001c) Manuscript in preparation.

Zhou TQ, Radaev S, Rosen BP, Gatti DL (2000) *EMBO Journal* **19**, 4838-4845.