Alma Mater Studiorum - Università di Bologna

Facoltà di Scienze Matematiche Fisiche e Naturali

Scuola di Dottorato di Ricerca in Scienze Biologiche, Biomediche e Biotecnologiche

Dottorato di Ricerca in Biologia Funzionale dei Sistemi Cellulari e Molecolari

Settore disciplinare di afferenza: BIO/04 - Fisiologia Vegetale

CP12: INTRINSICALLY UNSTRUCTURED PROTEINS REGULATING PHOTOSYNTHETIC ENZYMES THROUGH PROTEIN-PROTEIN INTERACTIONS

PhD Student:

Dott.ssa LUCIA MARRI

Tutor:

Chiar. mo Prof.

PAOLO BERNARDO TROST

PhD Coordinator:

Chiar. mo Prof.

VINCENZO SCARLATO

Key words: intrinsically unstructured proteins – supramolecular complex – redox regulation – Calvin-Benson cycle - photosynthetic enzymes

Academic Years 2004/2006

Index

Introduction 1

1.1 The reductive pentose phosphate cycle for photosynthetic CO_2 assimilation: enzyme modulation in light/dark transitions 1

1.1.1 Redox regulation via the ferredoxin-thioredoxin system 2

1.1.2 Regulation by other physiological parameters 7

1.1.3 Specific regulatory mechanisms for RubisCo 7

1.1.4 Regulation via protein-protein interactions: participation of chloroplast enzymes in supramolecular complexes 9

1.2 Intrinsically unstructured proteins 11

1.2.1 General characterization of intrinsic disorder 11

1.2.2 Experimental characterization of disordered proteins 15

1.2.3 Functions of intrinsic disorder in proteins 16

1.2.4 Plant IUPs 21

1.2.5 Intrinsically unstructured CP12 and its role in protein-protein interactions 23

1.3 Light-regulation of photosynthetic glyceraldehyde-3-phosphate dehydrogenase 26

1.3.1 Diversity and evolution of GAPDH subunits and isozymes 26

1.3.2 A_4-GAPDH: the simplest isoform of photosynthetic GAPDH as a paradigm of the fully active enzyme 29

1.3.3 Photosynthetic A_4-GAPDH isoform is a target of glutathionylation 35

1.3.4 CP12 and the ancient regulatory system of photosynthetic GAPDH 37

1.3.5 CP12-dependent vs. CTE-mediated regulation of GAPDH: alternative regulatory systems based on the same molecular mechanism? 41

Chapter 2: Co-ordinated gene expression of photosynthetic glyceraldehyde-3-phosphate dehydrogenase, phosphoribulokinase, and CP12 in *Arabidopsis thaliana* 47

2.1 Abstract 47

2.2 Introduction 48

2.3 Materials and methods 50

2.3.1 Plant material and treatments 50

2.3.2 Gene cloning 51

2.3.3 RNA extraction and northern blot analysis 51

2.3.4 Enzyme activity assays 53

2.4 Results 54

2.4.1 In silico search and molecular cloning of members of GAPDH and CP12 gene families, and of PRK from Arabidopsis thaliana 54

2.4.2 Organ-specific expression 55

2.4.3 Light-dependent expression 57

2.4.4 Effect of sugars 58

2.5 Discussion 60

Chapter 3: Reconstitution and properties of the recombinant glyceraldehyde-3-phosphate dehydrogenase/CP12/phosphoribulokinase supramolecular complex of Arabidopsis 65

3.1 Abstract 65

3.2 Introduction 66

3.3 Materials and Methods 69

3.3.1 Plant Material 69

3.3.2 Expression Vectors 69

3.3.3 Expression and Purification of Recombinant Proteins 70

3.3.4 Enzyme Activity Assays and Redox Titrations 71

3.3.5 Electrophoresis and Immunoblotting 71

3.3.6 Supramolecular Complex Reconstitution and Gel Filtration 72

3.3.7 Complex Dissociation and Recovery of Enzyme Activity 73

3.4 Results 73

3.4.1 Expression, Purification, and Redox Properties of Arabidopsis Recombinant GAPDH, CP12, and PRK 74

3.4.2 In Vitro Reconstitution of Binary and Ternary Complexes 75

3.4.3 In Vitro Dissociation of Binary and Ternary Complexes 79

3.5 Discussion 81

Chapter 4: Spontaneous assembly of photosynthetic supramolecular complexes mediated by the intrinsically unstructured protein CP12 87

4.1 Abstract 87

4.2 Introduction 88

4.3 Experimental procedures 91

4.3.1 Protein expression and purification 91

4.3.2 CP12-2 site specific mutants 91
4.3.3 NMR spectra 92
4.3.4 Analysis of thiol groups and redox titration of CP12-2 92
4.3.5 Isothermal titration calorimetry (ITC) 93
4.3.6 Multiangle light scattering (MALS-QELS) 94
4.4 Results 94
4.4.1 CP12-2 of *Arabidopsis thaliana* is an intrinsically unstructured protein 94
4.4.2 CP12-2 redox properties 96
4.4.3 Interaction between A_4-GAPDH, CP12-2 and PRK 97
4.4.4 CP12-2 site-specific mutants 100
4.5 Discussion 102

Chapter 5: The CP12 protein family in *Arabidopsis thaliana* 109
5.1 Introduction 109
5.2 Materials and methods 112
5.2.1 In silico analysis of CP12 isoforms 112
5.2.2 Protein expression and purification 112
5.2.3 Binary and ternary complex reconstitution 113
5.2.4 GAPDH and PRK activity assays 114
5.2.5 Redox titration of CP12 isoforms 114
5.3 Results 116
5.3.1 CP12 isoforms as hypothetical IUP proteins 116
5.3.2 Redox properties of the CP12 family 120
5.3.3 CP12 isoforms: the same linker function in supramolecular complex formation? 123
5.4 Discussion 127

Conclusions and future perspectives 131

References 139

Abbreviations

ACTR	activator for thyroid hormone and retinoid receptors
ATP	adenosine 5’-triphosphate
BPGA	1,3-bisphosphoglycerate
CD	Circular Dichroism-spectroscopy
CTE	C-terminal extension of subunits GAPDH subunit B
DTNB	5,5’-Dithiobis-(2-nitro-benzoic acid)
DTT	dithiotreitol
$E_{m,7.9}$	midpoint redox potential
EST	expressed sequence tag
FBPase	fructose-1,6-bisphosphatase
Fd	ferredoxin
FTR	ferredoxin-thioredoxin reductase
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GRX	glutaredoxin
ITC	isothermal titration calorimetry
IUP	Intrinsically Unstructured Proteins
K_D	dissociation constant
MALDI-TOF	matrix-assisted laser desorption ionization – time-of-flight
MALS-QELS	multiangle light scattering-quasi elastic light scattering
MDH	NADP-dependent malate dehydrogenase
M_r	molecular mass
MSP	manganese stabilizing protein
n	number of binding sites
NAC	non-Aβ component of the amyloid plaque
NAD(H)	nicotinamide adenine dinucleotide
NADP(H)	nicotinamide adenine dinucleotide phosphate
NCBD	nuclear-receptor co-activator binding domain of CREB-binding protein
NMR	Nuclear Magnetic Resonance
NTR	NADPH dependent TRX reductase
PGK	3-phosphoglycerate kinase
PONDR®	Predictor of Naturally Disordered Regions
PRK	phosphoribulokinase
PSII	photosystem II
R_H	hydrodynamic radius
RubisCo	ribulose-1,5-bisphosphate carboxylase/oxygenase
RuBP	ribulose bisphosphate
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TNB	2-nitro-5-thiobenzoate (thiolate) dianion
TRX	thioredoxin
ΔG	Gibbs energy change
ΔH	enthalpy change
ΔS	entropy change

Introduction

*based in part on:

Thioredoxin-dependent regulation of photosynthetic glyceraldehyde-3-phosphate dehydrogenase: autonomous vs. CP12-dependent mechanisms*

Paolo Trost[1], Simona Fermani[2], Lucia Marri[1], Mirko Zaffagnini[1], Giuseppe Falini[2], Sandra Scagliarini[1], Paolo Pupillo[1] and Francesca Sparla[1]
(2006) *Photosynth Res* **89:** 1-13

(1) Laboratory of Molecular Plant Physiology, Department of Evolutionary Experimental Biology, University of Bologna, Bologna, Italy

(2) Department of Chemistry "G. Ciamician", University of Bologna, Bologna, Italy

1.1 *The reductive pentose phosphate cycle for photosynthetic CO_2 assimilation: enzyme modulation in light/dark transitions*

Plants are sessile photosynthetic organisms living under continuously changing light regime and challenging environmental conditions. The Calvin–Benson pathway of CO_2 assimilation into carbohydrates is a major route for the chemical conversion of light energy in all photosynthetic eukaryotes and in cyanobacteria, often accounting for > 90% of total photosynthetic output. The reductive pentose phosphate cycle constitutes the interface between the transduction of sunlight energy in thylakoid membranes and the utilization of metabolites in other cellular compartments. By the simultaneous activation of biosynthetic enzymes of the cycle and inactivation of the enzymes that degrade carbohydrates, light drives chloroplasts to the production of reserve materials and concurrently minimizes the operation of futile cycles. Therefore, matching the rate of photosynthetic carbon assimilation with the rate of photosynthetic

production of ATP and NADPH is an absolute priority for plants. To achieve this, plants possess a complex regulatory network involving a considerable variety of proteins and molecular signals.

Two unique systems that link light-triggered events in thylakoid membranes with enzyme regulation are located in the soluble portion of chloroplasts (stroma): the ferredoxin-thioredoxin system and ribulose 1,5-bisphosphate carboxylase/oxigenase activase (Rubisco-activase) (Wolosiuk *et al.,* 1993).

1.1.1 *Redox regulation via the ferredoxin-thioredoxin system*

The inability of chloroplasts to fix CO_2 in the dark is not only due to lack of metabolic precursors but also to inactive enzymes. Upon illumination, the concentration of intermediates of the cycle increases in the chloroplast stroma and regulatory enzymes are converted to active forms. In chloroplast stroma, the ferredoxin activity, sensitive to light signal, is used for the generation of the modulator of enzyme activity (thioredoxin). In the presence of reduced ferredoxin, ferredoxin-thioredoxin reductase (FTR) catalyzes the reduction of the disulfide bridge of thioredoxin by photosynthetic electron transport. In turn, reduced thioredoxin modulates the activity of enzymes related to carbon, nitrogen and sulphur metabolism transferring the electrons to regulatory disulfides of target enzymes via thiol/disulfide interchange reactions. The ratio of the reductant (ferredoxin) to the oxidant (O_2) fixes the proportion of the reduced thioredoxin, which determines the concentration of the active enzyme. In the light, high concentration of reduced thioredoxin converts regulatory enzymes of the cycle to the active form. Alternatively, if the formation of reduced ferredoxin diminishes under illumination, enzymes of the cycle would return to the oxidized inactive form (Wolosiuk *et al.,* 1993; Buchanan and Balmer, 2005) (fig. 1.1).

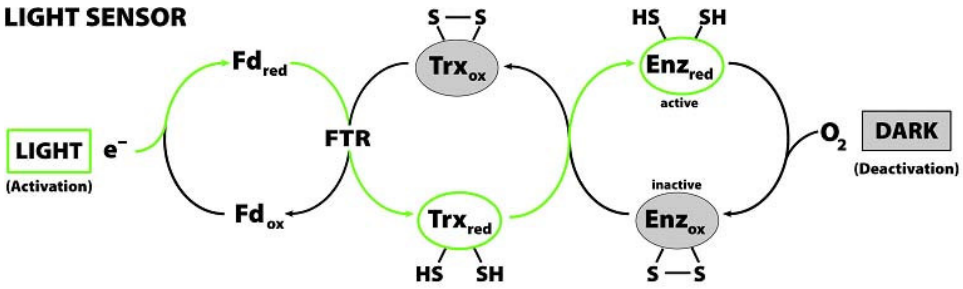


Fig. 1.1: Role of the ferredoxin/thioredoxin system in sensing light in chloroplasts (from Buchanan and Balmer, 2005).

At variance with ferredoxin-thioredoxin reductase, which is restricted to photosynthetic organisms, thioredoxins are widely distruibuted in most free-living organisms. Thioredoxins (TRX) are small redox proteins of ca. 12 kDa with the distinctive feature of the maintenance of their conserved redox center W-C-G(P)-P-C and a midpoint redox potential comprised between -270 and -300 mV (Hirasawa *et al.*, 1999; Collin *et al*., 2003; Michelet *et al*., 2006). Non-photosynthetic eukaryotes contain a limited number of citosolic and mitochondrial TRXs reduced by a NADPH dependent TRX reductase (NTR) present in these compartments. These TRX have been implicated in defense mechanisms against oxidative stress (Michelet *et al.,* 2006).

In *Arabidopsis thaliana*, a multigenic family of 19 TRXs grouped in six different TRX subfamilies has been recently identified, long since TRXs were first described in the late 70s (Schurmann and Jacquot, 2000) (fig. 1.3). The chloroplast thioredoxin system was long considered to be comprised of two thioredoxins (*f* and *m*) in addition to ferredoxin and FTR (Dai *et al.,* 2004) (Fig. 1.2). The first targets identified for this TRXs were key enzymes of the carbon metabolism (fructose-1,6-bisphosphatase; phosphoribulokinase, PRK; glyceraldehyde-3-phosphate dehydrogenase, GAPDH; sedeptuolose-1,7-bisphosphatase; NADP-malate dehydrogenase MDH; Rubisco activase and ATP-synthase γ-subunit). TRX *f* and *m* show

some degree of specificity *in vitro*: *f* being linked to the specific activation of Calvin cycle enzymes in the light, while *m* to either glucose 6-phosphate dehydrogenase or, under certain conditions, to NADP-malate dehydrogenase (Buchanan and Balmer, 2005).

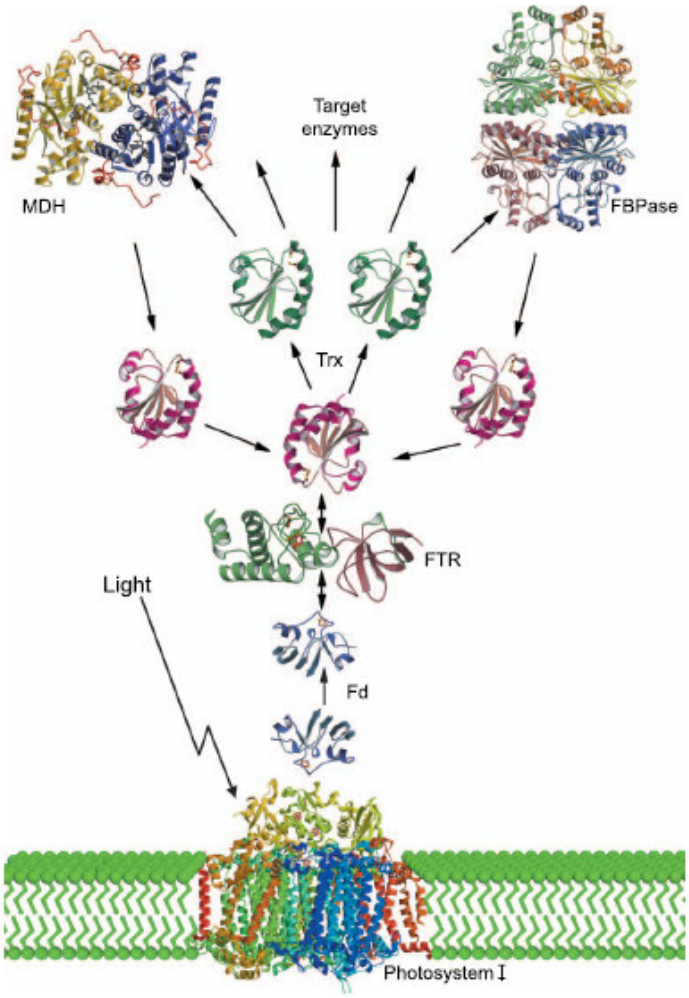


Fig. 1.2: Light-induced enzyme regulation in chloroplasts through the FTR system. Upon illumination, the photosynthetic electron-transfer chain is initiated by the reduction of ferredoxin (Fd) by Photosystem I. Fd can then reduce FTR, which in turn reduces the chloroplast thioredoxins (TRX). Finally, the TRX activate target enzymes, here exemplified by malate dehydrogenase (MDH) and fructose 1,6-bisphosphatase (FBPase), thereby changing the metabolism to anabolic pathways. In the figure, the oxidized thioredoxins are in red and reduced thioredoxins in green. (Dai *et al.*, 2004).

A third type of TRX was discovered in the cytosol, in the nucleus, in the endoplasmic reticulum and in mitochondria (fig. 1.3). These *h* type TRXs (*h* for heterothrophic), reduced by the NADPH/NTR system, are involved in the mobilization of seed reserve during germination, in self-incompatibility

mechanisms and ROS-detoxification (Gelhaye *et al.,* 2004). Recent genome research has added two additional chloroplastic TRX members (thioredoxins *x* and *y*) and mitochondrial TRX *o* (Laloi *et al.,* 2001). The experimental evidence suggests that thioredoxins *f* and *m* represent the major chloroplast forms and *x* and *y*, whose functions are involved in the defense mechanisms against oxidative stress, are of lower abundance (Collin *et al.,* 2003; Buchanan and Balmer, 2005).

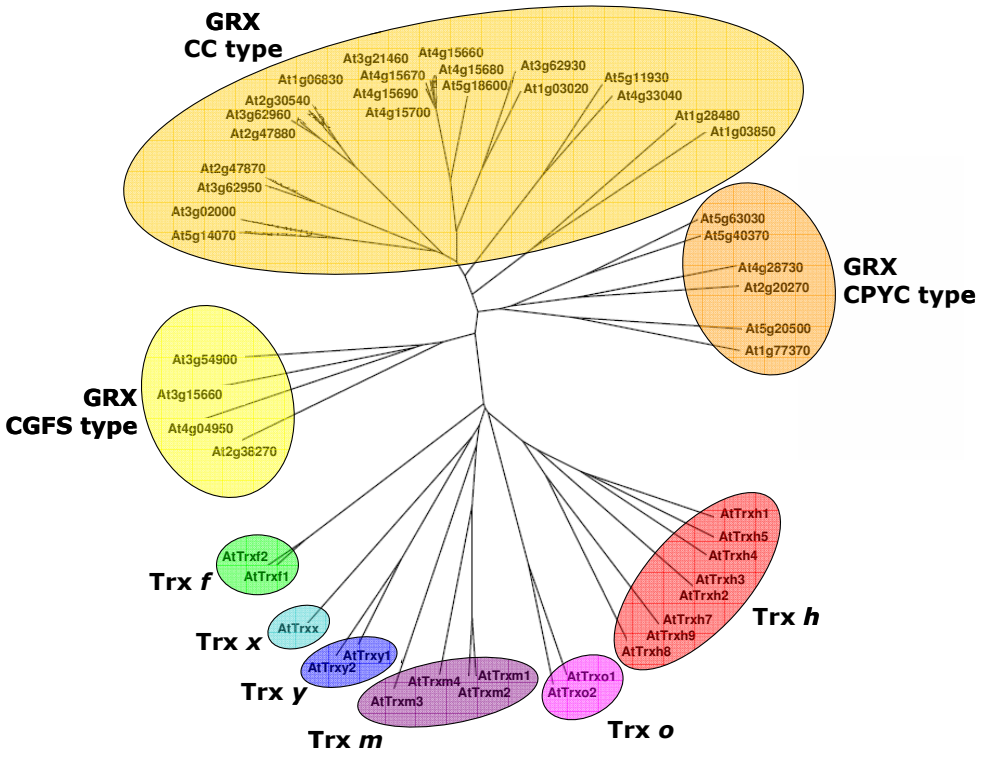


Fig. 1.3 The thioredoxin and glutaredoxin families in Arabidopsis. Alignment was obtained with Clustal W while, for the phylogenetic tree reconstruction, TreeTop-Phylogenetic Tree prediction of the GeneBee service and Phylodendron of the IUBio Archive at the Indiana University were used . http://www.genebee.msu.su/services/phtree_reduced.html, http://iubio.bio.indiana.edu/treeapp.

For Arabidopsis thaliana (At) the names indicated correspond to AGI codes for GRX. Accession numbers for TRX are as follows: h1, Z14084; h2, Z35475; h3, Z35474; h4, Z35473; h5, Z35476; h7, AAD39316; h8, AAG52561; h9, AAG51342; o1, AAC12840; o2, AF396650; m1, O48737; m2, Q9SEU8; m3, Q9SEU7; m4, Q9SEU6; f1, Q9XFH8; f2, Q9XFH9; x, AAF15952; y1, AAF04439; y2, AAM91085.

Closely related and similar to thioredoxins, glutaredoxins (GRX) are ubiquitous oxidoreductases using glutathione as electron donor and which possess a typical glutathione-reducible CxxC, CxxS or CC active site. GRXs are small proteins (ca. 10-15 kDa) which exhibit a TRX similar 3D structure, with a more positive redox potential comprised between -190 and -230 mV (Rouhier *et al.,* 2004; Michelet *et al.,* 2006). At least 31 members of the GRX superfamily were identified in Arabidopsis, grouped in 3 clusters with distinct active sites (the CPYC type, the CC type and the CGFS type) (Rouhier *et al.,* 2004; Buchanan and Balmer, 2005) (fig. 1.3). The specificity of the various apparently redundant forms within the GRX group or between GRX and TRX were analysed in terms of differential spatiotemporal expression of the genes, specificity vs. target proteins and mode of catalysis (glutathionylation/ deglutathionylation processes appear to be a specific function of glutaredoxin). GRXs seem to be involved in oxidative stress responses. Besides the redox-signalling mechanism mediated by dithiol/disulfide exchanges under TRX control, ROS and antioxidants could function as intracellular messengers being produced by cells under normal growth conditions and not only in stress conditions. Reversible oxidations of Cys residues appear to play a major role in ROS mediated signalling under the control of TRX, GRX and related proteins (Rouhier *et al.*, 2004, Michelet *et al.*, 2006).

In photosynthetic cells in the light, the chloroplast is the main site of ROS production. Several environmental parameters, high light for instance, can lead to enhanced ROS production and oxidative damage (Foyer and Noctor, 2005). With regard to cysteines, different oxidation states can exist: Cys thiols can be oxidized into sulfenic (SOH), sulfinic (SO_2H) or sulfonic (SO_3H) acids. Glutathione has multiple functions in plants, it could act in redox signalling through modification of Cys residues either directly by protein glutathionylation or indirectly, via GRX. This post-translational modification can constitute a reversible antioxidant mechanism that evolved to protect Cys residues from irreversible oxidation of –SH groups to sulfinic or sulfonic

acids and subsequent degradation of the protein (Buchanan and Balmer, 2005; Michelet *et al.,* 2006).

1.1.2 *Regulation by other physiological parameters*

Conditions in chloroplasts change dramatically in light vs. darkness. At the onset of illumination, stromal pH increases from 7 to 8 as protons are pumped into thylakoids (Kramer et al., 1999) and most intermediates of the Calvin–Benson cycle strongly increase concomitant with activation of photosynthetic metabolism (Gerhardt et al., 1987). Accordingly, it has been found that the optimum pH of most enzymes of the Benson-Calvin cycle is close to 8 (Heldt *et al.,* 1973). The proton traslocation across the membrane also causes the movement of a counter-ion, Mg^{2+}, whose concentration in the stroma has been shown to increase upon dark to light transition. The energy charge of adenylates increases in chloroplasts in the light and pyridine nucleotides become more reduced (Stitt *et al*., 1982; Heineke *et al*., 1991). Moreover, the pool of di-phosphorylated pyridine nucleotides (NAD and NADH) decreases in the light while tri-phosphorylated NADP and NADPH symmetrically increase (Muto *et al*., 1981; Heineke *et al.,* 1991), possibly as a result of a light-modulated NAD-kinase (Delumeau *et al.,* 2000; Chai *et al*., 2005).

1.1.3 *Specific regulatory mechanisms for RubisCo*

Subject to extensive processing and constant conformational changes is ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCo), for its functional role as a limiting factor in photosynthetic CO_2 assimilation. RubisCo has successfully served as a model for cellular processes centred on protein chemistry and amino acid modification (Houtz and Portis, 2003). Once translated, the large and small subunits of RubisCo undergo a co- and post-translational modifications accompanied by constant interaction with

structurally modifying enzymes. Extensive studies of the regulation of RubisCo have shown that its activity is regulated by variations in its activation state rather than by fluctuations in substrate concentration, ribulose bisphosphate (RuBP).

In vivo, both CO_2 and Mg^{2+} are required for the activation process: CO_2 for the carbamylation on Lys-201 and the formation of a Mg^{2+} binding site for the stabilization of the catalitically competent enzyme. This process is non-functional *in vivo* because ribulose bisphosphate binds tightly to native RubisCo preventing the carbamylation of the enzyme (Wolosiuk *et al.,* 1993). So RubisCo activation is mediated by the stromal protein RubisCo activase which acts by removing the otherwise inhibitory sugar phosphate, RuBP and subsequently facilitates the access of CO_2 and Mg^{2+} for the carbamylation of RubisCo (Houtz and Portis, 2003; Graciet *et al.,* 2004*a*).

RubisCo activase is a member of the ATPases associated with diverse cellular activities (AAA^+) protein family, that constitutes a wide variety of proteins with chaperone-like functions, typically involving the disruption of molecular and macromolecular structures and usually partecipating in macromolecular complexes composed of several AAA^+ proteins. It is a nuclear-encoded chloroplast protein that usually consists of two isoforms generated by alternative splicing of a pre-mRNA. The activity of RubisCo activase is regulated by ADP/ATP ratio and the response is modified by redox changes in the larger isoform that are mediated by thioredoxin f (Houtz and Portis, 2003; Graciet *et al.,* 2004*a*). The activity of RubisCo, therefore, can be regulated via the activase in response to light intensity and down regulated in response to triose phosphate use limitations associated with inadeguate sinks. When plants are placed under various types of environmental stress, a common response is accelerated senescence and a loss of RubisCo protein; RubisCo degradation during this process is dependent upon oxidative modification (Houtz and Portis, 2003). In transgenic plants, the major mechanism for a compensation of a decrease in the amount of RubisCo is an increase in the RubisCo activation

state, presumably involving RubisCO activase (Houtz and Portis, 2003; Graciet *et al.,* 2004*a*).

1.1.4 *Regulation via protein-protein interactions: participation of chloroplast enzymes in supramolecular complexes*

The idea that the enzymes involved in metabolic pathways, as Calvin-Benson cycle, are not randomly distributed in the chloroplast stroma, but interact to give multienzyme complexes, has been forwarded because some organized structures have been isolated. Inside these complexes, enzymes bind to other protein in a highly specific manner to form more organized structure and interact with many components of the cell, such as the membrane (Wolosiuk *et al.,* 1993; Graciet *et al.,* 2004*a*). For some pathways, the term "metabolon" has been introduced to describe supramolecular complexes of sequential metabolic enzymes and structural components. In the cell, protein are packed together and the mean distance between them is lower than the mean diameter of a protein (Goodsell, 1991). Thus, specific interactions between them, requiring both spatial and electrostatic complementarities, are very likely to occur *in vivo*. The formation of photosynthetic enzyme supercomplexes could facilitate concerted enzyme catalysis accompanied by direct channelling of substrate and metabolites (Süss *et al.,* 1993). There is a growing body of evidence that chloroplast enzymes form supramolecular complexes with thylakoid membranes or other component of the Benson-Calvin cycle.

Different authors have isolate multienzyme complexes from *Chlamydomonas reinhardtii* (Lebreton and Gontero, 1999; Gontero *et al.,* 2002; Graciet *et al.,* 2004a), pea and spinach (Muller, 1972; Süss *et al.,* 1993; Scheibe *et al.,* 2002) with different composition. Some evidence is provided that the Calvin cycle enzymes ribose-5-phosphate isomerase, ribulose-5-phosphate kinase, ribulose-1,5-carboxylase, glyceraldehyde-3-

phosphate dehydrogenase, sedoheptulose-1,7-bisphosphatase and electron transport protein ferredoxin-$NADP^+$ reductase are organized into stable CO_2-fixing multienzyme complexes with a molecular mass of 900 kDa (Süss *et al.,* 1993). Gontero *et al.* purified from higher plants the phosphoribulokinase activity associated with other four enzymes of the reductive pentose phosphate cycle: ribose-5-phosphate isomerase, RubisCo, 3-phosphoglycerate kinase and glyceraldehydes-3-phosphate dehydrogenase (Gontero *et al*., 1988). In some case, chloroplast enzymes are loosely held in supramolecular complexes through relative weak interactions, and dilution is sometimes sufficient for the dissociation from these structures. Sometimes, smaller complexes may be considered as subcomplexes of higher order structures. The interaction between two Calvin cycle enzymes: glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK) is one of the best documented examples, isolated from spinach leaves (Clasper *et al.*, 1991; Scheibe *et al.,* 2002), green algae as *Chlamydomonas reinhardtii* (Lebreton and Gontero, 1999), *Scenedesmus obliquus* and in the cyanobacterium *Synechocystis* PCC6803 (Wedel and Soll, 1998). This complex may, therefore, correspond to the core complex of a supercomplex involved in CO_2 assimilation.

In the last decade, evolving techniques and increasing studies permit the identification, inside huge complexes, also of small proteins with the hypothetic role of promoting aggregation. These small proteins, for their effect in promoting complex formation and/or modulation of partner enzymes, need proper characteristics, as the lack of a defined tertiary structure and a high flexibility. Their characteristics and properties clearly justified the recurrent difficulty in their identification inside greater complexes, and for many years some organized structures were supposed to be simpler that they really are.

1.2 *Intrinsically unstructured proteins*

A major challenge in the post-genomic era will be the determination of the functions of the encoded protein sequences. In the recent years the field of the structural biology has grown explosively and a huge sets of atomic coordinates have been deposited in the Protein Data Bank. In parallel with this dramatic growth in structural knowledge, prediction or experimental determination of the library of protein structures is a matter of high priority. There has been an increasing conviction that the biological function of proteins is intrinsically encoded in their detailed 3D structures. The central dogma of structural biology is that a folded protein structure is necessary for biological function. However, a large proportion of gene sequences in eukaryotic genomes appear to encode not for folded, globular proteins, but for long stretches of amino acids that are likely to be either unfolded in solution or adopt non globular structures of unknown conformation. For their characteristics, those proteins are known as Intrinsically Unstructured Proteins (IUPs) (Dyson and Wright, 1999; 2005; Tompa, 2002). Disordered regions can be highly conserved between species in both composition and sequences and, contrary to the traditional view, disordered regions are often functional.

1.2.1 *General characterization of intrinsic disorder*

There have been several attempts to predict disorder. Perhaps the earliest are methods finding regions of low complexity. Intrinsically unstructured proteins and regions, which are also known as natively unfolded and intrinsically disordered, differ from structured globular proteins and domains with regard to many attributes, including amino acid composition, sequence complexity, hydrofobicity, charge, flexibility and type and rate of amino acid substitution over evolutionary time. These differences can be utilized to predict intrinsic order and disorder from the amino acid sequence.

A signature of probable intrinsic disorder is the presence of low sequence complexity and amino-acid compositional bias, with a low content of bulky hydrophobic amino acids (Val, Leu, Ile, Met, Phe, Trp and Tyr), which would normally form the core of a folded globular protein, and a high proportion of particular polar and charged amino acids (Gln, Ser, Pro, Glu, Lys and, on occasion, Gly and Ala). The amino acid composition of IUPs results in their inability to fold due to the depletion of typically buried amino acids and enrichment of typically exposed ones, which promote an extended conformation by electrostatic repulsion. This implies that globular proteins have sequences with the potential to form a sufficient large number of favourable interactions, whereas IUPs do not (Tompa, 2002; Dyson and Wright, 2005). In general, the depletion in cysteine fits into this picture, as in globular proteins this amino acids often occurs in active sites or stabilizing disulfide bonds which are not required in IUPs. A high proline content is also linked with the lack of structure, because proline is known to disfavour a rigid secondary structure and has a strong preference for an open conformational motif (Tompa, 2002). Methods using hydrophobicity can give hints as to disordered regions, as they are typically exposed and rarely hydrophobic. Uversky et al. noticed that proteins disordered over their entire lengths can be separated from ordered proteins by considering their average net charge and hydropathy (Uversky *et al*., 2000).

Several tools are designed specifically for prediction of protein disorder (table 1.1; Radivojac *et al*., 2006), such as PONDR® (*P*redictor *o*f *N*aturally *D*isordered *R*egions; Romero *et al.,* 2001) the first tool designed specifically for prediction of protein disorder, based on artificial neural networks. Some alternative methods, more than 20 different disorder predictors, have been developed: for example GlobPlot (Linding *et al.,* 2003*b*), DisEMBL (Linding *et al.,* 2003*a*), DISOPRED (Ward *et al.*, 2004), IUPred (Dosztanyi *et al*., 2005) or FoldUnfold (Galzitskaya *et al*., 2006). These predictors are based on a spectrum of computational approaches relying on amino acids composition, derived properties (such as secondary structure prediction) or simple physicochemical properties (such as charge) of the local sequence

neighbourhood. Based on these predictors, an increasing number of proteins and regions were recently recognized as unfolded or partially disordered, and DisProt, a web server for IUPs, was developed to enable research by collecting and organizing knowledge regarding the experimental characterization and the functional associations of IUPs (http://www.disprot.org; Sickmeier *et al*., 2007).

Table 1.1: Summary of the web servers offering predictions of intrinsically disordered proteins, from Radivojac *et al*., 2006.

Server Name	URL	Approach
VLXT (PONDR®)	http://www.pondr.com	Feed-forward neural network with separate N-/C-terminus predictor. Based on amino acid compositions and physicochemical properties.
FoldIndex®	http://bip.weizmann.ac.il/fldbin/findex	Charge/hydrophobicity score based on a sliding window.
NORSp	http://rostlab.org/services/NORSp/	Rule-based using a set of several neural-networks. Amino acid compositions and sequence profiles used as features.
VL2/VL3	http://www.ist.temple.edu/disprot/predictor.php http://www.pondr.com	Ordinary least-squares linear regression (VL2) and bagged feed-forward neural-network (VL3). All models use amino acid compositions and sequence complexity. VL3 series uses sequence profiles.
DISOPRED	http://bioinf.cs.ucl.ac.uk/disopred/	Feed-forward neural network (DISOPRED) and linear support vector machine (DISOPRED2) based on sequence profiles.
GlobPlot	http://globplot.embl.de/	Autoregressive model based on amino acid propensities for disorder/globularity.
DisEMBL™	http://dis.embl.de/	Ensemble of feed-forward neural networks.
IUPred	http://iupred.enzim.hu/index.html	Linear model based on the estimated energy of pairwise interactions in a window around a residue.
PreLink	http://genomics.eu.org/spip/PreLink	Rule-based. Ratio of multinomial probabilities (for linker and structured regions) combined with the distance to the nearest hydrophobic cluster.
RONN	http://www.strubi.ox.ac.uk/RONN	Feed-forward neural network in the space of distances to a set of prototype sequences of known fold state.
DISpro	http://www.igb.uci.edu/servers/psss.html	Recursive neural network based on sequence profiles, predicted secondary structure and relative solvent accessibility.
VSL	http://www.ist.temple.edu/disprot/predictorVSL2.php	Logistic regression (VSL1) and linear support vector machine (VSL2) based on sequence composition, physicochemical properties and profiles. Combination of short and long disorder predictors.
DRIP-PRED	http://www.sbc.su.se/~maccallr/disorder/	Kohonen's self organizing maps based on sequence profiles.
SPRITZ	http://protein.cribi.unipd.it/spritz/	Non-linear support vector machine based on multiply aligned sequences. Separate predictors for short and long disorder regions.

From the general analysis of well-known protein structures present in the database, classification arises. Proteins fall into a structural continuum,

from tightly folded single domains, to multidomain proteins that might have flexible or disordered regions, to compact but disordered molten globules and, finally to highly extended, heterogeneous states (fig. 1.4) (Dyson and Wright, 2005). This continuum has been interpreted in terms of a "protein trinity" (ordered, molten globule and random coil; Dunker and Obradovic, 2001) or the protein-quartet (Uversky, 2002), with a number of different structural types within each of these subdivisions. From the ordered forms to the random coil-like fully extended forms, a collapsed molten-globule conformation and a pre-molten globule exist, which are distinguishable by the presence of unstable secondary structures (Radivojac *et al.,* 2006). It follows that protein function is associated with any of three (or four) distinct forms or with transitions between them.

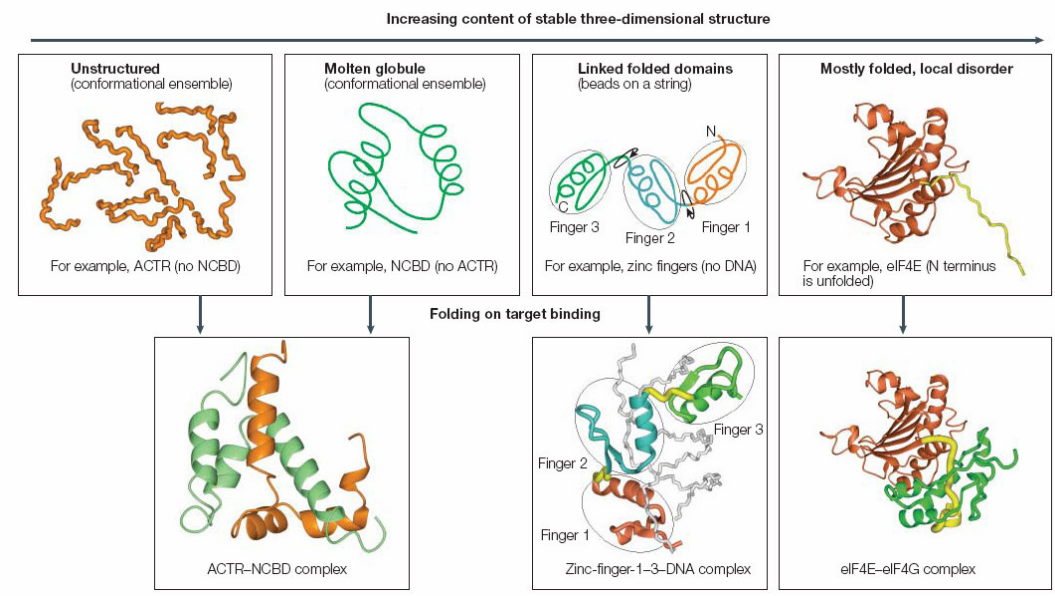


Fig. 1.4: The continuum of protein structure, from Dyson and Wright, 2005. From an unstructured conformational ensemble, as the interaction domain of activator for thyroid hormone and retinoid receptors (ACTR); a molten globule-like domain such as the nuclear-receptor co-activator binding domain (NCBD) of CREB-binding protein; linked folded domains such as the first three zinc-fingers of transcription factor-IIIA; to the eukaryotic translation-initiation factor eIF4E, which is mostly folded with only local disorder. The lower panels show the structure of the domains in the upper panels when they are folded onto their biological target domains or sequences. The mutually folded structure of the ACTR-NCBD complex; the well-ordered structure of the first three zinc-fingers of TFIIIA bound to DNA sequence; the complex between eIF4E and eIF4G, which highlights the mutual folding of the N-terminal tail of eIF4E and eIF4G.

1.2.2 *Experimental characterization of disordered proteins*

Protein disorder is only indirectly observed by a variety of experimental methods, such as X-ray crystallography, Nuclear Magnetic Resonance (NMR), Raman- and CD-spectroscopy, and hydrodynamic measurements. Each of these methods detects different aspects of disorder resulting in several operational definitions of protein disorder. Most protein structures known so far have been elucidated by X-ray crystallography. However. This technique is not suitable to characterize disordered regions of proteins, as corresponding regions are missing from the electron density map due to their intrinsic flexibility. NMR spectroscopy is the most powerful method to obtain site-specific information on the structure and dynamics of IUPs (Dyson and Wright, 2004). Other techniques, including CD and other spectroscopic methods, proteolytic degradation can also give important information about protein disorder. However, as each technique probes different aspects of protein structure, they do not necessarily correctly identify disorder. For example, loopy proteins, which have non repetitive secondary structure, would appear disordered by CD but ordered by the other technique; because a fundamental problem using CD is that the lack of regular secondary structure does not imply that the protein is disordered, merely that it is a "loopy protein". With NMR, disorder often is concluded from poor signal dispersion, which does not distinguish between random-coils and molten globule of high potential to fold in the presence of a partner. In X-ray crystallography, crystal packing may enforce certain disordered regions to become ordered and disordered binding segments are often crystallized in complex with their partner and are classified ordered despite their lack of structure in isolation. As a result, the inclusion of ordered segment in disordered databases and the exclusions of disordered protein/segment gave false-positive and false-negative classifications. In consequence, predictors trained on this dataset for assessing disorder reflect these uncertainties.

Unfolded conformations can also be detected and characterized by hydrodynamic techniques. Gel filtration (size-exclusion chromatography), small-angle X-ray scattering, sedimentation analysis and dynamic light scattering provide informations on hydrodynamic parameters, such as the Stokes radius (R_s, the apparent radius of a sphere with identical hydrodynamic behaviour to IUP). In these experiments IUPs resemble the denatured states of globular proteins. The extreme proteolytic sensitivity of IUPs results from protease cleaving substrates at sites that are sterically accessible an flexible enough to make productive contacts with the enzyme. Globular domains thus show significant resistance; IUPs are usually very sensitive along their entire length. A further method is sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) which is routinely used to assess the M_w of proteins. Because of their unusual amino acid composition, IUPs bind less SDS that usual and their apparent M_w is often 1.2-1.8 times higher than the real one calculated from sequence data or measured by mass spectrometry (Tompa, 2002).

1.2.3 *Functions of intrinsic disorder in proteins*

Analysis of sequence data for complete genomes indicates that intrinsically disordered proteins are highly prevalent, and that the proportion of proteins that contain such segments increases with the increasing complexity of an organism. The importance of protein disorder is further underlined by its prevalence in various proteomes. Ward has refined and systematized such an analysis and concluded that the fraction of proteins containing disordered regions of 30 residues or longer (predicted using DISOPRED) were 2% in archaea, 4% in bacteria and 33% in eukarya. In *Arabidopsis thaliana*, the percentage of chains with contiguous disordered segments of length greater than 30 and 50 residues, are 33.8% and 19.0% respectively (Ward *et al*., 2004). If indeed a great proportion of proteins contained long disordered regions, the natural question to ask was, what biological functions are carried out by these IUPs?

Protein disorder is important for understanding protein function as well as protein folding pathways. The functional importance of the unstructured state is underlined by the fact that most of these proteins have basic regulatory roles in key cellular processes. Protein disorder dominates in protein associated with signal transduction, cell-cycle regulation, gene expression, chaperon action and thus it is often implicated in cancer and diseases. Structural disorder actually predispose these proteins for special functional modes in which they take advantage of it. Their function cannot be fulfilled by a rigid structure but it is associated with the ability of the polypeptide chain to rapidly fluctuate among alternative states in a conformational ensemble (Tompa, 2005).

The lack of intrinsic structure in many case is relieved when the protein binds to its target molecule; IUPs are thought to become ordered only when bound to another molecule or owing to changes in the biochemical environment. For eukaryotic transcription factors, folding of locally disordered segments does accompany DNA binding in significant number of cases, such as *Xenopus laevis* transcription factor IIIA (TFIIIA, fig. 1.4; Wuttke *et al.,* 1997). Also the initiation of mRNA translation in eukaryotes is a highly regulated process involving assembly of a multisubunit protein-RNA complex that recruits the ribosome to the initiation codon. Coupled folding and binding have been identified in the recruitment of the small ribosomal subunit to the 5’ end of the mRNA, that is facilitated by eIF4G. eIF4G acts as a molecular bridge between eIF4E and the RNA helicase eIF4A. Interactions with eIF4E are mediated by a 98 residue domain of eIF4G that, according to both NMR and proteolysis experiments, is unfolded in the absence of eIF4E (fig. 1.4; Gross *et al*., 2003). Cell cycle regulation is another area in which unstructured states of proteins is functionally implicated. In particular, it is thought that rapid turnover of some of these proteins, presumably facilitated by their unstructured state in the absence of appropriate ligands, is a means of ensuring the sensitivity of the cell cycle to external conditions. Unfolded states play a functional role also in proteins associated with transport through membranes, folding and coupling

may help to direct the assembly of complexes bringing the membrane into close proximity and facilitating membrane fusion (Dyson and Wright, 1999). It is also been demonstrated that protein disorder play a central role in diseases mediated by protein misfolding and aggregation. Amyloid diseases such as Alzheimer's disease and the prion diseases are characterized by the deposition of protein plaques that arise from misfolding of proteins normally found in the brain and other tissues. Structure determination of fragments of the prion protein revealed that ~ 100 residues at the C terminus are folded into a largely helical domain; while the N-terminal half of the protein is completely unfolded under normal conditions. Partial folding occur in presence of Cu(II) ions. If this protein functions as a copper storage or transport protein, the extreme flexibility of the N terminus is probably of functional significance, as the membrane-anchored protein picks up copper ions from the extracellular fluid, giving a clue as to the overall physiological function of the prion protein, which is at present unknown (Viles *et al*., 1999). Another example of an unfolded amyloidogenic protein is the NACP protein implicated in Alzheimer's disease. The NACP protein is a precursor of the non-Aβ component (NAC) of the amyloid plaque and the 14 kDa precursor protein is found by a series of biophysical experiments to be unfolded under normal solution conditions (Weinreb *et al*., 1996). As for the prion protein, this characteristic may be an important potentiator of the protein-protein interactions that lead to plaque formation and disease.

A critical question that immediately arises is how unfolded proteins can survive in the cell, successfully avoiding the protein degradation machinery at least for long enough to perform their cellular functions. The most likely explanation is that proteolytic degradation is tightly regulated in eukaryotic cell. Clearly unfolded proteins can be functional under some circumstances, although the lifetime of such molecules might well be shorter than for other cellular components that are well-folded at all times. This reduced lifetime may also constitute a component of the regulation of these proteins, a possibility made more plausible by the vital roles many of them play in the cell cycle regulation, in transcriptional and translational processes.

Conceivably, the sensitivity and short response time implicit in many of these processes could be mediated by a system where the regulatory molecule are targeted for rapid turnover. Rapid turnover has a common means of regulation in the cell. In many cases it is mediated by a PEST sequence (rich in Pro, Glu, Ser and Thr) which targets the protein for degradation, or a preference for highly charged peptide sequences that predispose the protein for proteolysis. Both of these are characteristics of IUPs, so that rapid turnover may provide a level of control that allows rapid and accurate responses of the cell to changing environmental conditions (Dyson and Wright, 1999).

By considering details of their various modes of action, the many different functions of IUPs actually segregate into few general categories (fig. 1.5).

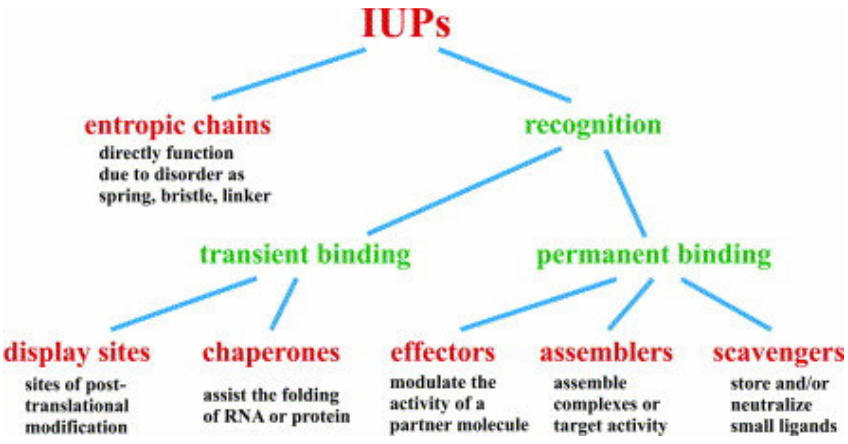


Fig. 1.5: Functional classification scheme of IUPs. The function of IUPs stems either directly from their capacity to fluctuate freely about a large configurational space (entropic chain function) or ability to transiently or permanently bind partner molecule(s).

The first general functional class of IUPs is that of *entropic chains,* with function that directly stem from their ensemble of structural states of similar conformational energies, influencing the localization of attached domains. In the other classes, IUPs function via molecular recognition; they permanently or transiently bind their target/partner that can be DNA, RNA, another macromolecule or a range of small ligands. On binding, IUPs can permanently alter the action of their partner in a variety of ways. Some of them, the *effectors*, modify the activity of a single partner protein or assembled proteins; so far, only inhibitors appear to belong to this second class. The third class is that of *scavengers*, which store and/or neutralize small ligands. The fourth class are the *assemblers,* which assemble, stabilize and regulate large multiprotein complexes such as the ribosome, cytoskeleton, transcription preinitiation complex, chromatin and even the extracellular matrix. A special case of molecular recognition is exhibited by the two classes of IUPs that can transiently bind their partner. The fifth subclass, *display sites*, which mediate regulatory posttranslational modification, such as phosphorylation or limited proteolysis. Such modifications often require intrinsic disorder, which enables transient but specific interaction with the active site of the modifying enzyme. A novel subclass within this category is *chaperones,* that assist the folding of RNA or proteins with their disordered segments, either recognizing, solubilizing or loosening the structure of the misfolded ligand. An important aspect of this classification is that the categories represent various functional modes but are in no way exclusive, as different domains within the same protein, or even the very same region may be involved in distinct functional modes (Tompa, 2005).

The current view on disorder is that IUPs are disordered to allow for more interaction partners and modification sites. A basic mechanism by which individual proteins can increase network complexity is moonlighting, that is the ability of a protein to fulfil more than one, apparent unrelated, function (Tompa *et al*., 2005). Eleven IUPs have been identified suggesting that the structural malleability of IUPs gives rise to unprecedented cases of

moonlighting by eliciting opposing (inhibiting and activating) action on different partners or even the same partner molecule. There are several potential molecular mechanisms that a moonlighting protein could use to switch between functions, for example, changes in cellular localization or ligand binding, expression in different cell types, or variations in oligomerization or complexation state. A disorder inhibitor/activator protein can bind to the same partner in different conformations and/or at different binding sites, resulting in two opposing effects. The opposite functions of the IUPs are often closely related to each other, either synergistically or antagonistically. It has also been suggested that disordered proteins exist to provide a simple solution to have large intermolecular interfaces while keeping smaller protein. For their capacity of binding to distinct partners, it might be the level of the partner molecule itself that controls the switching.

In terms of evolution, traditional moonlighting proteins are thought to have originally only one function, and then to have been recruited for another use later by virtue of their large, unused and evolutionarily unconstrained surfaces. On the contrary, IUPs must have followed a different evolutionary path. These proteins use overlapping, or the same, interaction surface(s) for distinct functions. IUPs versatility might be the result of evolution, that uses the same surface for more than one purpose (Tompa *et al*., 2005). It has already been raised that functional diversity ensuing from conformational diversity might facilitate the evolution of new proteins. This is particularly true for IUPs, which are more likely to have had multiple, less specific, activities than a single function at the outset.

1.2.4 *Plant IUPs*

The list of proteins that are unstructured under physiological conditions had grown rapidly in recent years, and there is a growing awareness that these proteins might represent a large portion of eukaryotic proteomes. Even if these increasing studies are often turned to human and animal kingdoms,

some examples of plant intrinsically unstructured proteins are emerging. Among them, the Arabidopsis HY5 protein is a basic leucine zipper (bZIP) transcription factor that promotes photomorphogenesis binding directly to the promoters of light responsible elements and thus regulating their transcriptional activity. Using proteomic approaches, it has been inferred that the overall structural features of full-length HY5 are dominated largely by the disordered N-terminal domain, despite the existence of a bZIP domain at its C-terminus. Yoon and colleagues proposed that HY5 is a member of the IUP family, and that HY5 functions as an unstructured protein and benefits *in vivo* from being the same (Yoon *et al.,* 2006). Proteins that must be translocated to specific organelles need to remain in a largely unfolded state in the cytoplasm, such as the photosynthetic protein plastocyanin (Koide *et al*., 1993). It appears to remain unfolded as a high-molecular mass precursor protein (apoplastocyanin) before being transported into the chloroplast. It folds extremely slowly *in vitro,* providing a rare opportunity to characterize the structure of intermediates into the folding pathway.

Structural plasticity of IUPs, that can undergo conformational rearrangements due to the binding to their partner, is common in plant IUPs. The Arabidopsis cryptochrome blue-light photoreceptor shows a conformational change in its C-terminal domain upon light exposure. Secondary structure predictions pointed out a domain largely devoid of ordered structure, with long stretches of predicted disorder interrupted by short minimal ordered regions, thought to correspond to regions important for function and/or protein-protein interactions. These short regions seem to act as nucleation sites for effector binding and induced folding into a stable tertiary structure (Partch *et al*., 2005).

Natively unfolded structure is proposed to provide the conformational flexibility necessary to permit assembly of proteins into multisubunit structure of which they are components. The manganese stabilizing protein (MSP) assembly into photosystem II involves a two-step process with a

reorganization of MSP tertiary structure. The conformational flexibility and its anomalous hydrodynamic properties may be necessary to expedite functional binding of MSP to PSII (Lydakis-Simantiris *et al*., 1999). It has also been demonstrated that the IUPs exist in a largely solvent-exposed structural state under physiological conditions. Plant dehydrin binds a significantly larger amount of water than globular proteins but can also bind a large amount of charged solute ions (Tompa *et al*., 2006). In accord, the dehydration stress function of this protein, probably results from its simultaneous action of retaining water in the drying cells and preventing an adverse increase in ionic strength, thus countering deleterious effects such as protein denaturation. In addition to this, also the dessication stress protein Dsp16, a dehydrin-related protein from the resurrection plant *Craterostigma plantagineum* (Lisse *et al*., 1996), is an example of IUP involved in the response to abiotic stress such as drought, low temperature and salinity.

1.2.5 *Intrinsically unstructured CP12 and its role in protein-protein interactions*

Genes coding for IUPs are predicted to represent a significant portion of eukaryotic genomes, but are much less common in prokaryotes (Ward *et al*., 2004) and CP12 is one of a few known in plants and cyanobacteria. Due to their inherent flexibility, IUPs are typically involved in protein–protein interactions and often promote the formation of supramolecular complexes (Dyson and Wright, 2005; Tompa *et al.*, 2005). The functioning of CP12 in chloroplasts does also seem to follow these general principles.

Studies on *Chlamydomonas reinhardtii* CP12 underlined that it is highly homologous to higher plant and cyanobacterium *Synechocystis PCC6803* CP12 (Graciet *et al.,* 2003*a*). *C. reinhardtii* CP12 contains four cysteine residues which form two intramolecular disulfide bridges. Oxidized and

reduced CP12 were analyzed in solution by CD and NMR to explain a possible structural transition upon reduction. The CD spectra show that oxidized CP12 has some secondary structure, mainly composed of α helices. The NMR experiments, however, indicate that CP12 does not adopt a rigid 3D structure in solution; even if CP12 has secondary structures, their relative mobility makes the resolution most unlikely and the NMR datasets obtained preclude the construction of a model. CP12 proposed structure changes dramatically upon reduction: it loses almost all its α-helices and becomes highly disorganized, almost completely unfolded and much more flexible and mobile than oxidized CP12. CP12 protein thus appears as one of the few known IUPs specific to plants (Graciet *et al.,* 2003*a*). Prediction of the secondary structure for the mature oxidized CP12 from *C. reinhardtii,* trough the JPRED server, indicates two long helices. Using comparative modelling approach, imposing the presence of the two disulfide bridges and the two helices confirmed by circular dichroism, a model of the tertiary structure of *C. reinhardtii* oxidized CP12 was derived (fig. 1.6, Gardebien *et al.*, 2006).

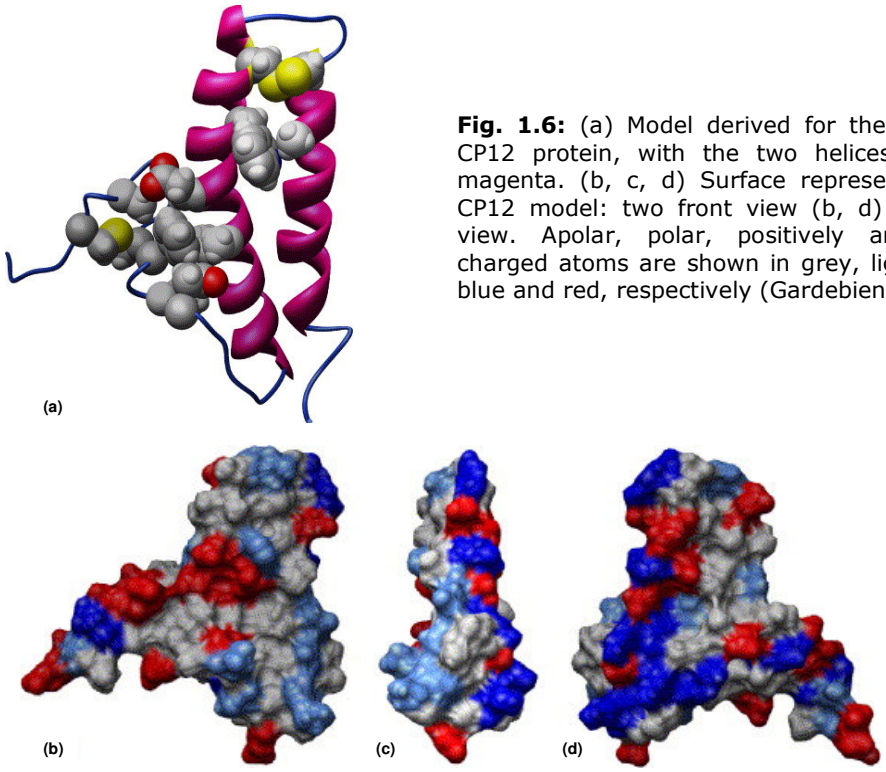


Fig. 1.6: (a) Model derived for the *C. reinhardtii* CP12 protein, with the two helices displayed in magenta. (b, c, d) Surface representation of the CP12 model: two front view (b, d) and one side view. Apolar, polar, positively and negatively charged atoms are shown in grey, light blue, deep blue and red, respectively (Gardebien et al., 2006).

The formation of the disulfide bridges upon oxidation plays a key role in the folding of CP12. These cysteines are highly, but not completely, conserved among prokaryotic and eukaryotic CP12 proteins. The lack of overall organization and the great flexibility of oxidized CP12 observed by NMR and CD are in good agreement with the role of this protein as a linker. The flexibility and the net negative charges of CP12 may increase its reactive area and its "stickiness", compared to rigid partners.

CP12 is a small nuclear-encoded chloroplast protein of 8.5 kDa, and in its oxidized form was shown to be involved in the regulation of the Calvin cycle by modulating the activity of both photosynthetic glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK) in the green unicellular alga *C. reinhardtii* (Graciet *et al*., 2003*a*; 2004*a*). It is a typical disordered protein linker that adopts locally regular secondary structures for protein-protein interactions. It exhibits two faces in the CP12 model with distinct electrostatic properties (fig. 1.6). This particular distribution of charges, with a cluster of positively and negatively charged residues, might be involved in the dimerization of *C. reinhardtii* CP12 in an antiparallel position, as proposed by Wedel (Wedel *et al.,* 1997), or in the binding with the surface of target enzymes (Gardebien *et al.*, 2006). It has been argued that Chlamydomonas CP12 binds to GAPDH and PRK through its C- and N-terminus cysteine-cysteine corresponding loops: the interaction between CP12 and GAPDH occurs at the C-terminal loop of CP12, while the N-terminal S-S bridge is used for the binding of PRK (Wedel and Soll, 1998; Graciet *et al*., 2003*a*). The presence of CP12 is necessary for the formation of the GAPDH/CP12/PRK sopramolecular complex, inside which the two enzymes are in fact inactivated, thus providing evidence for a novel way for the light regulation of photosynthetic enzymes in the Calvin cycle pathway (Graciet *et al.,* 2003*a*; 2004*a*).

1.3 *Light-regulation of photosynthetic glyceraldehyde-3-phosphate dehydrogenase*

Regulation of the Calvin–Benson cycle under varying light/dark conditions is a common property of oxygenic photosynthetic organisms and photosynthetic glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is one of the targets of this complex regulatory system. Chloroplast GAPDH is one of the eleven enzymes of the Calvin–Benson cycle and, together with phosphoglycerate kinase, catalyzes the reduction of 3-phosphoglycerate, the product of Rubisco reaction, into glyceraldehyde-3-phosphate, the first phosphorylated sugar produced in the cycle (Arnon *et al.*, 1954). Chloroplast GAPDH has been the first light-regulated enzyme to be discovered in plants (Ziegler and Ziegler, 1965). Although chloroplast GAPDH isoforms involved in the Calvin–Benson cycle are bispecific dehydrogenases able to accept either NADPH or NADH as electron donors, nonetheless their whole regulatory mechanism does only affect the NADPH-dependent reaction (or the $NADP^+$-dependent reverse reaction). The NAD(H)-reaction, which might possibly play a role in chloroplast dark metabolism, is constitutive and insensitive to any kind of regulation (Pupillo and Giuliani Piccari, 1973). This distinct biochemical feature has been exploited in recent studies aimed at elucidating the molecular basis of chloroplast GAPDH regulation (Sparla *et al.*, 2004; 2005).

1.3.1 *Diversity and evolution of GAPDH subunits and isozymes*

GAPDHs constitute a large and diverse family of dehydrogenases universally represented in living organisms. The family can be divided into two classes: members of the class II have only been found in archaebacteria and show very limited sequence similarities to class I GAPDHs encompassing all

eubacterial and eukaryotic GAPDH genes (Figge *et al.,* 1999). Class I is itself divided in two major groups exemplified by genes *Gap1* and *Gap2* of cyanobacteria (Martin and Schnarrenberger, 1997); a small third group with members only present in prokaryotes will not be dealt with here.

The first group includes the best-known glycolytic GAPDHs (GapC in eukaryotes, Gap1 in eubacteria) which are normally NAD-specific and cytosolic. Besides GapC, land plants (i.e. embryophytes, including bryophytes and higher (vascular) plants) contain a second type of glycolytic GAPDH targeted to plastids (GapCp, Petersen *et al.*, 2003). Both GapC and GapCp form NAD-specific homotetramers *in vivo* and are not subject to complex regulatory mechanisms.

The second group of Class I GAPDH is specifically present in organisms displaying oxygenic photosynthesis and catalyzing the unique reductive step of the Calvin–Benson cycle. In cyanobacteria, Calvin–Benson cycle GAPDH is encoded by *Gap2* genes which correspond to *GapA* genes in eukaryotes (Figge *et al.*, 1999). Both Gap2 and GapA subunits form homotetrameric enzymes, but differ from glycolytic GAPDH (Gap1, GapC, GapCp) by their ability to use both NADPH and NADH as cofactors, with a marked kinetic preference for NADPH (Koksharova *et al.*, 1998; Falini *et al.*, 2003). Similar to cyanobacteria, primitive photosynthetic eukaryotes including *Cyanophora*, red and green algae (except charophytes and related organisms) appear to contain a single type of photosynthetic GAPDH encoded by *GapA* genes (Petersen *et al.*, 2006). Although this type of photosynthetic GAPDH is not regulated *per se* (it is not affected by metabolites or pyridine nucleotides or thioredoxins unlike the higher plants enzyme), it is dark-inactivated *in vivo* through the interaction with CP12 and the Calvin–Benson enzyme PRK (Wedel and Soll, 1998). Thioredoxins and pyridine nucleotides in concert with metabolites cooperate in the modulation of this regulatory process (Wedel and Soll, 1998; Graciet *et al.,* 2004*a*; Tamoi *et al.,* 2005).

A further contribution to enzyme diversity within the group of photosynthetic GAPDHs derived from a gene duplication event which apparently occurred near the origin of Streptophyta (which include charophytes and land plants, Petersen *et al.,* 2006) giving rise to a novel GAPDH subunit named GapB. This protein appears to be a construct of a GapA moiety fused at the C-terminus with the C-terminal half of CP12 (Pohlmeyer *et al.,* 1996). The portion of CP12 acquired by GapB subunits confers original regulatory properties to GapB-containing GAPDH isozymes. These "modern" GAPDH isozymes are made up of GapA and GapB subunits in stoichiometric ratio and are directly regulated by thioredoxins, pyridine nucleotides and metabolites (Pupillo and Giuliani Piccari, 1975; Wolosiuk and Buchanan, 1976; 1978; Trost *et al.,* 1993; Baalmann *et al.,* 1995; Sparla *et al.,* 2002). Moreover AB-GAPDH isoforms have a peculiar propensity to form kinetically inefficient hexadecamers (A_8B_8) under conditions prevailing in chloroplasts in the dark (i.e. the NADPH-dependent activity is decreased to a lower level than the NADH-dependent one), and to dissociate into fully active tetramers (A_2B_2) under photosynthetic and hence activating conditions (Pupillo and Giuliani Piccari, 1973; 1975; Scagliarini *et al.,* 1993; Baalmann *et al*., 1994). This recently evolved type of autonomous GAPDH regulation co-exists in land plants with the ancient, CP12-based regulatory mechanism (Scheibe *et al.,* 2002).

In land plants the family of GAPDH is thus typically represented by four types of genes: *GapC* and *GapCp* (glycolytic, NAD-specific), *GapA* and *GapB* (photosynthetic, NAD(P)H-dependent). Due to further, highly conservative duplications, the actual number of GAPDH genes in a given plant may be higher. *Arabidopsis thaliana* for instance contains seven *Gap* genes, i.e. one *GapB* and three pairs each of *GapA*, *GapC* and *GapCp* genes.

1.3.2 *A_4-GAPDH: the simplest isoform of photosynthetic GAPDH as a paradigm of the fully active enzyme*

The nature of the non-regulatory A_4-GAPDH homotetramer in higher plants is somewhat controversial. Transcripts for GapA are prevailing in dark-grown seedlings, but a strong increase of both GapA and GapB-transcripts is observed during de-etiolation (Cerff and Kloppstech, 1982; Dewdney *et al.,* 1993). A_4-GAPDH may thus be the default GAPDH isoform of etiolated tissues, while AB-GAPDH would prevail in green tissues. Affimetrix ATH1 GeneChip analysis of Arabidopsis transcripts (Schmid *et al.,* 2005) shows that the sum of GapA-1 and GapA-2 transcripts (93% identical at the amino acid level) invariably exceeds the amount of GapB transcripts in leaves, flowers and developing seeds, suggesting the existence of GapA homotetramers in these organs (fig. 1.7). On the other hand, self-reassembly of A-subunits from degraded A_nB_n isozymes has been also suggested. B-subunits tend to be proteolyzed during storage (Scagliarini *et al*., 1998), and according to one extreme view, all NAD(P)-GAPDH forms, including the stable A_4 isoform, could derive from A_nB_n precursors (Brinkmann *et al.,* 1989; Scheibe *et al*., 1996).

At all events, very large amounts of A_4-GAPDH were obtained from spinach chloroplasts to grow crystals for structural analysis (Sabatino *et al*., 1999) and the tridimensional structure of native A_4-GAPDH from spinach is still the only one available among photosynthetic GAPDHs (Fermani *et al.,* 2001). The overall structure (fig. 1.8A) is similar to glycolytic GAPDH determined from many animal and bacterial sources (e.g. Skarzinsky *et al*., 1987; Song *et al.,* 1998). Each subunit of the tetramer is constituted by two distinct domains, an N-terminal coenzyme-binding domain and a C-terminal catalytic domain which includes the binding site for the substrates (BPGA in the photosynthetic reductive reaction, fig. 1.8B). A long and flexible loop of the catalytic domain, known as the S-loop, protrudes toward the coenzyme-binding domain of the adjacent subunit, viewed along the molecular axis R,

thereby contributing to formation of the coenzyme binding site (figs. 1.8A and 1.9A, Fermani *et al*., 2001).

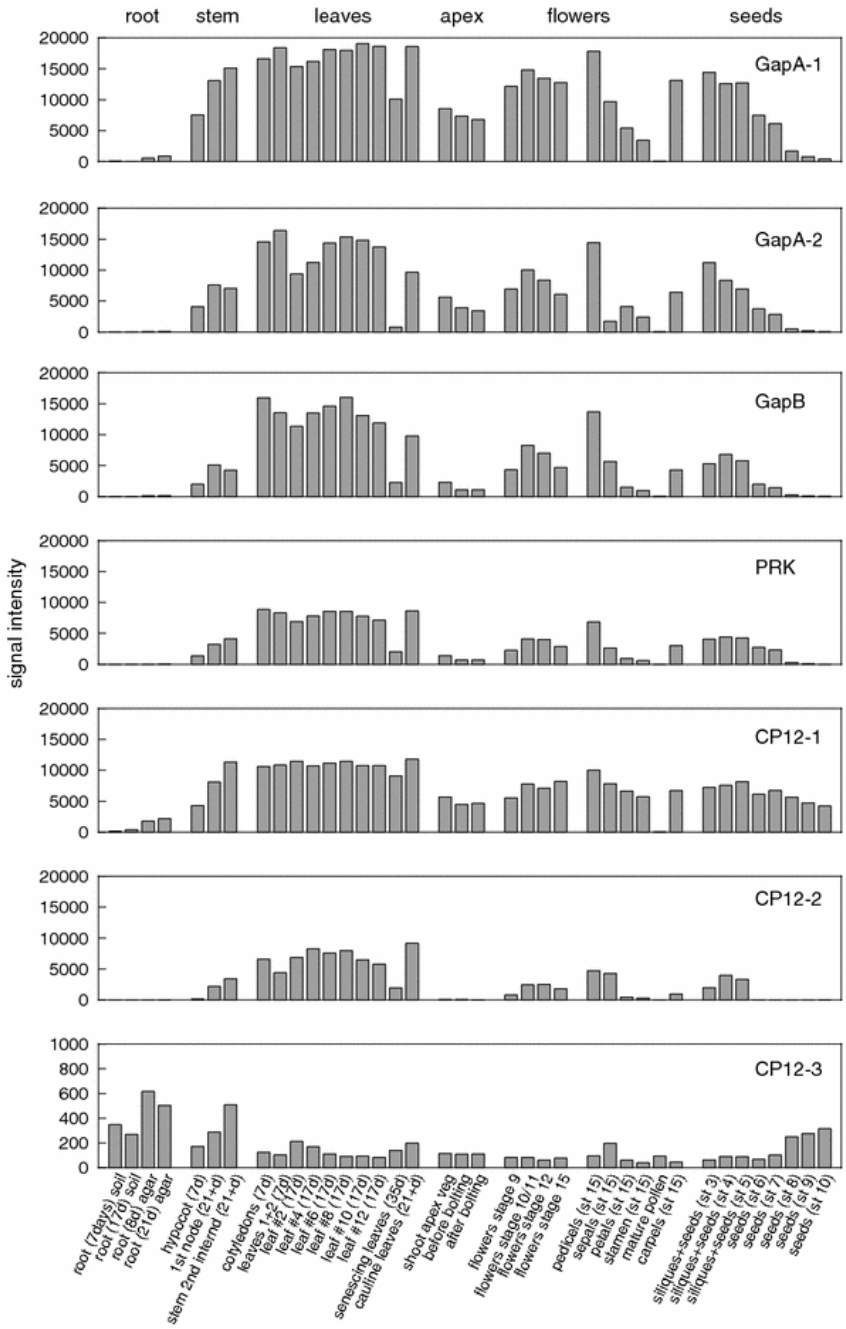
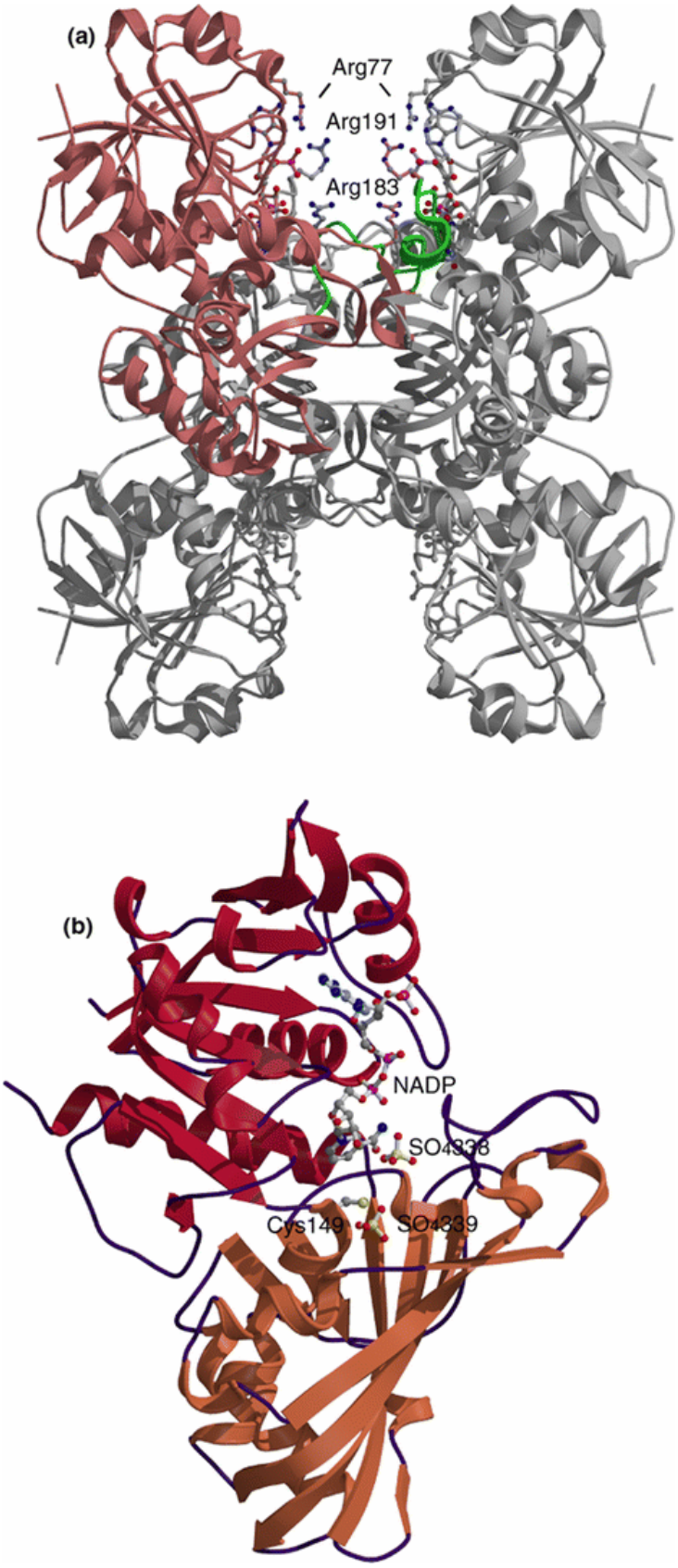


Fig. 1.7 Expression of *GapA-1* (At3g26650), *GapA-2* (At1g12900), *GapB* (At1g42970), *CP12-1* (At2g47400), *CP12-2* (At3g62410), *CP12-3* (At1g76560) and *PRK* (At1g32060) in different organs and growth stages of *Arabidopsis thaliana.* Note that in the CP12-3 panel the scale of the y-axis is 20-fold smaller than in other panels. Selected microarray data from the AtGenExpress developmental series (Schmid *et al*., 2005). Original data from triplicate arrays (Arabidopsis ATH1 chip, Affimetrix) are available at www.weigelworld.org. All RNA samples were extracted from wild type *Arabidopsis thaliana*, Columbia ecotype. A concise description of RNA samples is given in the figure. A complete description of RNA samples and experimental procedures can be found in Weigelworld website (www.weigelworld.org).

The two alternative coenzymes, NADPH or NADH, interact with each subunit at one and the same binding site. The coenzyme adopts an extended conformation with the nicotinamide group pointing toward the active site, where catalytic Cys-149 (fig. 1.8B) makes a transient thioester bond with C-1 of the substrate which is then reduced by NAD(P)H. Although NADPH and NADH both bind to the enzyme in the same position, the kinetic parameters of the NADPH-dependent reaction are different from the NADH-dependent ones. The K_m of A_4-GAPDH for NADPH is 5–10 fold lower than for NADH, and the V_{max} of the NADPH-dependent reaction is about twice as high as with NADH (Scagliarini *et al.,* 1998; Graciet *et al.,* 2003b, Sparla *et al.,* 2004). These kinetic parameters are similar to those of activated A_2B_2-GAPDH (Cerff, 1978; Scagliarini *et al.,* 1998; Sparla *et al*., 2004), supporting the view that A_4-GAPDH corresponds to a GAPDH isoform fixed in a fully active conformation.

Defining the structural basis of coenzyme specificity in GAPDH is necessary for understanding its regulation, as the latter process is strictly NADPH-specific no matter whether it is mediated by the C-terminal extension of GapB or by CP12. At the biochemical level, the hallmark difference between NADPH and NADH is given by the specific 2'-phosphate group of NADPH.

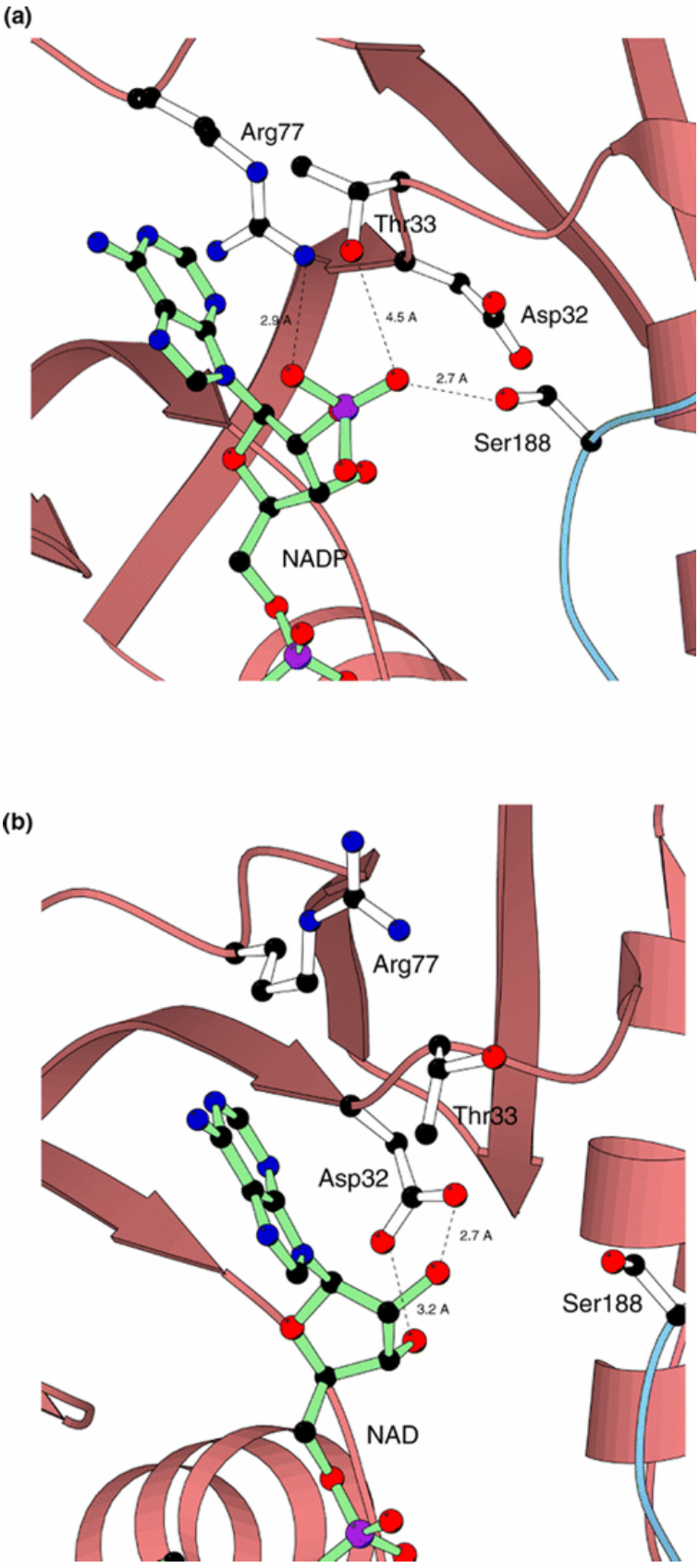
Fig. 1.8 Ribbon model of photosynthetic A_4-GAPDH NADP-complexed from spinach (Sparla *et al*., 2004). (a) Each subunit of the tetramer binds one NADP (represented as balls and sticks). The upper left subunit is coloured light salmon and the S-loop of this subunit is coloured green; the other three subunits are coloured grey. The side chains of Arg77, Arg183 and Arg191 of the upper left subunit (light salmon) and the upper right subunit (grey) are represented as balls and sticks. These basic residues are suggested to be involved in the interaction with acidic CTE (and possibly CP12) leading to enzyme regulation. Symmetrical residues in the lower subunits are not shown. (b) Overall fold of a single subunit of photosynthetic A_4-GAPDH (Sparla *et al*., 2004). The subunit is similarly oriented as in the tetramer, the coenzyme binding domain is coloured crimson and the catalytic domain is coloured salmon. NADP, sulphate ions and the side chain of catalytic Cys149 are represented as balls and sticks. Sulphate ions co-crystallize with the protein and mark the position of phosphates of the substrate BPGA in the catalytic site.


(a)
Arg77
Arg191
Arg183
(b)
NADP
SO4338
Cys149
SO4339

The high resolution structure of recombinant A_4-GAPDH from spinach shows that the 2'-phosphate of NADP is kept in place by a salt bridge with Arg-77 and a hydrogen bond with Ser-188, the latter belonging to the S-loop of the opposite subunit, viewed along the molecular axis R (fig. 1.9A, Sparla *et al*., 2004). When NAD substitutes for NADP, the side chain of Asp-32 stabilizes the 2'- and 3'-hydroxyls of NAD with its terminal carboxylate (fig. 1.9B, Falini *et al*., 2003). The capability of Asp-32 to rotate away from NADP, or close to NAD, depending on which type of coenzyme is bound to GAPDH is apparently a crucial feature of photosynthetic NAD(P)-dependent GAPDHs which evolved from a strictly NAD-specific ancestral form (Falini *et al.,* 2003).

Based on these structural data we postulated that any regulatory mechanism able to specifically down-regulate the NADPH-dependent activity of GAPDH would imply a specific role for residues Arg-77 and Ser-188, the committed residues of coenzyme specificity in photosynthetic GAPDH.

Fig. 1.9 Molecular basis of coenzyme recognition in photosynthetic GAPDH. Detailed view of the interactions between (a) the 2'-phosphate of NADP and (b) 2'- and 3'-hydroxyl groups of NAD with recombinant A_4-GAPDH from spinach (Falini *et al*., 2003; Sparla *et al*., 2004). Only relevant amino acids and interactions shorter than 5 Å are reported. Note that the coenzyme is bound to the subunit coloured light salmon, while Ser188 belongs to the adjacent subunit (light blue). Atom colour code: C black, N blue, O red, P magenta.


(a)
Arg77
Thr33
Asp32
2.9 A
4.5 A
2.7 A
Ser188
NADP
(b)
Arg77
Thr33
Asp32
2.7 A
3.2 A
Ser188
NAD

1.3.3 *Photosynthetic A_4-GAPDH isoform is a target of glutathionylation*

Glutathione represents the major low-molecular-weight thiol in most cells. In addition to its well-established role in cellular defense against oxidative stress, glutathione can also promote a reversible post-translational modification, termed protein glutathionylation (Ghezzi, 2005). This modification consists in the formation of a mixed disulfide between glutathione and cysteine residues of proteins. In mammals, glutathionylation occurs under oxidative stress conditions and may protect cysteines from oxidation to cysteine sulfinic ($-SO_2H$) or sulfonic ($-SO_3H$) acids. In fact, while oxidation into sulfinic or sulfonic groups is irreversible, 2-electron reduction of glutathionylated cysteines can regenerate protein thiols. Glutathionylation has been shown to alter, either positively or negatively, the activity of several proteins. Recent proteomic approaches allowed the identification of many proteins undergoing such post-translational modifications in mammalian and yeast cells (Ghezzi, 2005; Casagrande *et al.,* 2002). In animal cells, many proteins have been shown to undergo glutathionylation under conditions of oxidative stress. By contrast, very little is known about this post-translational modification in plants.

One of the prominent glutathionylated proteins in mammalian cells under stress is the glycolytic enzyme, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which is inactivated by glutathionylation, presumably of its active-site Cys149 (Cotgreave *et al*., 2002). All GAPDHs, including chloroplastic isoforms, share a common reaction mechanism based on a highly reactive cysteine (Cys149), which is made acidic by an interaction with His176. During the catalytic cycle, the highly reactive thiolated group of Cys149 ($Cys\text{-}S^-$) forms a thioacylenzyme intermediate by nucleophilic attack on the substrate. As a side-effect, the acidic nature of Cys149 makes it particularly prone to oxidation and to other redox modifications of its thiol group. In glycolytic mammalian GAPDH, these modifications include S-

glutathionylation, S-nitrosylation and formation of an intrasubunit disulfide with the neighbouring Cys153.

Chloroplasts are a major site of reactive oxygen species (ROS) production, particularly under conditions of oxidative stress, such as exposure to high light, cold or water stress (Foyer and Noctor, 2005), and recent studies have allowed the identification of a number of plant proteins undergoing glutathionylation (Dixon *et al.,* 2005). It has been recently shown that chloroplast f-type TRXs are modified by glutathionylation, resulting in less efficient activation of TRX-sensitive GAPDH and NADP-malate dehydrogenase in the light (Michelet *et al*., 2005).

A_4-GAPDH from *Arabidopsis thaliana* is glutathionylated with either oxidized glutathione or reduced glutathione and H_2O_2. The formation of a mixed disulfide between glutathione and A_4-GAPDH resulted in the inhibition of the enzyme activity. A_4-GAPDH was also inhibited by oxidants such as H_2O_2. However, the effect of glutathionylation was reversed by reductants, whereas oxidation resulted in irreversible enzyme inactivation. On the other hand, the major isoform of photosynthetic GAPDH of higher plants (i.e. the A_nB_n-GAPDH isozyme in either A_2B_2 or A_8B_8 conformation) was sensitive to oxidants but did not seem to undergo glutathionylation significantly. GAPDH catalysis is based on Cys149 forming a covalent intermediate with the substrate 1,3-bisphosphoglycerate (BPGA). In the presence of BPGA, A_4-GAPDH was fully protected from either oxidation or glutathionylation. From site-directed mutagenesis and MALDI-TOF mass spectrometry analysis, catalytic Cys149 is thus suggested to be the target of both glutathionylation and thiol oxidation. Glutathionylation could be an important mechanism of regulation and protection of chloroplast A_4-GAPDH from irreversible oxidation under stress (Zaffagnini *et al.*, 2007). In this hypothesis, glutaredoxins would be involved in a signaling pathway contributing to the redox regulation of the Calvin cycle by controlling deglutathionylation of A_4-GAPDH and other glutathionylated enzymes in the chloroplast.

1.3.4 *CP12 and the ancient regulatory system of photosynthetic GAPDH*

Similar to higher plants, also in cyanobacteria and lower photosynthetic eukaryotes the Calvin–Benson cycle is suppressed in the dark (Buchanan, 1992), and NAD(P)H-dependent GAPDH is light-modulated in these organisms despite the absence of regulatory GapB subunits (Figge *et al.,* 1999; Wedel and Soll, 1998; Lebreton *et al.,* 2003). In fact, regulation of homotetrameric GAPDH in lower photosynthetic organisms requires interaction with other partner proteins, namely CP12 and PRK (Wedel and Soll, 1998; Graciet *et al.*, 2004*a*). Essentially the same type of regulation has been conserved in higher plants where it co-exists with the autonomous regulation of AB-GAPDH (Scheibe *et al.,* 2002).

CP12 is universally distributed in oxygenic photosynthetic organisms and apparently absent in heterotrophs. Lower photosynthetic organisms generally contain a single copy gene for CP12 whereas CP12 proteins in seed plants are coded by small gene families (e.g. three members in Arabidopsis, all predicted to code for plastid proteins). In the case of Arabidopsis, CP12-1 and CP12-2 are 86% identical in amino acid sequence and the expression of *CP12-1* and *CP12-2* genes is generally coordinated to the expression of *GapA-1*, *GapA-2*, *GapB* and *PRK* in different organs and growth stages (fig. 1.7). On the other hand CP12-3 is less than 50% identical to CP12-1 and CP12-2, its expression is extremely low and follows a different pattern, suggesting unrelated and still undefined physiological roles (fig. 1.7).

CP12 are small proteins of about 80 amino acids which invariably include a pair of conserved cysteines in their C-terminal portion (Pohlmeyer *et al.,* 1996). A second pair of N-terminal cysteines is conserved in most CP12 proteins, but may be missing in some photosynthetic organisms, e.g. in rhodophytes, *Cyanophora* (Petersen *et al*., 2006) and in *Synechococcus*

PCC7942 (Tamoi *et al*., 2005). Both cysteine couples are subjected to dithiol/disulfide equilibria under the control of thioredoxins (Wedel and Soll, 1998; Scheibe *et al*., 2002; Graciet *et al.,* 2003*a*). Thioredoxins are ubiquitous, too, in oxygenic photosynthetic organisms (Buchanan, 1992).

For their amino acid composition, plant CP12 proteins seem to diverge from the average composition of globular proteins in the relative abundance of charged amino acids (especially Glu) and paucity of hydrophobic ones (especially Tyr, Phe, Ile and Leu), and they should be classified as intrinsically unstructured proteins too (IUP, Dyson and Wright, 2005). The involvement of CP12 as a scaffold protein in supramolecular complexes containing GAPDH and PRK has been demonstrated in several photosynthetic organisms including cyanobacteria (*Synechocystis* PCC6803, Wedel and Soll, 1998; *Synechococcus* PC7942, Tamoi *et al*., 2005), *Chlamydomonas reinhardtii* (Wedel and Soll, 1998; Graciet *et al.,* 2003*a*) and higher plants (spinach, Wedel *et al*., 1997; Scheibe *et al*., 2002). Reports on GAPDH/PRK interactions in bean leaves (Wara-Aswapati *et al.,* 1980) and in *Scenedesmus obliquus*, another green unicellular alga (Nicholson *et al.,* 1987), published long before the discovery of CP12 (Pohlmeyer *et al.,* 1996) can also be interpreted as CP12-dependent effects. Recurring features of GAPDH/CP12/PRK supramolecular complexes in different species include size of the complex (published values range between 460 and 640 kDa) and the inhibition of both GAPDH and PRK when embedded in the complex. Moreover, in most cases it was shown that complex formation was promoted by oxidizing conditions and by NAD(H), while reducing conditions and NADP(H) tend to dissociate the complex thereby activating the released enzymes. Consequently, the complex in chloroplasts or cyanobacteria is predicted to prevail *in vivo* under darkness while being dissociated in the light (fig. 1.10).

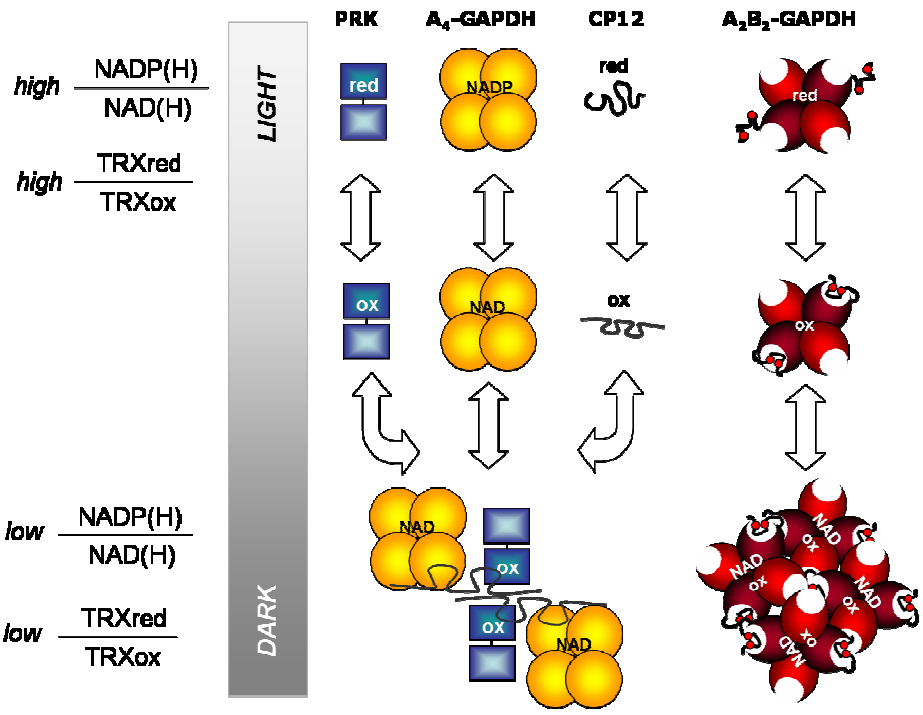


Fig. 1.10 A schematic view of A_4-GAPDH, CP12, PRK and AB-GAPDH interactions and regulation. "Light conditions" promote the dissociation of supramolecular complexes. Reduced PRK, A_4-GAPDH in the NADP-binding conformation and reduced A_2B_2-GAPDH are fully active while reduced CP12 does not interact with partner proteins. "Dark conditions" promote disulfide formation in PRK, CP12 and A_2B_2-GAPDH while A_4-GAPDH shifts to a NAD-binding conformation. These conditions promote the formation of supramolecular complexes A_4-GAPDH/CP12/PRK and A_8B_8-GAPDH. Within these complexes both NADPH–GAPDH and PRK activities are strongly inhibited.

Interestingly, a *Synechococcus* PC7942 mutant expressing no CP12 grew normally under continuous light but was impaired in growth under light/dark cycle (12 h/12 h), strongly suggesting that CP12 was required for a fine modulation of the Calvin–Benson cycle under variable light regime (Tamoi *et al.,* 2005).

The GAPDH/CP12/PRK system has been investigated more in depth in *Chlamydomonas* (Graciet *et al.,* 2004*a*) but the system works similarly also in cyanobacteria or higher plants (Wedel and Soll, 1998; Scheibe *et al*.,

2002). In *Chlamydomonas* (Graciet *et al*., 2003*a*) the interaction between A_4-GAPDH(NAD) and oxidized CP12 is quite strong (K_D 0.4 nM) and GAPDH activity is inhibited by CP12 (mostly a k_{cat} effect, Graciet *et al*., 2003*b*). CP12 has little affinity for PRK, but PRK avidly binds the GAPDH/CP12 binary complex (K_D 60 nM in *Chlamydomonas*, Graciet *et al.,* 2003*a*). The role of CP12 is crucial in this context because the affinity of PRK for GAPDH is poor in the absence of CP12 (Graciet *et al.,* 2003*a*; Scheibe *et al*., 2002). Inhibition of catalytic activities by complex formation is a reversible process and different conditions (reduced thioredoxin, NADP(H), ATP, BPGA) are able to dissociate the complex and release free enzymes (Wedel and Soll, 1998; Scheibe *et al*., 2002; Graciet *et al.,* 2004*a*; Tamoi *et al.,* 2005). Interestingly neither for *Chlamydomonas* PRK nor for GAPDH does complex dissociation necessarily implicate full enzyme activation (Graciet *et al*., 2004*a*), in some cases it is only a step towards full kinetic competence. It is therefore apparent that free A_4-GAPDH, after release from the complex, preserves an imprinting effect of its previous inactive state. In this view, BPGA seems to reset this conformational memory and promote the stabilization of a fully active conformation (Graciet *et al.,* 2004*a*).

In conclusion, CP12 provides a means for a robust light-regulation of GAPDH and PRK in oxygenic photosynthetic organisms. In cyanobacteria and lower photosynthetic eukaryotes, which do not contain autonomously regulated GAPDH based on GapB subunits, this is apparently the only possibility to decrease NADPH-dependent GAPDH activity in the dark. Higher plants have inherited this ancient regulatory mechanism which coexists here with a more recently developed, autonomously regulated AB-GAPDH as the major GAPDH isoform of advanced photosynthetic organisms.

1.3.5 *CP12-dependent vs. CTE-mediated regulation of GAPDH: alternative regulatory systems based on the same molecular mechanism?*

Upon evolution of GAPDH from unicellular green algae to higher plants, characterized by the acquisition of GapB chimeric subunits, redox regulation of photosynthetic GAPDH has become a highly sophisticated process. GapB subunits contain a predominant glyceraldehyde-3-phosphate dehydrogenase moiety which is nearly 80% identical in amino acid sequence to GapA subunits of the same species, and a minor C-terminal extension (CTE) of about 30 amino acids clearly related to the C-terminal end of CP12 (fig. 1.11, Pohlmeyer *et al.,* 1996). As a result, GapB subunits become autonomously redox-sensitive. These subunits are associated in vivo with GapA to form the major GAPDH isoform of higher plants chloroplasts, consisting of GapA and GapB in stoichiometric ratio (AB-GAPDH; Cerff and Chambers, 1979; Brinkmann *et al*., 1989; Ferri *et al.,* 1990).

At variance with A_4-GAPDH, the activity of purified AB-GAPDH is strongly redox regulated such that the NADPH-dependent activity of this isozyme can be titrated by varying the ambient redox potential in the presence of thioredoxin *f* (Sparla *et al*., 2002). Under reducing (i.e. activating) conditions, the catalytic activity (k_{cat}) and the affinity (K_m) for the substrates of AB-GAPDH are similar to those of A_4-GAPDH (Sparla *et al*., 2004). However, under oxidizing conditions AB-GAPDH displays a specific drop in k_{cat} of the NADPH-dependent reaction, while K_m values remain fairly constant. The redox-potential vs. activity plot was characterized by a midpoint potential around −355 mV at pH 7.9 (Hutchinson *et al*., 2000; Sparla *et al*., 2002), slightly more reducing than thioredoxin *f* (Hirasawa *et al.*, 1999).

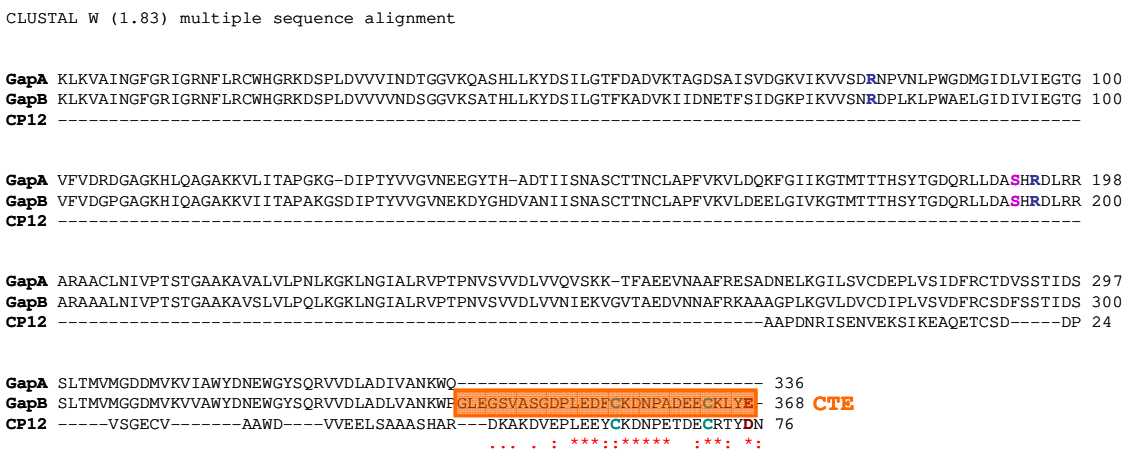


Fig. 1.11 Sequence alignment of GapA, GapB and CP12 from *Spinacia oleracea*. C-terminal extension of GapB (CTE) is framed in an orange box. Conserved cysteines of the CTE and CP12 are bold green. Other residues important for the structure are coloured and bold, blue for Arg-77 and Arg-191, violet for Ser-188, red for Glu-362.

The redox regulation of native AB-GAPDH is closely mimicked by an artificial recombinant isozymes comprised of GapB subunits alone (Baalmann *et al*., 1996; Li and Anderson, 1997; Sparla *et al*., 2002). By using GapB as a model system, it was shown that the drop in $k_{cat(NADPH)}$ depended on the formation of an internal disulfide between the two cysteines of the CTE (Sparla *et al.,* 2002). Interestingly, the CTE could be artificially transferred from a GapB to a GapA subunit with the result of generating a redox insensitive GapB(*min*CTE) (similar to GapA) and a redox-sensitive GapA(*plus*CTE) (similar to GapB, Sparla *et al*., 2005). These redox-sensitive GAPDH isoforms are also prone to aggregate in the presence of NAD(H) giving rise to multimers with low NADPH-dependent activity (Pupillo and Giuliani Piccari, 1975; Baalmann *et al.,* 1996; Li and Anderson, 1997; Sparla *et al.,* 2002; 2005). Redox-insensitive GapB mutants or GapA, on the contrary, are also insensitive to pyridine nucleotide regulation, strongly suggesting that redox regulation (depending on the CTE) and NAD(P)(H)-regulation (depending on the coenzyme binding site) are two strictly interacting mechanisms (Baalmann *et al*., 1996; Sparla *et al.,* 2002).

Given the 80% sequence identity between GapA and GapB (excluding the CTE) it was speculated that A_4-GAPDH may represent a structural model for A_2B_2-GAPDH under reductive (non-inhibited) conditions. Preliminary results on the crystallographic structure of A_2B_2-GAPDH support this assumption (Fermani S and Sparla F, personal communication). As mentioned above, Arg-77 and Ser-188 (fully conserved in both GapA and GapB) are supposed to be responsible for the kinetic preference of reduced A_2B_2-GAPDH for NADPH over NADH (fig. 1.9; Carugo and Argos, 1997). Mutants of these residues are thus reminiscent of oxidized AB-GAPDH and suggest that redox regulation of GAPDH isozymes that are directly regulated by thioredoxins depends on the ability of oxidized CTE to interfere with the recognition of bound NADPH mediated by Arg-77 and Ser-188. In this view oxidized AB-GAPDH would be lowly active with NADPH because NADP is not properly recognized and the enzyme would adopt a kinetically inefficient conformation characterized by a $k_{cat(NADPH)}$ similar to $k_{cat(NADH)}$ (Sparla *et al.,* 2004). Interestingly, these subtle conformational changes may be the cause of the observed kinetic inhibition.

How oxidized CTE would cause this effect is not yet clear. The CTE is rich in charged amino acids and bears a net negative charge (6 Glu, 2 Asp and 2 Arg among the last 20 residues). In neighbourhood of the coenzyme binding site of A_4-GAPDH, two arginines belonging to the S-loop (Arg-183 and Arg-191, fig. 1.8A and fig. 1.11) are exposed to the medium and represent likely candidates to be involved in electrostatic interactions (fig. 1.12) with negative charges of the CTE. In principle, the stability of the interactions between the 2'-phosphate of NADP and the side chains of Arg-77 (salt bridge) and Ser-188 (hydrogen bond, fig. 1.9A) may also be affected by the close proximity of the negatively charged CTE. As a support to this hypothesis, the redox regulation of recombinant GapB was completely abolished in a mutant having Glu-362 (the last residue of the CTE) exchanged into Gln (Sparla *et al.,* 2005), indicating a crucial role for this single negative residue of the CTE in the whole regulatory process.

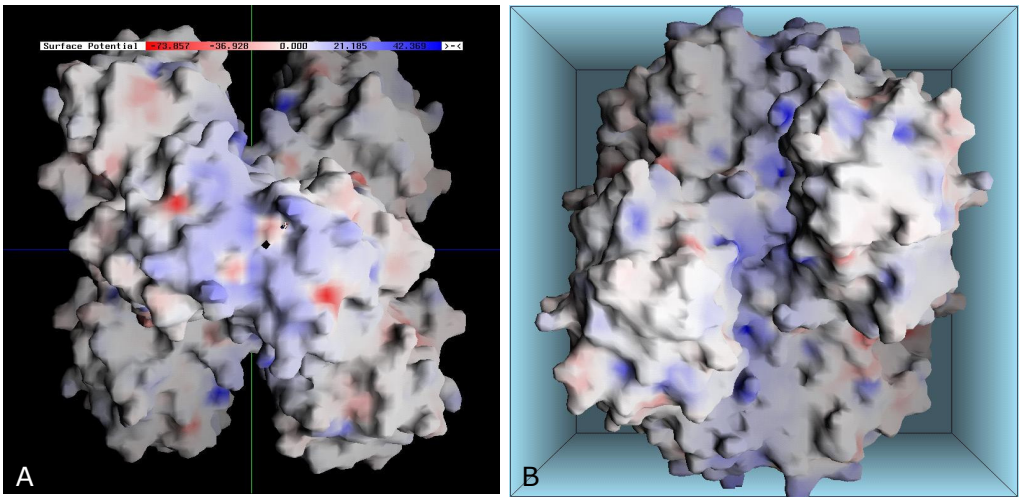


Fig. 1.12 Surface potential of A_4-GAPDH. A, A_4-GAPDH tetramer viewed along the molecular P axis. B, the tetramer viewed along the molecular Q axis. Charge distribution is displayed: negative surface potential in red, blue for positive ones, as in legend.

Although we do not know yet the exact mechanism by which oxidized CTE, through its terminal Glu-362, would specifically inhibit the NADPH-reaction of AB-GAPDH, it is interesting to observe that light-regulated NADP-malate dehydrogenase (NADP-MDH) shows some analogies with AB-GAPDH (Miginiac-Maslow and Lancelin, 2002). As shown by the crystal structures of oxidized NADP-MDH from *Sorghum* (Johansson *et al.*, 1999) and *Flaveria* (Carr *et al*., 1999), the side chain of the penultimate glutamate docks the tip of the oxidized C-terminal extension of NADP-MDH into the active site, thereby preventing the access of the substrate. As a major difference with GAPDH, however, NADP-MDH is specific for NADP and is completely inactivated when oxidized (Miginiac-Maslow and Lancelin, 2002). The peculiarity of GAPDH is that redox-inhibition converts an active NADPH-preferring enzyme into a catalytically inefficient dehydrogenase with no marked coenzyme preference.

The high similarity between the CTE and the C-terminal end of CP12 includes a strict conservation of most of the charged amino acids (including the C-terminal Glu of the CTE which corresponds to a penultimate Glu or

Asp in most CP12 proteins, fig. 1.11, Petersen *et al.*, 2006) suggesting that the function of CTE may resemble that of CP12 at molecular level. In *Chlamydomonas*, binding of CP12 to GAPDH is prevented when Arg-197 of GAPDH (belonging to the S-loop and corresponding to Arg-191 in spinach, fig. 1.8A) is mutated into a glutamate (Graciet *et al.,* 2004*b*), in full agreement with the model of a negative CTE (or CP12) interacting with a positive S-loop in the GAPDH core (fig. 1.12).

Both CP12-dependent and CTE-dependent regulatory systems of GAPDH will eventually result in the formation of large complexes stabilized by NAD(H). It is tempting to speculate that interaction of CP12/CTE with the core of GAPDH induces similar conformational effects such that novel protein/protein interactions are triggered. In the case of GapA/CP12 the final supramolecular complex recruits PRK with a possible stoichiometry of $(A_4\text{-GAPDH})_2\text{-CP12}_2\text{-PRK}_2$ (Wedel and Soll, 1998; Graciet *et al.,* 2004*a*), while in the case of AB-GAPDH the "nocturnal" complex is a hexadecamer (A_8B_8-GAPDH, fig. 1.10; Pupillo and Giuliani Piccari, 1975; Wolosiuk and Buchanan, 1976).

Hopefully forthcoming tridimensional structures and the analysis of further mutants will soon clarify whether the autonomous regulation of AB-GAPDH by CTE and the CP12-dependent regulation of A_4-GAPDH are based on a similar molecular mechanism, as evolutionary considerations and experimental data would suggest.

Chapter 2

*Co-ordinated gene expression of photosynthetic glyceraldehyde-3-phosphate dehydrogenase, phosphoribulokinase, and CP12 in Arabidopsis thaliana**

*published as:
Lucia Marri, Francesca Sparla, Paolo Pupillo, and Paolo Trost (2005) *J. Exp. Bot.* **56:** 73-80

2.1 *Abstract*

Photosynthetic glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK) interact in the chloroplast stroma through the action of the small peptide CP12. This supramolecular complex concurs with the light-dependent modulation *in vivo* of GAPDH and PRK activities. The expression patterns of several genes potentially involved in the formation of the complex have been studied. The genome of *Arabidopsis thaliana* includes seven genes for phosphorylating GAPDH isozymes, one PRK gene, and three genes for CP12. The expression of four GAPDH genes was analysed, i.e. *GapA-1* and *GapB* for photosynthetic GAPDH of chloroplasts (NAD(P)-dependent), *GapC-1* for cytosolic GAPDH, and *GapCp-1* for plastid GAPDH (both NAD-dependent). A similar analysis was performed with *PRK* and two *CP12* genes (*CP12-1*, *CP12-2*). The expression of *GapA-1, GapB, PRK*, and *CP12-2* was found to be co-ordinately regulated with the same organ specificity, all four genes being mostly expressed in leaves and flower stalks, less expressed in flowers, and little or not expressed in roots and siliques. The expression of all these genes in leaves was terminated during

prolonged darkness or following sucrose treatments, and their transcripts decayed with similar kinetics. At variance with *CP12-2*, gene *CP12-1* appeared to be expressed more in flowers, it was totally insensitive to darkness, and less affected by sucrose. The expression of glycolytic *GapC* was strong and ubiquitous, insensitive to dark treatments, and unaffected by sucrose. GapCp transcripts were also found to be ubiquitous at lower levels, slowly decreasing in the dark and stable in sucrose-treated leaves. The co-ordinated expression of genes *GapA-1*, *GapB*, *PRK*, and *CP12-2* is consistent with their specific involvement in the formation of the photosynthetic regulatory complex of chloroplasts.

Key words: Calvin cycle, light regulation, sugar sensing, supramolecular complexes, transcriptional control

2.2 *Introduction*

Calvin cycle enzymes exist at high concentration in the stroma and several of these enzymes have been suggested to be components of organized supramolecular complexes (Süss *et al.*, 1993; Gontero *et al.*, 2002). The interaction between glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK) is one of the best documented examples (Müller, 1972; Wara-Aswapati *et al.*, 1980; Clasper *et al.*, 1991; Lebreton and Gontero, 1999; Mouche *et al.*, 2002). A specific explanation for this long-known multienzyme interaction came when the chloroplast peptide CP12 was discovered (Pohlmeyer *et al.*, 1996) and its role as a linker between GAPDH and PRK demonstrated (Wedel and Soll, 1998; Scheibe *et al.*, 2002; Graciet *et al.*, 2003a, 2004a). CP12 is a small peptide present in chloroplasts of higher plants and green algae and in cyanobacteria. It contains two pairs of fully conserved cysteines which can form disulphide bridges. Following oxidation of both cysteine pairs, the *in vivo* supramolecular complex appears to include two PRK dimers linked to two

GAPDH tetramers by means of a CP12 dimer (Wedel and Soll, 1998; Graciet *et al.*, 2003b, 2004a).

Photosynthetic GAPDH consists of two subunits, GapA and GapB. The latter is 30 amino acids longer than GapA, but apart from this C-terminal extension (CTE) the two subunits are 80% identical (Brinkmann *et al.*, 1989). The CTE is implicated in kinetic enzyme regulation (inhibition in the dark) and aggregation of GAPDH into high-molecular weight polymers such as A_8B_8 (Zapponi *et al.*, 1993; Baalmann *et al.*, 1996; Li and Anderson, 1997; Sparla *et al.*, 2002). Since the C-terminal portion of CP12 is co-linear in sequence with the last 20 amino acids of the CTE, it has been proposed that GapB was evolutionary derived by fusion between GapA and part of the CP12 (Pohlmeyer *et al.*, 1996) and both domains may interact with GAPDH with similar effects. Less is known of CP12 binding to PRK. There is evidence from *Chlamydomonas* that complex formation with PRK requires a previous interaction between CP12 and GAPDH (Graciet *et al.*, 2003a).

Both GAPDH and PRK are kinetically activated by light *in vivo.* The activity of both enzymes is regulated by thioredoxins (Wolosiuk and Buchanan, 1978; Brandes *et al.*, 1996; Sparla *et al.*, 2002) and GAPDH activity is also dependent on metabolites including NAD(P)(H) and BPGA (Pupillo and Giuliani Piccari, 1975; Trost *et al.*, 1993; Baalmann *et al.*, 1995). Moreover, both GAPDH and PRK are inhibited when embedded in the supramolecular complex involving CP12 (Scheibe *et al.*, 2002). During dark-to-light transitions, GAPDH and PRK activation *in vivo* seems to depend on the accumulation of reduced thioredoxins and metabolites in the chloroplast stroma (Buchanan, 1980; Scagliarini *et al.*, 1993; Baalmann *et al.*, 1994). Besides activating freely soluble GAPDH and PRK, these effectors promote supramolecular complex dissociation (Wara-Aswapati *et al.*, 1980; Clasper *et al.*, 1991; Wedel and Soll, 1998; Scheibe *et al.*, 2002).

GAPDH and PRK gene expression has been investigated (Shih and Goodman, 1988; Raines *et al.*, 1989; Yang *et al.*, 1993; Lemaire *et al.*,

1999). The expression of *GapA* and *GapB* is co-ordinately regulated by light at the transcriptional level and several *cis* acting elements and cognate binding factors have been identified (Chan *et al.*, 2002). Both cryptochrome and phytochrome are involved in the light signal (Conley and Shih, 1995). *GapA* and *GapB* were reported to be induced by anaerobiosis but unaffected by sucrose in arabidopsis leaves or tobacco callus (Shih and Goodman, 1988; Yang *et al.*, 1993). *PRK* expression is also regulated by light under the control of the circadian clock and down-regulated during senescence (Raines *et al.*, 1989; Lemaire *et al.*, 1999). The existence of a supramolecular complex including GAPDH, PRK, and CP12 in chloroplasts also suggests that the expression of the relevant genes might be co-ordinately regulated. The expression pattern of five genes of *A. thaliana* potentially involved in the formation of the GAPDH/CP12/PRK complex (*GapA*, *GapB*, *PRK*, and two *CP12* genes; no data are available for the latter), and two further genes coding for GAPDH isozymes unlikely to be involved in the complex (cytosolic *GapC* and plastid *GapCp*) are reported here. Results indicate that the expression of *GapA-1*, *GapB*, *PRK*, and one *CP12-2* genes are co-ordinately regulated under manifold conditions, whereas different regulatory mechanisms may govern the expression of the other genes tested.

2.3 *Materials and methods*

2.3.1 *Plant material and treatments*

Arabidopsis thaliana plants (ecotype Columbia) were grown on a sterile mix of humus/perlite (3:1, v:v) at 22 °C under 14/10 h light/dark cycle in the growth chamber.

For dark treatments, 8-week-old plants were placed in the dark for up to 5 d at 22 °C. For sugar treatments, leaves were harvested from 6-week-old plants after 16 h adaptation to continuous light or continuous dark.

Detached leaves were put in Petri dishes containing either bidistilled water or solutions of sucrose, glucose or sorbitol (200–300 mM) and incubated under gentle shaking for up to 72 h at 22 °C, either in continuous light or continuous dark. To prevent contaminations, solutions were renewed every 24 h.

2.3.2 *Gene cloning*

A. thaliana cDNAs corresponding to plastidial and cytosolic GAPDH isoforms have been cloned by RT-PCR. Total RNA was extracted from *A. thaliana* plants (Nawrath and Mètraux, 1999) and DNA complementary strands were synthesized using M-MLV reverse transcriptase (Sigma) and Oligo dT(15) primers (Promega). Specific DNA primers (see below), derived from genomic sequences (At3g26650 for *GapA-1*, At1g42970 for *GapB*, At3g04120 for *GapC-1*, and At1g79530 for *GapCp-1*; table 2.1), were used to clone GAPDH fragments into a p-Drive cloning vector (QIAGEN PCR cloning kit). Resulting recombinant plasmids were sequenced to confirm the nature of cloned fragments. Plasmids containing the cDNA sequence for CP12-1 (U17174), CP12-2 (U10360), and PRK (C63669) were kindly provided by Arabidopsis Biological Resource Centre, Ohio State University.

2.3.3 *RNA extraction and northern blot analysis*

Total RNA was extracted from different plant organs according to Nawrath and Métraux (1999). Leaves, stalks, flowers, and siliques were harvested from 6-week-old plants and immediately frozen in liquid nitrogen. Total RNA from roots was obtained from 4-week-old plants grown on sterile MS agar pots at 22 °C with a 14/10 h light/dark cycle. RNA pellets were resuspended in H_2O-diethylpyrocarbonate. Concentration and purity were spectrophotometrically determined. For northern blot analysis 10 µg of total RNA were heat denaturated, separated by electrophoresis on a 6%

formaldehyde-1% agarose gel, and transferred by capillarity to a nylon membrane (Hybond-N^+, Amersham Pharmacia Biotech). Electrophoresis buffer and the blotting buffer consisted of 20 mM MOPS, 5 mM Na-acetate, pH 6.0 and 1.5 M NaCl, 150 mM Na-citrate, pH 7.0, respectively.

Table 2.1 Nomenclature and some relevant features of *Arabidopsis thaliana* genes and proteins belonging to the families of phosphorylating GAPDH, CP12, and PRK.

Gene acronym	Locus (TAIR)[a]	No. of EST's (TAIR)[b]	% Identity among closely related genes[c]	Transit peptide for plastid location	Coenzyme specificity (GAPDH isozymes)
GapA-1[d]	At3g26650	211	–	Yes	NAD(P)
GapA-2	At1g12900	136	92.6 (336)	Yes	NAD(P)
GapB	At1g42970	158	81.1 (338)	Yes	NAD(P)
GapC-1	At3g04120	202	–	No	NAD
GapC-2	At1g13440	486	97.9 (338)	No	NAD
GapCp-1	At1g79530	31	–	Yes	NAD
GapCp-2	At1g16300	5	93.8 (337)	Yes	NAD
PRK	At1g32060	141	–	Yes	–
CP12-1	At2g47400	32	–	Yes	–
CP12-2	At3g62410	27	85.9 (78)	Yes	–
CP12-3	At1g76560	3	49.3 (75)	Yes	–

[a] TAIR, The Arabidopsis Information Resource (www.arabidopsis.org).
[b] Deposited before 15 July 2004.
[c] Transit peptides, when present, were not included in alignments. The value in parentheses indicates the number of amino acids which have been aligned.
[d] Genes written in bold have been cloned and studied in this work.

Blotted membranes were hybridized with about 50 ng of gene-specific probes obtained by PCR with different plasmids as template and specific primer sets were designed as follows: *GapA-1*: up 5'-AGGTGGCCATTAATGG-3' and down 5'-TATCATACCAAGCAATCACC-3'; *GapB*: up 5'-TTAGGTGTTGGCATGGT-3' and down 5'-CTCGTTATCATACCAAGC-3'; *GapC-1*: up 5'-ATGGCTGACAAGAAGATTAG-3' and down 5'-ACATGTGGACGATCAAGTC-3'; *CP12-1*: up 5'-ATGACAACCATAGCTGCAGC-3' and down 5'-TTAATTATCATAAGTACGACACTC-3'; *CP12-2*: up 5'-ATGGCAACTATAGCTACTGGTC-3' and down 5'-TCAGTTGTCGTAAGTACGGCAC-3'; *GapCp-1*: up 5'-ATCGAGGTTGTAGCAGTC-3' and down 5'-GTCATACCAGGAGACAAG-3'; *PRK*: up 5'-ATGGCTGTCTCAACTATCTAC-3' and down 5'-TTAGGCTTTAGCTTCTGCACG-3'. PCR programs included a hot start at 95 °C for 5 min followed by 30 cycles of 1 min at 95 °C plus 1 min at 48 °C (*GapA-1*), or 54 °C (*GapB, GapC-1, GapCp-1*, and *PRK*) or 58 °C (*CP12-1* and *CP12-2*), and finally 72 °C for 1.5 min.

Probes were radiolabelled with the kit Ready-To-Go DNA labelling beads (Amersham Pharmacia Biotech) according to the manufacturer's instructions. Hybridizations were performed overnight at 55 °C in 1% w/v BSA, 1 mM EDTA, 0.5 M sodium phosphate buffer pH 7.2, and 7% w/v sodium dodecylsulphate.

2.3.4 *Enzyme activity assays*

Enzyme activities were assayed on fresh leaf homogenates, obtained as for western blotting experiments. Both GAPDH and PRK activities were assayed by following the oxidation of NAD(P)H at 340 nm ($\varepsilon_{NAD(P)H}$=6.23 mM^{-1}) in an Uvikon 941 plus spectrophotometer (Kontron). NAD(P)H-GAPDH activity assay was performed at 25 °C in a reaction mixture containing 50 mM TRIS–HCl pH 7.5, 1 mM EDTA, 3 mM 3-phosphoglyceric acid, 5 mM $MgCl_2$,

0.5 mM KCN, 2 mM ATP, 5 U ml^{-1} 3-phosphoglycerate kinase, and 5 mM dithiothreitol. Reactions were started by the addition of 0.2 mM NAD(P)H. Activated GAPDH activity was measured after incubation of the sample for 5 min at room temperature in the reaction mixture (without NAD(P)H).

Fully activated PRK activity was assayed at 25 °C after incubation for 30 min at room temperature in the assay mixture including 50 mM TRIS-HCl pH 7.5, 1 mM EDTA, 40 mM KCl, 10 mM $MgCl_2$, 0.5 mM KCN, and 5 mM DTT. Baseline was acquired after addition of 2 mM ATP, 2.5 mM phosphoenolpyruvate, 5 U ml^{-1} pyruvate kinase, 6 U ml^{-1} lactate dehydrogenase, 0.2 mM NADH, and the reaction was initiated by the addition of 0.5 mM ribulose-5-phosphate.

Protein content was determined according to Bradford (1976). Activity data were analysed with the ANOVA procedure (CoStat, CoHort Software, Monterey, CA) for the existence of significantly different means (P <0.05).

2.4 *Results*

2.4.1 *In silico search and molecular cloning of members of GAPDH and CP12 gene families, and of PRK from Arabidopsis thaliana*

The genome of *Arabidopsis thaliana* contains seven nuclear genes of the family of phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPDH; table 2.1). Photosynthetic GAPDH isoforms catalysing the reductive step of the Calvin cycle and using either NADPH or NADH as coenzyme are encoded by three genes, all including an N-terminal transit peptide for plastid location. Two of these (*GapA-1* and *GapA-2*) code for GapA subunits and are 93% identical in terms of amino acid sequence, the third one codes for a unique GapB subunit. Cytosolic NAD-specific GAPDH,

involved in glycolysis and gluconeogenesis, is encoded by two closely related genes in *A. thaliana* (*GapC-1, GapC-2*; table 2.1). Finally, the primary sequence of a third type of GAPDH closely resembling GapCp of *Pinus sylvestris* (Meyer-Gauen *et al.*, 1998) features an NAD-specific isozyme located in *A. thaliana* plastids. It is encoded by two similar genes (*GapCp-1 and GapCp-2*; table 2.1) both including putative transit peptides. Based on the relative abundance of ESTs in the TAIR database (http://www.arabidopsis.org; table 2.1) it was observed that *GapCp-2* is much less expressed than *GapCp-1*. All NAD-specific GAPDH isozymes are less than 50% identical to photosynthetic counterparts. In this work, an RT-PCR approach has been followed to clone the entire coding sequences of *GapA-1*, *GapB*, *GapC-1*, and *GapCp-1* of *A. thaliana*. Sequencing of the resulting cDNA clones confirmed the coding sequences predicted by *in silico* analysis of respective genes.

The PRK of *A. thaliana* is encoded by a single gene, while three genes (*CP12-1*, *CP12-2*, *CP12-3*; table 2.1) have been identified for CP12. The three CP12 genes include putative N-terminal transit peptides and are predicted by TargetP software (Emanuelsson *et al.*, 2000) to encode plastid proteins. *CP12-1* and *CP12-2* encode peptides which are 86% identical between them and 78–80% identical to CP12 of spinach (Pohlmeyer *et al.,* 1996). Within the *CP12* gene family, *CP12-3* has the lowest degree of conservation with *CP12* of spinach (47% identical in amino acids) and, as judged from the relative abundance of ESTs in TAIR, it seems to be very little expressed in *A. thaliana* (table 2.1). The cDNAs for *PRK*, *CP12-1*, and *CP12-2* were cloned and sequenced and their *in silico* predicted sequences were confirmed.

2.4.2 *Organ-specific expression*

The similarity of expression levels of *GapA-1*, *GapB, CP12-1*, *CP12-2*, and *PRK* genes in different plant organs has been investigated by northern blot

analysis (fig. 2.1). High expression levels of all these genes have been detected in photosynthetic organs (leaves and stalks) while transcripts were hardly detectable in roots and siliques. In flowers, the expression of *GapA-1*, *GapB*, *PRK*, and *CP12-2* was very low, but the expression of *CP12-1* was higher and almost approached the expression in leaves. *GapC-1* was abundantly expressed in all organs, with higher transcript levels in leaves, stalks, and siliques. The expression of *GapCp-1* was generally low but ubiquitous.

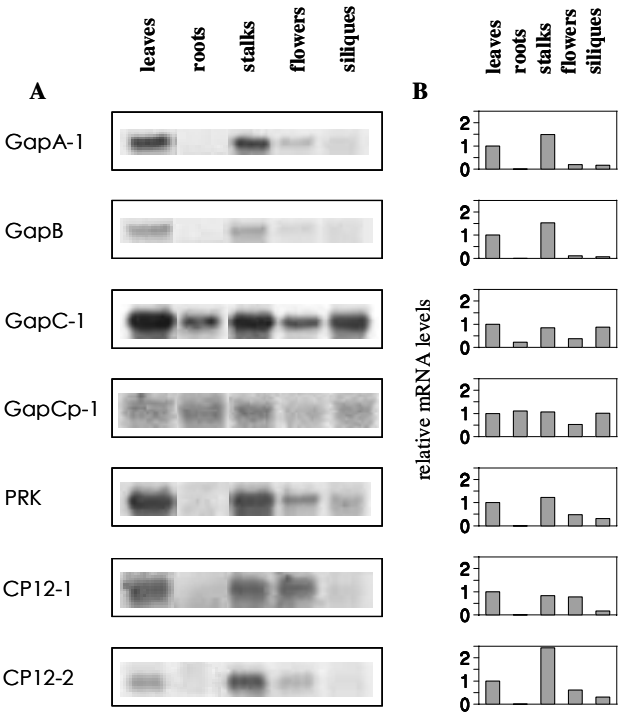


Fig. 2.1 Differential gene expression of *GAPDH, CP12*, and *PRK* in different *Arabidopsis* organs. (A) Northern blot analyses of RNA extracted from leaves, stalks, flowers, and siliques of 6-week-old plants grown on soil, and from roots of 4-week-old plants grown on agar plates. Ten µg of total RNA were loaded per lane and probed with specific radiolabelled probes as indicated. (B) Densitometric analysis of repeated experiments as in (A). Densitometric values for each type of transcript in each experiment have been normalized to the signal of an extra lane loaded with 10 µg of total RNA from leaves as internal control. Data shown are means ±SD (n=3).

2.4.3 *Light-dependent expression*

The expression of photosynthetic genes is regulated by light (Terzaghi and Cashmore, 1995). Accordingly, total RNA was extracted from leaves of 8-week-old plants either grown under a normal light–dark regime or after transfer to total darkness for 5 d (fig. 2.2).

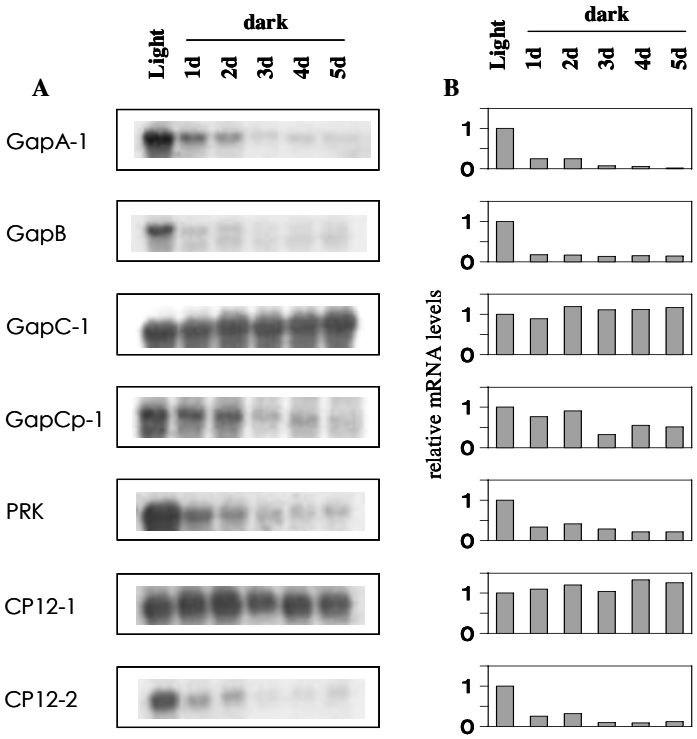


Fig. 2.2 Effect of prolonged dark treatment on expression of *GAPDH*, *CP12*, and *PRK* genes in *Arabidopsis*. (A) Northern blot analysis of total RNA (10 µg lane^{-1}) extracted from leaves. *Arabidopsis* plants were grown for 8 weeks under 14/10 h light/dark conditions, then in complete darkness for 5 d. (B) Densitometric analysis of northern blotting results. Densitometric values are normalized as in fig. 2.1. Data shown are means ±SD (n=3).

Northern blot analysis showed the expression of Calvin cycle genes *GapA-1*, *GapB*, and *PRK* to be declining after 24 h in the dark, and to drop to very low levels during the following days.

Interestingly, *CP12-2* was also down-regulated in darkness with similar kinetics, whereas the amounts of *CP12-1* transcripts failed to change until the end of dark treatment (5 d). *GapC-1* expression was also independent of the light/dark conditions. The expression of *GapCp-1* was inhibited in the dark, although a decrease of the transcripts was slow and GapCp-1 messengers were easily detectable after as much as 5 d of darkness.

2.4.4 *Effect of sugars*

To check whether the genes under study were sugar responsive, mature leaves from 6-week-old plants were harvested and incubated either with sterile water or 300 mM sucrose solution. Experiments using a 300 mM sorbitol solution as a control failed to influence the expression of these genes, thereby excluding the possibility of mere osmotica-dependent effects (not shown). To avoid interference between sugar-dependent and light-dependent control of gene expression, all experiments were performed either under continuous light or darkness (16 h acclimation, 48 h treatment). During the experiments, protein content in the control or treated leaves did not change significantly on a fresh weight basis, whereas chlorophyll decreased dramatically in sucrose-treated leaves but not in control leaves. Purple pigmentation of cells surrounding major veins, due to enhanced synthesis of anthocyanins, was apparent in sucrose-treated leaves following prolonged incubation.

Expression of three Calvin cycle genes (*GapA*, *GapB*, *PRK*) and two *CP12* genes in the light was strongly inhibited by sucrose (fig. 2.3). The transcripts of *GapA-1*, *GapB*, *PRK*, and *CP12-2* rapidly declined with similar time-courses, down to undetectable levels between 6 h and 24 h of treatment. The response of *CP12-1* was slower, but inhibition was complete 48 h after the start of the treatment.

By contrast, neither *GapC-1* nor *GapCp-1* expression were significantly influenced by sucrose. When similar experiments were performed in the dark, lowering of *GapA* and *GapB* transcript levels were again observed with slower kinetics than in the light (not shown), whereas *GapC* and *GapCp* remained constant. The expression of GAPDH genes was also investigated using glucose as the permeant sugar and the results confirmed the effect, but 300 mM glucose proved to be somewhat less effective than 300 mM sucrose (not shown).

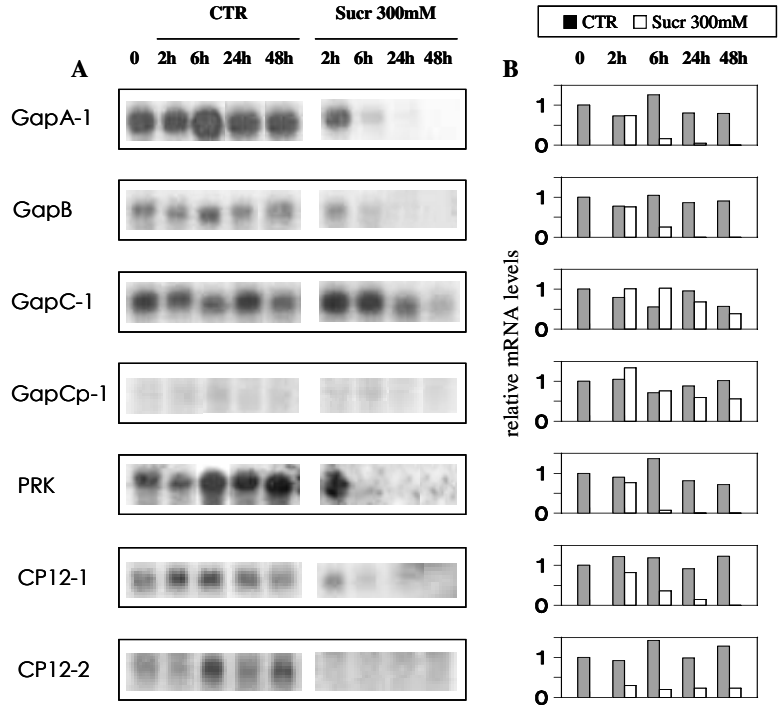


Fig. 2.3 Effect of sucrose incubation on GAPDH, CP12, and PRK gene expression in detached leaves. Six-week-old Arabidopsis plants grown in a 14/10 h light/dark regime, were harvested after adaptation to 16 h continuous light and placed in Petri dishes containing either bidistilled water (control) or 300 mM sucrose solution. Incubation al 22 °C lasted for up to 48 h under gentle shaking in continuous light. Solutions were renewed after the first 24 h. (A) Northern blot analyses of RNA extracted from leaves after different incubation times. (B) Mean densitometric values derived from experiments as in (A). Normalization of the signal was performed as in Fig. 1. Data shown are means ±SD (n=3).

To test whether the dramatic decrease of GapA and GapB transcript abundance caused by sucrose had consequences on enzyme activities, protein extracts were obtained for enzyme assays. Photosynthetic GAPDH activity (NADPH-dependent) stayed costant in control leaves, but showed a slow decline in sucrose-incubated leaves. After 72 h treatments the activity of sucrose-incubated leaves (0.36±0.16 μmol min–1 mg–1) was half the activity of control leaves (0.71±0.11 μmol min–1 mg–1). The effect was statistically significant ($P < 0.05$). A very similar inhibition was observed when photosynthetic GAPDH was assayed under activating conditions, indicating that the activation state of the enzyme was not affected. In the presence of NADH rather than NADPH, the overall GAPDH activity referred to all GAPDH isozymes (A4, AnBn, GapC, and GapCp) was not significantly affected by sucrose treatments. The activity of PRK steadily decreased during the experiments in both control and sucrose-treated leaves and no significant variations could be ascribed to the treatment (not shown).

2.5 Discussion

The expression of genes involved in photosynthesis is influenced by several factors including organ and developmental specificity, nutrient status of the plant, environmental conditions, and stress (Terzaghi and Cashmore, 1995; Koch, 1996). In some instances, for example, during greening of etiolated leaves, the regulation of expression was found to be strictly co-ordinated (Raines *et al.*, 1991). Co-ordinated transcription of light-regulated genes may depend on common light-regulatory elements, although these elements are not always conserved among promoters (Arguello-Astorga and Herrera-Estrella, 1998).

The family of GAPDH genes in *Arabidopsis thaliana* includes seven members. Here the focus has been on four genes representing the four types of GAPDH isozymes: *GapA-1* and *GapB* for the dominant photosynthetic GAPDH isozyme (NAD(P)-dependent), *GapC-1* for cytosolic GAPDH involved

in glycolysis and gluconeogenesis, and *GapCp-1* for NAD-specific GAPDH of plastids (Meyer-Gauen *et al.*, 1998). PRK is encoded by a single gene, while three genes constitute the CP12 family with *CP12-1* and *CP12-2* being more strongly expressed than *CP12*-3 (table 2.1). Calvin cycle *GapA*, *GapB*, and *PRK* transcripts were especially abundant in *Arabidopsis* leaves and flower stalks, less prominent in flowers, and virtually undetectable in roots and siliques. The basic expression pattern of these photosynthetic genes was also found for *CP12* genes, supporting a role of CP12 proteins in photosynthesis. Minor differences in *CP12* expression patterns with respect to Calvin cycle genes include the relatively higher expression of *CP12-1* in flowers and *CP12-2* in stalks.

While the absence of PRK transcripts in roots and its low levels in flowers and siliques are consistent with an exclusive role of this enzyme in the Calvin cycle, the analogous patterns of *GapA-1* and *GapB* gene expression are unexpected since GAPDH activity is required for the glycolytic pathway of non-photosynthetic plastids leading to lipids, isoprenoids, and other end-products. The NADPH-dependent GAPDH activity of non-photosynthetic plastids is generally low, but the NADH-dependent activity can be substantial (Neuhaus *et al.*, 1993) and is apparently catalysed by *GapCp* product (Backhausen *et al.*, 1998; Petersen *et al.*, 2003). Accordingly, *GapCp-1* is expressed in *Arabidopsis* roots and siliques where *GapA-1* and *GapB* transcripts are undetectable. The amino acid sequence of GapCp-1 is 82.5% identical to GapCp of *Pinus sylvestris*, which has been shown to be chloroplast located and NADH-dependent (Meyer-Gauen *et al.*, 1998). *Arabidopsis* GapCp has a putative transit peptide for plastid localization and the typical sequence signature of NADH-specific GAPDH (Koksharova *et al.*, 1998; Falini *et al.*, 2003).

In leaves of adult plants transferred from a normal light/dark cycle to continuous darkness, the transcripts of *GapA-1*, *GapB*, *PRK*, and *CP12-2* slowly decreased with similar kinetics. On the other hand, *CP12-1* transcript abundance remained unaffected, indicating a different type of transcriptional

control reminiscent of *GapC*. Pohlmeyer *et al.* (1996) found that spinach *CP12* was equally expressed in green and etiolated cotyledons, confirming the light-independent expression of certain *CP12* genes.

Sucrose treatment of detached leaves in the light also led to a rapid loss of *GapA-1*, *GapB*, and *PRK* transcripts, again with parallel time-courses. Both *CP12* genes behaved similarly, but *CP12-2* closely matched the drop of Calvin cycle transcipts, while the response of *CP12-1* was somewhat delayed. In this study, *GapC* was steadily expressed during incubations with sucrose. By contrast, Shih and colleagues reported that *Arabidopsis GapA-1* and *GapB* genes were not influenced by sucrose and *GapC* expression was stimulated (Shih and Goodman, 1988; Yang *et al.*, 1993). The discrepancy with the present results is obvious and unexplained. However, it is commonly observed that those plant genes whose expression is increased by sugars are involved in the synthesis of carbohydrate reserves or in sugar catabolism, whereas sugar-repressed genes are usually involved in photosynthesis or the mobilization of storage compounds (Koch, 1996). Among Calvin cycle genes only *rbcs* has been clearly shown to be repressed by sugars, and the present demonstration that *GapA-1*, *GapB*, and *PRK* also behave in this way corroborates the sugar-sensing concept outlined above.
The effect of sucrose on photosynthetic GAPDH of chloroplasts was also notable in that the corresponding enzyme activities appeared to be far less affected than transcript levels. Although NADPH-GAPDH activity slowly decreased during lengthy treatments, it was still high in leaf material which had been depleted of either *GapA-1* or *GapB* messengers for 48 h at least. The delay between the decrease of transcript levels and significant changes in GAPDH activity is remarkable. However, Chan *et al.* (2002) detected normal levels of GapA and GapB subunits and normal NADPH-GAPDH activity in an *Arabidopsis* mutant impaired in light-stimulated *GapA-1* expression, i.e. deficient of *GapA-1* transcripts, suggesting that post-transcriptional control was involved in the up-regulation of the GapA protein in the mutant.

In conclusion, the expression of two CP12 genes in organs of *A. thaliana* has been examined under different conditions, in parallel with genes (*GapA*, *GapB*, and *PRK*) whose products are known to interact with CP12 proteins *in vivo*. The expression of cytosolic *GapC* and plastid *GapCp* which do not interact with CP12 was also investigated for comparison. The expression of *CP12-2*, but not *CP12-1*, was found to be co-ordinately regulated with that of *GapA-1*, *GapB*, and *PRK* under all conditions tested. On the other hand, the expression patterns of *GapC-1* and *GapCp-1* did not correlate with photosynthetic genes, and the latter finding is especially worth noting. The co-ordinated regulation of *GapA-1*, *GapB*, *CP12-2*, and *PRK* at the gene expression level is fully consistent with the existence and likely physiological relevance of a supramolecular complex involving GAPDH, CP12, and PRK proteins in chloroplasts (Wedel and Soll, 1998; Scheibe *et al.*, 2002; Graciet *et al.*, 2004a).

Chapter 3

*Reconstitution and properties of the recombinant glyceraldehyde-3-phosphate dehydrogenase/CP12/phosphoribulokinase supramolecular complex of Arabidopsis.**

*published as:

Lucia Marri, Paolo Trost, Paolo Pupillo, and Francesca Sparla (2005) *Plant Physiol.* **139:** 1433-1443

3.1 Abstract

Calvin cycle enzymes glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK) form together with the regulatory peptide CP12 a supramolecular complex in Arabidopsis (*Arabidopsis thaliana*) that could be reconstituted in vitro using purified recombinant proteins. Both enzyme activities were strongly influenced by complex formation, providing an effective means for regulation of the Calvin cycle in vivo. PRK and CP12, but not GapA (A_4 isoform of GAPDH), are redox-sensitive proteins. PRK was reversibly inhibited by oxidation. CP12 has no enzymatic activity, but it changed conformation depending on redox conditions. GapA, a bispecific NAD(P)-dependent dehydrogenase, specifically formed a binary complex with oxidized CP12 when bound to NAD. PRK did not interact with either GapA or CP12 singly, but oxidized PRK could form with GapA/CP12 a stable ternary complex of about 640 kD (GapA/CP12/PRK). Exchanging NADP for NAD, reducing CP12, or reducing PRK were all conditions that prevented formation of the complex. Although GapA activity was little affected by CP12 alone, the NADPH-dependent activity of GapA embedded in the GapA/CP12/PRK complex was 80% inhibited in respect to the free enzyme.

The NADH activity was unaffected. Upon binding to GapA/CP12, the activity of oxidized PRK dropped from 25% down to 2% the activity of the free reduced enzyme. The supramolecular complex was dissociated by reduced thioredoxins, NADP, 1,3-bisphosphoglycerate (BPGA), or ATP. The activity of GapA was only partially recovered after complex dissociation by thioredoxins, NADP, or ATP, and full GapA activation required BPGA. NADP, ATP, or BPGA partially activated PRK, but full recovery of PRK activity required thioredoxins. The reversible formation of the GapA/CP12/PRK supramolecular complex provides novel possibilities to finely regulate GapA ("non-regulatory" GAPDH isozyme) and PRK (thioredoxin sensitive) in a coordinated manner.

3.2 Introduction

Life of photosynthetic organisms depends on a finely controlled balance between light reactions of photosynthesis and photosynthesis-dependent metabolism. Signals involved in the maintenance of such balance include thioredoxins and a number of biochemical factors, including pyridine nucleotides, several metabolites, as well as pH and magnesium ions (Wolosiuk *et al*., 1993). In the Calvin cycle, five out of 11 enzymes are prominently sensitive to these signals (glyceraldehyde-3-phosphate dehydrogenase [GAPDH], Fru bisphosphate and sedoheptulose bisphosphate phosphatases, phosphoribulokinase [PRK], and Rubisco via Rubisco activase), and their activities are strongly regulated in vivo by light/dark conditions (Ruelland and Miginiac-Maslow, 1999; Schürmann and Jacquot, 2000; Dai *et al.,* 2004; Buchanan and Balmer, 2005). Moreover, redox regulation in chloroplasts is not restricted to carbon reduction, and the whole plastid metabolism seems to be controlled by the redox status of thioredoxins, as suggested by redox proteomic analysis (for review, see Buchanan and Balmer, 2005).

Besides the fine and complex tuning of individual enzyme activities, protein-protein interactions also contribute to the overall regulation of

photosynthetic metabolism. Supramolecular complexes of Calvin cycle enzymes with different compositions and stoichiometries, possibly interacting with thylakoids, have been widely documented (Müller, 1972; Wara-Aswapati *et al*., 1980; Clasper *et al*., 1991; Rault *et al*., 1993; Süss *et al*., 1993; Graciet *et al*., 2004a; Dani and Sainis, 2005). The presence of GAPDH and PRK in the aggregates was a common trait of all these reports, which were eventually supported by discovery of the small peptide CP12 (Pohlmeyer *et al*., 1996; Wedel and Soll, 1998).

CP12 is a small protein of 75 amino acids including four Cys separated by eight residues such that two short loops are believed to be generated by the formation of two internal disulfides. A supramolecular complex including GAPDH, CP12, and PRK has been described in different photosynthetic organisms (Wedel *et al*., 1997; Wedel and Soll, 1998; Scheibe *et al*., 2002; Graciet *et al.,* 2003a; Tamoi *et al*., 2005). CP12 has no enzymatic activity and seems to act as a scaffold protein within these complexes.

In higher plants photosynthetic GAPDH is formed by two types of subunits (A and B), giving rise to two isozymes A_4 and A_nB_n (Cerff, 1979; Pupillo and Faggiani, 1979; Brinkmann *et al.,* 1989). B-subunits are very similar to A-subunits but contain an additional C-terminal extension (CTE) homologous to the C-terminal end of CP12 (Pohlmeyer *et al*., 1996). This 30-amino acid tail contains a conserved couple of redox-sensitive Cys and is responsible for the regulatory properties of A_nB_n isoforms, including the capacity to form stable A_8B_8 oligomers in the presence of NAD(H) (Baalmann *et al*., 1996; Li and Anderson, 1997; Sparla *et al*., 2002). Consistently, the A_4 isozyme has no CTE and is constitutively activated (Scagliarini *et al.,* 1998). In spinach (*Spinacia oleracea*) chloroplasts, hexadecameric GAPDH coexists with GAPDH/CP12/PRK complexes, and both regulatory systems seem to contribute to dark/light regulation of Calvin cycle in vivo (Scheibe *et al*., 2002).

Several genes potentially implicated in the GAPDH/CP12/PRK complex are present in the genome of Arabidopsis (*Arabidopsis thaliana*). GAPDH is coded by two duplicated genes (*GapA-1* and *A-2*) and one *GapB* gene, PRK is coded by a single *PRK* gene, and CP12 by two closely related genes (*CP12-1* and *CP12-2*) and a third divergent one (*CP12-3*; Marri *et al*., 2005*a*; fig. 3.1).

```
CP12-1   MTTIAAAGLNVATPRVVVRP-VAR---VLGPVRLNYPWKFGS-----MKRMVVVKATS-- 49
CP12-2   MATIATG-LNIATQRVFVTS-ENRPVCLAGPVHLNNSWNLGSRTTNRMMKLQPIKAAP-- 56
CP12-3   ---MISGSATASHGRVLLPSQRERRPVSTGSNILRFRETVPRQFSLMMVTKATAKYMGTK 57
            : :.  . :  **.: .   *     *.  *.   ..       *      *

CP12-1   --EGEISEKVEKSIQEAKETCADDPVSGECVAAWDEVEELSAAASHARDKKKAGGSDPLE 107
CP12-2   --EGGISDVVEKSIKEAQETCAGDPVSGECVAAWDEVEELSAAASHARDKKKADGSDPLE 114
CP12-3   MREEKLSEMIEEKVKEATEVCEAEEMSEECRVAWDEVEEVSQARADLRIKLKLLNQDPLE 117
           *  :*: :*:.::** *.*  : :* ** .*******:* * :. * * *  ..****

CP12-1   EYCNDNPETDECRTYDN 124
CP12-2   EYCKDNPETNECRTYDN 131
CP12-3   SFCQENPETDECRIYED 134
         .:*::****:*** *::
```

Figure 3.1 Multiple alignment of Arabidopsis CP12s (CP12-1, At2g47400; CP12-2, At3g62410; CP12-3, At1g76650; according to Marri et al., 2005a). The alignment was performed by ClustalW. The transit peptide as predicted by ChloroP is shown on a gray field. Conserved Cys are on a black field. The sequence of recombinant CP12-2 retained at the N terminus four residues from the His tag. Therefore, the N terminus of recombinant CP12-2 started with GSHM (not shown), followed by AAPEGG....

GapA-1, *GapB*, *PRK*, and *CP12-2* were found to be coordinately expressed under different conditions in Arabidopsis leaves (Marri *et al*., 2005a). In this work, we used recombinant A_4-GAPDH (*GapA-1*), CP12 (*CP12-2*), and PRK of Arabidopsis to reconstitute the GAPDH/CP12/PRK complex under given conditions. Enzyme activities in the complex were found to be strongly inhibited, demonstrating that a constitutively activated enzyme such as A_4-GAPDH could be regulated by the reversible formation of a supramolecular complex with PRK and CP12. It is proposed that GAPDH/CP12/PRK supramolecular complexes occur in chloroplasts in the dark to ensure strong down-regulation of the Calvin cycle. Conditions leading to complex

destabilization and enzyme reactivation were also identified and appear to be representative of the general resurgence of photosynthetic metabolism at the onset of light.

3.3 Materials and Methods

3.3.1 *Plant Material*

Arabidopsis (*Arabidopsis thaliana*) plants were grown on sterile soil (humus:perlite 3:1) for 1 month at 22°C, under 15-h-dark/9-h-light cycle in a growth chamber. Leaves were collected after 15 h of dark. One gram of green tissue was homogenized on ice with 30 mL of Xpl buffer (330 mM sorbitol, 50 mM HEPES, pH 7.5, 2 mM EDTA, 1 mM $MgCl_2$), filtered through a Miracloth net, and centrifuged at 1,200*g* for 10 min. The pellet was resuspended in 1.5 mL of Xpl buffer, and chloroplasts were separated on a Percoll gradient, as described by Weigel and Glazebrook (2002). Arabidopsis chloroplast proteins were precipitated with 80% ammonium sulfate and resuspended in Xpl buffer without sorbitol but with the addition of 0.2 mM NAD. The sample was incubated for 2 h on ice before loading on the gel filtration column equilibrated with 50 mM Tris-HCl, pH 7.5, 150 mM KCl, 1 mM EDTA, and 0.2 mM NAD.

3.3.2 *Expression Vectors*

Arabidopsis cDNAs coding for the A-subunit of GAPDH (GapA-1, At3g26650), CP12 (CP12-2, At3g62410), and PRK (At1g32060) were transferred into a pET28a(+) expression vector (Novagen) using the following PCR primers: GapA forward (*Nco*I site), 5'-TGTGACCATGGCCAAGC-3', reverse (*Bam*HI site), 5'-CAAGGATCCCTCACTTC-3'; CP12 forward (*Nde*I site), 5'-CGCATATGGCAGCACCGG-3', reverse (*Bam*HI site), 5'-AGGATCCTGATCGCTTCAG-3'; and PRK forward (*Nco*I site),

5'-AGAAACCATGGTGATCGGAC-3', reverse (*Bam*HI site), 5'-TTGGATCCGTTTGTTTTAGGC-3'.

Specific endonuclease sites (underlined) were introduced at the 5' and 3' ends of the cDNA sequences. PCR-amplified fragments were digested with endonucleases, purified, and ligated into a predigested pET28a(+) vector. In the CP12 construct only, the cDNA sequence for CP12 was in frame with a His tag and a cleavable thrombin site. The coding sequence for mature CP12-2 was established after alignment of the three Arabidopsis CP12s (CP12-1, At2g47400; CP12-2, At3g62410; CP12-3, At1g76560; Marri *et al*., 2005*a*) and prediction of transit peptides by ChloroP (Emanuelsson *et al*., 1999; fig. 3.1).

Recombinant plasmids, amplified into *Escherichia coli* HB101 cells, were sequenced before transformation of *E. coli* BL21(DE3) cells.

3.3.3 *Expression and Purification of Recombinant Proteins*

Heterologous expression and purification of recombinant GapA was performed as described by Sparla *et al*. (2002). Expression and purification of CP12-2 was performed as by Sparla *et al*. (1999). After cleavage of the His tag by thrombin, four amino acids of the thrombin cleavage site (GSHM) remained attached to the N terminus of CP12 (AAPEG...; fig. 3.1). Resulting pure proteins were desalted in 100 mM Tricine-NaOH, pH 7.9, and stored at –20°C.

An overnight culture of *E. coli* BL21(DE3) cells, harbouring the pET28-PRK expression plasmid, was transferred to fresh LB medium supplied with kanamycin (50 *μ*g/mL) and grown for 6 to 8 h at 30°C. When the optical density at 600 nm reached 0.7, expression was induced by addition of 0.4 mM isopropylthio-β-galactoside. Fifteen hours after induction, the cells were

collected by centrifugation. The resulting pellet was washed with 25 mM potassium phosphate, pH 7.5, 1 mM EDTA, 10 mM β-mercaptoethanol (buffer A), and spun down again before storage at −70°C. The same GapA-purification protocol comprising two anion-exchange chromatographic steps (Sparla *et al*., 2005) was adopted to purify recombinant PRK. Pure PRK was desalted in 100 mM Tricine-NaOH, pH 7.9, and stored at −20°C.

Purified proteins were quantified by absorbance at 280 nm. Molar extinction coefficients at 280 nm were derived from the sequence of each monomer: ε_{GapA} = 36,250 M^{-1}, ε_{CP12} = 8,370 M^{-1}, and ε_{PRK} = 29,360 M^{-1}.

3.3.4 *Enzyme Activity Assays and Redox Titrations*

GAPDH activity was assayed as described by Sparla *et al*. (2002). PRK activity was assayed as described by Porter *et al.* (1986). Redox titrations of purified GapA and PRK were performed according to Hutchinson and Ort (1995) after desalting the purified proteins in 100 mM Tricine-NaOH, pH 7.9. In this buffer, enzymes were incubated for 3 h at 25°C with equimolar concentrations of recombinant thioredoxin from *E. coli* (Sigma) and 20 mM DTT in different dithiol/disulfide ratios, as described by Sparla *et al.* (2002). Following incubation, GAPDH and PRK activities were assayed and results were fit by nonlinear regression (CoStat, CoHort Software) to the Nernst equation with an *n* value of 2 (Hirasawa *et al.,* 2000) and analyzed as described (Sparla *et al*., 2002).

3.3.5 *Electrophoresis and Immunoblotting*

Purified samples of GapA, CP12, and PRK were examined by vertical SDS-PAGE on 12.5% acrylamide gels.

Reduced and oxidized CP12 were obtained by incubating the samples for 2 h at 25°C with equimolar concentrations of prereduced or preoxidized

thioredoxin, respectively. Prereduced and preoxidized thioredoxins were prepared by incubation for 2 h at 25°C with 20 mM reduced or oxidized DTT followed by washing out the DTT by ultrafiltration (Centricon YM3). Samples were then boiled for 3 min in sample buffer with no reductants and the proteins were separated on 15% acrylamide gels. Gels were stained with Coomassie Brilliant Blue R-250.

Fractions obtained from gel filtration columns were concentrated by ultrafiltration (Centricon YM3), run on denaturing 12.5% acrylamide gels, and electroblotted (Sammy-dry cell; Schleicher-Schuell) on nitrocellulose membranes. The membranes were stained with Red Ponceau before incubation with rabbit antiserum raised against spinach (*Spinacia oleracea*) CP12, spinach GapA, and Arabidopsis PRK, kindly provided by Renate Scheibe (University of Osnabrueck), and peroxidase-conjugated secondary antibodies. Primary and secondary antibodies were diluted 1:2,000 and 1:10,000, respectively. Blots were developed by chemiluminescence according to standard procedures.

3.3.6 *Supramolecular Complex Reconstitution and Gel Filtration*

Samples containing different combinations of purified recombinant GAPDH, PRK, and CP12 under different conditions were analyzed by gel filtration to detect the *in vitro* reconstitution of binary and ternary complexes. Oxidized and reduced forms of PRK and CP12 were obtained by incubation for 3 h at 25°C with 25 mM oxidized DTT or 25 mM reduced DTT, respectively, followed by buffer exchange to 100 mM Tricine-NaOH, pH 7.9, through ultrafiltration (Centricon YM10 and YM3 for PRK and CP12, respectively). GapA was also equilibrated in buffer 100 mM Tricine-NaOH, pH 7.9. In order to *in vitro* reconstitute the supramolecular complexes, purified proteins were incubated for 2 h at 4°C at equimolar ratios (subunit basis) under different conditions as described in the text.

Gel filtration analysis was performed on a Superdex 200 HR10/30 column connected to an ÅKTA Purifier system (General Electric Healthcare). The column was equilibrated with 50 mM Tris-HCl, pH 7.5, 150 mM KCl, 1 mM EDTA, and 0.2 mM NAD(P) or 2 mM DTT as specified in the figure legends. The volume of loaded samples was 0.2 mL, and fractions of 0.35 mL were collected. The column was calibrated as done by Sparla *et al.* (2002).

3.3.7 *Complex Dissociation and Recovery of Enzyme Activity*

The GapA/CP12/PRK supramolecular complex eluted from the Superdex 200 column in the presence of 0.2 mM NAD was collected and equilibrated with 100 mM Tricine-NaOH, pH 7.9, in the absence of NAD. Potential dissociating agents as described in "Results" were incubated with the complex for 1 h at 25°C. A steady-state concentration 43 µM BPGA was obtained in a mixture of 3 mM 3-phosphoglycerate, 2 mM ATP, and 5 units mL^{-1} of rabbit muscle phosphoglycerate kinase (Sigma). Following incubation, samples were reloaded on the Superdex 200 column equilibrated with 100 mM Tricine-NaOH, pH 7.9, in the absence of effectors. GAPDH and PRK activities were measured immediately before the addition of effectors and immediately before loading the samples on the gel filtration column.

3.4 *Results*

As previously shown in other plant species (Wedel *et al*., 1997; Scheibe *et al.*, 2002), GAPDH/CP12/PRK complexes could be detected in Arabidopsis by gel filtration of NAD-treated chloroplast proteins. GAPDH and PRK activities coeluted at high molecular mass (600–750 kDa) together with a 20-kDa protein recognized by antibodies raised against spinach CP12 (data not shown). These fractions are likely to contain a mixture of complexes of similar size, including A_4-GAPDH/CP12/PRK, A_2B_2-GAPDH/CP12/PRK, and

A_8B_8-GAPDH oligomers (Scheibe *et al*., 2002). In order to study the GAPDH/CP12/PRK complex in a simplified system, we tried to reconstitute the supramolecular complex from isolated recombinant components.

3.4.1 *Expression, Purification, and Redox Properties of Arabidopsis Recombinant GAPDH, CP12, and PRK*

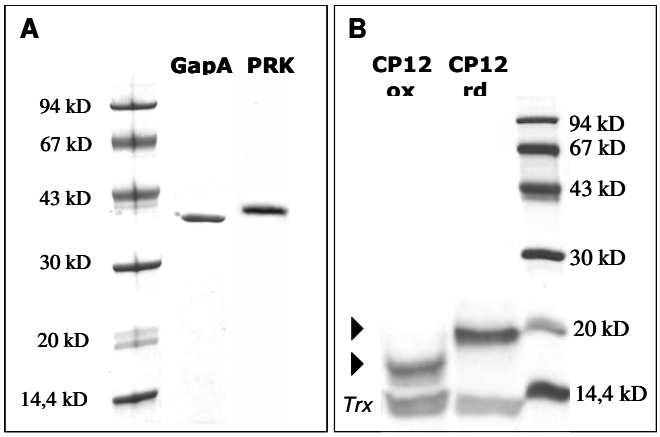
Each single component of the GAPDH/CP12/PRK complex was heterologously expressed in *Escherichia coli*. For the sake of simplicity, GAPDH was expressed as the simplest "non-regulatory" isozyme constituted by A-subunits only (GapA). Recombinant GapA of Arabidopsis was purified until it displayed a single band of 36 kDa in SDS-PAGE (fig. 3.2A). Purified GapA was redox insensitive when treated with thioredoxins but was found to be inactivated by oxidized dithiothreitol (DTT) and other oxidants, including H_2O_2. This effect could not be reverted by reductants and seemed to depend on irreversible oxidation of catalytically essential Cys-149 (Zaffagnini *et al*., 2007).

Although spinach PRK was reported to be reluctant to expression as an active enzyme in *E. coli* (Brandes *et al.,* 1996), about 2 mg of active Arabidopsis PRK per liter of liquid culture were routinely obtained when bacterial cells were grown at 30°C. Purified PRK migrated as a single 38-kDa band in SDS-PAGE (fig. 3.2A). Redox sensitivity of Arabidopsis PRK showed a midpoint redox potential of –330 mV at pH 7.9, within the range of E_m measured for PRK of other plant species (Hirasawa *et al.,* 1999; Hutchinson and Ort, 2000). Residual activity of the enzyme equilibrated with oxidized DTT and thioredoxin accounted for 25% of the activity of the fully reduced enzyme (see fig. 3.4).

Recombinant CP12-2 was expressed as a fusion protein with a His tag at the N terminus of the transit peptide cleavage site as predicted by ChloroP (Emanuelsson *et al*., 1999). The protein was purified by nickel affinity chromatography and the tag removed by proteolysis. Despite the calculated

molecular mass of 8.5 kDa, purified CP12 migrated in SDS-PAGE as a peptide of 20 kDa after incubation with reduced thioredoxins and 16 kDa after incubation with oxidized thioredoxins, indicating that disulfide formation led to a major conformational change that appeared to be retained even under denaturating conditions (fig. 3.2B).

Figure 3.2 SDS-PAGE of GapA, PRK, and CP12 of Arabidopsis expressed in *E. coli* and purified to homogeneity. Gels were stained with Coomassie Brilliant Blue R-250. A, GapA and PRK were treated with sample buffer including reductant (5 mM DTT) and loaded on a 12.5% polyacrylamide gel. B, Reduced (rd) and oxidized (ox) CP12 samples were obtained by incubation for 2 h at 25°C with equimolar concentrations of prereduced or preoxidized thioredoxin, respectively (see "Materials and Methods"). Samples were then boiled for 3 min in sample buffer with no reductants, and the proteins were separated on 15% polyacrylamide gels. Thioredoxin migrated below the 14.4-kD marker. The apparent molecular mass of CP12 (arrowheads) shifted from 20 to 16 kD depending on redox conditions.



3.4.2 *In Vitro Reconstitution of Binary and Ternary Complexes*

The capability of isolated GAPDH, CP12, and PRK to bind to each other protein was tested, and supramolecular complexes were detected by gel filtration chromatography. Different possible conformations of single protein components were compared: Holo-GapA was tested as either NADP or NAD complex, and CP12 and PRK were tested as either reduced or oxidized

proteins. On the whole, 12 possible combinations of binary complexes and eight combinations of ternary complexes could be envisaged, but only a few were found to be productive in terms of formation of supramolecular complexes.

Free GapA is a tetramer of identical subunits with an apparent mass of 120 kDa in the presence of either NAD or NADP (fig. 3.3A). PRK is a dimer with an estimated mass of 110 kDa under oxidizing conditions (fig. 3.3A), and apparently smaller (97 kDa) when reduced (data not shown). Although the peaks of GapA and PRK overlapped, both were fully separated from the 29-kDa peak of oxidized CP12 and the 35-kDa peak of reduced CP12 (fig. 3.3A). Since CP12 is an intrinsically unstructured protein (Graciet *et al.,* 2003*a*) while the gel filtration column was calibrated with globular proteins, these determinations of highly overestimated, making the polymerization state of native CP12 uncertain.

Incubation of GapA with PRK at equimolar ratio (subunit basis) failed to result in formation of a binary complex, regardless of the type of pyridine nucleotide bound to GapA or the redox state of PRK (data not shown). No complex was also detected when GapA was incubated with NADP and oxidized CP12 (1:1 subunit ratio). Both proteins eluted as isolated moieties and specific antibodies could not detect any CP12 interacting with GapA (fig. 3.3B). On the contrary, in the presence of NAD instead of NADP, the peak of GapA displayed an apparent increase in size of 30 kDa, while the peak of free CP12 dropped to low levels (fig. 3.3C). Immunoblots confirmed that under these conditions most of CP12 coeluted with GapA, although inhibition of GapA activity was negligible (fig. 3.4A). Replacing oxidized CP12 with reduced CP12 had the effect of abolishing the interaction between the two proteins, even in the presence of NAD (data not shown).

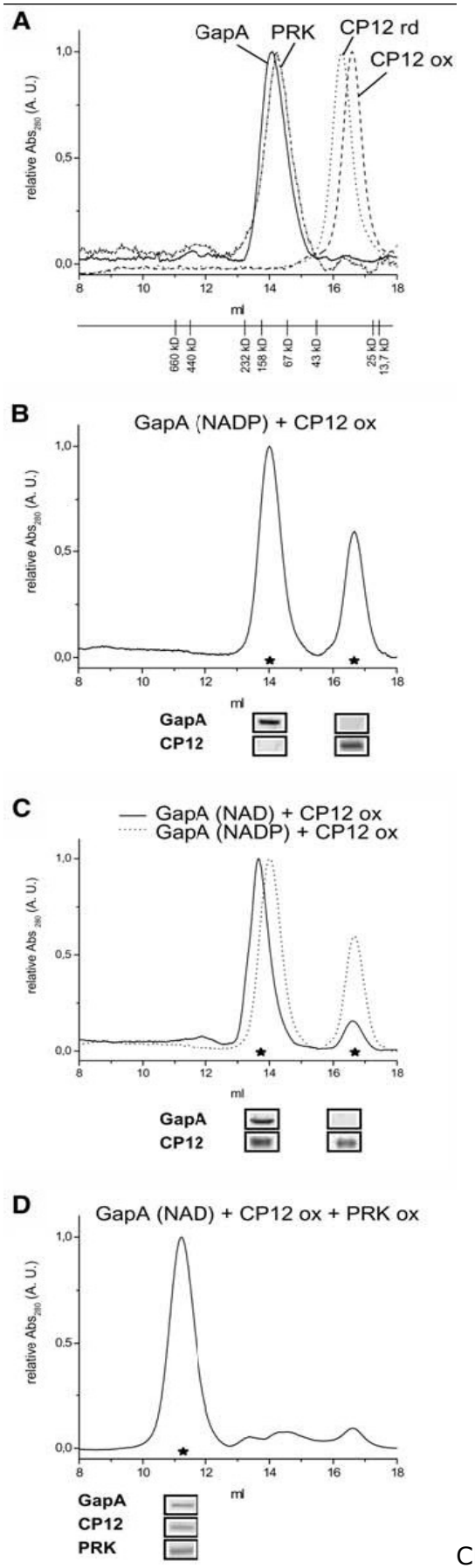


Figure 3.3 In vitro reconstitution of binary (GapA/CP12) and ternary (GapA/CP12/PRK) complexes. A, Gel filtration (Superdex 200) of purified GapA, PRK (oxidized), and CP12 (either oxidized or reduced). Oxidized and reduced proteins were obtained by 3 h incubation at 25°C in the presence of 25 mM reduced or oxidized DTT. The four samples were individually loaded on the column and run under identical conditions. The absorption patterns at 280 nm were normalized and superimposed. Column equilibration buffer was 50 mM Tris-HCl, pH 7.5, 150 mM KCl, 1 mM EDTA; volume of loaded samples was 0.2 mL; and flow rate was 0.5 mL min^{-1}. Column calibration is reported at the bottom of the figure. Estimated molecular masses of samples were 120 kDa (GapA), 110 kDa (PRK, oxidized), 35 kDa (CP12, reduced), and 29 kDa (CP12, oxidized). B, Same type of chromatography as in A, except that the sample loaded on the column was a mixture of equimolar GapA and CP12 (oxidized) on subunit basis, incubated for 2 h at 4°C in the presence of 0.5 mM NADP before loading. Column equilibration buffer was as in A with the addition of 0.2 mM NADP. Insert, Western blots showing that anti-GAPDH polyclonal antibodies recognize GapA only in the peak corresponding to 120 kDa, and anti-CP12 antibodies recognize CP12 only in the low M_r peak (29 kDa). Stars indicate the column fractions (0.35 mL) which were concentrated and loaded on the gel for western blotting. C, Same experiments as in B except that NAD substituted NADP in both incubation and column buffers. The GapA peak shifted to the left to an apparent molecular mass of 150 kDa. The elution pattern of B is superimposed for easy comparison. Insert, Western blots showing that anti-CP12 antibodies recognized a 16-kDa peptide coeluting with GapA tetramers (36 kDa in SDS-PAGE). D, Same experiment as in C except for the addition, in the incubation buffer, of equimolar PRK (oxidized) on subunit basis. Insert, Immunoblots showing that the peak at 640 kDa contained GapA, CP12, and PRK.

CP12

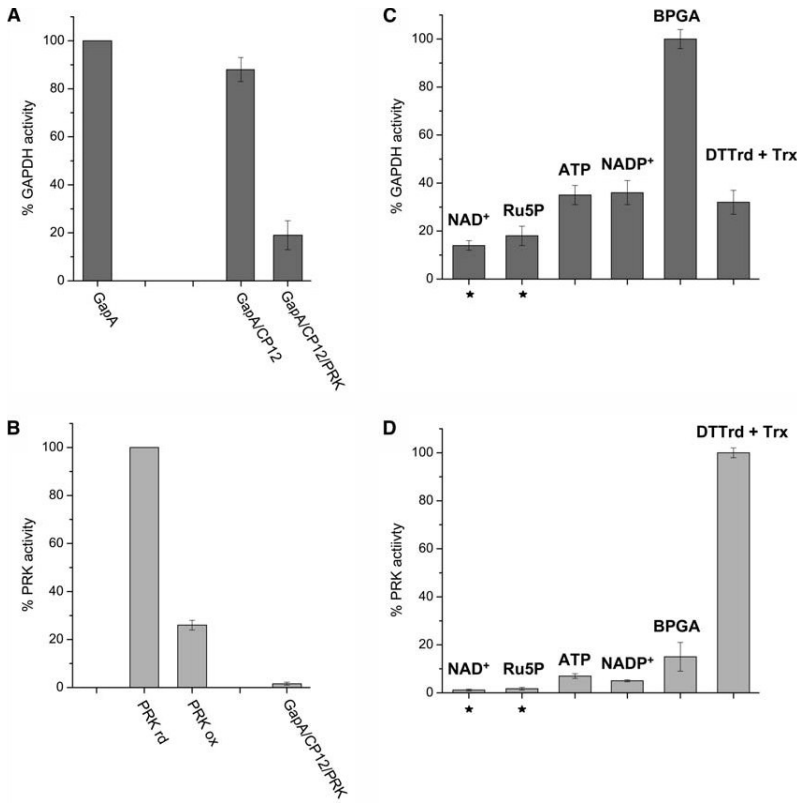


Figure 3.4 Activities of GapA and PRK as free enzymes, as enzymes embedded in complexes, and after complex treatment with several potential effectors. A, Activity expressed as percentage of the full activity of free tetrameric GapA (NADPH dependent). GapA/CP12 and GapA/CP12/PRK complexes were obtained under the same conditions as described in legends to figures 3.3C and 3.3D, respectively. B, PRK activity was assayed and expressed as a percentage of fully reduced PRK. PRK oxidation was obtained by 3 h incubation at 25°C with 25 mM oxidized DTT. The ternary complex with GapA and CP12 was obtained as in figure 3.3D. C, The GapA/CP12/PRK complex was reconstituted and chromatographed as in figure 3.3D, re-equilibrated with 100 mM Tricine-NaOH, pH 7.9, in the absence of NAD, and incubated under different conditions (0.2 mM NAD; 0.5 mM ribulose-5-P; 2 mM ATP; 0.2 mM NADP; 43 μM BPGA; 5 mM reduced DTT plus 1 μg/mL thioredoxin). BPGA (43 μM) was produced at equilibrium by the reaction of phosphoglycerate kinase with 3 mM 3-phosphoglycerate and 2 mM ATP (initial concentrations). After 1 h incubation at 4°C with different effectors, the NADPH activity of GapA was assayed and expressed as percentage of the activity of GapA before complex reconstitution. An aliquot of the incubated sample was also loaded on a gel filtration column to check the aggregation state of the proteins. The stars indicate those conditions that did not lead to complex dissociation (NAD and ribulose-5-P; see also fig. 3.5). D, Same conditions as in C except that PRK activity was assayed and expressed as percentage of the activity of the fully reduced enzyme. Ru5P, Ribulose-5-P.

PRK was itself unable to bind CP12 under any redox conditions (data not shown). On the other hand, oxidized PRK quantitatively formed a supramolecular complex of about 640 kD when incubated with GapA-NAD and oxidized CP12 (fig. 3.3D). Immunoblots demonstrated that all three

partner proteins coeluted in the high M_r peak. Formation of the complex was prevented by reduction of PRK (data not shown). Interestingly, formation of the GapA/CP12/PRK complex led to dramatic inhibition of the activity of both enzymes. Within the complex, PRK was 12-fold less active than the free oxidized counterpart and 50-fold less active than the reduced enzyme (fig. 3.4B). In a similar mood, the NADPH-dependent activity of GapA embedded in the complex was 5-fold lower than for free enzyme (fig. 3.4A), whereas the NADH-activity remained unchanged.

3.4.3 *In Vitro Dissociation of Binary and Ternary Complexes*

The supramolecular GapA/CP12/PRK complex isolated in the presence of NAD proved to be quite stable. Buffer exchange to remove excess of NAD did not affect complex stability, and reloading of the complex on the column equilibrated without NAD led to negligible dissociation.

Ligands of either GapA or PRK were tested as possible effectors of complex dissociation. The PRK substrate ribulose-5-P had no significant effect on complex stability or enzyme activities (figs. 3.4 and 3.5). Incubations with NADP or ATP dissociated the complex, and a major peak of about 120 kDa, including both GapA and PRK free proteins, was observed. The peak of free CP12 was hardly detectable under these conditions, partially due to the low molar extinction coefficient of this small protein (fig. 3.5). While dissociating the complex, ATP and NADP stimulated the activity of both GapA (NADPH dependent, 2-fold) and PRK (3- to 4-fold). However, the activity of both GapA and PRK released from the complex was much lower than the activity displayed by the enzymes before complex formation (fig. 3.4).

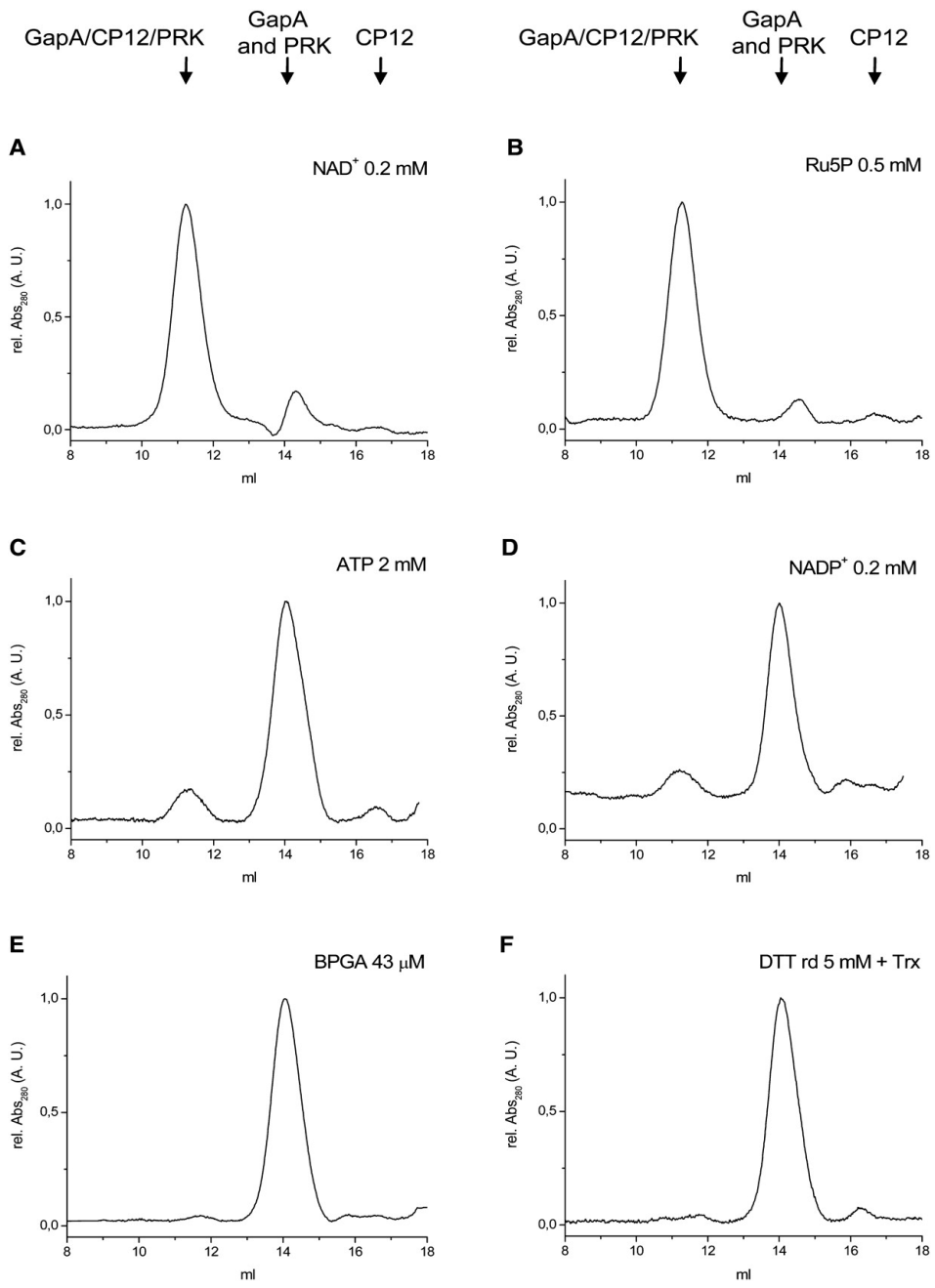


Figure 3.5 Effect of GapA and PRK ligands on GapA/CP12/PRK complex stability. The GapA/CP12/PRK complex was reconstituted and chromatographed as in figure 3.3D, re-equilibrated with 100 mM Tricine- NaOH, pH 7.9, in the absence of NAD, and incubated under different conditions (A, 0.2 mM NAD; B, 0.5 mM ribulose-5-P; C, 2 mM ATP; D, 0.2 mM NADP; E, 43 μM BPGA; F, 5 mM reduced DTT plus 1 μg/mL thioredoxin) before loading on the gel filtration column (Superdex 200). The column equilibration buffer included different effectors as reported in the figure. The elution volumes of GapA/CP12/PRK (640 kDa), GapA and PRK (110–120 kDa), and CP12 (29–35 kDa) are indicated. Under conditions of complex disruption (ATP, NADP, BPGA, or reductants), the CP12 peak was hardly detectable, partly due to the lower molar extinction coefficient in comparison with either GapA or PRK (see "Materials and Methods").

Full recovery of GapA activity could be achieved by further incubation with the substrate 1,3-bisphosphoglycerate (BPGA; produced by ATP, 3-phosphoglycerate, and phosphoglycerate kinase). The BPGA-producing mixture also dissociated the complex directly (fig. 3.5), thereby activating GapA (6-fold) at maximal levels. Complex dissociation by BPGA also resulted in PRK activation (7-fold) yet without reaching the activity of the oxidized free enzyme (fig. 3.4, B and D). Reduced DTT and thioredoxins quantitatively dissociated the complex and fully activated PRK. GapA was activated only 2-fold by reducing conditions, similar to the effect of NADP or ATP. Independent of the effector used to destabilize the GapA/CP12/PRK complex, full activity of GapA was always recovered by further incubation with BPGA, while full PRK activity required DTT (fig. 3.4). In no case did dissociation of the ternary complex lead to binary complexes of whatsoever composition, as indicated by the elution volumes of released proteins. Complex dissociation invariably gave rise to GapA, CP12, and PRK free proteins (fig. 3.5).

3.5 *Discussion*

Enzymatic supramolecular complexes in photosynthetic organisms have long been investigated, but their physiological meaning is still a matter of debate (Gontero *et al*., 2002). Like in other cell compartments characterized by strong metabolic activity (Goodsell, 1991), the protein concentration of the stroma is very high and chloroplastic enzymes are necessarily subjected to continuous interactions in vivo. However, we do not know whether protein-protein interactions commonly result in the formation of supramolecular complexes with specific organization and function. The existence of a complex network of interactions within the proteome has recently been documented in many organisms (Gavin and Superti-Furga, 2003) and analogous networks of protein interactions may be suspected to exist in plants as well. Among photosynthetic enzymes, GAPDH and PRK catalyze

two non-consecutive reactions of the Calvin cycle and have been proposed to physically interact in vivo together with a scaffold protein known as CP12 (Wedel *et al*., 1997; Graciet *et al*., 2004*a*). The reversible formation of this complex has been proposed to contribute to the regulation of the Calvin cycle in vivo (Wedel and Soll, 1998; Scheibe *et al*., 2002; Graciet *et al*., 2004*a*; Tamoi *et al*., 2005).

In oxygenic photosynthetic organism thioredoxins, pyridine nucleotides and metabolites play an important role in regulating the Calvin cycle in dark/light transitions (Wolosiuk *et al*., 1993; Buchanan and Balmer, 2005). PRK is strongly regulated by thioredoxins in higher plants and green algae (Porter *et al*., 1988; Graciet *et al*., 2004*a*), but apparently is less redox sensitive in cyanobacteria and diatoms (Kobayashi *et al*., 2003; Michels *et al.*, 2005). GAPDH is finely regulated by thioredoxins, NAD(P)(H), and BPGA in higher plants (Pupillo and Giuliani Piccari, 1975; Wolosiuk and Buchanan, 1978; Trost *et al*., 1993; Baalmann *et al*., 1995; Sparla *et al*., 2002), but it does not appear to be regulated in lower photosynthetic organisms. In fact, in green unicellular algae and in cyanobacteria, the GAPDH is composed of a single type of subunit related to subunit A of higher plants (Figge *et al*., 1999). This subunit lacks the pair of redox-sensitive Cys of subunits B and is therefore insensitive to regulatory effectors including thioredoxins (Sparla *et al*., 2002).

CP12 is a redox-sensitive protein widely distributed in oxygenic photosynthetic organisms that is able to interact with both GAPDH and PRK (Pohlmeyer *et al*., 1996; Wedel *et al.,* 1997). In those organisms that do not contain an autonomously regulated GAPDH, CP12 may provide a means to regulate this activity in concert with PRK (Wedel and Soll, 1998; Graciet *et al.,* 2004*a*; Tamoi *et al*., 2005). In *Chlamydomonas reinhardtii*, for instance, native GAPDH could be purified as a stable complex with CP12 (Graciet *et al.,* 2003*b*). PRK steadily interacted with the GAPDH/CP12 binary complex, and activities of both enzymes were inhibited compared to free (activated) counterparts (Graciet *et al.,* 2003*a*, 2003*b*).

The relevance of the GAPDH/CP12 interaction in lower photosynthetic organisms is strengthened by the existence of an autonomously regulated GAPDH in higher plants. This GAPDH isoform contains B-subunits resulting from the fusion of GapA with the C-terminal end of CP12 (Pohlmeyer *et al*., 1996). As a result, B-subunits are regulated by thioredoxins, pyridine nucleotides, and BPGA, while A-subunits are not (Baalmann *et al*., 1996; Li and Anderson, 1997; Sparla *et al*., 2002). The C-terminal extension of GapB is suspected to interact with the coenzyme binding site of the protein, thereby leading to a specific down-regulation of the NADPH-dependent enzyme activity (Sparla *et al*., 2002). The high sequence similarity between CTE and CP12 (Graciet *et al*., 2004*a*) suggests that both peptides interact with GAPDH in the same way.

Despite the evolution of the autonomous regulation of GAPDH (CTE independent) from an ancient CP12-dependent system, CP12 genes are present in multiple copies in higher plant genomes, and CP12-dependent regulation of GAPDH and PRK seems to be conserved up to higher photosynthetic organisms. In this work, we show that in Arabidopsis the homomeric A_4 isozyme of photosynthetic GAPDH (GapA) can form a complex with CP12 (GapA/CP12) and this binary complex can further polymerize by interacting with PRK to give rise to the ternary complex GapA/CP12/PRK (fig. 3.6). The molecular mass of the reconstituted GapA/CP12/PRK complex was about 640 kDa, similar to the 550- to 600-kDa complexes previously detected in spinach chloroplasts (Clasper *et al*., 1991; Wedel *et al*., 1997; Scheibe *et al*., 2002). A complex composition of two GapA tetramers (2 x 120 kDa), two PRK dimers (2 x 110 kDa), and two CP12 (29 or 2 x 29 kDa, depending on whether native CP12 is a monomer or a dimer; fig. 3.3), as tentatively proposed by others (Wedel and Soll, 1998; Graciet *et al*., 2004*a*), might apply to the Arabidopsis complex as well, but different technical approaches are needed to precisely define this stoichiometry. The formation of complexes involving CP12 in Arabidopsis was promoted by specific conditions. Only the interaction between GapA bound to NAD and CP12 in the oxidized state led to a stable binary complex (fig. 3.6). The

effect of CP12 on GapA activity was negligible, in contrast with the inhibition observed in the GapA/CP12 complex of Chlamydomonas (Graciet *et al*., 2003*b*). In Arabidopsis, strong inhibition of GapA activity was only observed in the presence of oxidized PRK, when a stable GapA/CP12/PRK complex could be reconstituted (fig. 3.6). Albeit intrinsically insensitive to oxidized thioredoxins and pyridine nucleotides (Baalmann *et al*., 1996; Scagliarini *et al.,* 1998; Sparla *et al*., 2002), GapA could thus be effectively regulated by a CP12/PRK-dependent mechanism, provided NAD and oxidizing conditions were applied. Moreover, although oxidation of free PRK by oxidized thioredoxins resulted in 80% loss of activity, complexation with GapA/CP12 caused a further, almost complete PRK inhibition (fig. 3.6).

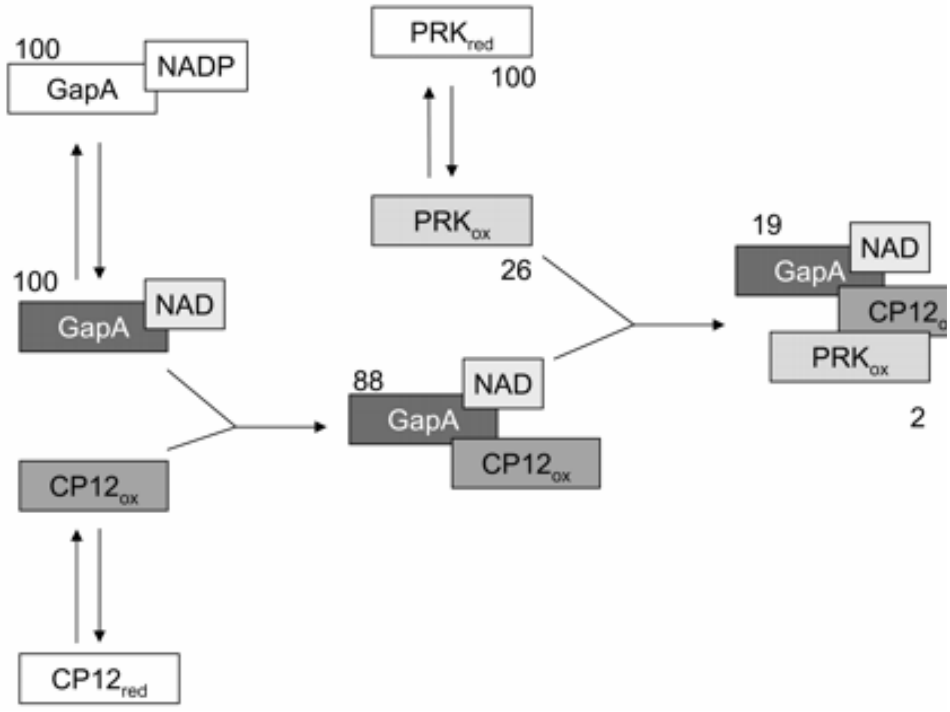


Figure 3.6 Schematic representation of the sequential formation of GapA/CP12 and GapA/CP12/PRK complexes. The percentage activity of GapA (NADPH dependent) and PRK in respect to free, fully activated enzymes is reported by numbers close to each enzyme form. The scheme indicates which enzyme forms participate in complex formation but does not represent the stoichiometry of complexes.

Both oxidized/reduced thioredoxins and NAD/NADP ratios increase in chloroplasts in the dark (Muto *et al*., 1980; Heineke *et al.,* 1991; Buchanan and Balmer, 2005), suggesting that the formation of the GapA/CP12/PRK complex in vivo may contribute to the kinetic down-regulation of the Calvin cycle in the dark. Similar oscillations of pyridine nucleotides were also measured under dark/light conditions in *Synechococcus* PCC7942. In this cyanobacterium, a GapA/CP12/PRK complex was detected in vitro under conditions reproducing the cellular NAD/NADPH ratio measured in the dark. In this organism at least, pyridine nucleotides seem to play a major role in regulating the Calvin cycle via CP12 (Tamoi *et al*., 2005).

The contrasting effects of NAD and NADP on GapA/CP12/PRK complex formation and consequent enzyme inhibition are consistent with the structural similarity between CP12 and the C-terminal extension of GAPDH B-subunits. In A_2B_2-GAPDH, binding of NAD leads to enzyme aggregation to A_8B_8 oligomers and specific inhibition of the NADPH-dependent activity. The process is slow, reverted by NADP, and strictly dependent on the C-terminal extension of B-subunits (Pupillo and Giuliani Piccari, 1975; Baalmann *et al*., 1996; Sparla *et al*., 2002). Clearly, the process of GapA/CP12/PRK complex formation shares several features with this system, further supporting the evolutionary derivation of regulated B-subunits of photosynthetic GAPDH from a fusion between ancient A-subunits and the C-terminal portion of CP12 (Pohlmeyer *et al.,* 1996).

Dissociation of the GapA/CP12/PRK complex of Arabidopsis occurred under several conditions with variable effects on enzyme activities. Reduced thioredoxins led to complex disruption and total recovery of full PRK activity. Although GapA was found as a free tetramer after complex dissociation by reductants, the activity was only partially recovered. NADP and ATP behaved similarly to reductants in dissociating the complex and partially activating GapA. Full activation of GapA required BPGA incubation. This result was puzzling since GapA, which lacks the CTE, was known to be insensitive to common activators of AB-GAPDH, including BPGA (Cerff, 1979; Baalmann *et*

al., 1996; Scagliarini *et al*., 1998; Sparla *et al*., 2002). Therefore, we propose that dissociation of the GapA/CP12/PRK complex by reductants, NADP, or ATP released tetrameric GapA in a state reminiscent of the inhibited conformation the protein had within the complex, being rescued by its substrate BPGA alone. PRK behaved somehow similarly, as the activity of the enzyme released by the complex by NADP, ATP, or BPGA was still much lower than the activity of free oxidized enzyme. A similar "imprinting" effect was reported for both GAPDH and PRK of Chlamydomonas (for review, see Graciet *et al.,* 2004*a*).

Although in higher plants both AB-GAPDH and PRK can be directly regulated by thioredoxins and metabolites in the absence of CP12, the formation of a GapA/CP12/PRK supramolecular complex provides new potentialities for the regulation of the Calvin cycle in dark/light conditions. First, in the absence of CP12, GapA would be constitutively activated, and this might not be compatible with the need to silence the Calvin cycle in the dark. Second, CP12-mediated regulation of GAPDH and PRK provides a novel way to coordinately regulate both enzyme activities. Within the complex, GapA becomes sensitive to molecules (primarily BPGA, but also thioredoxins, ATP, and NADP) that do not affect the activity of free, isolated enzyme at all; much in the same way, complexed PRK becomes sensitive to GapA substrates (NADP, BPGA) while the free enzyme is only sensitive to thioredoxins. The CP12-dependent, coordinated regulation of GAPDH and PRK may be a major requirement for an effective modulation of the Calvin cycle in light/dark conditions. As a first confirmation of this hypothesis, a Synechococcus mutant in which the CP12 gene was disrupted showed limited growth in light/dark cycle but normal growth under continuous light (Tamoi *et al*., 2005).

Chapter 4

*Spontaneous assembly of photosynthetic supramolecular complexes as mediated by the intrinsically unstructured protein CP12.**

*submitted as:

Lucia Marri [1], Paolo Trost [1], Xavier Trivelli [2], Leonardo Gonnelli [3], Paolo Pupillo [1] and Francesca Sparla [1]

(1) Laboratory of Molecular Plant Physiology, Department of Experimental Evolutionary Biology, University of Bologna, Bologna, Italy
(2) UGSF, UMR 8576 CNRS, IFR147, Science and Technology University of Lille, 59655 Villeneuve d'Ascq, France;
(3) Magnetic Resonance Center (CERM), University of Florence, Sesto Fiorentino (Firenze), Italy

4.1 *Abstract*

CP12 is a regulatory protein of 8.7 kDa which was proposed to work as a scaffold element in promoting the formation of a supramolecular complex with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK) in different photosynthetic organisms. Solution NMR studies of recombinant CP12 (isoform 2) of *Arabidopsis thaliana* show that CP12-2 is poorly structured. Dynamic light scattering analysis demonstrates that CP12-2 is monomeric in solution although it behaves as a 29 kDa protein in size-exclusion chromatography. CP12-2 contains four cysteines, which can form two intramolecular disulfides characterized by midpoint redox potentials of -326 and -352 mV, respectively, at pH 7.9. Site specific mutants showed that the C-terminal disulfide is involved in the interaction between CP12-2 and GAPDH (isoform A_4), while the N-terminal

disulfide is involved in the interaction between this binary complex and PRK. In the presence of NAD and under oxidizing conditions CP12-2 and A_4-GAPDH interact with a K_D of 0.18 μM to form a binary complex with (A_4-GAPDH)-(CP12-2)$_2$ stoichiometry, as determined by isothermal titration calorimetry. PRK interacts with this binary complex (K_D 0.17 μM) leading to the formation of a 498 kDa ternary complex made of two binary complexes bound to two PRK dimers, i.e. [(A_4-GAPDH)-(CP12-2)$_2$–(PRK)]$_2$. Thermodynamic parameters indicate that assembly of both binary and ternary complexes are both exothermic and exoergonic though penalized by a decrease in entropy which may reflect induced folding of CP12-2 upon binding to its partners.

4.2 *Introduction*

The photosynthetic reduction cycle for carbon organication (Calvin cycle) is a finely regulated metabolism of plants that is kept in balance, under oscillating light supply, with light reactions of photosynthesis. Thioredoxins and metabolic intermediates play essential signalling roles within this regulatory system (Wolosiuk *et al*., 1993). Two non-consecutive enzymes of the Calvin cycle, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK), are both individually regulated and subject of reversible formation of a supramolecular complex with perturbed kinetic properties (Wedel *et al.,* 1997; Wedel *et al.,* 1998; Scheibe *et al.,* 2002; Graciet *et al.,* 2004*b;* Tamoi *et al.,* 2005; Marri *et al.,* 2005*b*; Trost *et al.,* 2006). Both GAPDH and PRK catalyze energy consuming reactions and account as a whole for most of the energetic needs of the Calvin cycle. Reversible formation of supramolecular complexes under variable environmental conditions is envisioned as a powerful means for regulating individual enzyme activities in a tightly coordinated manner (Marri *et al.,* 2005*a*; Trost *et al.,* 2006).

Chloroplast GAPDH is mainly heteromeric in higher plants, being constituted by A and B subunits in stoichiometric ratio (Cerff *et al.,* 1979; Fermani *et*

al., unpublished). B-subunits confer regulatory properties to the whole enzyme which oscillates between a fully active A_2B_2 tetramer, at one extreme, and a partially inhibited A_8B_8 hexadecamer on the other extreme (Pupillo and Giuliani Piccari, 1975). Partially polymerized intermediates like A_4B_4 have also been reported (Scagliarini *et al.,* 1993; 1998; Baalmann *et al.,* 1994). Thioredoxins and metabolites directly regulate AB-GAPDH activity and strongly affect the equilibrium between active tetramers and aggregated forms (Trost *et al.,* 2006; Fermani *et al.,* unpublished).

A second isoform of chloroplast GAPDH is a stable homotetramer of A-subunits (A_4-GAPDH; Cerff *et al.,* 1979; Scagliarini *et al.,* 1998; Fermani *et al.,* 2001) similar to Calvin cycle's GAPDH of lower photosynthetic organisms (Figge *et al.,* 1999). Due to the absence of B-subunits, A_4-GAPDH is not directly regulated by thioredoxins and metabolites and shows different regulatory properties as compared to AB-GAPDH (Trost *et al.,* 2006). Reversible glutathionylation of the active site cysteine-149 provides a mechanism of A_4-GAPDH regulation which may be relevant under stress (Zaffagnini *et al.,* 2007). Alternatively, reversible down-regulation of A_4-GAPDH activity can be attained through the formation of a supramolecular complex with PRK and the regulatory peptide CP12, particularly under varying light/dark conditions (Wedel *et al.,* 1997; Wedel and Soll, 1998; Scheibe *et al.,* 2002; Graciet *et al.,* 2003*a,b*; Graciet *et al.,* 2004*a*; Marri *et al.,* 2005*b*; Tamoi *et al.,* 2005; Trost *et al.,* 2006). PRK itself is subject to light/dark modulation mediated by thioredoxins (Porter *et al.,* 1988), but once embedded in the complex, both A_4-GAPDH and PRK are strongly inhibited in a tightly coordinated manner (Marri *et al.,* 2005*b*). Similar to A_8B_8-GAPDH, the stability of the supramolecular complex involving A_4-GAPDH, CP12 and PRK is also controlled by thioredoxins, coenzymes (NAD(P)(H), ATP) and 1,3-bisphosphoglycerate (Wedel and Soll, 1998; Scheibe *et al.,* 2002; Graciet *et al.,* 2004*a*; Tamoi *et al.,* 2005; Marri *et al.,* 2005*b*). In general, it is accepted that aggregated forms of GAPDH (A_8B_8) and GAPDH/CP12/PRK complexes are prevailing in chloroplasts in the dark while illumination favours the accumulation of fully active GAPDH tetramers

(A_2B_2, A_4) and PRK dimers (Scagliarini *et al.,* 1993; Baalmann *et al.,* 1994; Baalmann *et al.,* 1995; Scheibe *et al.,* 2002; Graciet *et al.,* 2004*a;* Tamoi *et al.,* 2005; Trost *et al.,* 2006).

The acronym CP12 refer to small proteins of nearly 80 amino acids, widespread in oxygenic photosynthetic organisms (Pohlmeyer *et al.,* 1996; Petersen *et al.,* 2006), apparently lacking an ordered structure in solution (Graciet *et al.,* 2003*b*). Interestingly, the C-terminal half of CP12 is closely related to the C-terminal extension (CTE) of GAPDH B-subunits (Pohlmeyer *et al.,* 1996). Both the CTE and the C-terminal part of CP12 bear a couple of conserved cysteines which can form an intramolecular disulfide and are potential targets of thioredoxin regulation (Baalmann *et al.,* 1996; Wedel and Soll, 1998; Sparla *et al.,* 2002; Graciet *et al.,* 2004*a*; Trost *et al.,* 2006). In addition, most CP12 proteins contain a second couple of conserved cysteines in the N-terminal half of the molecule also able to form a disulfide bond (Wedel *et al.,* 1997; Petersen *et al.,* 2006). However, as long as redox properties of CP12 disulfide bonds are not known it is difficult to predict whether CP12 could exist in different redox states *in vivo*. Studies on the unicellular green algae *Chlamydomonas reinhardtii* (Graciet *et al.,* 2003*b*) and the flowering plant *Arabidopsis thaliana* (Marri *et al.,* 2005*b*) concurred in defining a sequence of events involved in CP12-mediated supramolecular complex formation: A_4-GAPDH (complexed with NAD) first interacts with oxidized CP12, then oxidized PRK can participate in the assembly of a ternary complex with still uncertain stoichiometry.

Here we describe molecular and redox properties of *Arabidopsis thaliana* CP12-2, the product of one of the three CP12 genes known for this species (Marri *et al.,* 2005*a*) and report thermodynamic parameters and stoichiometry of the assembly of supramolecular complexes with A_4-GAPDH and PRK. The results further support the view that formation of supramolecular complexes is a means for photosynthetic organisms to safely and reversibly store photosynthetic enzymes in inactive conformation during the night.

4.3 *Experimental procedures*

4.3.1 *Protein expression and purification*

Heterologous expression and purification of recombinant A_4-GAPDH (At3g23650), PRK (At1g32060), CP12-2 (At3g62410) and CP12-2 site-specific mutants of *Arabidopsis thaliana* was performed as described (Marri *et al.,* 2005*b*). NMR analysis was performed on uniformly ^{15}N-labelled His-tagged CP12-2 samples obtained by transformed *E. coli* BL21(DE3) cells grown in M9 minimal medium containing 1 g L^{-1} of $^{15}NH_4Cl$ (Euriso-top) as the sole nitrogen source. An overnight culture of 25 mL in M9 medium was transferred to fresh 500 mL of M9 medium, both supplied with kanamycin (50 µg/mL) and grown at 37°C under shaking. When optical density at 600 nm reached 0.6-1.0 units, expression was induced by addition of 0.4 mM isopropyl-β-D-thiogalactopyranoside. About 15 hours after induction, cells were collected by centrifugation (10.000 rpm; 15 min). CP12-2 was purified from the resulting pellet as previously described (Marri *et al.,* 2005*b*). Purified proteins were quantified by absorbance at 280 nm (Marri *et al.,* 2005*b*), desalted in appropriate buffers and stored at -20°C. Concentrations of purified proteins are all referred to native conformations (CP12-2 monomers, A_4- GAPDH tetramers, PRK dimers).

4.3.2 *CP12-2 site specific mutants*

Site-specific mutants of recombinant CP12-2 were obtained as described in ref Sparla *et al.,* 2002. PCR primers were as follows: C22S(up): 5'-AAGCTCAGGAGACTTCTGCGGGCGATCC-3'; C22S(down): 5'-ATCGCCCGCAGAAGTCTCCTGAGCTTCC-3'; C73S(up): 5'-TTGTCGTAAGTACGGGACTCGTTGGTCTCAGG-3'; C73S(down): 5'-ACAATCCTGAGACCAACGAGTCCCGTACTTACG-3'.

The presence of mutations was confirmed by DNA sequence analysis.

4.3.3 *NMR spectra*

Uniformly ^{15}N-labelled CP12-2, provided with His-tag, was oxidized by addition of 20 mM oxidized DTT (Sigma). After 16-18 h of incubation at 4°C, the sample was desalted in 25 mM potassium phosphate buffer, pH 7.0 and concentrated. NMR samples were typically 300 or 600 µL of 1 mM CP12-2 solution in 25 mM potassium phosphate buffer, pH 7.0; 5% (v/v) $^{2}H_2O$; 0.05% Na-azide.

2D ^{1}H-^{15}N HSQC (Bodenhausen and Ruben, 1980) spectra were recorded with 256 and 2048 complex points in F1 and F2 dimensions, respectively, at 20°C on a Bruker AvanceII 800 MHz spectrometer equipped with a triple-resonance (^{1}H, ^{13}C, ^{15}N) probe, including field xyz-gradients.

Spectra were processed using Topspin™ version 1.3 (Bruker). Chemical shifts were referenced to internal d_4-TSPA, according to ref Wishart *et al.,* 1995.

4.3.4 *Analysis of thiol groups and redox titration of CP12-2*

Free thiol content and redox titration were performed with pure CP12-2, desalted in 100 mM Tricine-NaOH, pH 7.9. Redox titration experiments were performed with 70 µM CP12-2 incubated for 3 hours at 25°C with variable ratios of reduced and oxidized DTT (20 mM total concentration) in a final volume of 500µL. After incubation, samples were desalted by PD10 columns (GEhealthcare) equilibrated with 100 mM Tricine-NaOH, pH 7.9. In order to avoid any possible DTT contamination, only the first 2 mL of eluted samples were collected. Control experiments were performed under same conditions but in the absence of CP12-2. Absorbance at 280 and 412 nm was recorded immediately before and after the addition of 0.5 mM 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB). The number of solvent-accessible thiol groups

under different redox conditions were calculated from the ratio between the absorbance at 412 nm (molar extinction coefficient of 14,150 M^{-1} for 2-nitro-5-thiobenzoate (thiolate) dianion (TNB); Conway *et al.,* 2004) and the absorbance at 280 nm (molar extinction coefficient of 8,370 M^{-1} for CP12-2).

Redox titration results were fit by non linear regression (CoStat, CoHort Software) to the Nernst equation for two redox components (Sparla *et al.,* 2002). Midpoint redox potential are reported as average values ± standard deviations of triplicate experiments.

4.3.5 *Isothermal titration calorimetry (ITC)*

Calorimetric measurements were carried out using a VP-ITC MicroCalorimeter (MicroCal Inc., Northampton, MA). Each experiment was performed at a constant temperature of 30°C and consisted of 25 injections of 10 µL aliquots, repeated every 200 sec. All samples were degassed by stirring under vacuum before use. Heat of dilution, measured by control experiments in which samples were injected into a buffer-filled cell, was subtracted. Signals recorded in each experiments were integrated using OriginPro 7.5 software supplied with the instrument. The thermodynamic binding parameters (dissociation constant, K_D; enthalpy change, ΔH; Gibbs energy change, ΔG; entropy change, ΔS; and the number of binding sites, n) were obtained by non-linear regression of the integrated heat plots, using the "One set of sites" model of the Origin software.

Calorimetric titrations of binary complex formation were carried out with 15 µM oxidized CP12-2 in 25 mM potassium phosphate, 0.2 mM NAD, pH 7.5 in both sample and reference cells, while the syringe was filled with 52 µM A_4-GAPDH in the same buffer. Although CP12-2 behaved as a ligand it was not filled in the syringe because of its very high heat of dilution. The presence of ligand (CP12-2) in the cell and macromolecule (A_4-GAPDH) in the syringe was taken into account in the elaboration of primary data.

Calorimetric titrations of ternary complex formation were performed with 5 µM preformed binary complex in 25 mM potassium phosphate, 0.2 mM NAD, pH 7.5 in sample and reference cells, while the syringe was filled with 70 µM PRK dissolved in the same buffer.

Thermodynamic parameters of binary and ternary complex formation are reported as average values ± standard deviations of triplicate experiments

4.3.6 *Dynamic light scattering (MALS-QELS)*

Purified single proteins, and pre-formed binary and ternary complexes (Marri *et al.,* 2005*b*) were analyzed by size exclusion chromatography connected to a multiangle light scattering equipped with QELS module (quasi elastic light scattering) for Rh measurements. Hundred-µL samples were loaded on a Superdex 200HR column (GEhealthcare) equilibrated in 25 mM potassium phosphate, pH 7.5, 1 mM EDTA, 150 mM KCl, with 0.2 mM NAD (for A_4-GAPDH, binary and ternary complexes) or without NAD (for PRK and CP12-2). A constant flow rate of 0.6 mL min^{-1} was applied. Elution profiles were detected by an Optilab rEX interferometric refractometer and a Dawn EOS multiangle laser light scattering system at 690 nm (Wyatt Technology Corp., USA). Data acquisition and processing were carried out using ASTRA 5.1.9.1 software (Wyatt Technology Corp., USA).

4.4 Results

4.4.1 *CP12-2 of Arabidopsis thaliana is an intrinsically unstructured protein*

NMR analysis of CP12-2 in the oxidized state revealed that most of the amide proton resonances are localized between 8.5 and 8.0 ppm, the so-called random-coil region, strongly suggesting that CP12-2 is mainly

unstructured (fig. 4.1). Only few residues, belonging to the C-terminus, exhibit amide chemical shifts outside the random-coil range, indicating structuration (Yao *et al.,* 1997). Consistent with the structural relevance of disulfide bridges, the reduction of oxidized CP12-2 by dithiotreitol (DTT), led to typical random-coil signals (not shown), as previously described for *Chlamydomonas* CP12 (Graciet *et al.,* 2003*b*).

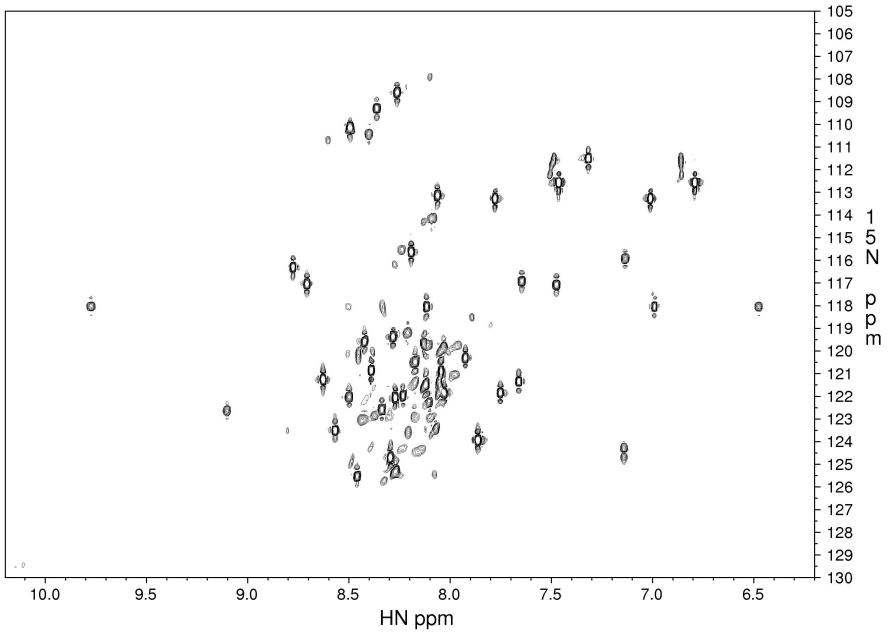


Figure 4.1 2D-[^{1}H-^{15}N]-HSQC spectra of the His-tagged oxidized Arabidopsis thaliana CP12-2 recorded at ^{1}H=800 MHz and 20 °C with 16 transients by t1 increment.

Oxidized CP12-2 behaves as a protein of 29 kDa in size exclusion chromatography (Marri *et al.,* 2005*b*). Although the calculated molecular mass of recombinant CP12-2 (after proteolytic cleavage of the His-tag) was 8.7 kDa, the higher apparent molecular mass of native CP12-2 was not due to polymerization. In fact multiangle light scattering (MALS-QELS) analysis of the protein eluted from the size exclusion chromatographic column estimated a molecular mass of 9±1 kDa, conclusively demonstrating that

under native conditions CP12-2 is a monomer (Table 4.1) and suggesting that the estimate of the molecular mass by size exclusion chromatography (29 kDa) was strongly biased by its unstructured nature.

4.4.2 *CP12-2 redox properties*

To get insights into the redox properties of CP12-2, redox titrations were performed in the presence of DTNB as a probe to reveal free protein thiols under varying redox conditions. Fully reduced and fully oxidized samples were obtained after equilibration with 20 mM reduced or oxidized DTT, which was then removed by desalting to prevent interactions with DTNB. While reduced CP12-2, whose amino acid sequence includes four Cys, was experimentally found to contain four reactive thiols (4.5±0.5), oxidized CP12-2 had none (-0.3±0.1), indicating that both CP12-2 disulfides could be redox titrated by DTT plus DTNB. Data from redox titrations of pure CP12-2 were therefore fitted to a Nernst equation for two different thiol/disulfide equilibria equally contributing to the total redox response. At pH 7.9, the midpoint redox potentials ($E_{m,7.9}$) of CP12-2 disulfides were -326±2 and -352±6 mV, respectively (fig. 4.2).

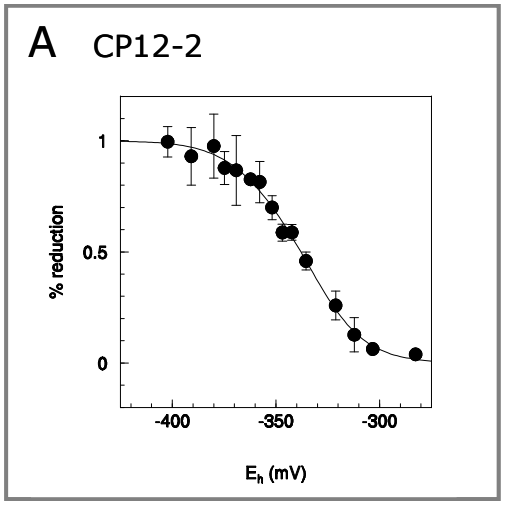


Figure 4.2 Redox titration of CP12-2. Free protein thiols were quantified by DTNB. The number of reacting thiols was 4.5 ± 0.5 for fully reduced CP12-2 (after 3h incubation with 20 mM reduced DTT) and -0.3 ± 0.1 for fully oxidized CP12-2 (3 h incubation with 20 mM oxidized DTT; means ± SD of three experiments). Results were fit by non linear regression to the Nernst equation for two redox components (Spa02). Data points are means of triplicate determinations ± SD.

4.4.3 *Interaction between A_4-GAPDH, CP12-2 and PRK*

Size exclusion chromatography, carried out in the presence of NAD, combined with MALS-QELS analysis, demonstrated that a stable binary complex of 170±14 kDa was reconstituted by incubating oxidized CP12-2 (9±1 kDa, MALS) with A_4-GAPDH complexed with NAD (146±2 kDa, MALS; tab. 4.1). These data suggest that two CP12-2 molecules could bind one A_4-GAPDH tetramer giving rise to an (A_4-GAPDH)-(CP12-2)$_2$ complex (calculated M_r 166 kDa, tab. 1).

Table 4.1 Determination of relative molecular masses (M_r) and hydrodynamic radii (R_H) of purified single proteins and reconstituted complexes by dynamic light scattering (MALS-QELS). Samples were loaded on a Superdex 200 column connected on line to a MALS detector equipped with a QELS module. Data shown are mean values ± standard deviations of triplicate runs.

Sample	Calculated		Measured	
	M_r (kDa)	R_H (nm)	MW (kDa)	R_H (nm)
CP12-2	8.7	1.63	9±1	*nd*[1]
A_4-GAPDH	148.8[2]	4.21	146±2	3.9±0.3
PRK	77.6	3.39	85±7	3.3±0.3
(A_4-GAPDH)+(CP12-2)	166.2[2,3]	4.36	170±14	4.3±0.5
(A_4-GAPDH)+(CP12-2)+(PRK)	487.7[2,4]	6.25	498±6	7.0±0.1

[1] the R_H of CP12-2 was close to the lower detection limit of the QELS module (about 1 nm)
[2] since chromatographic runs were performed in the presence of 0.2 mM NAD, relative masses M_r were calculated on the assumption that each A_4-GAPDH tetramer bound four NAD molecules (Fal03).
[3] calculated M_r for a binary complex with stoichiometry (A_4-GAPDH)-(CP12-2)$_2$
[4] calculated M_r for a ternary complex with stoichiometry [(A_4-GAPDH)-(CP12-2)$_2$-(PRK)]$_2$

The affinity between A_4-GAPDH and CP12-2 was analyzed by isothermal titration calorimetry (ITC) (fig. 4.3). Such interaction was characterized by a K_D of 0.18±0.02 µM and two binding sites for CP12-2 were detected per A_4-GAPDH tetramer (n=1.9±0.2) with no evidence for binding sites with

different affinity. The binding of two CP12-2 molecules to each A_4-GAPDH tetramer was exothermic (ΔH -15 kcal mol^{-1}) and though leading to a simultaneous decrease in entropy (TΔS -5 kcal mol^{-1}), the process was overall highly exoergonic (ΔG -9.4 kcal mol^{-1}, tab. 4.2).

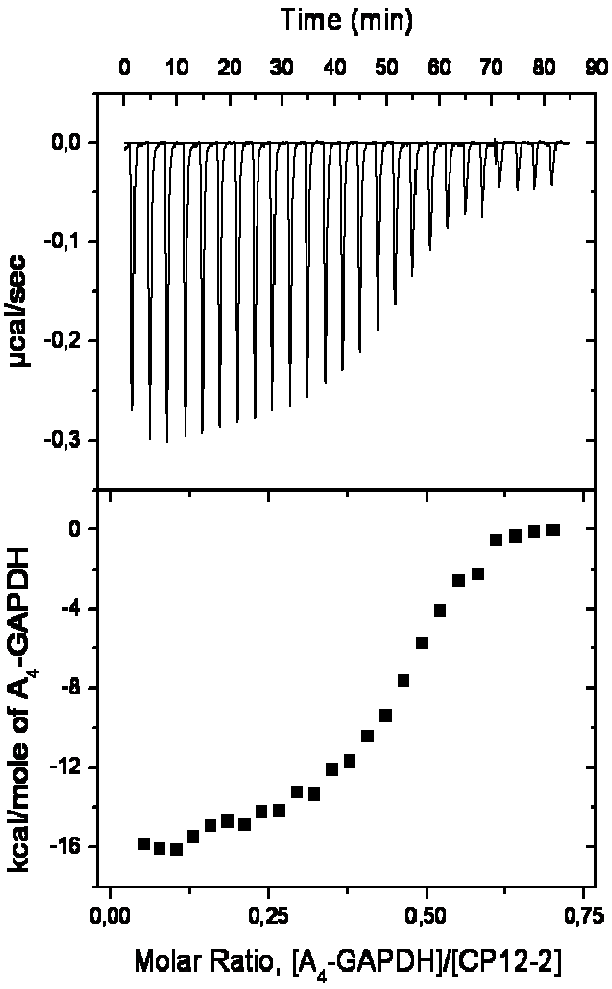


Figure 4.3 Calorimetric titration of CP12-2 with A_4-GAPDH at 30°C. A solution of 52 µM A_4-GAPDH in 25 mM potassium phosphate, 0.2 mM NAD, pH 7.5 was injected into the sample cell containing 15 µM oxidized CP12-2 in the same buffer. The reference cell was filled exactly like the sample cell. The experiment consisted of 25 injection of 10 µL each with 200 sec intervals between subsequent injections. (A) Heat effects recorded as a function of time during successive injections. The heat of dilution, comparable to the last peaks of the titration, was independently measured and subtracted for calculations. (B) Enthalpy per mole of A_4-GAPDH injected versus A_4-GAPDH/CP12-2 molar ratio in the sample cell. The thermodynamic binding parameters were obtained by non-linear regression of the integrated heat plots as in panel B, using the "One set of sites" model of the Origin software.

As previously demonstrated (Marri *et al.,* 2005*b*), PRK in the oxidized state can bind preformed (A_4-GAPDH)-$(CP12\text{-}2)_2$ binary complexes. Here we show that binding of PRK (85±7 kDa, MALS) to the binary complex gave rise to a ternary complex with a molecular mass of 498±6 kDa (MALS-QELS) suggesting a composition of two dimers of PRK and two (A_4-GAPDH)-$(CP12\text{-}2)_2$ binary complexes (calculated M_r: 488 kDa, tab. 1). The stoichiometric

ratio of two A_4-GAPDH subunits per PRK subunit was confirmed by densitometric analysis of Coomassie-stained bands separated by denaturating gel-electrophoresis (the ratio of A_4-GAPDH monomers to PRK monomers was 2.3, not shown).

Thermodynamic parameters of the interaction between PRK and (A_4-GAPDH)-(CP12-2)$_2$ were also determined by ITC (fig. 4.4). As for experiments of binary complex formation, data were fitted to a simple model describing the interaction on *n* "ligands" (PRK) to one "macromolecule" (A_4-GAPDH)-(CP12-2)$_2$ with *n* identical binding sites.

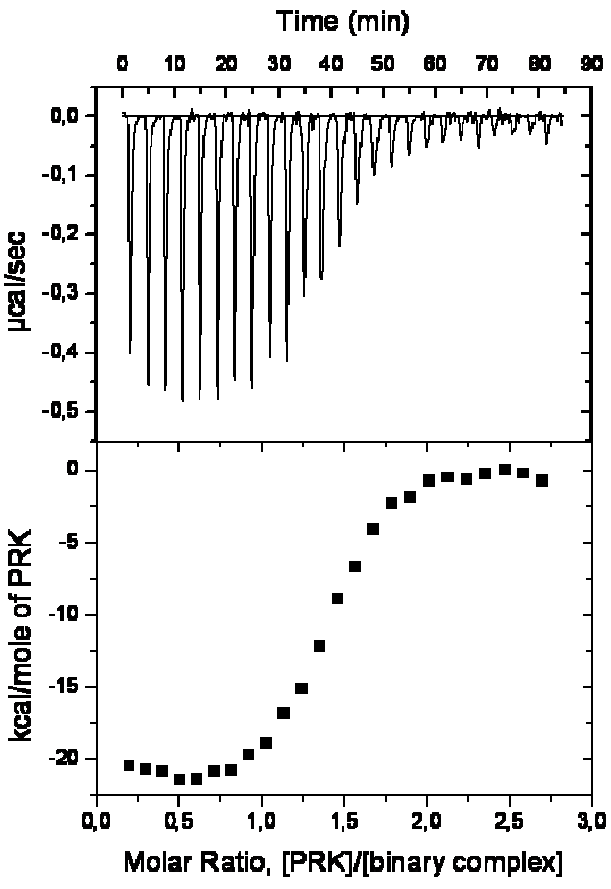


Figure 4.4 Calorimetric titration of (A_4-GAPDH)-(CP12-2)$_2$ binary complex with PRK at 30°C. A solution 70 µM PRK dissolved in 25 mM potassium phosphate, 0.2 mM NAD, pH 7.5 was injected into the sample cell containing 5 µM preformed binary complex in the same buffer. The reference cell contained the same solution as the sample cell. The experiment was conducted under same conditions as in Fig. 3. (A) Heat effects recorded as a function of time during successive injections. The heat of dilution, comparable to the last peaks of the titration, was independently measured and subtracted for calculations. (B) Enthalpy per mole of PRK injected versus (PRK)/[(A_4-GAPDH)-(CP12-2)$_2$] molar ratio in the sample cell. The thermodynamic binding parameters were obtained by non-linear regression of the integrated heat plots as in panel B, using the "One set of sites" model of the Origin software.

With this model, a K_D of 0.17±0.09 μM with n=1.3±0.1 was estimated by non linear regression analysis. Thermodynamic parameters indicated that ternary complex assembly was exothermic (ΔH -20 kcal mol^{-1}) and exoergonic (ΔG -9.3 kcal mol^{-1}), though penalized by a significant decrease in entropy (TΔS -11 kcal mol^{-1}, tab. 4.2).

Table 4.2 Thermodynamic parameters of binary and ternary complexes formation/dissociation at 30°C, as determined by isothermal titration calorimetry.

Complex	*n*	K_D (μM)	ΔH^2 (kcal mol^{-1})	$T\Delta S^2$ (kcal mol^{-1})	ΔG^2 (kcal mol^{-1})
(A_4-GAPDH)+CP12-2	1.9±0.2[1]	0.18±0.02	-15±2	-5±2	-9.4±0.1
(A_4-GAPDH)-($CP12\text{-}2)_2$ + PRK	1.3±0.1[3]	0.17±0.09	-20±4	-11±3	-9.3±0.2

[1] CP12-2 molecules per A_4-GAPDH.
[2] Thermodynamic parameters of association referred to moles of A_4-GAPDH (first line) or (A_4-GAPDH)-($CP12\text{-}2)_2$ (second line).
[3] PRK dimers per (A_4-GAPDH)-($CP12\text{-}2)_2$ binary complexes.

4.4.4 *CP12-2 site-specific mutants*

CP12-2 was shown to bind A_4-GAPDH only under oxidizing conditions (Wedel and Soll, 1998; Graciet *et al.,* 2003*b*; Marri *et al.,* 2005*b*). In order to verify the individual role of each CP12-2 disulfide in stabilizing the optimal conformation for the interaction with A_4-GAPDH and PRK, two site-specific mutants have been made in which either the first (Cys22) or the last (Cys73) cysteine of CP12-2 were mutated into serine. Under reducing conditions, both mutants (C22S, C73S) behaved in size exclusion chromatography as wild type CP12-2. Under oxidizing conditions both mutants tended to form dimers (not shown), possibly mediated by

intermolecular disulfides between cysteines not engaged in intramolecular disulfides (Cys31 in C22S, and Cys64 in C73S).

Mutant C22S interacted with A_4-GAPDH similarly to wild type CP12-2 (Fig. 5). On the contrary, mutant C73S did not form any complex with A_4-GAPDH (Fig. 5). In the presence of PRK, no CP12-2 mutant was able to promote the formation of a ternary complex analogous to the 498 kDa complex observed with wild type CP12-2.

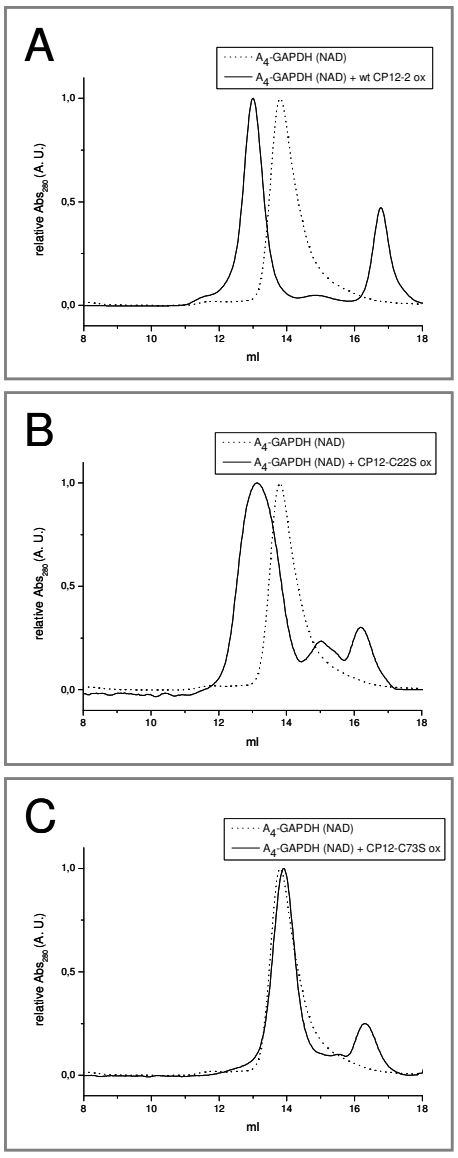


Figure 4.5 In vitro reconstitution of (A_4-GAPDH)/(CP12-2) binary complexes with wild type and site-specific mutants of CP12-2. Overlapped elution profiles after size-exclusion chromatography with a Superdex 200 column equilibrated with 50 mM Tris-HCl, pH 7.5, 150 mM KCl, 1 mM EDTA. Volume of loaded samples was 0.2 mL and flow rate was 0.5 mL/min. Concentrations of A_4-GAPDH and CP12-2 in each loaded sample were equimolar on a subunit basis. (A) Samples loaded on the column contained either A_4-GAPDH alone (dotted line) or A_4-GAPDH plus wild type CP12-2 in the oxidized state (full line). The elution peak at about 13 ml contained (A_4-GAPDH)-$(CP12\text{-}2)_2$ binary complex while the peak at about 17 ml contained free monomeric CP12-2 (see also Marri et al., 2005). (B) Same conditions as in panel A except that wt CP12-2 was substituted by site specific mutant C22S. The elution profile was similar as in panel A. (C) Same conditions as in panel B, except that CP12-2 site specific mutant was C73S. The binary complex (13 ml) does not form.

4.5 Discussion

Solution NMR studies show that CP12 of *Arabidopsis thaliana* (CP12-2) is a poorly structured protein even under oxidizing conditions, when it bears two interchain disulfide bridges (fig. 1). As expected, reduction of disulfides led to increased disorder. The amino acid sequence of CP12-2 includes signatures of potential intrinsic disorder such as low content of bulky hydrophobic amino acids (*e.g.* Tyr, Phe, Leu, Ile) and relative abundance of charged residues (*e.g.* Glu) (Marri *et al.,* 2005*b*; Dyson and Wright, 2005; Trost *et al.,*2006). Several bioinformatic predictors, including PONDR (Romero *et al*., 2001), GlobPlot (Linding *et al*., 2003*b*) and IUPred (Dosztanyi *et al.,* 2005), agree in predicting CP12-2 as poorly structured. As a whole, these findings strongly point to CP12-2 as an intrinsically unstructured protein (IUP, Dyson and Wright, 2005), similar in this respect to CP12 of *Chlamydomonas reinhardtii*, recently shown to be very flexible even in the oxidized state (Graciet *et al.,* 2003*b*; Gardebien *et al.,* 2006). Consistent with this disordered nature, *Arabidopsis* CP12-2 did not associate into polymers in the absence of binding partners. Dynamic light scattering analysis (MALS-QELS) showed indeed that CP12-2 was a monomer of 9 kDa under native conditions and higher estimates of CP12 molecular mass based on size exclusion chromatography (29-70 kDa; Wedel and Soll, 1998; Marri *et al.,* 2005*b*; Tamoi *et al.,* 2005) were thus possibly biased by the non-globular and disordered nature of CP12. In the light of current evidence, conjectures about CP12 capability to form dimers (Wedel and Soll, 1998; Gardebien *et al.,* 2006) seem therefore unjustified.

The formation of the binary complex between A_4-GAPDH and CP12-2 was found to be exoergonic (ΔG -9 kcal mol^{-1}) but associated to an entropic penalty ($T\Delta S$ -5 kcal mol^{-1}), indicating that complex formation resulted in a general decrease of disorder, an effect which might be attributable to CP12-2. Coupled folding and binding to partner proteins is a typical, physiological important process of intrinsically unstructured proteins and is generally associated to an entropic cost (Dyson and Wright, 2005). It seems

likely that in spite of the disordered structure of CP12-2 in solution, it might fold, partially at least, upon binding to its partner protein A_4-GAPDH thus leading to a decrease in entropy of the whole system (Perozzo *et al.,* 2004).

The stoichiometry of the binary complex A_4-GAPDH/CP12-2 was assessed by a combined approach based on isothermal titration calorimetry (ITC) and dynamic light scattering (MALS-QELS) analysis. ITC experiments clearly indicated that up to two CP12-2 molecules (n=1.9±0.1) could bind each A_4-GAPDH tetramer and this result was confirmed by MALS-QELS determination of the molecular mass of the binary complex (experimental 170±14 kDa, calculated 166 kDa). The stoichiometry of the binary complex seems consistent with the homology between the C-terminal part of CP12 and the C-terminal extension (CTE) of subunits B of A_nB_n-GAPDH isoforms. The recently solved structure of spinach A_2B_2-GAPDH in the oxidized state shows that each CTE, partially structured by a disulfide bridge, is allocated in a wide cleft delimited a pair of A/B-subunits (Fermani *et al.,* unpublished). In general, GAPDH tetramers (either A_2B_2, Fermani *et al.,* unpublished, or A_4, Fermani *et al.,* 2001) contain two symmetrical clefts of this type and it is tempting to speculate that CP12-2 may bind to A_4-GAPDH in a similar manner as CTE is allocated in A_2B_2-GAPDH. This conclusion is also supported by site-specific mutant C73S of CP12-2 (and similar site-specific mutants of CP12 from other species; Wedel and Soll, 1998; Tamoi *et al.,* 2005; Lebreton *et al.,* 2006) which could not form the C-terminal disulfide and was consequently unable to form a stable complex with GAPDH. This result is reminiscent of the impaired redox regulation in the analogous CTE-mutant (C349S) of B_4-GAPDH (Sparla *et al.,* 2002).

The affinity of *Arabidopsis* CP12-2 for A_4-GAPDH was in the submicromolar range (K_D 0.18 µM, ITC), i.e. 450-fold lower than in *Chlamydomonas* (K_D 0.4 nM, as measured by surface plasmon resonance, Graciet *et al.,* 2003*b*) and no other studies can be referred to for comparison. The high stability of the complex in *Chlamydomonas* might be related to the inhibitory effect

that CP12 exerts on GAPDH activity in this species (Graciet *et al.,* 2003*a*; Lebreton *et al.,* 2006). On the contrary, no analogous inhibition was observed in *Arabidopsis* as long as PRK is also recruited (Marri *et al.,* 2005*b*). It should be noted that A_4-GAPDH is the only Calvin cycle's GAPDH isoform in green algae like *Chlamydomonas* (Petersen *et al.,* 2006) and it seems reasonable that CP12 may directly regulate GAPDH activity in this species. Quite differently, higher photosynthetic organisms also contain an autonomously-regulated AB-GAPDH isoform (Trost *et al.,* 2006), and the essential role of CP12 in higher plants seems rather to promote the interaction between A_4-GAPDH and PRK, thereby leading to the down-regulation of both enzyme activities in a tightly coordinated manner (Marri *et al.,* 2005*b*).

Although a ternary complex of uncertain molecular mass (460-640 kDa) including GAPDH, CP12 and PRK was described in different photosynthetic organisms (Wedel *et al.,* 1997; Wedel and Soll, 1998; Scheibe *et al.,* 2002; Mouche *et al.,* 2002; Graciet *et al.,* 2004*a*; Tamoi *et al.,* 2005; Marri *et al.,* 2005*b*) its exact composition remained elusive. Here we show that MALS-QELS determinations of ternary complex molecular mass (498±6 kDa) are fully consistent with a composition including two binary complexes (A_4-GAPDH)-$(CP12\text{-}2)_2$ and two PRK dimers, which corresponds to a calculated molecular mass of 488 kDa (table 4.1). No other possible stoichiometries were so nicely supported by the light scattering data. Electrophoretic analysis also indicated a stoichiometry ratio of two GAPDH subunits per PRK subunit within the complex.

Based on the symmetry of both (A_4-GAPDH)-$(CP12\text{-}2)_2$ binary complexes and PRK homodimers, the stoichiometry of the ternary complex strongly suggests that each binary complex had two binding sites for PRK, and each PRK dimer had two binding sites for each binary complex. As a result, a toric supramolecular complex would finally be assembled, as depicted in fig. 4.6. This conclusion substitutes an alternative stoichiometry with only two CP12 molecules (instead of four) being part of the complex which was early

suggested and widely accepted on the basis of the purported capability of CP12 to dimerize (Wedel and Soll, 1998; Graciet *et al.,* 2003*b*; Graciet *et al.,* 2004*a*; Lebreton *et al.,* 2006).

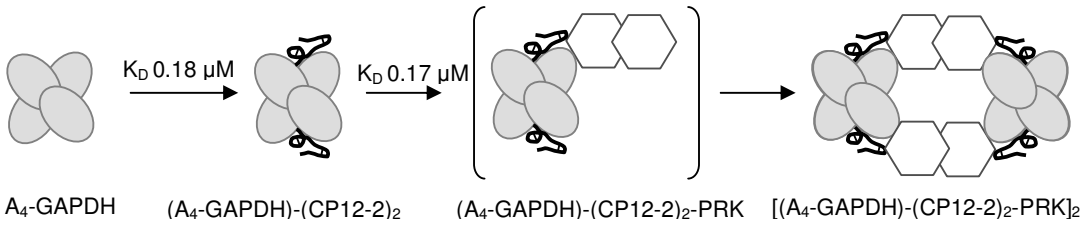


Figure 4.6 A schematic diagram of binary and ternary complex formation compatible with current evidencies. A_4-GAPDH is represented in gray, CP12-2 in black and PRK in white.

ITC analysis of ternary complex formation was also in line with our conclusion. By fitting ITC data with a simple model equation in which only one type of interaction between binary complex and PRK was postulated, an estimate of one PRK dimer (n=1.3±0.1) binding one binary complex was obtained. It should be noted however that ITC experiments could not discriminate between the interaction of PRK with binary complex (second step in fig. 4.6) and the following dimerization of the transient complex (last step in fig. 4.6). The thermodynamic parameters obtained (tab. 4.2) are therefore to be referred to the whole process of ternary complex assembly.

Arabidopsis CP12-2 contains 4 cysteines in its sequence and under oxidizing conditions two internal disulfides are formed (see also Wedel and Soll, 1998 and Graciet *et al.,* 2003*b*). In analogy with AB-GAPDH it is generally accepted that CP12 forms a C-terminal disulfide (Cys64-Cys73, homologous to Cys349-Cys358 regulatory disulfide of A_2B_2-GAPDH, Sparla *et al.,* 2002) and an N-terminal disulfide Cys22-Cys31 (Pohlmeyer *et al.,* 1996; Wedel and Soll, 1998; Lebreton *et al.,* 2006). Redox titration analysis of CP12-2 revealed a disulfide bond, likely the C-terminal one, with midpoint redox potential ($E_{m,7.9}$ -352 mV) identical to that of spinach A_2B_2-GAPDH (-353 mV, Sparla *et al.,* 2002). Surprisingly, the $E_{m,7.9}$ of the second disulfide of CP12-2 (-326 mV) was otherwise identical to that of spinach PRK (-330 mV, Marri *et al.,* 2005*b*), though no sequence similarities are displayed by the two proteins.

The N-terminal disulfide was not required for binary complex formation with GAPDH (see also Wedel and Soll, 1998; Lebreton *et al.,* 2006), but without this disulfide the ternary complex of 498 kDa did not form (see also Wedel and Soll, 1998). Since PRK does not bind CP12-2 alone, it is likely that the N-terminal disulfide of CP12-2 forms together with A_4-GAPDH a binding site for PRK. However, this conclusion should not be generalized to all oxygenic phototrophs since in the cyanobacterium *Synechococcus* PC7942 a GAPDH/CP12/PRK ternary complex of about 500 kDa was detected even if CP12 of this species has no N-terminal cysteines (Tamoi *et al.,* 2005).

The different redox properties of the two disulfides of *Arabidopsis* CP12-2 are fully consistent with the sequence of formation of the supramolecular complex (Graciet *et al.,* 2003*a*; Marri *et al.,* 2005*b*) and with a plausible physiological scenario. During the transition from light to dark, the incipient oxidation of the chloroplast thioredoxins ($E_{m,7.9}$ -351 mV for thioredoxin *f*, Collin *et al.,* 2003) would cause in parallel the initial oxidation of C-terminal cysteines of CP12-2 (-352 mV) and partial inactivation of AB-GAPDH (-353 mV, Sparla *et al.,* 2002; Trost *et al.,* 2006); under these conditions, two molecules of partially oxidized CP12-2 would make a complex with A_4-

GAPDH tetramers bound to NAD(H) (Marri *et al.,* 2005*b*). In full darkness, further oxidation of thioredoxins would lead to simultaneous formation of the more positive, the N-terminal disulfide of CP12-2 (-326 mV) and partial inactivation via disulfide formation of PRK (-330 mV). A supramolecular complex of two A_4-GAPDH-CP12-2 binary complexes bound to two PRK dimers could then form (fig. 4.6), providing a means for safely and reversibly store two Calvin cycle enzymes in a strongly inhibited conformation during the dark period.

Chapter 5

The CP12 protein family in Arabidopsis thaliana

5.1 *Introduction*

CP12 is an intrinsically unstructured protein universally distributed in oxygenic photosynthetic organisms, from cyanobacteria up to higher plants (fig. 5.1). Sequences similar to Arabidopsis protein CP12-2 were identified, assuming a common evolution from lower photosynthetic organisms.

CP12 occurs in many cyanobacteria; it was studied in *Synechocystis* PCC6803 (P73654 in fig. 5.1) suggesting that NADPH-mediated light regulation of Calvin cycle activity occurred already in cyanobacteria (Wedel and Soll, 1998). CP12-mediated regulation of Calvin cycle may be conserved in all photosynthetic organisms. From multiple alignment of CP12s amino acids sequences, a C-terminal pair of conserved cysteine was observed (Pohlmeyer *et al.,* 1996), while the N-terminal cysteines may be missed in some cyanobacteria (as *Synechococcus* PCC7942, *Synechococcus* PCC6301), in rhodophytes, or in *Cyanophora paradoxa (*Petersen *et al.*, 2006). Lower photosynthetic organisms, as red algae (*Galdieria sulphuraria*), green algae (*Chlamydomonas reinhardtii* and *Volvox carteri)* and bryophyta (*Ceratodon purpureus)* generally contain a single copy gene of CP12. In land plants, CP12 is whole distributed in single copy in the genome of tobacco (*Nicotiana tabacum),* spinach (*Spinacea oleracea*), pea (*Pisum sativum*), peanut (*Arachis hypogea),* field mustard (*Brassica campestris),* artichoke (*Cynara scolymus).*

On the other hand, rice (*Oryza sativa*) and *Arabidopsis thaliana* seems to be the unique species in which CP12 proteins are coded by small gene families (three members in each, fig. 5.1).

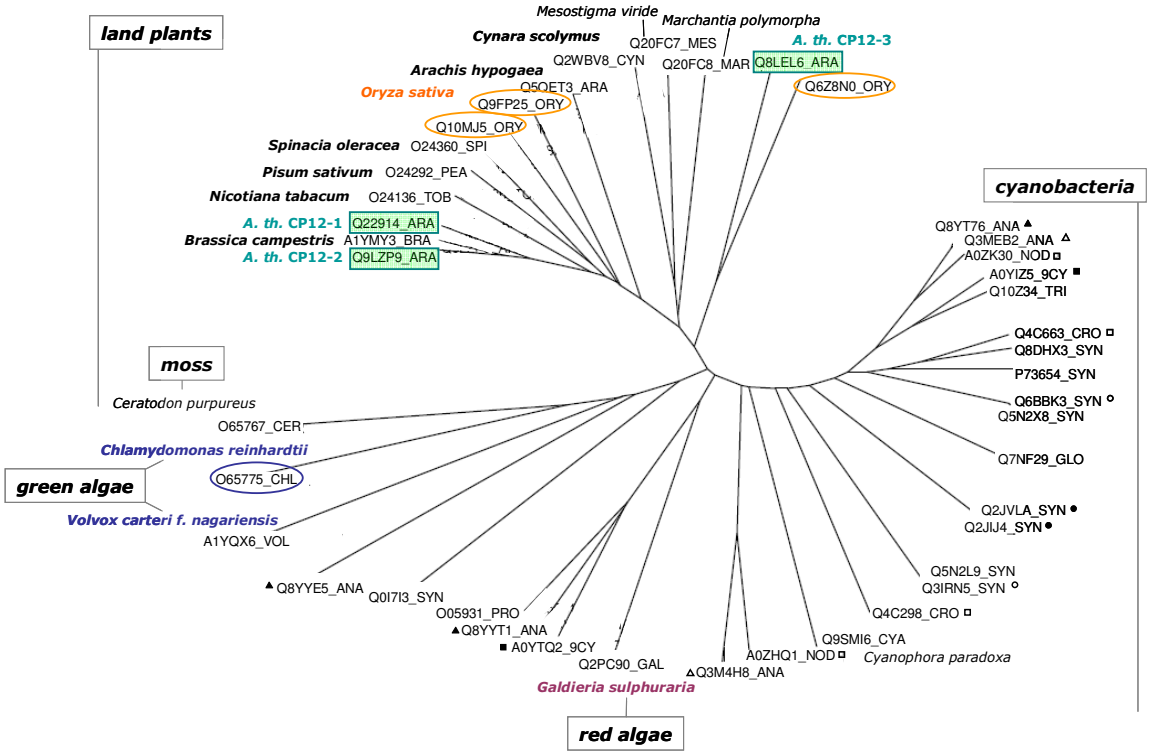


Fig. 5.1 Phylogenetic distribution of CP12 sequence. BlastP search (NCBI BLAST on ExPASy) was obtained using *Arabidopsis thaliana* CP12-2 (At3g62410) protein sequence. Tree was created using TreeTop- Phylogenetic Tree prediction of the GeneBee service and Phylodendron of the IUBio Archive at the Indiana University. (http://www.genebee.msu.su/services/phtree_reduced.html, http://iubio.bio.indiana.edu/treeapp, http://www.expasy.org/tools/blast). Arabidopsis CP12 (isoforms 1, 2 and 3) were pointed out in green boxes. O*ryza sativa* is the only other plant species that contain different CP12 isoforms (in orange). Green algae (*Chlamydomonas reinhardtii* and *Volvox carteri*) are displayed in blu, red algae (*Galdieria sulphuraria*) in magenta. Inside cyanobacteria, CP12 sequences belonging to same species are marked: black triangle for *Anabaena sp.*(strain PCC 7120), white triangle for *Anabaena variabilis* (strain ATCC 29413/ PCC 7937), black circle for *Synechococcus sp.* (bacterium Yellowstone), white circle for *Synechococcus sp.* (strain PCC 7942), black square for *Lyngbya sp.*(PCC 8106), white square for *Crocosphaera watsonii*, grey square for *Nodularia spumigena.*

The three CP12 sequences present in Arabidopsis are predicted to code for three chloroplastic proteins: besides the well studied At3g62410 for CP12-2,

At2g47400 for CP12-1 and At1g76560 for CP12-3 are also part of the Arabidopsis proteome (Marri *et al.,* 2005*a*; Trost *et al.,* 2006).

Fig. 5.2 Multiple alignment of Arabidopsis CP12 proteins using ClustalW at EBI (http://www.ebi.ac.uk/clustalw). Transit peptides were not considered. Conserved cysteines in the sequence were displayed in green boxes. High content of charged aminoacids is shown: acidic amino acids and their amides are bold blue, while basic amino acids are bold orange. Conserved proline are red.

CP12-1 and CP12-2 are 86% identical in their amino acid composition, with similar charge distribution, while their sequence homology to CP12-3 is under 50% (Trost *et al*., 2006).

Since Arabidopsis is one of the few species containing more than a CP12 isoform, it should be assumed that these isoforms might have different functions, or might be implicated in different environmental conditions. Regarding to this, it should be important to characterize them, understanding their nature as hypothetical IUP proteins, and their supposing roles in cellular processes.

5.2 *Materials and methods*

5.2.1 In silico *analysis of CP12 isoforms*

Protein sequences of mature Arabidopsis CP12-1 (At2g47400), CP12-2 (At3g62410) and CP12-3 (At1g76560) were used to determine amino acids frequencies, thus compared to Swiss-Prot and globular proteins presents in databases. The same protein sequences were applied to disorder predictors: PONDR® (http://www.pondr.com, Romero *et al*., 2001), GlobPlot™ (http://wglobplot.embl.de/, Linding *et al*., 2003*b*) and IUPred (http://iupred.enzim.hu/index.html, Dosztanyi *et al.,* 2005).

5.2.2 *Protein expression and purification*

Expression and purification of recombinant A_4-GAPDH (At3g23650), PRK (At1g32060) and CP12-2 (At3g62410) of *Arabidopsis thaliana* was performed as described in Marri *et al*. (2005*b*). cDNAs coding for CP12-1 (At2g47400), and CP12-3 (At1g76560) were transferred into a pET28a(+) expression vector (Novagen) using the following PCR primers: CP12-1 forward (*Nde*I site), 5'-GTGCATATGGCTACATCG-3', reverse (*Hind*III site), 5'-AAGCTTATTTTTTAATTATC-3'; CP12-3 forward (*Nde*I site), 5'-GGACGCATATGAGGGAGG-3', reverse (*Hind*III site), 5'-GAAAGCTTAAATTTTAGTCTTCG-3'.

Specific endonuclease sites used to clone PCR-amplified fragments were underlined. In the CP12 constructs, the cDNA sequence was *in frame* with a His tag and a cleavable thrombin site.

Heterologous expression and purification of CP12-1 and CP12-3 were carried out as CP12-2 (Marri *et al*., 2005*b*). Purified proteins were desalted

in appropriate buffers and stored at -20°C. Protein quantification was performed by absorbance at 280 nm, using molar extinction coefficients at 280 nm derived from the sequence of each monomer (ε_{GapA} = 36,250 M^{-1}, ε_{PRK} = 29,360 M^{-1}, $\varepsilon_{CP12\text{-}1}$ = 8,730 M^{-1}, $\varepsilon_{CP12\text{-}2}$ = 8,730 M^{-1} and $\varepsilon_{CP12\text{-}3}$ = 7,240 M^{-1}).

5.2.3 *Binary and ternary complex reconstitution*

Samples containing different combinations of purified recombinant GAPDH, PRK, and CP12 (isoforms 1, 2 or 3) were analyzed by gel filtration to detect the *in vitro* reconstitution of binary and ternary complexes, as described in Marri *et al.* (2005*b*).

Oxidized and reduced forms of CP12s were obtained by incubation for 3 h at 25°C with 20 mM oxidized DTT or 20 mM reduced DTT, respectively. Gel filtration analysis was performed on a Superdex 200 HR10/30 column connected to an ÅKTA Purifier system (GE Healthcare). The column was equilibrated with 50 mM Tris-HCl, pH 7.5, 150 mM KCl, 1 mM EDTA (plus 2 mM DTT in reduced samples), adding 200 µg of ferritin as internal standard to detect elution volume of oxidized and reduced CP12s. For the binary and ternary complexes reconstitution, purified proteins were incubated for 2 h at 4°C, in presence of 0.2 mM NAD, before loading on size exclusion chromatography. The volume of loaded samples was 0.2 mL, and fractions of 0.35 mL were collected. The column was calibrated as done by Sparla *et al.* (2002).

Fractions obtained from gel filtration columns were concentrated by ultrafiltration (Centricon YM3), and run on denaturing 12.5% acrylamide gels for Western blot analysis as in Marri *et al*. (2005*b*).

5.2.4 *GAPDH and PRK activity assays*

GAPDH and PRK activities were assayed as described by Marri *et al*. (2005*b*). Activities were assayed immediately after the addition of purified proteins for the reconstitution of the binary and ternary complex, and before loading the samples on the gel filtration column to determine inactivation by complex formation.

Fractions eluted from the Superdex 200 column, in the presence of 0.2 mM NAD, were collected and enzyme activities were assayed in each chromatographic peak. For the full recovery of enzymatic activities, GAPDH were assayed after 5 minutes incubation in 43 *μ*M BPGA, while full PRK activity was obtained after 30 minutes incubation in 20mM reduced DTT (Marri *et al*., 2005*b*).

5.2.5 *Redox titration of CP12 isoforms*

Pure CP12 proteins for analysis of thiol groups and redox titrations were desalted in 100mM Tricine-NaOH, pH 7.9. Data points were prepared as in Marri *et al*. (2007) (see paragraph 4.3.2), CP12s were present in the equilibration mixture at a concentration of 70 μM in a final volume of 500 μl. Oxidation/reduction equilibria was obtained using mixtures of oxidized and reduced DTT to define the ambient redox potential ($E_{h,7.9}$) (table 5.1). Nernst equation with an *n* value of 2 according to Hirasawa *et al.,* 2000 was applied,

$$E_{h,7.9} = E_{m,7.9} + 59/n \log [DTT_{ox}]/[DTT_{rd}]$$

where $E_{m,7.9}$ is the midpoint redox potential of the DTT_{rd}/DTT_{ox} mixture at pH 7.9 (-380mV), with a pH dependence of -59 mV/pH unit over the pH range

from 5.5 to 8.2 (corresponding to the uptake of 2 protons/disulfide reduced). Redox equilibration was carried out for 3 hours at 25°C, then samples were desalted by PD10 columns (GEHealthcare) equilibrated with 100mM Tricine-NaOH, pH 7.9. Absorbance at 280 nm and 412 nm was recorded as described in Marri *et al*. (2007) (see paragraph 4.3.2).

DTT_{rd} (mM)	DTT_{ox} (mM)	$E_{h,7.9}$ (mV)
20	---	---
17	3	-402.223
14	6	-390.855
12	8	-385.195
10	10	-380.000
6	14	-369.145
4	16	-362.239
3	17	-357.777
2	18	-351.850
1.4	18.6	-346.860
1	19	-342.277
0.6	19.4	-335.465
0.2	19.8	-321.129
0.1	19.9	-312.184
0.05	19.95	-303.271
0.01	19.99	-282.626
---	20	---

Table 5.1 Oxidized and reduced DTT mixtures (mM) used to define the ambient redox potential ($E_{h,7.9}$) according to the Nernst equation.

Redox titration results were fit by nonlinear regression (CoStat, CoHort Software) using the equation:

$$\% \text{ reduction} = k(10^{[(E_h - E_m ; 7.9)/29.5]}+1)^{-1}+(1-k)\,(10^{[(E_h - E_m ; 7.9)/29.5]}+1)^{-1}$$

CP12s redox responses were calculated as % of reduction, by a ratio between solvent–accessible thiol groups and free thiols in fully reduced conditions (in 20 mM reduced DTT mixture). K is a parameter varying from 0 to 1, the two extreme conditions of the totally reduced free thiols and the fully oxidized disulfides. $E_{m,7.9}$ is the midpoint redox potential. Good fits were obtained with the Nernst equation containing two different thiol/disulfide equilibria contributing to the total redox response.

Contributions of each disulfide bridge is k and (1-k) respectively. In these cases, the two $E_{m,7.9}$ values and the contribution of the two components have been optimized by the software. Midpoint redox potentials are reported as average values ± standard deviations of triplicate experiments.

5.3 *Results*

5.3.1 *CP12 isoforms as hypothetical IUP proteins*

Intrinsically unstructured proteins share some characteristics that make them typical and permit to distinguish these from ordered and well-defined globular proteins. Various factors have been suggested to be important in terms of protein disorder, including flexibility, aromatic content, secondary structure preferences and various scale associated with hydrophobicity. Besides low mean hydrophobicity, high net charge was also suggested to contribute to disorder. Amino acids composition is a crucial point in the determination of disordered behaviours. On this basis, amino acids are catalogued in order promoting and disorder promoting residues.

While Arabidopsis CP12-2, similar to CP12 previously described in *C. reinhardtii,* was classified as a IUP protein (Marri *et al*., 2007, see also paragraph 4.3.1), CP12-1 and CP12-3 were not studied before. For the three Arabidopsis CP12, analysis of the amino acid composition was performed. Percentage of each amino acid present in the CP12 sequences, was compared to mean amino acid frequencies given for globular proteins, Swiss-Prot proteins and for the IUPs (fig. 5.3, Tompa, 2002). Relative to the ordered protein, the intrinsically disordered segments were significantly depleted in Trp, Cys, Phe, His, Ile, Met, Tyr, Leu and Asn, and significantly enriched in Pro, Ser, Gln, Glu and Lys, suggesting that the first set should be called order-promoting and the second set disorder-promoting.

CP12-2 shows composition similar to IUPs for the depletion in Phe, Gly, His, Ile, Leu, Met, Thr and Tyr, and a significant enrichment in Ser, Lys and Glu (fig. 5.3). Consistent with a very high sequence similarity between CP12-1 and CP12-2, also CP12-1 is enriched and depleted in these amino acids in respect to globular proteins (fig. 5.3). If, from this behaviour, CP12-1 should be considered intrinsically disordered, characteristics of CP12-3 are less certain.

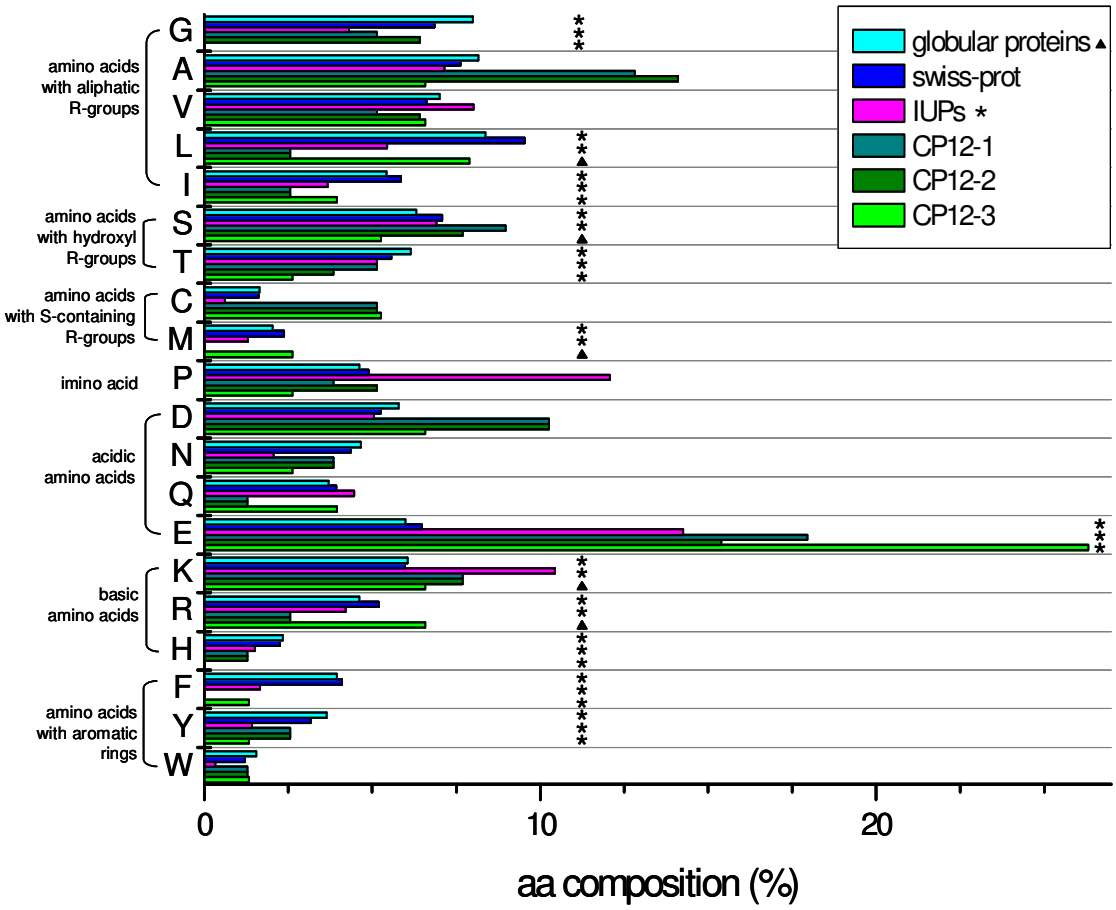


Fig. 5.3 Amino acids frequencies in % of ordered (Swiss-Prot and globular proteins) and disordered (IUPs) proteins, compared to Arabidopsis CP12. Asterics identify frequencies in CP12 that are similar to IUPs (more or less frequent than in an average globular protein in the Protein Data Bank). On the contrary, black triangle indicates where a CP12-3 frequency is similar to globular proteins.
For Swiss-Prot data are from http://us.expasy.org/sprot; for globular proteins data "globular-3D" are from http://disorder.chem.wsu.edu as in Tompa *et al.,* 2002.

As indicated in fig. 5.3, a major content of Glutamic Acid and a minor content of F, G, H, I, T and Y is present in CP12-3, comparing it to unstructured proteins. But surprisingly, CP12-3 moves away from IUPs in the frequencies of L, S, M, K, and R, questioning about its ordered or disordered structure (fig. 5.3).

For this purpose, CP12 sequences were investigated with computational approaches able to predict disorder. Different disorder predictors were developed in recent years, using the differences in compositional complexity to discriminate between globular and non-globular regions (Radivojac *et al.,* 2006) (table 1.1). Three of them: PONDR® (Romero *et al*., 2001), GlobPlot™ (Linding *et al*., 2003*b*) and IUPred (Dosztanyi *et al.,* 2005) were applied to predict disorder in CP12 isoforms (fig. 5.4).

PONDR® was the first predictor to be constructed. It is a two-layer feed-forward neural network based on amino acid composition and physiochemical properties, that achieved a surprising accuracy of about 70%. In the initial studies, disorder was partitioned according to length, with the development of different predictors for short, medium, and long disordered regions (PONDR VL1). Later, disorder was partitioned according to position, with the development of different predictors for N-terminal, internal, and C-terminal regions. To enable prediction from the first to the last residue in a protein, the PONDR VL1 and the predictors for the N- and C-terminal regions were integrated with overlapping predictions. This integrated predictor is herein called PONDR VL-XT (Romero *et al*., 2001). Using this, the three CP12 isoforms were all predicted to be unstructured, with a PONDR score > 0.5 in almost all the proteins length (fig. 5.4).

Besides this, also GlobPlot™ and IUPred were applied to CP12 sequences. GlobPlot™ is an autoregressive model based on amino acid propensities for order/globularity and disorder. These parameters are based on the hypothesis that the tendency for disorder depends on the Russell-Linding propensities for amino acids to be either in regular secondary structures (α-

helices or β-strands) or outside of them (‘random coil’, loops, turns etc.). Even if, in general, GlobPlot™ and PONDR® were said to predict about the same on the disordered proteins (Linding *et al.*, 2003*b*), here CP12-1 is similar to CP12-2 as disordered, while CP12-3 shows a tendency to globularity in a low complexity structure (fig. 5.4).

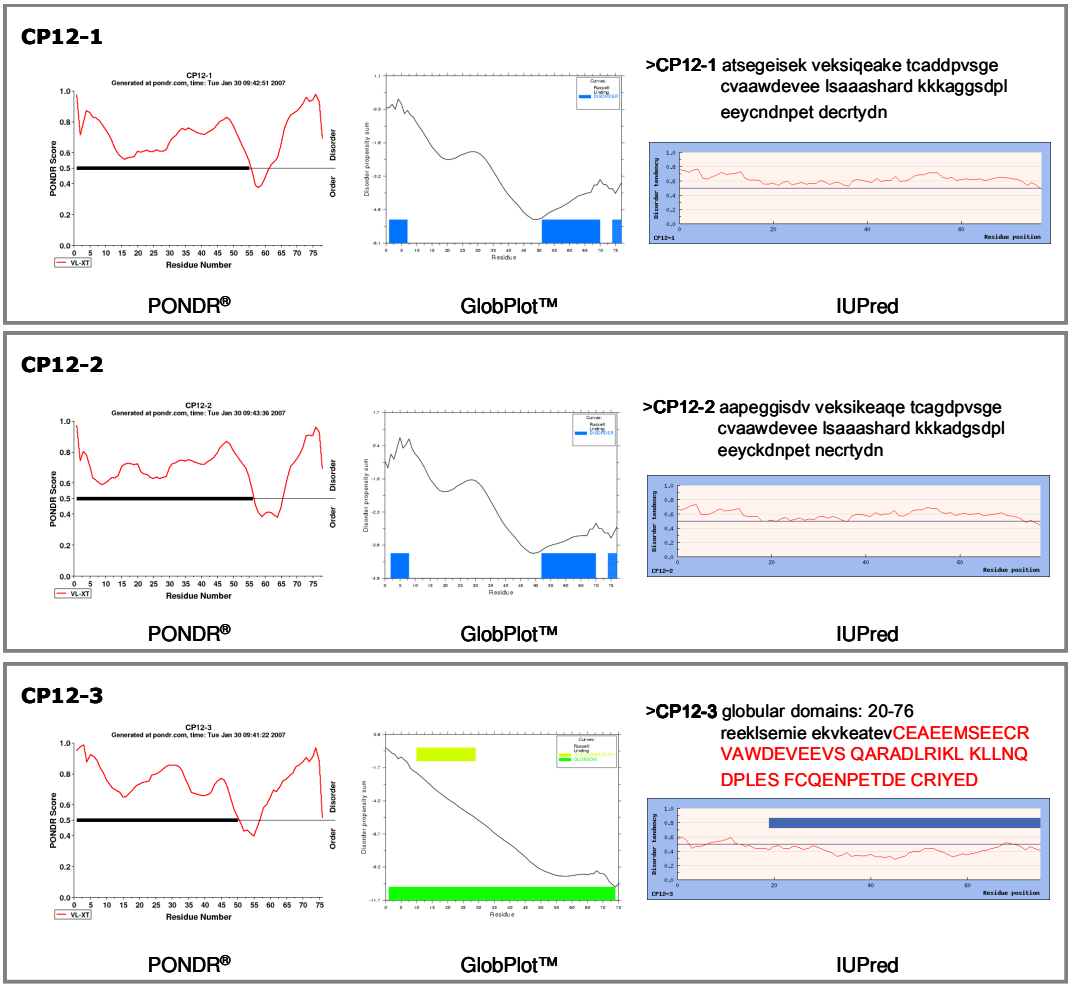


Fig. 5.4 Analysis of CP12 order/disorder using predictors. Disorder predictors (PONDR®, GlobPlot™ and IUPred) were applied to CP12 protein sequences lacking the transit peptides. GlobPlot™ legend put the evidence on clear disordered segments in blue, globular proteins in green and low complexity sequences in yellow. IUPred shows disordered regions, written in small letters, while predicted globular domains are marked with blue bar and written in red capital letters.

The third, IUPred is a linear model presenting a novel algorithm for predicting unstructured regions from amino acid sequences by estimating their total pairwise interresidue interaction energy. This program is based on the assumption that IUP sequences do not fold due to their inability to form sufficient stabilizing interresidue interactions. In general, proteins with intrinsically disordered sequences cannot bury sufficient hydrophobic core to fold into the highly organized 3D structures that characterize the proteins that are represented in the Protein Data Bank. Due to this, in our study CP12-1 and CP12-2 were predicted to be completely unfolded in their entire sequence. Also here, CP12-3 seems to be quite different from the other isoforms, with a potential globular domain in amino acids 20-76. This disagreeing prediction using different algorithms does not clearly identify CP12 behaviours in their potentially structured or unstructured domains. At the same time, while CP12-1 and CP12-2 sequence similarity and their disorder propensity should be index of similar unfolding and potential interchanging functions to be investigated, CP12-3 still remains doubtful.

5.3.2 *Redox properties of the CP12 family*

CP12-2 is a small redox protein, containing four cysteines able to form two disulfide bridges with midpoint redox potentials ($E_{h,\ 7.9}$) of -326±2 mV and -352±6 mV respectively (Marri *et al.,* 2007, see paragraph 4.4.2). The great proportion of IUP proteins does not contain structural cysteines, due to their low structural complexity and high flexibility and disorder. Differently, CP12 is an example of IUP protein sensitive to changing redox conditions in the stroma.

Despite the calculated molecular mass of 8.5 kDa, CP12-2 shows an apparent greater molecular mass, migrating in SDS-PAGE as a peptide of 20 kDa when reduced or 16 kDa if oxidized. The same in size exclusion chromatography, where a 29 kDa peak of oxidized CP12-2 and a 35 kDa peak of reduced CP12-2 were detected, indicating a major conformational change due to disulfide formation (fig. 5.5A, Marri *et al*., 2005*b*). CP12-1

and CP12-3 were first examined in gel filtration, displaying similar behaviours. CP12-1 was detected as a 30kDa peak when oxidized, while a 36kDa represents the reduced form (fig. 5.5B). Also an oxidized 30kDa peak for CP12-3 and a reduced 40kDa CP12-3 peak elute from gel filtration (fig. 5.5C).

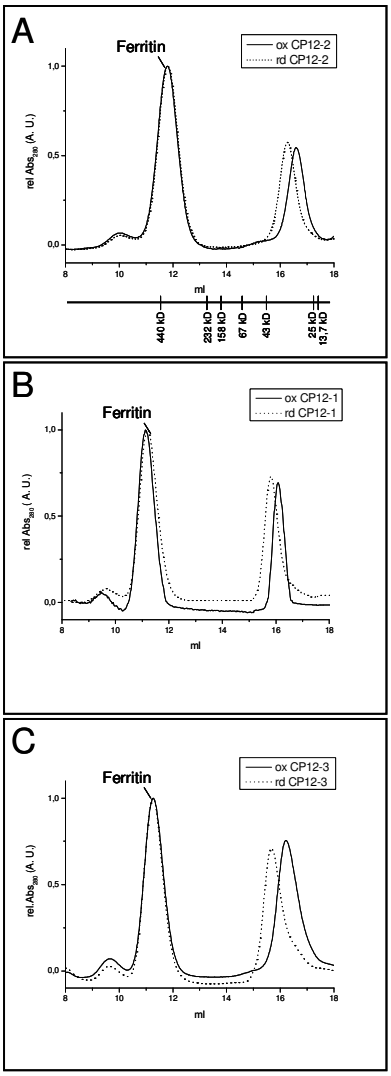


Fig. 5.5 Size exclusion chromatography of oxidized and reduced CP12 isoforms. (A) CP12-2 well known behaviour (Marri *et al*., 2005*b*) was compared to CP12-1 (B) and CP12-3 (C).
Oxidized and reduced CP12 forms were obtained by incubation with 20 mM oxidized or reduced DTT, respectively, for 3 h at 25°C in 25 mM potassium phosphate buffer, pH 7.5. To each sample, before loading in gel filtration, 200 µg of ferritin was added as a standard. The absorption patterns at 280 nm were normalized to ferritin content.
Gel filtration analysis were performed on a Superdex 200 HR10/30 column connected to an ÅKTA Purifier system (General Electric Healthcare). The column was equilibrated with 50 mM Tris-HCl, pH 7.5, 150 mM KCl, 1 mM EDTA (plus 2 mM DTT in reduced samples).

To explain transitions from these totally reduced to totally oxidized CP12 forms, redox titrations were performed in the presence of DTNB as probe under varying redox conditions. Fully reduced and fully oxidized samples

were obtained incubating CP12s in 20 mM reduced or oxidized DTT, subsequently removed to prevent interactions with DTNB. For each CP12 isoform, number of free protein thiols were determined. Each reduced CP12-2 was experimentally found to contain four reactive thiols, while oxidized CP12s had none (table 5.2, Marri *et al*., 2007; see paragraph 4.4.2). These data suggest both disulfides in the three CP12 isoforms could be redox sensitive and titrated with DTT plus DTNB (fig. 5.6).

Table 5.2 Number of reactive thiols for each CP12s in the reduced and oxidized forms. Number of free protein thiols were determined by the ratio between DTNB concentration and CP12s concentration in samples incubated with 20 mM reduced or oxidized DTT, respectively. Data ± standard deviations are means of three independent experiments.

	No. of reduced thiols	No. of oxidized thiols
CP12-2	4.5 ± 0.5	-0.3 ± 0.1
CP12-1	4.3 ± 0.3	0.6 ± 0.1
CP12-3	3.5 ± 0.8	0.8 ± 0.5

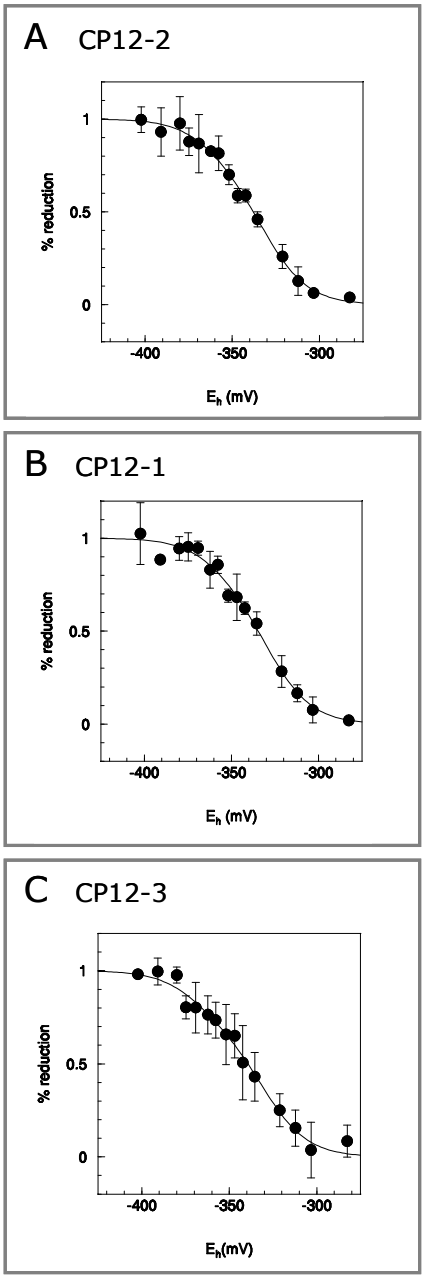


Fig. 5.6 Redox titrations of the three CP12 isoforms. Pure CP12 proteins in 100 mM Tricine-NaOH pH 7.9 buffers were used in each data points.
Absorbance values at 412 nm of blanks (100 mM Tricine-NaOH pH 7.9 plus DTNB) were subtracted to each data points. Then % reduction was a ratio between DTNB concentration (determined at 412 nm using a molar extinction coefficient of 14150 M^{-1}) and CP12 concentration (from absorbance at 280 nm on the basis of molar extinction coefficients of 8730 M^{-1} for CP12-2 (A) and CP12-1 (B) or 7240 M^{-1} for CP12-3 (C)).
Presented data ± standard deviations are means of three independent experiments.

Midpoint redox potentials		
	N-terminal S-S bridge	C-terminal S-S bridge
CP12-2	-326 ± 2	-352 ± 6
CP12-1	-326 ± 11	-347 ± 11
CP12-3	-332 ± 6	-373 ± 14

Table 5.3 Midpoint redox potentials for the three different CP12s, derived by fitting experimental values presented in fig. 5.6 with non-linear regression (CoStat, CoHort Software).
Presented data ± standard deviations are means of three independent experiments.

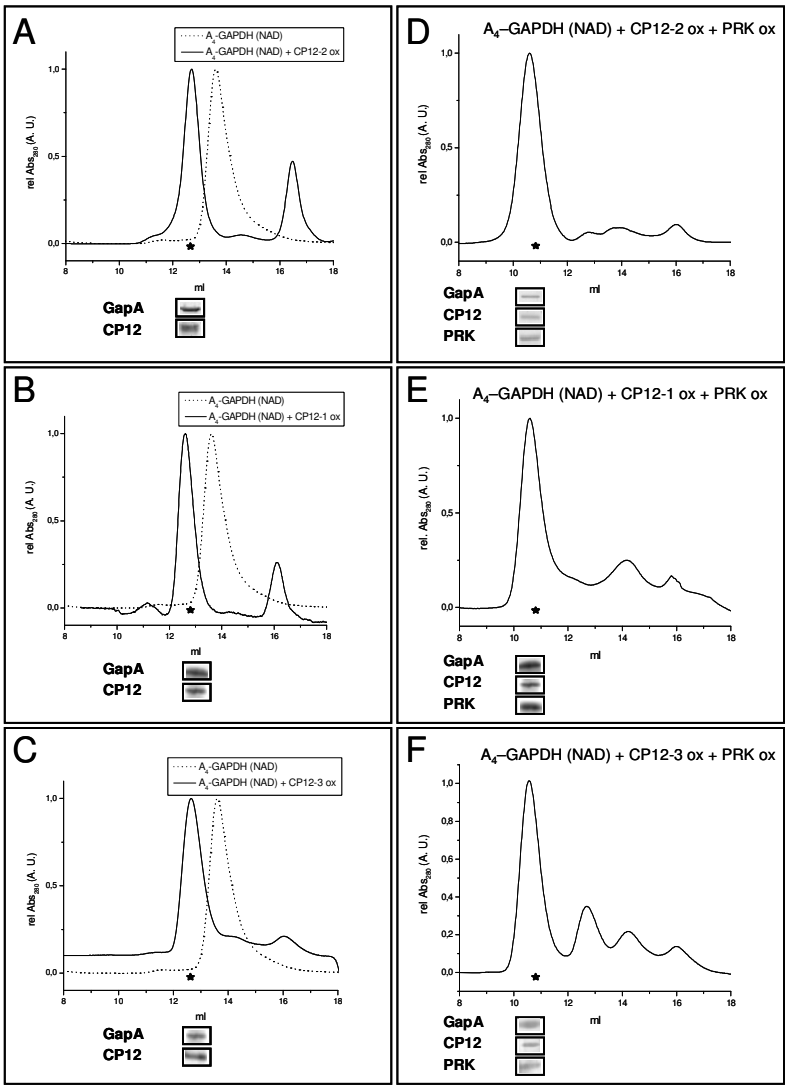
Experimental values from the three redox titrations at pH 7.9 (fig. 5.6) were fitted to a Nernst equation for two different thiol/disulfide equilibria. With these criteria, the midpoint redox potentials ($E_{h,\ 7.9}$) for the two disulfide formation were compared to known potentials for CP12-2 (table 5.3, Marri *et al.,* 2007, see paragraph 4.4.2). Importance of redox values of CP12-1 and CP12-3 compared to CP12-2, should be in explaining the hypothetical CP12-1 and CP12-3 capability to bind GAPDH and PRK. CP12-2 redox properties, on the basis of the behaviour of CP12-2 cysteine mutants, identifies the C-terminal disulfide (with a midpoint redox potential of -352 mV), able to bind GAPDH. In fact AB-GAPDH shows an analogous midpoint redox potential of -353 mV (Marri *et al*., 2007, Sparla *et al.,* 2002). On the other hand, the CP12-2 N-terminal disulfide (-326 mV) and PRK (-330 mV) share this characteristic, explaining the importance of N-terminal S-S bridge in PRK binding. Based on this, similar CP12-1 and CP12-3 midpoint redox potentials might be implicated in the same chloroplastic function.

5.3.3 *CP12 isoforms: the same linker function in supramolecular complex formation?*

To get insights into CP12 function, the three Arabidopsis CP12s were tested in their capability to bind A_4–GAPDH and PRK for the reconstitution of the supramolecular complex. For CP12-2 the formation of a kinetically inhibited

complex with A_4–GAPDH and PRK led to a dark-inhibition of the Calvin-Benson cycle. Conditions promoting the complex formation are typical conditions prevailing in chloroplasts in the dark, i.e. oxidation of thioredoxins and low NADP(H)/NAD(H) ratio (Marri *et al*., 2005*b*).

Fig. 5.7 Binary and ternary complex reconstitution by CP12 isoforms. Equimolar A_4-GAPDH and oxidized CP12-2 (A), (on subunit basis), were incubated for 2 h at 4°C in the presence of 0.2 mM NAD before loading on Superdex 200. Same samples preparation for detecting GAPDH binding to CP12-1 (B) and CP12-3 (C). Stars indicate the column fractions (0.35 mL) which were concentrated and loaded on the gel for western blotting analisys. Insert, Western blots showing that anti-GAPDH and anti-CP12 polyclonal antibodies recognize the presence of GapA and CP12 in the starred peak.
D, E, F: same experiments, except for the addition, in the incubation buffer, of equimolar oxidized PRK on subunit basis. CP12-2 (D), CP12-1 (E) and CP12-3 (F) complex formation were compared. Insert, Immunoblots showing that the starred peaks contained A_4-GAPDH, CP12, and PRK.



CP12-2 is known to bind A_4–GAPDH when oxidized and in presence of NAD, conditions that mimic dark state, resulting in a peak shift in gel filtration. While A_4–GAPDH bound to NAD eluted as a 120 kDa protein, addition of oxidized CP12-2 results in a shift to an apparent molecular mass of 150 kDa (fig. 5.7A, Marri *et al*., 2005*b)*. CP12s presence in 150 kDa peaks, together with A_4–GAPDH, was confirmed by Western Blot analysis with antibodies against GapA and CP12. The same conditions were used to test potential binding of oxidized CP12-1 and CP12-3 to A_4–GAPDH. Incubation of oxidized CP12-1 and CP12-3 with A_4-GAPDH (NAD) gave identical peak shift (155 kDa for CP12-1 and 148 kDa for CP12-3) (fig. 5.7B, C). CP12-2 binding to A_4–GAPDH was understand to have little inhibitory effect on NADPH-dependent GAPDH activity (fig. 5.8A, Marri *et al*., 2005*b*), causing about a 12% of activity loss. At the same time, CP12-1 and CP12-3 binding to GAPDH result in a residual NADPH-dependent GAPDH activity of about 90% (fig. 5.8 B and C respectively).

Since both CP12-1 and CP12-3 can bind GAPDH as CP12-2 does, the capability to reconstitute the complex with PRK was tested. The formation of the Arabidopsis supramolecular complex using CP12-2 as a scaffold protein gave rise to an apparent 640 kDa complex in gel filtration (fig. 5.7D, Marri *et al*., 2005*b*). Interestingly, formation of the GAPDH/CP12-2/PRK complex led to dramatic inhibition of the activity of both enzymes. Within the complex, PRK was 50-fold less active than the reduced enzyme while NADPH-dependent activity of GAPDH embedded in the complex was 5-fold lower than for free enzyme (fig. 5.8A, Marri *et al*., 2005*b*). Figure 5.7E shows as CP12-1 could form an inhibited 680 KDa complex with GAPDH and PRK. Activities inside this complex were 5-fold lower than the full active enzyme for NADPH-dependent GAPDH activity, while PRK was 20-fold less active (fig. 5.8B). Moreover, CP12-3 was able to bind GAPDH and PRK in a 670 kDa ternary complex (fig. 5.7F), which formation inhibits NADPH-dependent activity of GAPDH and PRK. GAPDH into the complex was 4-fold less active than free enzyme, while PRK was 10-fold less then the full active ones (fig. 5.8 C). On the contrary, GAPDH being part of the complexes with

alternative CP12 isoforms retains its NADH-activity unchanged (data not shown). In any case, inactivation of GAPDH and PRK was reversible, as demonstrated for CP12-2 (Marri *et al*., 2005*b*). While these enzymes were inhibited inside the complex, their activities could be totally recovered by incubation with the substrate, BPGA, for GAPDH or reductants for PRK, conditions that are known to completely activate enzymes in complex with CP12-2. In fact, the formation of GAPDH/CP12/PRK complex is a transient inactivation adapting to dark conditions in chloroplast stroma. Enzymes blocked in the complex and strongly inhibited recover their active conformation in dark to light transitions, showing a reversible modulation of photosynthetic enzymes.

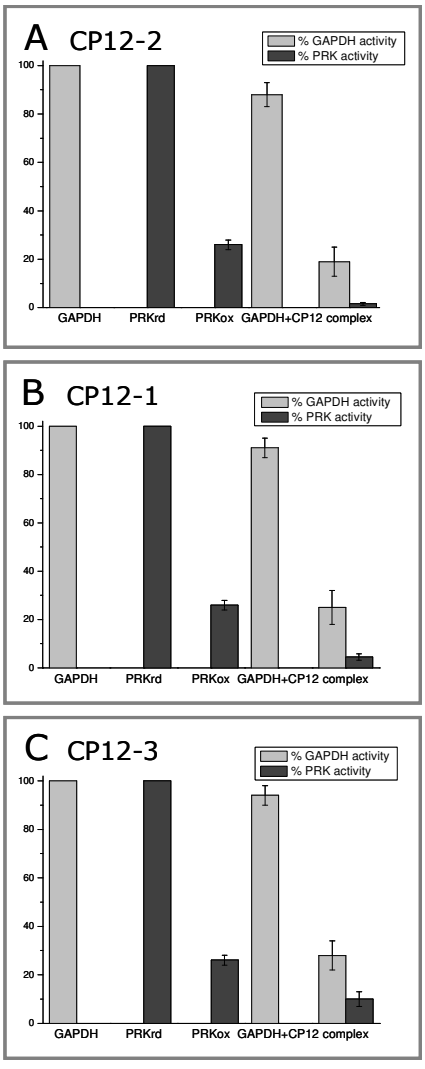


Figure 5.8. Activities of GAPDH and PRK as free enzymes and as enzymes embedded in complexes with different isoforms of CP12 (CP12-2 in box A, CP12-1 in B and CP12-3 in C). Activity expressed as percentage of the full activity of free tetrameric GAPDH (NADPH dependent) in light grey. GAPDH/CP12 and GAPDH/CP12/PRK complexes were obtained under the same conditions as described in legends to figure 5.7A, B and C respectively. PRK activity, in dark grey, was assayed and expressed as a percentage of fully reduced PRK. PRK oxidation was obtained by 3 h incubation at 25°C with 25 mM oxidized DTT. The ternary complex with GAPDH and CP12 was obtained as in figure 5.7 D, E and F.

5.4 *Discussion*

Our recent studies pointed out the basic role of *Arabidopsis thaliana* CP12-2 (At3g62410) in the reversible regulation of Calvin cycle enzymes in dark/light transitions. From its coordinate expression to GapA (At3g26650) and PRK (At1g32060) genes (Marri *et al*., 2005*a*) and from the strictly inhibition that CP12-2 could exert on glyceraldehyde-3-phosphate dehydrogenase and phosphoribulokinase activities into a supramolecular complex, CP12-2 was suggested to act as a modulator of both enzymes (Marri *et al*., 2005*b*). CP12-2, due to its characteristics of intrinsically unstructured protein, in addition to the ability to sense changes in the redox state of chloroplast stroma (Marri *et al*., 2007), acts as the fundamental linker for the complex assembly. Besides this, the whole Arabidopsis genome contains three genes supposed to code for different CP12 isoforms. Based on this assumptions, a characterization of the alternative CP12 isoforms (CP12-1, At2g47400 and CP12-3, At1g76560) becomes not obvious but necessary.

In order to arrange the hypothetical traits of IUP (Tompa, 2002), their sequences were examined and disorder predictors were applied. CP12-1 appeared very similar to CP12-2 in its amino acids composition, with a 86% of homology. CP12-1 amino acids distribution reflects the characteristic composition of unstructured protein, with a major proportion of disorder promoting residues and a depletion in order promoting ones. Predictors applied to CP12-1, as PONDR®, GlobPlot™ and IUPred, unanimously identified this as an intrinsically unstructured protein, lacking organized tertiary structures. More discordant was the position of CP12-3. Its low amino acids sequence similarity to other CP12 isoforms (under 50%) put it far form the other Arabidopsis CP12s in a phylogenetic analysis. Moreover, some evidences seems to take it away from IUP features. Composition in amino acids was sometimes closer to globular protein than IUPs, and disorder predictors add perplexity to this scenario. Use of computational

algorithms to predict order/disorder structures was sometimes controversial, due to their different *modus operandi* leading to possible opposite results. Here, while PONDR® saw the unstructured nature of CP12-3, putting the great proportion of its residues in a disorder state, GlobPlot™ and IUPred disagreed. A long globular domain at the C-terminus of CP12-3 was identified by IUPred; same structured amino acids result from GlobPlot™ with the characteristic of a low complexity region. Possible is that the sequence diversity of CP12-3 in respect to other Arabidopsis isoforms explains a different structure, maybe ordered, arranged in domains with a potential different function. It is realistic that, from the common evolution of the three CP12 genes, CP12-3 spacing out assuming a hypothetical tertiary structure where CP12-1 and CP12-2 have not, perhaps linked to a requested function. Otherwise, also CP12-1 and CP12-2 could be evolved lacking imposed structures and assuming a high flexible and adaptable disordered nature, that permits them to modulate partner enzymes in a regulative or assembling function. This flexible structure, for its characteristic of fast and high adaptability, is often used by those proteins that need very specific and fast responsiveness to changing environmental conditions.

For the sake of understanding Arabidopsis CP12-1 and CP12-3 functions inside the chloroplast, the already known CP12-2 role was taken into consideration. No doubt was on the CP12-2 ability to bind photosynthetic A_4-GAPDH in presence of NAD and under oxidized conditions, nor on the following recruitment of oxidized PRK into the final $[(A_4\text{-GAPDH})\text{-}(CP12\text{-}2)_2\text{-}(PRK)]_2$ supramolecular complex (Marri *et al.*, 2007, 2005*b*). CP12-1 and CP12-3 were tested for this function, with the result that both CP12 isoforms could bind GAPDH and PRK in the same conditions. Besides this, their effects on enzyme activities were examined, cause the formation of the CP12-2 complex has strong inhibitory role (Marri *et al.*, 2005*b*). In parallel with the CP12-2 behaviour, the binding of CP12-2 to A_4-GAPDH gave only a partial activity inhibition (about 10% with all the three isoforms). The subsequent formation of the ternary complex with PRK

results in a comparable and drastic effect on NADPH-dependent GAPDH activity (with a residual activity equal to 18% to 28% of the free full active enzyme), and on PRK activity (from 2 to 10% the reduced fully active ones). Surprisingly the three CP12s are able to reconstitute the complex, even if the experiments were conducted *in vitro* and CP12s have unclarified affinities and times of action. In this hypothesis, it is not certain which could be their major function and if the regulation of the Calvin cycle could be of minor importance.

Properties of CP12-2 dwell also in its capability to sense redox chloroplast conditions, with a regulative and maybe conformational transition between oxidized and reduced forms, and typical chloroplastic midpoint redox potentials (Marri *et al.*, 2007). CP12-2 redox titration, as well as for CP12-1 and CP12-3, gave as a result the formation of two disulfide bridge with their specific midpoint redox potentials. The C-terminal disulfide has similar values for the three isoforms, and it seems to be involved in GAPDH binding, due to the comparable midpoint redox potential of AB-GAPDH and to the behaviour of the cysteine-mutants of CP12-2. The same for the N-terminal S-S bridge of CP12-1 and CP12-2, necessary for PRK binding, no sensible differences were revealed. Some questions should arise from CP12-3 titration, in which the midpoint redox potential of the N-terminal disulfide is more negative, with a slightly distant value from thioredoxins typical redox responsiveness and from enzymes objects of this study.

Deeply interesting will be the identification of proper compartimental localization and functions of the alternative CP12 isoforms and/or their hypothetical partner enzymes. Due to their redox sensitivity, understand the role of chloroplastic thioredoxins with their specificity towards CP12s will be also important.

Conclusion and future perspectives

Through million of years, evolution tends to preserve what is beneficial and useful and to prevent what is futile, spacing out from non convenient situations and overcoming functions that are not required. The acquisitions of mitochondria and plastids were important events in the evolution of the eukaryotic cell. Ancient invasions by eubacteria through symbiosis more than a billion years ago initiated these processes (Dyall *et al*., 2004). By that time, the internalized cyanobacteria had turned into what we know as chloroplasts. Chloroplasts retained a small degree of genetic autonomy, a large degree of their biochemistry, but lost some of their original functions and also acquired ones they did not possess when freeliving (Timmis *et al.*, 2004). The chloroplast requirements to carry out photosynthesis would determine the land plant development and its need to adapt such development to environmental signals, such as light or the availability of raw materials (Lopez-Juez and Pyke, 2005). Molecular events that drove the evolution of endosymbionts into contemporary organelles caused useful protein structures to be maintained, required pathways to be adapted, while efficient strategies were presumably transferred to regulate different processes in the same way.

To some extent it is reminiscent of the debate concerning the old philosophical paradigm, which was first, the chicken or the egg? IUP proteins evolved with their extremely characterizing structure to fulfill more than one function, often with the same part of the protein (Dyson and Wright, 2005; Tompa *et al*., 2005). It is not clear if this flexible feature helped IUPs to survive adapting to different and changing situations or evolution selects IUPs for complying different needs. First was the egg? Is evolution that imposed on proteins this disorder forsaking ordered structures to adapt to fast and specific modality of regulation? Or chicken comes first and their flexibility and disorder help unstructured proteins shaping into different conformations and causing IUPs to interact with

alternative partners? Millions of years could allow evolution to "extinguish" intrinsically unstructured proteins if not favourable or to change their characteristics into preferred ones. It does not happens. Evolution did not. IUPs might be extremely profitable for their potentiality. In this contest, evolution takes the advantages of disorder and flexibility to treat important functions that need accurate control, to bind and assemble what is unable to bind, or to modulate what can not regulate itself.

Evolution takes the C-terminal part of CP12, binds it to A_4-GAPDH creating a GapB subunit capable of self-regulation in higher plants (fig. 6.1; Pohlmeyer *et al*., 2005; Trost *et al*., 2006; Petersen *et al*., 2003). Evolution needs IUPs also to modulate enzymes that are unable to regulate themselves to the same level. From cyanobacteria up to green algae, GAPDH is present in its simplest isoform A_4-GAPDH. It catalyzes a crucial step in carbon organication. As in all oxygenic photosynthetic organisms, A_4-GAPDH, the only dehydrogenase of the Calvin Benson cycle, catalyzes the reduction of bisphosphoglycerate into glyceraldehydes-3-phosphate using NAD(P)H. Since enzymes involved in key processes need to be strongly regulated, the existence of CP12, with its very high affinity for GAPDH, guarantees to A_4-GAPDH a reliable kind of regulation (Graciet *et al*., 2003*b,* 2004*b)*. Through the action of CP12 as scaffold protein, the formation of the GAPDH/CP12/PRK complex is involved in the regulation of the Calvin cycle, modulating the activity of both enzymes in the green unicellular alga *C. reinhardtii* (Graciet *et al*., 2004*a)*. Activation of the enzymes is triggered upon thioredoxin light-mediated reduction of both disulfide bridges in CP12 whereby the complex dissociates.

In land plants, the C-terminal portion of CP12 and the C-terminal extension of B-subunits (CTE) have similar sequences, reflecting a common evolutionary origin and suggesting a similar regulatory function down to the molecular level (Pohlmeyer *et al*., 2005; Trost *et al*., 2006). It seems that evolution has recruited the small CP12 to ensure A_4-GAPDH autonomous regulation by light, and this ancestral choice led to GapB formation.

Exclusively higher plants contain AB-GAPDH isoform (Sparla *et al.,* 2002; Scheibe *et al*., 1996). The CTE is responsible for thioredoxin dependent light/dark regulation of GapB into the major heterotetrameric GAPDH isozyme (A_2B_2-GAPDH) present in higher plants chloroplast stroma. After formation of a disulfide within the CTE (fig. 6.1), this protruding regulatory domain is proposed to fold nearby the active site of the protein and interact with essential residues for coenzyme recognition and catalysis. As a consequence, the NADPH-dependent activity of the enzyme is specifically slowed down (Sparla *et al*., 2002). Dark conditions in chloroplast stroma (as low NADP(H)/NAD(H) ratio and oxidized thioredoxins pool) lead to the assembly of a strongly inhibited A_8B_8-GAPDH form, capable of dissociation and reactivation when light conditions are restored. Mechanistic details of CTE-dependent, A_2B_2-GAPDH regulation are now revealed by the crystallographic structure of oxidized A_2B_2-GAPDH combined with mutant analysis.

In higher plants, light-dark transitions implies a fine regulation of Calvin cycle enzyme mediated by thioredoxins, pyridine nucleotides and metabolites (Buchanan and Balmer, 2005; Wolosiuk *et al.,* 1993). Up to land plants, chloroplasts conserved A_4-GAPDH in its CP12 based regulatory mechanism. Although A_4-GAPDH does not contain CTE, it responds to thioredoxins and metabolites via the small regulatory peptide CP12, causing both A_4-GAPDH and A_2B_2-GAPDH enzymes to be regulated by light through mechanisms which display unique features and common traits. A_2B_2-GAPDH is an "evolved" enzyme which belongs the capability of self-regulation. Since A_2B_2-GAPDH represents the model with its acquired mechanism of modulation, evolution could have forsake previously existing mechanism based on A_4-GAPDH and CP12, evolving into an "all AB-GAPDH" type. Why this ancestral regulation is kept? The other enzyme involved in the supramolecular complex (fig. 6.1), phosphoribulokinase, is a second member of the Calvin cycle, catalyzing a crucial step for the reconstitution of the primary acceptor in Calvin-Benson cycle. PRK, essential for the phosphorilation of ribulose-5-phopshate into ribulose-bisphosphate using

ATP, is itself sensitive to redox conditions in the stroma. In its oxidized form PRK is about 30% active, while a stronger down-regulation is reached only with assembly to the previously formed binary complex A_4-GAPDH/CP12 (Marri *et al*., 2005*b*). In *Arabidopsis thaliana*, a key role for the CP12-mediated regulation consists in the severe inhibitory effect on both enzymes involved in the complex, to a level that can not be reached anyhow.

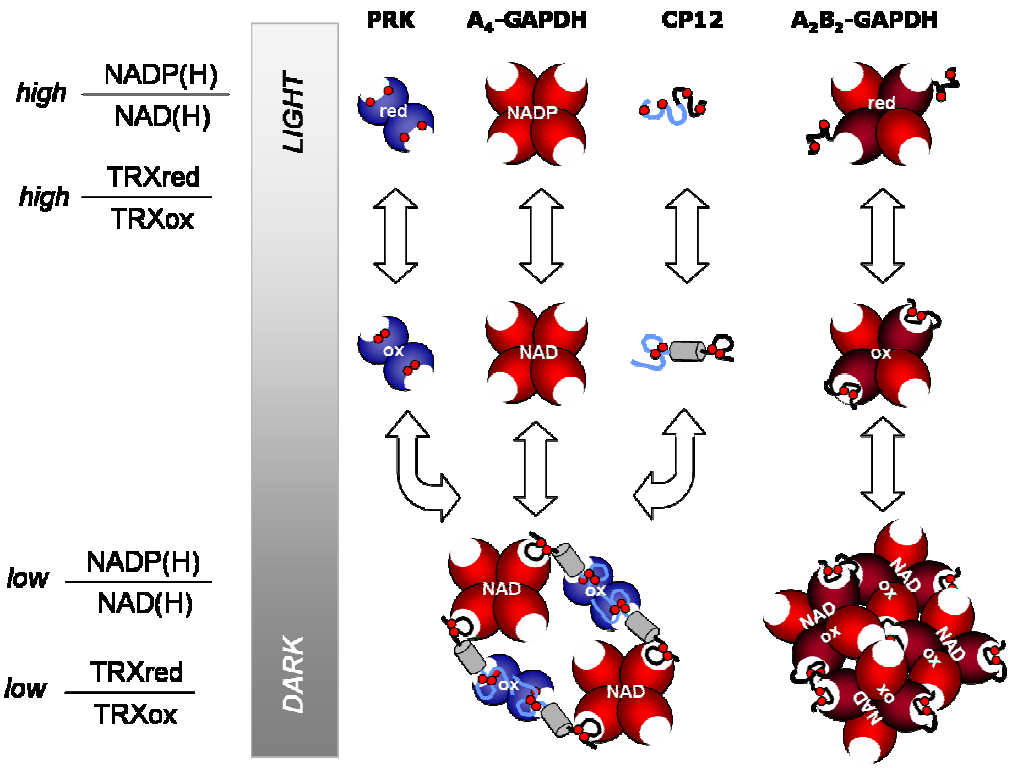


Fig. 6.1 Model of A_4-GAPDH, CP12, PRK and AB-GAPDH interactions and regulation in Arabidopsis chloroplast. The model is based on the integration between known AB-GAPDH self-regulation and our data on A_4-GAPDH/CP12/PRK supramolecular complex. The final stoichiometry of the ternary complex $[(A_4\text{-GAPDH})\text{-}(CP12\text{-}2)_2\text{-}(PRK)]_2$ reflects the toric model proposed in Marri et al., 2007 (see paragraph 4.5), compared to the extremely inhibited A_8B_8-GAPDH isoform.

Organized assembly into supramolecular complexes is a well-know strategy to regulate some metabolic pathways, such as Calvin cycle enzymes (Wolosiuk *et al.,* 1993; Süss *et al.,* 1993). Complex formation was previously described, but recently the frequency and the dynamic behaviour of macromolecular complexes trigger new question as to their composition, the cellular functions they are involved in and the advantage of such complexes compare with individual proteins. In recent years, the use of very sensitive analytical techniques has also revealed the presence of small proteins in macromolecular complexes. The identification of such small unstructured proteins, as CP12, in higher order structures raises questions about their role in the formation and regulation of these complexes through protein-protein interactions.

The involvement of CP12 as a scaffold protein in supramolecular complexes containing GAPDH and PRK has been demonstrated in spinach (Wedel *et al*., 1997; Scheibe *et al.,* 2002). Arabidopsis genome contains three genes for CP12 (*CP12-1*, At2g47400; *CP12-2*, At3g62410; *CP12-3*, At1g76560), together with two GapA genes (*GapA-1,* At3g26650; *GapA-2*, At1g12900), a GapB gene (At1g42970) and a PRK gene (At1g32060). The existence of such a kind of complex regulation in *Arabidopsis thaliana* was suggested by the coordinate expression of the relative genes (*GapA-1*, *GapB*, *PRK*, and *CP12-2*) under different environmental conditions (Marri *et al.*, 2005*a*). In our study, light/dark regulation, sugar repression and the organ-specific expression of involved genes display a similar pattern, fully consistent with the physiological relevance of GAPDH, CP12-2, and PRK proteins into a supramolecular complex in Arabidopsis chloroplasts. From these emerging data, the characterization of Arabidopsis complex was required.

Arabidopsis CP12-2 is a small 8.5 kDa protein without enzymatic activity. The amino acid sequence of CP12-2 includes signature of potential intrinsic disorder. Predictors of intrinsic disorder, such as PONDR® (Romero *et al*., 2001), GlobPlot™ (Linding *et al*., 2003*b*) and IUPred (Dosztanyi *et al.,* 2005), applied to CP12-2 sequence, classified it as a IUP (intrinsically

unstructured protein; Tompa, 2002; Graciet *et al*., 2003*a*; Marri *et al*., 2007). The oxidized CP12-2 partially loses its flexibility and shows a limited structural organization in NMR analysis due to the formation of disulphide bridges (Marri *et al*., 2007). As expected, reduction of disulfides led to increased disorder. Consistent with its unstructured nature, CP12-2 was demonstrated by multiangle light-scattering analysis (MALS-QELS) to be a monomer of 9 kDa under native conditions (Marri *et al*., 2007), even if it was largely estimated in size-exclusion chromatography and gel-electrophoresis (Marri *et al*., 2005*b*). Besides this, CP12-2 is a redox sensitive linker inside the supramolecular complex and it seems to change conformation depending on redox conditions (Marri *et a*l., 2007). It bears two internal disulfides responding to thioredoxin: a C-terminal Cys64-Cys73 bridge shows a midpoint redox potential, $E_{m,7.9}$ -352 mV, identical to spinach A_2B_2-GAPDH (-353 mV; Sparla *et al*., 2002); while the N-terminal Cys22-Cys31, $E_{m,7.9}$ -326 mV, was otherwise similar to PRK (-330 mV; Marri *et al*., 2005*b*). Analyzing mutants on the two CP12-2 disulfide bridges (C22S and C74S), it appears that the C-terminal S-S bond is necessary for GAPDH binding in the A_4-GAPDH/CP12 binary complex, arranging to the subsequent binding of PRK at the N-terminal bridge. Waiting for obtaining a crystallographic structure of binary and ternary complexes, to understand complex stoichiometry, combining dynamic light scattering analysis with binding affinity determination was our pursued approach.

Studies on complexes stoichiometry and binding affinity through MALS-QELS light scattering and isothermal titration microcalorimetry identify a $[(A_4\text{-GAPDH})\text{-}(CP12)_2\text{-}(PRK)]_2$ ternary complex of 498 kDa. Binding affinity in $(A_4\text{-GAPDH})\text{-}(CP12)_2$ was in the submicromolar range (k_D 0.18 μM), reflecting its physiological relevance. Moreover, its exoergonic (ΔG -9 kcal mol^{-1}) and esothermic nature (ΔH -15 kcal mol^{-1}), along with its entropic penalty (TΔS -5 kcal mol^{-1}) indicate that complex formation resulted in a general decrease of disorder attributable to CP12 (Marri *et al*., 2007). Coupling folding and binding to partner proteins is a typical, physiological important characteristic of intrinsically unstructured proteins, generally

associated to an entropic cost (Dyson and Wright, 2005). It seems that CP12-2, poorly structured in solution, might fold upon binding to its partner GAPDH, leading to a decrease in entropy of the whole system. This folding upon binding is in our intent to be confirmed by NMR analysis, identifying residues potentially responsible of this conformational changes. Binding affinity (k_D 0.17 µM), and thermodynamic parameters (ΔG -9 kcal mol^{-1}, ΔH -20 kcal mol^{-1}, TΔS -11 kcal mol^{-1}) confirmed a favourable interaction of PRK to preformed binary complex, into a toric sopramolecular complex (fig. 6.1; Marri *et al*., 2007). Integrating our data from light scattering analysis to confirm complex molecular mass, with affinity results on protein-protein interactions, combining to behaviour of CP12-2 mutants, no other possible stoichiometries were so nicely supported. Therefore, our interpretation of symmetric toric sopramolecular complex is the only possible. In this scenery, the reversible formation of the $[(A_4\text{-GAPDH})\text{-}(CP12)_2\text{-}(PRK)]_2$ complex provides novel possibilities to finely regulate A_4-GAPDH and PRK in a coordinate manner. Individually, not able to regulate itself nor to sense redox conditions is the former, while insensitive to pyridine nucleotides and metabolites, nor to be completely inactivated by oxidation, the latter. GAPDH, together with PRK, is subject to a combined modulation (fig. 6.1). Their separate regulation pathways are converging into supramolecular complex formation, explaining the physiological relevance into chloroplast stroma.

Moreover, the presence of three CP12 genes in Arabidopsis genome, has to be taken into consideration. A comparative analysis of their characteristic amino acids composition and their intrinsic disorder was conducted. The identification of CP12-1 as an intrinsically unstructured protein, pointed out diverging characteristics of CP12-3, maybe more ordered. According to redox sensitivity, i*n vitro* reconstitution of binary and ternary complexes seems to be a function of every CP12, even if this role should not be the primary for the other isoform of CP12. Structural malleability of IUPs makes them moonlighting or multi-tasking (Tompa et al., 2005), using several mechanisms to switch between functions, i.e. changes in cellular localization

or ligand binding, expression in different cell types or variations in complexation state. It seems that evolution has recruited their unused surface to alternative purpose, ensuing capacity to fulfil more than one function. Evolution conserved three CP12s in *Arabidopsis thaliana.* Their existence should not be viewed as redundant, nor random. In this contest, our purpose to identify their potential functions is of great relevance. CP12 roles in Arabidopsis chloroplasts could be the same or, on the opposite, strongly different, mutually exclusive or coordinate. Identifying their hypothetical targets, together with the mechanisms or pathways they are implied in, should be a further step to deeply investigate Arabidopsis intricate regulatory network in light regulation of photosynthetic enzymes.

References

Arguello-Astorga G, Herrera-Estrella L (1998) Evolution of plant regulated plant promoters. *Annual Review of Plant Physiology and Plant Molecular Biology* **49:** 525–555

Arnon DI, Rosenberg LL, Whatley FR (1954) A new glyceraldehyde phosphate dehydrogenase from photosynthetic tissues. *Nature* **173**:1132–1134

Baalmann E, Backhausen JE, Kitzmann C, Scheibe R (1994) Regulation of NADP-dependent glyceraldehyde-3-phosphate dehydrogenase in spinach chloroplasts. *Botanica Acta* **107:** 313–320

Baalmann E, Backhausen JE, Rak C, Vetter S, Scheibe R (1995) Reductive modification and non-reductive activation of spinach chloroplast NADP-glyceraldehyde-3-phosphate dehydrogenase. *Archives of Biochemistry and Biophysics* **324:** 201–208

Baalmann E, Scheibe R, Cerff R, Martin W (1996) Functional studies of chloroplast glyceraldehyde-3-phosphate dehydrogenase subunits A and B expressed in *Escherichia coli*: formation of highly active A4 and B4 homotetramers and evidence that aggregation of the B4 complex is mediated by the B subunit carboxy terminus. *Plant Molecular Biology* **32:** 505–513

Backhausen JE, Vetter S, Baalmann E, Kitzmann, Scheibe R (1998) NAD-dependent malate dehydrogenase and glyceraldehyde-3-phosphate dehydrogenase isoenzymes play an important role in dark metabolism of various plastid types. *Planta* **205:** 359–366

Bodenhausen G, Ruben DJ (1980) Natural abundance nitrogen-15 NMR enhanced heteronuclear spectroscopy. *Chem Phys Lett* **69:** 185-189

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Analytical Biochemistry* **72:** 248–254

Brandes HK, Larimer FW, Hartman FC (1996) The molecular pathway for the regulation of phosphoribulokinase by thioredoxin *f*. *Journal of Biological Chemistry* **271:** 3333–3335

Brinkmann H, Cerff R, Salomon M, Soll J (1989) Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits GapA and GapB of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach. *Plant Molecular Biology* **13:** 81–94

Buchanan BB (1980) Role of light in the regulation of chloroplast enzymes. *Annual Review of Plant Physiology* **31:** 341–374

Buchanan BB (1992) Carbon dioxide assimilation in oxygenic and anoxygenic photosynthesis. *Photosynth Res* **33**:147–162

Buchanan BB, Balmer Y (2005) Redox regulation: a broadening horizon. *Annu Rev Plant Biol* **56:** 187–220

Carr P, Verger D, Ashton AR, Ollis D (1999) Chloroplast NADP-malate dehydrogenase: structural basis of light-dependent regulation of activity by thiol oxidation and reduction. *Structure* **7**:461–475

Carugo O, Argos P (1997) NADP-dependent enzymes. I: Conserved stereochemistry of cofactor binding. *Proteins: Struct.Funct Genet* **28**:10–28

Casagrande S, Bonetto V, Fratelli M, Gianazza E, Eberini I, Massignan T, Salmona M, Chang G, Holmgren A, Ghezzi P (2002) Glutathionylation of human thioredoxin: a possible crosstalk between the glutathione and thioredoxin systems. *Proc Natl Acad Sci USA* **99:** 9745–9749

Cerff R (1978) Glyceraldehyde-3-phosphate dehydrogenase (NADP) from Sinapis alba: steady state kinetics. *Phytochemistry* **17**:2061–2067

Cerff R (1979) Quaternary structure of higher plant glyceraldehyde-3-phosphate dehydrogenases. *Eur J Biochem* **94:** 243–247

Cerff R, Chambers S (1979) Subunit structure of higher plant glyceraldehyde-3-phosphate dehydrogenases. *J Biol Chem* **254**:6094–6098

Cerff R, Kloppstech K (1982) Structural diversity and differential light control of mRNAs coding for angiosperm glyceraldehyde-3-phosphate dehydrogenases. *Proc Natl Acad Sci USA* **79**:7624–7628

Chai MF, Chen QJ, An R, Chen YM, Chen J, Wang XC (2005) NADK2, an Arabidopsis chloroplastic NAD kinase, plays a vital role in both chlorophyll synthesis and chloroplast protection. *Plant Mol Biol* **59**:553–564

Chan CS, Peng H-P, Shih M-C (2002) Mutations affecting light regulation of nuclear genes encoding chloroplast glyceraldehyde-3-phosphate dehydrogenase in *Arabidopsis*. *Plant Physiology* **130:** 1476–1486

Clasper S, Easterby JS, Powls R (1991) Properties of two high-molecular mass forms of glyceraldehyde-3-phosphate dehydrogenase from spinach leaf, one of which also possesses latent phosphoribulokinase activity. *European Journal of Biochemistry* **202:** 1239–1246

Collin V, Issakidis-Bourget E, Marchand C, Hirasawa M, Lancelin JM, Knaff DB, Miginiac-Maslow M (2003) The Arabidopsis plastidial thioredoxins: new functions and new insights into specificity. *J Biol Chem* **278**: 23747-23752

Conley TR, Shih M-C (1995) Effects of light and chloroplast functional state on expression of nuclear genes encoding chloroplast glyceraldehyde-3-phosphate dehydrogenase in long hypocotyls (*hy*) mutants and wild-type *Arabidopsis thaliana*. *Plant Physiology* **108:** 1013–1022

Conway ME, Poole LB, Hutson SM (2004) Roles of cysteines residues in the regulatory CXXC motif of human mitochondrial branched chain aminotransferase enzyme. *Biochem* **43:** 7356-7364

Cotgreave IA, Gerdes R, Schuppe-Koistinen I, Lind C (2002) S-glutathionylation of glyceraldehyde-3-phosphate dehydrogenase: role of thiol oxidation and catalysis by glutaredoxin. *Methods Enzymol* **348:** 175–182

Dai S, Johansson K, Miginiac-Maslow M, Schürmann P, Eklund H (2004) Structural basis of redox signalling in photosynthesis: structure and function of ferredoxin:thioredoxin reductase and target enzymes. *Photosynth Res* **79:** 233–248

Dani DN, Sainis JK (2005) Isolation and characterization of a thylakoid membrane module showing partial light and dark reactions. *Biochim Biophys Acta* **1669:** 43–52

Delumeau O, Renard M, Montrichard F (2000) Characterization and possible redox regulation of the purified calmodulin-dependent NAD^+ kinase from *Lycopersicon pimpinellifolium*. *Plant Cell Environ* **23**: 1267–1273

Dewdney J, Conley TR, Shih M-C, Goodman H (1993) Effects of blue and red light on expression of nuclear genes encoding chloroplast glyceraldehyde-3-phosphate dehydrogenase of *Arabidopsis thaliana*. *Plant Physiol* **103**:1115–1121

Dixon DP, Skipsey M, Grundy NM & Edwards R (2005) Stress-induced protein S-glutathionylation in Arabidopsis. *Plant Physiol* **138:** 2233–2244

Dosztanyi Z, Csizmok V, Tompa P, Simon I (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21:** 3433-3434

Dunker AK, Obradovic Z (2001) The protein trinity - linking function and disorder. *Nat. Biotechnol.* **19:** 805-806

Dyall SD, Brown MT, Johnson PJ (2004) Ancient Invasions: From Endosymbionts to Organelles. *Science* **304:** 253 - 257

Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6:**197–208

Dyson HJ, Wright PE (2004) Unfolded proteins and protein folding studied by NMR. *Chem Rev* **104:** 3607-3622

Dyson HJ, Wright PE (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *JMB* **293:** 321-331

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* **8:** 978–984

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of Molecular Biology* **300:** 1005-1016

Falini G, Fermani S, Ripamonti A, Sabatino P, Sparla F, Pupillo P, Trost P (2003) The dual coenzyme specificity of photosynthetic glyceraldehyde-3-phosphate dehydrogenase interpreted by the crystal structure of A_4 isoform complexed with NAD. *Biochemistry* **42:** 4631–4639

Fermani S, Ripamonti A, Sabatino P, Zanotti G, Scagliarini S, Sparla F, Trost P, Pupillo P (2001) Crystal structure of the non-regulatory A_4 isoform of spinach chloroplast glyceraldehyde-3-phosphate dehydrogenase complexed with NADP. *J Mol Biol* **314**:527–542

Ferri G, Stoppini M, Meloni ML, Zapponi MC, Iadarola P (1990) Chloroplast glyceraldehyde-3-phosphate dehydrogenase (NADP): amino acid sequence of the subunits from isoenzyme I and structural relationship with isoenzyme II. *Biochim Biophys Acta* **1041**:36–42

Figge RM, Schubert M, Brinkmann H, Cerff R (1999) Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eukaryotes: evidence for intra- and inter-kingdom gene transfer. *Mol Biol Evol* **16:** 429–440

Foyer CH, Noctor G (2005) Redox homeostasis and antioxidant signalling: a metabolic interface between stress perception and physiological responses. *Plant Cell* **17**; 1866-1875

Galzitskaya OV, Garbuzynskiy SO, Lobanov MY (2006) FoldUnfold: web server for the prediction of disordered regions in protein chain. *Bioinformatics* 22: 2948-2949

Gardebien F, Thangudu RR, Gontero B, Offmann B (2006) Construction of a 3D model of CP12, a protein linker. *J Mol Graph Model* **25**: 186–195

Gavin AC, Superti-Furga G (2003) Protein complexes and proteome organization from yeast to man. *Curr Opin Chem Biol* **7:** 21–27

Gelhaye E, Rouhier N, Jacquot JP (2004) The thioredoxin h system of higher plants. *Plant Physiol Biochem* **42**: 265-271

Gerhardt R, Stitt M, Heldt HW (1987) Subcellular metabolite levels in spinach leaves. *Plant Physiol* **83**: 399–407

Ghezzi P (2005) Oxidoreduction of protein thiols in redox regulation. *Biochem Soc Trans* **33:** 1378–1381

Gontero B, Cardenas ML, Ricard J (1988) A functional five-enzyme complex of chloroplasts involved in the Calvin cycle. *Eur J Biochem* **173:** 437-443

Gontero B, Lebreton S, Graciet E (2002) Multienzyme complexes involved in the Benson–Calvin cycle and in fatty acid metabolism. In: McManus MT, Laing W, Allan AC, eds. *Protein–protein interactions in plant biology*. Annual Plant Reviews, Vol. 7. Sheffield Academic Press, 120–150

Goodsell DS (1991) Inside a living cell. *Trends Biochem Sci* **16:** 203–206

Graciet E, Gans P, Wedel N, Lebreton S, Camadro J-M, Gontero B (2003a) The small protein CP12: a protein linker for supramolecular protein assembly. *Biochemistry* **42:** 8163–8170

Graciet E, Lebreton S, Camadro JM, Gontero B (2003b) Characterization of native and recombinant A4 glyceraldehyde 3-phosphate dehydrogenase. Kinetic evidence for conformation changes upon association with the small protein CP12. *Eur J Biochem* **270:** 129–136

Graciet E, Lebreton S, Gontero B (2004a) Emergence of new regulatory mechanisms in the Benson-Calvin pathway via protein-protein interactions: a glyceraldehyde-3-phosphate dehydrogenase/CP12/phosphoribulokinase complex. *J Exp Bot* **55:** 1245–1254

Graciet E, Mulliert G, Lebreton S, Gontero B (2004b) Involvement of two positively charged residues of *Chlamydomonas reinhardtii* glyceraldehyde-3-phosphate dehydrogenase in the assembly process of a bi-enzyme complex involved in CO_2 assimilation. *Eur J Biochem* **271**:4737–4744

Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JE, Wagner G (2003) Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* **115:** 739-750

Heineke D, Riens B, Grosse H, Hoferichter P, Heldt HW (1991) Redox transfer across the inner chloroplast envelope membrane. *Plant Physiol* **95:** 1131–1137

Heldt WH, Werdan K, Milovancev M and Geller G (1973) Alkalization of the chloroplast stroma caused by light-depedent proton flux into the thylakoid space**.** *Bioch Biophy Acta* **314:** 224-241

Hirasawa M, Ruelland E, Schepens I, Issakidis-Bourguet E, Miginiac-Maslow M, Knaff D (2000) Oxidation-reduction properties of the regulatory disulfides of sorghum chloroplast nicotinamide adenine dinucleotide phosphate-malate dehydrogenase. *Biochemistry* **39:** 3344–3350

Hirasawa M, Schürmann P, Jacquot J-P, Manieri W, Jacquot P, Keryer E, Hartman F, Knaff D (1999) Oxidation-reduction properties of chloroplast thioredoxins, ferredoxin:thioredoxin reductase, and thioredoxin *f*-regulated enzymes. *Biochemistry* **38:** 5200–5205

Houtz RL, Portis AR Jr (2003) The life of ribulose-1,5-bisphosphate carboxylase/oxygenase–posttranslational facts and mysteries. *Arch Biochem Biophys* **414:** 150-158

Hutchison RS, Groom Q, Ort DR (2000) Differential effects of chilling-induced photooxidation on the redox regulation of photosynthetic enzymes. *Biochemistry* **6**:6679–6688

Hutchinson RS, Ort DR (1995) Measurement of equilibrium midpoint potentials of thiol/disulfide regulatory groups on thioredoxin-activated chloroplast enzymes. *Methods Enzymol* **252:** 220–228

Johansson K, Ramaswamy S, Saarinen M, Lemaire-Chamley M, Issakidis-Bourguet E, Miginiac-Maslow M, Eklund H (1999) Structural basis for light-activation of a chloroplast enzyme: the structure of NADP-malate dehydrogenase in its oxidized form. *Biochemistry* **38**:4319–4326

Kobayashi D, Tamoi M, Iwaki T, Shigeoka S, Wadano A (2003) Molecular characterization and redox regulation of phosphoribulokinase from the cyanobacterium Synechococcus sp. PCC 7942. *Plant Cell Physiol* **44:** 269–276

Koch KE (1996) Carbohydrate-modulated gene expression in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47:** 509–540

Koide S, Dyson HJ, Wright PE (1993) Characterization of a folding intermediate of apoplastocyanin trapped by proline isomerization. *Biochemistry* **32:** 12299-12310

Koksharova O, Schubert M, Shestakov S, Cerff R (1998) Genetic and biochemical evidence for distinct key functions of two highly divergent GAPDH genes in catabolic and anabolic carbon flow of the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Molecular Biology* **36:** 183–194

Kramer DM, Sacksteder CA, Cruz JA (1999) How acidic is the lumen? *Photosynth Res* **60**: 151–163

Laloi C, Rayapuram N, Chartier Y, Grienenberger JM, Bonnard G, Meyer Y (2001) Identification and characterization of a mitochondrial thioredoxin system in plants. *Proc Natl Acad Sci USA* **98**:14144-14149

Lebreton S, Graciet E, Gontero B (2003) Modulation via protein–protein interactions of glyceraldehyde-3-phosphate dehydrogenase activity through redox phosphoribulokinase regulation. *J Biol Chem* **278**: 12078–12084

Lebreton S, Gontero B (1999) Memory and imprinting in multienzyme complexes. Evidence for information transfer from glyceraldehyde-3 phosphate dehydrogenase to phosphoribulokinase under reduced state in *Chlamydomonas reinhardtii*. *Journal of Biological Chemistry* **274:** 20879–20884

Lemaire SD, Stein M, Issakidis-Bourguette M, Keryer E, Benoit V, Pineau B, Gerard-Hirne C, Miginiac-Maslow M, Jacquot J-P (1999) The complex regulation of ferredoxin/thioredoxin-related genes by light and the circadian clock. *Planta* **209:** 221–229

Li AD, Anderson LE (1997) Expression and characterization of pea chloroplastic glyceraldehyde-3-phosphate dehydrogenase composed of only the B-subunit. *Plant Physiology* **115:** 1201–1209

Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB (2003*a*) Protein disorder prediction: implications for structural proteomics. *Structure* **11:**1453-9

Linding R, Russell RB, Neduva V, Gibson TJ (2003*b*) GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic Acids Res* **31:** 3701-3708

Lisse T, Bartels D, Kalbitzer HR, Jaenicke R (1996) The recombinant dehydrin-like desiccation stress protein from the resurrection plant Craterostigma plantagineum displays no defined three-dimensional structure in its native state. *Biol Chem* **377:** 555-561

Lopez-Juez E, Pyke KA (2005) Plastids unleashed: their development and their integration in plant development. *Int J Dev Biol* **49:** 557-577

Lydakis-Simantiris N, Hutchison RS, Betts SD, Barry BA, Yocum CF (1999) Manganese stabilizing protein of photosystem II is a thermostable, natively unfolded polypeptide. *Biochemistry* **38:** 404-414

Marri L, Sparla F, Pupillo P, Trost P (2005a) Coordinated gene expression of photosynthetic glyceraldehyde-3-phosphate dehydrogenase, phosphoribulokinase and CP12 in *Arabidopsis thaliana. J Exp Bot* **56:** 73–80

Marri L, Trost P, Pupillo P, Sparla F (2005b) Reconstitution and properties of the recombinant glyceraldehyde-3-phosphate dehydrogenase/CP12/phosphoribulokinase supramolecular complex of Arabidopsis. *Plant Physiol.* **139**: 1433-43

Marri L, Trost P, Trivelli X, Gonnelli L, Pupillo P, Sparla F (2007) Spontaneous assembly of photosynthetic supramolecular complexes mediated by the intrinsically unstructured protein CP12. *submitted*

Martin W, Schnarrenberger C (1997) The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. *Curr Genet* **32**: 1–18

Meyer-Gauen G, Herbrandt H, Pahnke J, Cerff R, Martin W (1998) Gene structure, expression in *Escherichia coli* and biochemical properties of the NAD^+-dependent glyceraldehyde-3-phosphate dehydrogenase from *Pinus sylvestris* chloroplasts. *Gene* **209:** 167–174

Michelet L, Zaffagnini M, Massot V, Keryer E, Vanacker H, Miginiac-Maslow M, Issakidis-Bourget E, Lemaire SD (2006) Thioredoxins, glutaredoxins, and glutathionylation: new crosstalk to esplore. *Photosynth Res* **89**:225-245

Michelet L, Zaffagnini M, Marchand C, Collin V, Decottignies P, Tsan P, Lancelin JM, Trost P, Miginiac-Maslow M, Noctor G, Lemaire SD (2005) Glutathionylation of chloroplast thioredoxin f is a redox signaling mechanism in plants. *Proc Natl Acad Sci USA* **102:** 16478–16483

Michels AK, Wedel N, Kroth PG (2005) Diatom plastids possess a phosphoribulokinase with an altered regulation and no oxidative pentose phosphate pathway. *Plant Physiol* **137:** 911–920

Miginiac-Maslow M, Lancelin JM (2002) Intrasteric inhibition in redox signalling: light activation of NADP-malate dehydrogenase. *Photosynth Res* **72**: 1–12

Mouche F, Gontero B, Callebaut I, Mornon J-P, Boisset N (2002) Striking conformational change suspected within the phosphoribulokinase dimer induced by interaction with GAPDH. *Journal of Biological Chemistry* **277:** 6743–6749

Müller B (1972) A labile CO_2-fixing enzyme complex in spinach chloroplasts. *Zeitschrift für Naturforschung* **27b:** 925–932

Muto S, Miyachi S, Usuda H, Edwards GE, Bassham JA (1980) Light-induced conversion of nicotinamide adenine dinucleotide to nicotinamide adenine dinucleotide phosphate in higher plant leaves. *Plant Physiol* **68:** 324–328

Nawrath C, Mètraux JP (1999) Salicylic acid induction-deficient mutants of *Arabidopsis* express PR-2 and PR-5 and accumulate high levels of camalexin after pathogen inoculation. *The Plant Cell* **8:** 1393–1404

Neuhaus HE, Batz O, Thom E, Scheibe R (1993) Purification of intact plastids from various heterotrophic plant tissues: analysis of enzymic equipment and precursor dependency for starch biosynthesis. *Biochemical Journal* **296:** 395–401

Nicholson S, Easterby JS, Powls (1987) Properties of a multimeric protein complex from chloroplasts possessing potential activities of NADPH-dependent glyceraldehyde-3-phosphate dehydrogenase and phosphoribulokinase. *Eur J Biochem* **162**: 423–431

Partch CL, Clarkson MW, Ozgur S, Lee AL, Sancar A (2005) Role of structural plasticity in signal transduction by the cryptochrome blue-light photoreceptor. *Biochemistry* **44:** 3795-3805

Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK, Obradovic Z (2005) Optimizing long intrinsic disorder predictors with protein evolutionary information. *J Bioinform Comput Biol* **3**: 35–60

Perozzo R, Folkers G, Scapozza L (2004) Thermodynamics of protein-lignad interactions: History, presence, and future aspects. *J Recept Signal Transduct Res* **24:** 1-52

Petersen J, Brinkman H, Cerff R (2003) Origin, evolution, and metabolic role of a novel glycolytic GAPDH enzyme recruited by land plant plastids. *Journal of Molecular Evolution* **57:** 16–26

Petersen J, Teich R, Becker B, Cerff R, Brinkmann H (2006) The GapA/B gene duplication marks the origin of Streptophyta (charophytes and land plants). *Mol Biol* Evol **23:** 1109-1118

Pohlmeyer K, Paap BK, Soll J, Wedel N (1996) CP12: a small nuclear-encoded chloroplast protein provides novel insights into higher-plant GAPDH evolution. *Plant Mol Biol* **32:** 969–978

Porter MA, Milanez S, Stringer CD, Hartman FC (1986) Purification and characterization of ribulose-5-phosphate kinase from spinach. *Arch Biochem Biophys* **245:** 14–23

Porter MA, Stringer CD, Hartman FC (1988) Characterization of the regulatory thioredoxin site of phosphoribulokinase. *J Biol Chem* **263:** 123–129

Pupillo P, Faggiani R (1979) Subunit structure of three glyceraldehyde-3-phosphate dehydrogenases of some flowering plants. *Arch Biochem Biophys* **154:** 475–482

Pupillo P, Giuliani Piccari G (1973) The effect of NADP on the subunit structure and activity of spinach chloroplast glyceraldehyde-3-phosphate dehydrogenase. *Arch Biochem Biophys* **154**: 324–331

Pupillo P, Giuliani Piccari G (1975) The reversible depolymerization of spinach chloroplast glyceraldehyde-3-phosphate dehydrogenase. Interaction with nucleotides and dithiothreitol. *Eur J Biochem* **51:** 475–482

Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK (2006) Intrinsic Disorder and Functional Proteomics. *Biophys J* [Epub ahead of print]

Raines CA, Lloyd JC, Dyer TA (1991) Molecular biology of the C_3 photosynthetic carbon reduction cycle. *Photosynthesis Research* **27:** 1–14

Raines CA, Longstaff M, Lloyd JC, Dyer TA (1989) Complete coding sequence of phosphoribulokinase: developmental and light-dependent expression of the mRNA. *Molecular and General Genetics* **220:** 43–48

Rault M, Giudici-Orticon M-T, Gontero B, Ricard J (1993) Structural and functional properties of a multienzyme complex from spinach chloroplasts. I. Stoichiometry of the polypeptide chains. *Eur J Biochem* **217:** 1065–1073

Romero P, Obradovic Z, Li X, Garner E, Brown C, Dunker AK (2001) Sequence complexity of disordered protein. *Proteins: Struct Funct Gen* **42:**38-48

Rouhier N, Gelhaye E, Jacquot JP (2004) Plant glutaredoxins: still mysterious reducing systems. *Cell Mol Life Sci* **61**:1266-1277 .

Ruelland E, Miginiac-Maslow M (1999) Regulation of chloroplast enzyme activities by thioredoxin: activation or relief from inhibition? *Trends Plant Sci* **4:** 136–141

Sabatino P, Fermani S, Ripamonti A, Cassetta A, Scagliarini S, Trost P (1999). Crystallization and preliminary X-ray study of chloroplast glyceraldehyde-3-phosphate dehydrogenase. *Acta Crystallog D* **55**: 566–567

Scagliarini S, Trost P, Pupillo P (1998) The non-regulatory isoform of NAD(P)-glyceraldehyde-3-phosphate dehydrogenase from spinach chloroplasts*. J Exp Bot* **49:** 1307–1315

Scagliarini S, Trost P, Pupillo P, Valenti V (1993) Light activation and molecular mass changes of NAD(P)-glyceraldehyde-3-phosphate dehydrogenase of spinach and maize leaves. *Planta* **190:** 313–319

Scheibe R, Baalmann E, Backhausen JE, Rak C, Vetter S (1996) C-terminal truncation of spinach chloroplast NAD(P)-dependent glyceraldehyde-3-phosphate dehydrogenase prevents inactivation and reaggregation. *Biochim Biophys Acta* **1296**: 228–234

Scheibe R, Wedel N, Vetter S, Emmerlich V, Sauermann SM (2002) Co-existence of two regulatory NADP-glyceraldehyde 3-P dehydrogenase complexes in higher plant chloroplasts. *Eur J Biochem* **269:** 5617–5624

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann J (2005) A gene expression map of Arabidopsis development. *Nat Genet* **37**: 501–506

Schürmann P, Jacquot JP (2000) Plant thioredoxin systems revisited. *Annu Rev Plant Physiol Plant Mol Biol* **51:** 371–400

Shih M-C, Goodman HM (1988) Differential light-regulated expression of nuclear genes encoding chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenase in *Nicotiana tabacum*. *EMBO Journal* **7:** 893–898

Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A, Szabo B, Tompa P, Chen J, Uversky VN, Obradovic Z, Dunker AK (2007) DisProt: the Database of Disordered Proteins. *Nucleic Acids Res* **35:** D786-93

Skarzynski T, Moody PCE, Wonacott AJ (1987) Structure of the holo-glyceraldehydde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 Å of resolution. *J Mol Biol* **193**: 171–187

Song S, Li J, Lin Z (1998) Structure of the holo-glyceraldehyde-3-phosphate dehydrogenase from *Palinurus versicolor* refined at 2.0 Å resolution. *Acta Crystallog D* **54**: 558–569

Sparla F, Fermani S, Falini G, Zaffagnini M, Ripamonti A, Sabatino P, Pupillo P, Trost P (2004) Coenzyme site directed mutants of photosynthetic A_4-GAPDH show selectively reduced NADPH-dependent catalysis, similar to regulatory AB-GAPDH inhibited by oxidized thioredoxin. *J Mol Biol* **340**: 1025–1037

Sparla F, Pupillo P, Trost P (2002) The C-terminal extension of glyceraldehyde-3-phosphate dehydrogenase subunit B acts as an autoinhibitory domain regulated by thioredoxins and nicotinamide adenine dinucleotide. *J Biol Chem* **277:** 44946–44952

Sparla F, Tedeschi G, Pupillo P, Trost P (1999) Cloning and heterologous expression of NAD(P)H:quinone reductase of *Arabidopsis thaliana*, a functional homologue of animal DT-diaphorase. *FEBS Lett* **463:** 382–386

Sparla F, Zaffagnini M, Wedel N, Scheibe R, Pupillo P, Trost P (2005) Regulation of photosynthetic GAPDH dissected by mutants. *Plant Physiol* **138:** 2210–2219

Stitt M, Lilley RM, Heldt HW (1982) Adenine nucleotide levels in the cytosol, chloroplasts, and mitochondria of wheat leaf protoplasts. *Plant Physiol* **1982**: 971–977

Süss K-H, Arkona C, Manteuffel R, Adler K (1993) Calvin cycle multienzyme complexes are bound to chloroplast thylakoid membranes of higher plants *in situ. Proc Natl Acad Sci USA* **90:** 5514–5518

Tamoi M, Myazaki T, Fukamizo T, Shigeoka S (2005) The Calvin cycle in cyanobacteria is regulated by CP12 via NAD(H)/NADP(H) ratio under light/dark conditions. *Plant J* **42:** 504–513

Terzaghi WB, Cashmore AR (1995) Light-regulated transcription. *Annual Review of Plant Physiology and Plant Molecular Biology* **46:** 445–474

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* **5:** 123-135

Tompa P, Banki P, Bokor M, Kamasa P, Kovacs D, Lasanda G, Tompa K (2006) Protein-water and protein-buffer interactions in the aqueous solution of an intrinsically unstructured plant dehydrin: NMR intensity and DSC aspects. *Biophys J* **91:** 2243-2249

Tompa P (2005) The interplay between structure and function in intrinsically unstructured proteins. *FEBS Letters* **579:** 3346-3354

Tompa P, Szasz C, Buday L (2005) Structural disorder throws new light on moonlighting. *Trends Biochem Sci* **30**:484–489

Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* **27:** 527-533

Trost P, Fermani S, Marri L, Zaffagnini M, Falini G, Scagliarini S, Pupillo P, Sparla F (2006) Thioredoxin-dependent regulation of photosynthetic glyceraldehyde-3-phosphate dehydrogenase: autonomous vs. CP12-dependent mechanisms. *Photosynth Res* **89:** 1-13

Trost P, Scagliarini S, Valenti V, Pupillo P (1993) Activation of spinach chloroplast glyceraldehyde-3-phosphate dehydrogenase: effect of glycerate 1,3-bisphosphate. *Planta* **190:** 320–326

Uversky VN (2002) Natively unfolded proteins: a point were biology waits for physics. *Protein Sci.* 11:739-756

Uversky VN, Gillespie JR, Fink AL (2000) Why are "natively unfolded" proteins under physiological conditions? *Proteins* **41:** 415-427

Viles JH, Cohen FE, Prusiner SB, Goodin DB, Wright PE, Dyson HJ (1999) Copper binding to the prion protein: structural implications of four identical cooperative binding sites. *Proc Natl Acad Sci U S A* **96:** 2042-2047

Wara-Aswapati O, Kemble RJ, Bradbeer JW (1980) Activation of glyceraldehyde-phosphate dehydrogenase (NADP) and phosphoribulokinase in *Phaseolus vulgaris* leaf extracts involves the dissociation of oligomers. *Plant Physiology* **66:** 34–39

Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* **337**: 635–645

Wedel N, Soll J (1998) Evolutionarily conserved light regulation of Calvin cycle activity by NADPH-mediated reversible phosphoribulokinase/CP12/glyceraldehyde-3-phosphate dehydrogenase complex dissociation. *Proceedings of the National Academy of Sciences, USA* **95:** 9699–9704

Wedel N, Soll J, Paap BK (1997) CP12 provides a new mode of light regulation of Calvin cycle activity in higher plants. *Proc Natl Acad Sci USA* **94:** 10479–10484

Weigel D, Glazebrook J (2002) Arabidopsis: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Weinreb PH, Zhen W, Poon AW, Conway KA, Lansbury PT Jr (1996) NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry* **35:** 13709-13715

Wishart DS, Bigam CG, Yao JFA, Dyson HJ, Oldfield E, Markley JL, Sykes BD (1995) 1H, 13C and 15N chemical shift referencing in biomolecular NMR. *J Biomol NMR* **6:** 135-140

Wolosiuk RA, Ballicora MA, Hagelin K (1993) The reductive pentose phosphate cycle for photosynthetic CO2 assimilation: enzyme modulation. *FASEB J* **7:** 622–637

Wolosiuk RA, Buchanan BB (1976) Studies on the regulation of chloroplast NADP-linked glyceraldehyde 3-phosphate dehydrogenase. *J Biol Chem* **251**: 6456–6461

Wolosiuk RA, Buchanan BB (1978) Activation of chloroplast NADP-linked glyceraldehyde 3-phosphate dehydrogenase by the ferredoxin/thioredoxin system. *Plant Physiology* **61:** 669–671

Wuttke DS, Foster MP, Case DA, Gottesfeld JM, Wright PE (1997) Solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA sequence: determinants of affinity and sequence specificity. *J Mol Biol* **273:** 183-206

Yao J, Dyson HJ, Wright PE (1997) Chemical shift dispersion and secondary structure prediction in unfolded and partly folded proteins. *FEBS Lett* **419:** 285-289

Yang Y, Kwon H-B, Peng H-P, Shih M-C (1993) Stress responses and metabolic regulation of glyceraldehyde-3-phosphate dehydrogenase genes in *Arabidopsis*. *Plant Physiology* **101:** 209–216

Yoon MK, Shin J, Choi G, Choi BS (2006) Intrinsically unstructured N-terminal domain of bZIP transcription factor HY5. *Proteins* **65:** 856-866

Zaffagnini M, Michelet L, Marchand C, Sparla F, Decottignies P, Le Maréchal P, Miginiac-Maslow M, Noctor G, Trost P, Lemaire SD (2007) The thioredoxin-independent isoform of chloroplastic glyceraldehyde-3-phosphate dehydrogenase is selectively regulated by glutathionylation. *FEBS Journal* **274:** 212-226

Zapponi MC, Iadarola P, Stoppini M, Ferri G (1993) Limited proteolysis of chloroplast glyceraldehyde-3-phosphate dehydrogenase (NADP) from *Spinacia oleracea*. *Biological Chemistry Hoppe-Seyler* **374:** 395–402

Ziegler H, Ziegler I (1965) Der Einfluss der Belichtung auf die $NADP^+$-abhängige Glycerinaldehyd-3-phosphat-dehydrogenase. *Planta* **65**: 369–380