S5-023

Some aspects of carboxyl-terminal processing of precursor D1 protein in photosystem II

Y Yamamoto[1, 3], N Inagaki[2], F Sato[3] and K Satoh[1]

[1] *Department of Biology, Faculty of Science, Okayama University, Okayama 700-8530, Japan, kimiyuki@cc.okayama-u.ac.jp*

[2] *National Institute of Agrobiological Science, Tsukuba 305-8602, Japan*

[3] *Division of Integrated Life Science, Graduate School of Biostudies, Kyoto University, Kyoto 606-8502, Japan, yumiya@lif.kyoto-u.ac.jp*

Keywords: C-terminal extension, CtpA, D1 protein, dinoflagellate, photosystem II

Introducion

The cleavage of carboxyl-terminal (C-terminal) extension of precursor D1 protein (pD1) by an endopeptidase (CtpA), a process that is absolutely required for the assembly of water oxidation center (Mn-cluster) in the photosystem II (PSII), is a typical example of nuclear regulation of chloroplast gene expression in eukaryotic organisms. The C-terminal extension, which is cleaved off during the processing, is variable in terms of chain length and is less conserved compared with the mature part, among different classes of organisms ranging from cyanobacteria to higher plants. A piece of evidence difficult to interpret the processing was that *Euglena gracilis* lacks information on the C-terminal extension in its *psbA* gene sequence, and that in truncated mutants of *Synechocystis* (Nixon et al., 1992) and *Chlamydomonas* (Schrader and Johanningmeier, 1992), the absence of C-terminal extension is not fatal to the photoautotrophic growth, although the cleavage is absolutely required if it is present. Thus the meaning of existence of the C-terminal extension has not been established. A recent mixed-culture growth experiment using *Synechocystis* 6803 (Ivleva et al., 2000), however, clearly demonstrated an advantage of the presence of C-terminal extension in the long-term viability. In addition, some other examples of organisms, which have no C-terminal extension in the D1, were recently supplemented.

On the other hand, the enzymatic process of C-terminal cleavage has been analyzed extensively, both *in vivo* and *in vitro*. The involved enzyme has been identified in several organisms and its three-dimensional structure has been determined (Liao et al., 2000). This is classified into a novel type of serine protease with Ser/Lys catalytic dyad, as also supported by our recent results on its inhibitor specificity (Yamamoto et al., 2001) and the site-directed mutational analysis (Inagaki et al., in press). We have thoroughly investigated this enzymatic process using various systems of different integrities (*e.g.*, Yamamoto et al., 2001), in order to understand the mechanism of proteolytic processing of interest. A fortunate situation for the analysis is that approximately a hundred of deduced sequence data are now available on substrate pD1 for a wide variety of organisms ranging from prokaryotic cyanobacteria to higher plants. In the present study, we intended to correlate the results obtained in our enzymatic study both *in vivo* and *in vitro* with the sequence data for pD1 obtained from protein databases, to discuss C-terminal processing from the viewpoint of evolution.

Materials and methods

Most of sequence data for D1 were obtained from SwissProt, PIR, PDB, PRF and GenBank databases. The deduced sequences of C-terminal extension from various organisms are summarized in Table 1. The secondary structure of C-terminal region of D1 was predicted by the GOR IV method (Garnier et al., 1996).

Results and discussion

Cleavage site:

The proteolytic cleavage of pD1 by CtpA occurs at C-side of Ala-344 in spinach (Takahashi et al., 1990) and in *Synechocystis* sp. PCC 6803 (Nixon et al., 1992), as demonstrated by direct protein sequencing of mature D1. The recognition of cleavage site by the enzyme is very strict, as shown by our *in vitro* analysis using substituted synthetic oligopeptides (Taguchi et al., 1995). Since the C-terminal sequence of mature D1 is highly conserved between different organisms, the prediction of cleavage site can be done without ambiguity. Ala at the cleavage site is conserved throughout all D1 proteins deduced from 86 *psbA* gene species (with only one exception in *Alexandrium tamarense*). Ala at this position is required in the substrate recognition and cleavage. On the other hand, a recent structural analysis provided evidence that this residue may participate in the ligation of Mn-cluster, as previously proposed by Nixon et al. (1992).

In most cases, the residue at the C-side of cleavage site is either Ala or Ser; Ala in 52 species and Ser in 21 species, out of 76 gene species collected in this study (three exceptional cases are Val in *Chlorella vulgaris* C-27 (green alga), Asp in *Cyanidium caldarium* (red alga), and Cys in *Gymnodinium mikimotoi* (dinoflagellate)). In our analysis using synthetic oligopeptides, either as the substrate (Taguchi et al., 1995) or the competitive inhibitor (Yamamoto and Satoh, 1998), as well as genetically engineered *Chlamydomonas* (Taguchi et al., 1998), this position was shown to be crucial. The residues found in nature at this position are the ones demonstrated to be efficient in our system, *i.e*., Ala, Ser Cys and Val (except for *Cyanidium*), and the cleavage site tends to form α-helical structure.

Chain length:

The oxygenic photosynthesis organisms can be divided into three groups based on the chain length of C-terminal extension of D1. In group I, the extension consists of 16 amino acid residues (with few exceptions where it is 12, 13 or 15). This group covers both prokaryotic and eukaryotic organisms, *i.e*., cyanobacteria (21 species) and marine algae such as Glaucophyta (1 specie), Cryptophyta (1 specie), Rhodophyta (5 species) and Heterokontophyta (4 species), suggesting phylogenic connection between them. Group II organisms comprise Prochlorophyta and Chlorophyta including land plants and green algae. They are chlorophyll *b*-containing green organisms.

Table 1. Sequence alignment for the predicted C-terminal extension of D1 protein.

Organism	Sequence of C-terminal extension	Length
Prokaryotes		
Cyanophyta (21)		
Synechocystis 6803 (*psbA2*/*3*)	SGEQAPVALTAPAVNG	16
Anabaena 7120 (*psbA2*/*3*/*4*)	AGEVAPVAISAPAING	16
Synechococcus WH7803	AAESTPVALQAPAI-G	15
Cyanothece ATCC51142	SAE---PV-SAPVING	12
Prochlorophyta (3)		
Prochlorococcus marinus	AAESTSVALVAPSI-G	15
Prochlorothrix hollandica (*psbA1*/*2*)	AVK-------APSIIG	9
Eukaryotes		
-Primary endosymbiosis-		
Chlorophyta-land plants (34)		
Spinacia oleracea	AIE-------APSTNG	9
Chlorophyta-green algae (6)		
Chlamydomonas moewusii	AFE-------APSINA	9
Chlamydomonas reinhardtii	S--------TNSSSNN	8
Chlorella ellipsoidea	SVE-------APSI-A	8
Chlorella vulgaris C-27	VVE-------APAVNG	9
Mesostigma viride	SVE-------APAVNG	9
Nephoroselmis olivacea	SVD-------APAVQG	9
Glaucophyta (1)		
Cyanophora paradoxa	SGEVMPVALTAPSINA	16
Rhodophyta (5)		
Odontella sinensis	SGDVLPVALNAPAVNG	16
Palmaria palmata	SGDSCPVALVAPSING	16
Porphyra purpurea	SGESLPVALTAPAVNG	16
Antithamnion sp.	SNESLPLALVAPAING	16
Cyanidium caldarium	DNSLLPVASSSPSINS	16
-Secondary endosymbiosis-		
Cryptophyta (1)		
Guillardia theta	SGESLPVALTAPAVIG	16
Heterokontophyta (4)		
Bumilleriopsis filiformis	AGEVLPVAVSAPAVHA	16
Vaucheria litorea	AGEILPVAVTAPVIAG	16
Ectocarpus siliculosus	SNEILPVAISAPSVVG	16
Heterosigma carterae	SNEVLPVAVNAPAVNG	16
Euglenophyta (1)		
Euglena gracilis	*	0
-Tertiary endosymbiosis-		
Dinophyta (11)		
Amphidinium carterae	*	0
Amphidinium operculatum	*	0
Gonyaulax polyedra	*	0
Heterocapsa niei		0
Heterocapsa pygmaea	*	0
Heterocapsa rotundata		0
Heterocapsa triquetra	*	0
Lingulodinium polyedra	*	0
Prorocentrum micans	*	0
Gymnodinium mikimotoi	*	13
Alexandrium tamarense	*	?
	CAN---CLLSLWPMVGL	

The gaps in the alignment analysis are indicated by the hyphen. Asterisks represent the lack of C-terminal extension. In cases of Cyanophyta and land plants, only typical sequences are shown. The number of gene species is indicated in parentheses.

The C-terminal extension of D1 in this group consists of 9 amino acids (but, 8 amino acids in two green algae, *Chlamydomonas reinhardtii* and *Chlorella ellipsoidea*). Only 3 species have been sequenced in Prochlorophyta; one in *Prochlorococcus* and two in *Prochlorothrix*. In the latter case(s), the chain length of C-terminal extension is 9 as in green algae and land plants, suggesting another line of endosymbiosis with the C-terminal extension of D1 consisting of 9 amino acids, in chlorophyll *b*-containing organisms. In group III, the C-terminal extension of protein is absent. Until recently only *Euglena gracilis* has been classified into this group. However, recent analysis revealed that dinoflagellates have no C-terminal extension in the D1, although there are few exceptions, probably reflecting the heterogenous origin of plastids in these organisms (Takishita et al., 1999; Barbrook and Howe, 2000). It is noteworthy that these organisms have chloroplasts with three envelope membranes thought to be generated by a result of secondary (Euglenophyta) or tertiary (Dinophyta) endosymbiosis. Thus, higher-order endosymbiosis seems to give a pressure to loose the C-terminal extension, although the extension is speculated to be important in protecting the C-terminus of functional importance in the ligation of Mn-cluster.

Acknowledgements

We thank Drs. A. Murakami (Kobe University), A. Uchida (Kyoto University) and M. Yamaguchi (National Research Institute of Fisheries and Environment of Inland Sea) for their helpful discussion.

References

Barbrook AC, Howe CJ (2000) *Mol. Gen. Genet.* **263**, 152-158.
Garnier J, Gibrat JF, Robson B (1996) *Methods Enzymol*. **266**, 540-543.
Inagaki N, Maitra R, Satoh K, Pakrasi HB, *J. Biol. Chem*., in press.
Ivleva NB, Shestakov SV, Pakrasi HB (2000) *Plant Physiol*. **124**, 1403-1411.
Liao D-I, Qian J, Chisholm DA, Jordan DB, Diner BA (2000) *Nature Struc. Biol.* **7**, 749-753.
Nixon PJ, Trost JT, Diner BA (1992) *Biochemistry* **31**, 10859-10871.
Schrader S, Johanningmeier U (1992) *Plant Mol. Biol.* **19**, 251-256.
Taguchi F, Takahashi Y, Satoh K (1998) *Plant Cell Physiol.* **39**, 1324-1329.
Taguchi F, Yamamoto Y, Satoh K (1995) *J. Biol. Chem.* **270**, 10711-10716.
Takahashi Y, Nakane H, Kojima H, Satoh K (1990) *Plant Cell Physiol.* **31**, 273-280.
Takishita K, Nakano K, Uchida A (1999) *Phycol. Res.* **47**, 257-262.
Yamamoto Y, Inagaki N, Satoh K (2001) *J. Biol. Chem.* **276**, 7518-7525.
Yamamoto Y, Satoh K (1998) *FEBS Lett.* **430**, 261-265.