Chapter 10

A Protein Family Saga: From Photoprotection to Light-Harvesting (and Back?)

Stefan Jansson*
Umeå Plant Science Centre,
Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden

Summary 145
I. Introduction 146
II. The Light-Harvesting Complexes (LHCs) of Higher Plants 146
A. Ten, Twelve, or Fourteen LHC Proteins? 146
B. Secondary and Tertiary Structure 146
C. The Cousins: One, Two, and Four-Helix Proteins 147
III. The Light-Harvesting Antenna of Lower Plants 147
A. Cyanobacteria 147
B. Green Algae 147
C. Liverworts, Mosses, and Ferns 148
D. Conifers 148
IV. The Evolution of LHC Proteins 148
A. One, Two, Four, Three 148
B. From Photoprotection to Light Harvesting? 150
V. The Evolution of Feedback De-Excitation 150
A. Cyanobacteria 151
B. Xanthophyll Conversions are Old 151
C. What is the "Quenching Protein" in *Chlamydomonas*? 152
VI. Conclusions 152
Acknowledgments 152
References 152

Summary

Photoprotection seems to be an intrinsic property of light-harvesting systems, and an interesting question to address is whether the light-harvesting or the photoprotection function was the "original" function, and which function evolved subsequently. It appears that the cyanobacterial one-helix proteins, the presumed ancestors to the LHC proteins, were not designed as antenna proteins but were involved in photoprotection and/or pigment metabolism. Some intermediate steps (two- and four-helix proteins) also seem to have photoprotective functions. The antenna function appeared later in evolution, and many different LHC proteins with somewhat diversified functions arose. To some extent, this happened before the lineages leading to *Chlamydomonas* and higher plants separated, but further diversification also took place following the split, and some of the proteins may have evolved in a direction away from optimizing light harvesting. When the evolution of feedback de-excitation is put into this evolutionary scheme, it is likely that xanthophyll conversions, that evolved previously to optimize photoprotection, were starting to be used as indicators of light stress and regulators of antenna function.

*Author for correspondence, email: stefan.jansson@plantphys.umu.se

B. Demmig-Adams, William W. Adams III and A.K. Mattoo (eds), Photoprotection, Photoinhibition, Gene Regulation, and Environment, 145–153.

I. Introduction

Novel genes do not fall from heaven, but evolve from pre-existing genes. This means that when a new biological function arises through evolution, this occurs through mutations that make a protein with one enzymatic or structural function gain a new function. A classical example of this is the evolution of crystallins, the structural proteins of eyes, where different enzymes have become major lens components in different vertebrate species (Wistow and Piatigorsky, 1987), probably as a consequence of a need for different ocular optics. The eye is a structure that has evolved several times during animal evolution, since the ability to interpret light signals from the surroundings and to change behaviour is, of course, expected to confer a huge evolutionary advantage to an animal. Higher plants use the information in light as signals to change developmental patterns but, and perhaps more importantly, the energy of light to drive the photosynthetic reactions. Photosynthetic reaction centers seem to have evolved only once; PS I, PS II, and the corresponding reaction centers of photosynthetic prokaryotes share all structural elements and are likely to be homologous structures. In contrast, the proteins of the light-harvesting systems that are present in the different taxa share no sequence, and very little structural, similarity and it is most likely that the light-harvesting systems of higher plants, cyanobacteria, purple bacteria, and green sulphur bacteria evolved independently from each other (Green, 2001). In this process, different proteins have been recruited to fulfill the function of coordinating the photosynthetic pigment molecules into ordered arrays that enable efficient transfer of excitation energy into the reaction centers where charge separation takes place. In the following, the evolution of the higher plant light-harvesting antenna structure into its present form, where light harvesting and light dissipation are intimately coupled processes, will be discussed.

II. The Light-Harvesting Complexes (LHCs) of Higher Plants

A. Ten, Twelve, or Fourteen LHC Proteins?

The proteins of the higher plant photosynthetic antenna, the light-harvesting chlorophyll a/b-binding (LHC) proteins, make up a protein family of ten principal members, plus a couple of related proteins (Jansson, 1994). The gene products of the Lhca1-4 genes associate with PS I and the Lhcb1-6 genes primarily with PS II, although the Lhcb1 and Lhcb2 proteins that, at least in low and intermediate light conditions, make up the bulk of the antenna, distribute between PS I and PS II to balance the flow-through of electrons in the photosystems. Several names have been used to designate the gene products (see Jansson, 1994 for a compilation). In the following, the names relating to gene names (e.g. Lhca1) will primarily be used, although e.g. Lhcb4 is commonly designated as CP29 in the photosynthesis literature, an informative name reflecting the mobility of the protein during electrophoresis. In terms of evolution, the focus of this paper, the Lhc acronyms are more consistent. Several of the LHC proteins are encoded by multiple genes; in *Arabidopsis thaliana* there are e.g. 5 and 3 genes encoding Lhcb1 and Lhcb2, respectively (Leutweiler et al., 1986; McGrath et al., 1992; Legen et al., 2001; Andersson et al., 2003). Although the individual Lhcb1 proteins have slightly different amino acid sequences, these differences are not conserved among plant species and thus they probably do not represent differences in functions, which would likely be conserved. This distinction is, however, not straightforward to make based on sequence differences alone. For instance, an *Arabidopsis* gene encoding a protein originally named Lhca2 (Zhang et al., 1992) encodes a distinct protein (Lhca6), and the same is probably true for the "third *Lhcb4* gene", *Lhcb4.3* (Jansson, 1999). Recently, a few additional related genes have been identified in *Arabidopsis*. Under normal conditions, these have a much lower expression level than the "normal" proteins, and these will be more thoroughly discussed later in the chapter.

B. Secondary and Tertiary Structure

Although LHC II (in this case a mixture of Lhcb1 and Lhcb2) is the only LHC protein whose tertiary structure has been experimentally determined (Kühlbrandt et al. 1994), all LHC proteins are believed to fold in an identical way. They consist of three membrane-spanning helices (MSHs), of which the first and the third are homologous to each other, and form a dimeric "core" of the protein. The LHC proteins bind pigment in various stoichiometries; chlorophyll a, chlorophyll b, xanthophylls, and the pigments of the xanthophyll cycle associate with most, if not all, of the polypeptides. The pigments are bound by non-covalent bonds and, although there is some conservation of the

Abbreviations: ELIP – Early light inducible proteins; HLIP – High light inducible proteins; LHC – Light-harvesting Complex; MSH – Membrane spanning helices; SCP – Small cab-like proteins