Functional proteomics of circadian expressed proteins from *Chlamydomonas reinhardtii*

Volker Wagner[a], Monika Fiedler[a], Christine Markert[b], Michael Hippler[b], Maria Mittag[a,*]

[a] *Institute of General Botany and Plant Physiology, Friedrich Schiller-University, Am Planetarium 1, D-07743 Jena, Germany*
[b] *Institute of General Botany and Plant Physiology, Friedrich Schiller-University, Dornburger Str. 159, D-07743 Jena, Germany*

Received 30 October 2003; revised 10 January 2004; accepted 12 January 2004

First published online 27 January 2004

Edited by Lukas Huber

Abstract In this study, functional proteomics was successfully applied for the characterization of circadian expressed, basic proteins. For this purpose, we have chosen the green model alga *Chlamydomonas reinhardtii* since its entire nuclear genome is available and it is ideally suited for biochemical enrichment procedures. Proteins from cells harvested during subjective day and night were heparin affinity purified. They were separated by two-dimensional gel electrophoresis suited for basic proteins and analyzed after tryptic digestion by electrospray ionization mass spectrometry. We can show for the first time that the expressions of a protein disulfide isomerase-like protein and a tetratricopeptide repeat protein change in a circadian manner. Interestingly, both proteins are known to be interaction partners in multiprotein complexes including RNA binding proteins.

Key words: Circadian rhythm; Functional proteomics; Protein disulfide isomerase; Tetratricopeptide repeat protein; *Chlamydomonas reinhardtii*

1. Introduction

Circadian rhythms, which regulate a variety of cellular processes, can be found in some prokaryotes and most likely all eukaryotes. They are defined as biological rhythms that persist under constant conditions of light and temperature with a period of about 24 h. The molecular basis of the circadian system has been the object of intensive research for the last 20 years [1–3]. In the so far characterized model systems it has been shown that nucleic acid binding proteins and their temporal involvement in multiprotein complexes represent a key feature of the central oscillator mechanism [1–3]. So far, all characterized clock components have been identified by classical and molecular genetic approaches. However, functional proteome approaches are also suitable to identify such components given that the genome and/or numerous EST (expressed sequence tag) sequences of an organism are available and that it is suitable for efficient biochemical purification procedures.

In this context, the green unicellular alga *Chlamydomonas reinhardtii* emerges as an eukaryotic model system since its entire chloroplast, mitochondrial and nuclear genomes have been sequenced and there are more than 180 000 ESTs available. In *C. reinhardtii*, several circadian rhythms are known and physiologically well characterized including phototaxis, chemotaxis, ultraviolet (UV) sensitivity and cell division (summarized in [4]). Notably, this alga is ideally suited for proteomic analysis with prior purification steps, since it can be easily and quickly grown up in high amounts [5]. First proteome studies on *C. reinhardtii* with high-resolution two-dimensional gel electrophoresis (2-DE) have already been carried out [6].

In order to characterize and identify novel or conserved circadian DNA and RNA binding proteins and/or their interaction partners, the analysis of the heparin binding proteome of *C. reinhardtii* by 2-DE and the identification of circadian expressed proteins by mass spectrometry (MS) seemed a promising approach. Heparin affinity chromatography achieves an enhancement of potential nucleic acid binding proteins including complex partners from a soluble crude extract [7,8]. This is especially important when these proteins are present in low abundance in a cell extract.

Here we have used a 2-DE procedure for the efficient separation of basic proteins. We show that the selected functional proteome approach results in the identification of two proteins (protein disulfide isomerase (PDI) and tetratricopeptide repeat (TPR) protein), whose amounts peak during night phase. Both proteins have been found in multiprotein complexes in other systems partially in association with RNA binding proteins [9–12].

2. Materials and methods

2.1. Cell culture

C. reinhardtii wild-type strain SAG 73.72 was grown in high salt acetate (HSA) medium under a 12-h light/12-h dark cycle with a light intensity of 71 μE per m^2 per s at 24°C and then put under constant conditions of dim light (LL: 15 μE per m^2 per s) before harvesting [13]. The beginning of the dim light period is defined as time zero (LL0). The numbers (LL25, LL29, LL37, LL41) indicate how many hours the cells have been kept under LL conditions. Cells were harvested at different circadian times corresponding to early subjective day (LL25), the middle of the subjective day (LL29), the beginning of

*Corresponding author. Fax: (49)-3641-949202.
E-mail address: m.mittag@uni-jena.de (M. Mittag).

Abbreviations: CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; DTT, dithiothreitol; ESI, electrospray ionization; EST, expressed sequence tag; HOP, heat shock protein (HSP)70/HSP90 organizing protein; HPLC, high-pressure liquid chromatography; IPG, immobilized pH gradient; LC, liquid chromatography; LL, constant conditions of dim light; MS, mass spectrometry; PDI, protein disulfide isomerase; S.E.M., standard error of the mean; SDS, sodium dodecyl sulfate; TCA, trichloroacetic acid; TPR, tetratricopeptide repeat; WD-40 repeat, tandem repeats of about 40 residues, each containing a central Trp[W]-Asp[D] motif; 2-DE, two-dimensional gel electrophoresis

0014-5793/04/$30.00
doi:10.1016/S0014-5793(04)00051-1

the subjective night (LL37) and the middle of the subjective night (LL41).

2.2. Preparation of crude extracts

Crude protein extracts were prepared as described [13] in a buffer of 10 mM phosphate pH 7.0/14 mM dithiothreitol (DTT) and 4% (v/v) Complete™ protease inhibitor cocktail (Roche Molecular Biochemicals).

2.3. Heparin affinity chromatography

3 mg of soluble protein (S 90000) was applied to a FPLC (fast-performance liquid chromatography) heparin affinity chromatography column (HiTrap™, Amersham Biosciences) at a flow rate of 1 ml/min. After sample application, the column was washed with five volumes of starting buffer (10 mM phosphate buffer pH 7/2 mM DTT) and five volumes of washing buffer (150 mM NaCl/10 mM phosphate buffer pH 7/2 mM DTT). Proteins with a high affinity to heparin were then eluted by 10 mM phosphate buffer, pH 7/2 mM DTT/1 M NaCl.

2.4. 2-DE procedure for efficient separation of basic proteins

The high salt-eluted proteins were dialyzed for 3 h against a 500-fold volume of 10 mM phosphate pH 7.0/2 mM DTT at 4°C under constant stirring in QuickSep® micro dialyzing capsules (ROTH) followed by a trichloroacetic acid (TCA)/acetone precipitation. Thereby, one volume of dialyzed sample was mixed with four volumes of 10% TCA dissolved in acetone (−20°C) containing 0.1% DTT and kept at −20°C overnight. Following centrifugation (35 000×*g*, 30 min, 4°C), the supernatant was discarded and the pellet resuspended in acetone containing 0.1% DTT. After a 1-h incubation at −20°C the sample was centrifuged again (35 000×*g*, 30 min, 4°C). The supernatant was discarded and the pellet dried under vacuum. The protein pellet was solubilized for 2 h at room temperature in standard lysis buffer (9 M urea/2% (w/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS)/1% DTT/0.8% (v/v) ampholyte pH 3–10) [14]. The resuspended sample was then centrifuged for 30 min at 26 800×*g*. Rehydration of the IPG (immobilized pH gradient) strip (pH 3–10, 24 cm separation distance) was carried out with 450 µl of a rehydration solution (8 M urea/0.5% (w/v) CHAPS/20 mM DTT/10% (w/v) glycerol/0.2% (v/v) ampholyte pH 3–10) [14]. The rehydrated strip was removed from the strip holder, any excessive rehydration solution was discarded and 150 µg of the resuspended sample was applied into the lateral sample application wells of the strip holder near the anode. Alongside the cathodic end of the IPG strip an extra paper strip soaked with 15 mM DTT was placed on the surface [15], before the rehydrated IPG strip was put again into the strip holder and covered with paraffin. After sample loading, the IPG strip was placed again in the strip holder and isoelectric focussing was started with the following IPGphor program: 200 V (1 h), 500 V (1 h), 1000 V (1 h), linear gradient from 1000 to 8000 V (1 h) and 8000 V (6 h). The current was limited to 50 µA/IPG strip. After the end of this program, the strip was removed from the strip holder and equilibrated for 15 min in solution 1 (6 M urea/50 mM Tris–HCl pH 8.8/2% (w/v) sodium dodecyl sulfate (SDS)/30% (w/v) glycerol/1% (w/v) DTT) and 15 min in solution 2 (6 M urea/50 mM Tris–HCl pH 8.8/2% (w/v) SDS/30% (w/v) glycerol/4% (w/v) iodoacetamide) [14]. The strips were placed on the top of a 10% polyacrylamide gel, which contained 0.02% Na-thiosulfate and was cross-linked with 0.27% piperazine diacrylamide [6]. The strips were fixed on the gel with 0.5% agarose in electrode buffer (0.1% SDS, 25 mM Tris and 0.192 M glycine) and the electrophoresis was performed at 4°C and 2.5 W per gel overnight (~12 h) using an Ettan Dalt 6 electrophoresis unit (Amersham Biosciences). After electrophoresis, silver staining of the gels was carried out according to the manufacturer's instructions (Amersham Biosciences).

2.5. Comparative quantitative analysis of proteins from 2-DE gels

The analyzed heparin-purified proteins were derived from two cultivation and purification procedures for each LL time point. From each of the two purification procedures and per LL time point, proteins were checked in two different 2-DE gels, so that four data sets per LL time point were available for each protein. The analysis of the stained 2-DE gels including the quantitative analysis of the spot volume of all proteins were carried out with the ImageMaster-2D v4.01 software from Amersham Biosciences.

At first, all protein spots in each gel were normalized by the following procedures: The volume of each spot was divided by the total

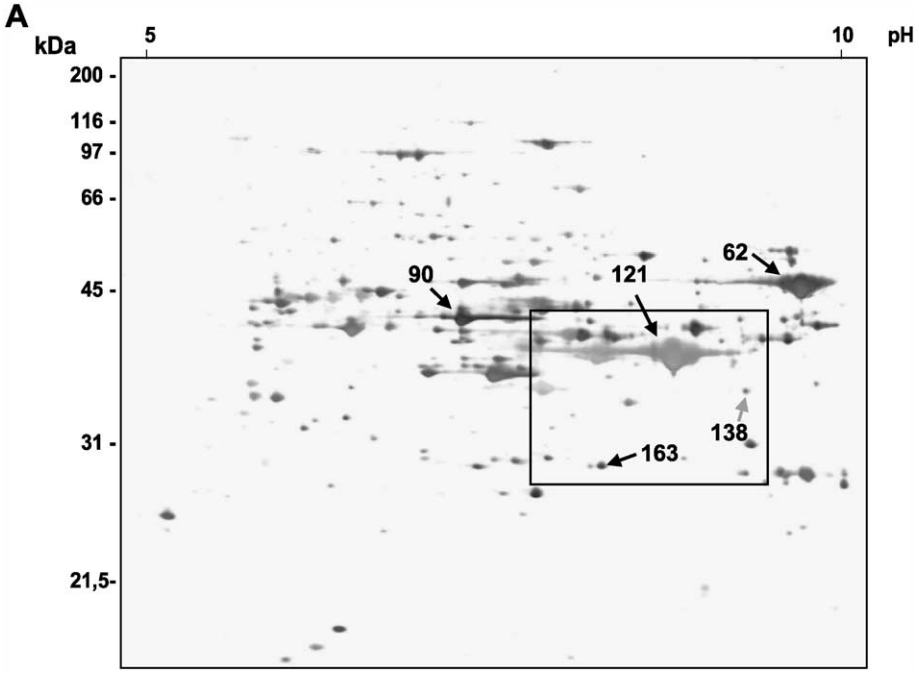

Fig. 1. 2-DE of heparin-bound proteins. A: Heparin-bound proteins were enriched from *C. reinhardtii* crude extracts, which were prepared from cells harvested at LL41 and applied to 2-DE (see Section 2). Proteins were visualized by silver staining. The big box in the figure represents the area where two protein spots (138 and 163) were identified, which change their expression pattern in a circadian manner with an amplitude ≥4 in four independent data sets. The protein spots 62, 90 and 121 have a volume of >200 000, which was taken into account for normalization method B. B: Circadian changes of protein spots 138 and 163: Cells were harvested at LL25, LL29, LL37 and LL41 and further treated as described under A. C: Average strength of expression on the basis of normalized spot volume (method A) for gel spots from four data sets using ImageMaster-2D software. Quantification was carried out as described under Section 2. The values based on method A are also listed in Table 1 (for spot 163) and Table 2 (for spot 138). There, they were compared to values obtained with normalization method B.

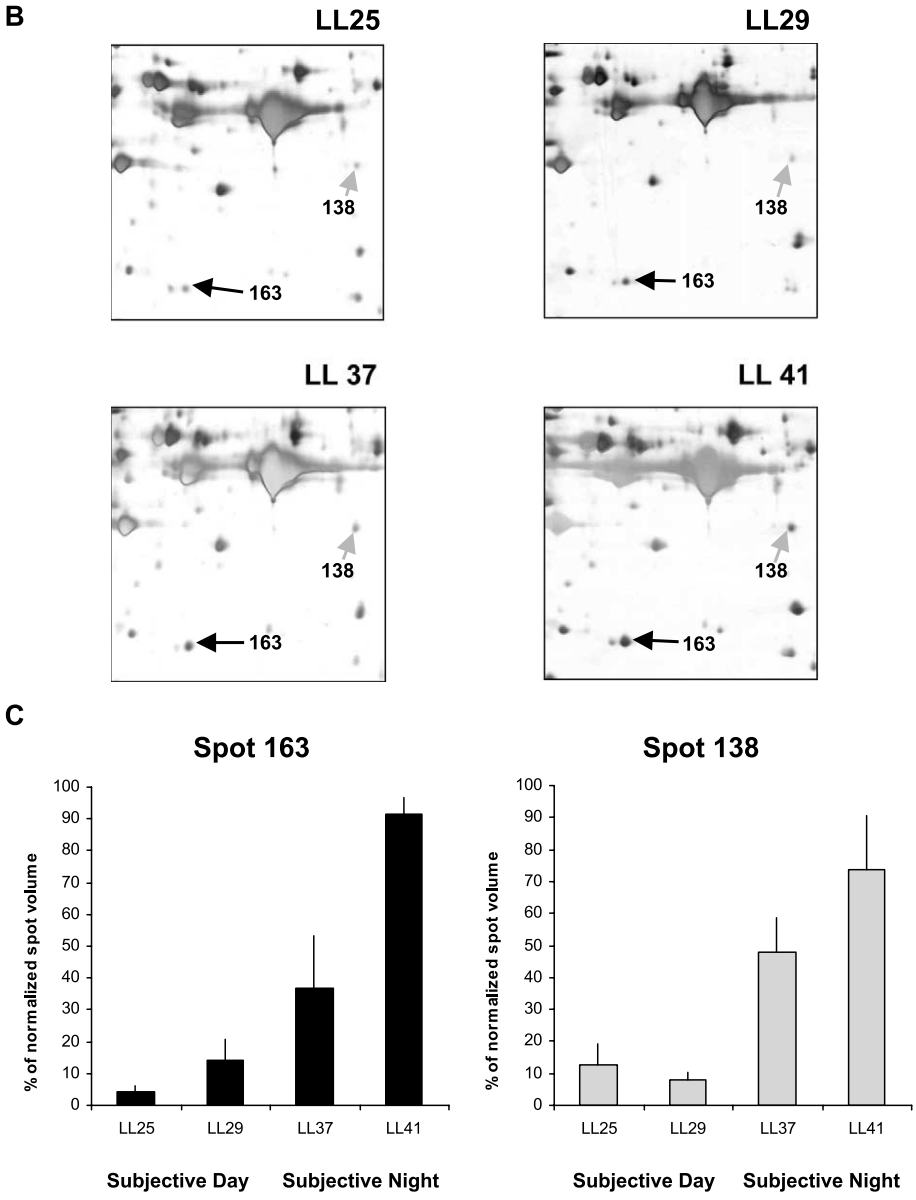

Fig. 1 (*Continued*).

volume of all non-saturated spots in the gel (normalization method A) or by the total volume of all spots having a spot volume less than 200 000 (normalization method B). Saturation was determined by the saturation map feature of the ImageMaster-2D software. As suggested by the ImageMaster-2D manual, the resulting values were then multiplied by a scaling factor representing the total area of all non-saturated spots in the gel (method A) or the total area of all spots with a volume $< 200\,000$ (method B).

Protein spots having a normalized volume, which shows an amplitude of variance ≥ 4 during the chosen circadian cycle (LL25, LL29, LL37, LL41), were further considered. The highest volume of a circadian expressed protein spot with regard to the different LL time points was set to 100%. The average percentage and the standard error of the mean (S.E.M.) of the normalized spot volumes from the four data sets per LL time point were determined for each selected protein.

2.6. Capillary liquid chromatography-electrospray ionization (LC-ESI)-MS analysis

Spots were cut out of silver-stained 2-DE gels, in-gel digested with trypsin and purified by high-pressure liquid chromatography (HPLC) [16]. The HPLC system consisted of a Dionex Ultimate HPLC pump with Nanoflow Setup, a PepMap C18 reversed phase capillary HPLC column and a Dionex Famos autosampler with 1 μl injection possibility, directly connected to the MS. All instruments were controlled via Xcalibur System software (ThermoFinnigan, San Jose, CA, USA) and were operated as described [17].

LC-MS analysis was performed using an ion trap mass spectrometer (LCQ Deca XP, ThermoFinnigan, San Jose, CA, USA) with ESI [17]. For the collision-induced dissociation the collision energy in the ion trap was set to 35% of 5 V. The mass spectrograms were analyzed by using the Sequest program (ThermoFinnigan) and the *C. reinhardtii* EST and genomic databases [17]. Only high scores of peptides ($X_{corr} > 2$ if the charge z of the peptide is 2) were considered. The cross-correlation factor 'X_{corr}' describes the cross-correlation between the experimentally measured MS/MS spectrum and an in situ generated MS/MS spectrum of candidate peptides in the databases [17].

3. Results and discussion

3.1. Identification of heparin-enriched circadian expressed proteins by 2-DE

Regulatory proteins, which are part of the circadian feedback loop, often include nucleic acid binding proteins and

Table 1
Circadian changes of spot 163, based on normalization method A or B

Normalization method (see Section 2.5)	LL25	LL29	LL37	LL41
Method A	4(±2)%	14(±7)%	37(±17)%	92(±5)%
Method B	8(±3)%	15(±7)%	38(±17)%	92(±5)%

The values indicate the percentages of normalized spot volumes as well as the S.E.M.

their interaction partners [1–3]. Such protein complexes can be efficiently enriched by heparin affinity chromatography [7,8].

We used a 2-DE procedure (see Section 2) for high resolution of heparin-purified proteins from *C. reinhardtii* in the pH range from 3 to 10 (Fig. 1A). The 2-DE procedure was applied to heparin affinity-purified proteins from cells harvested at four different LL time points, corresponding to the day (LL25, LL29) and night phase (LL37, LL41) (Fig. 1A, B). Thereby, about 230 protein spots per gel were detectable (Fig. 1A). All protein spot volumes of each gel were normalized according to methods A and B as described in Section 2.5. In method A, the volume of each spot was divided by the total volume of all non-saturated spots. The determination of saturated spots with the ImageMaster-2D software saturation map feature gave the result that no saturated spots were present in the 16 analyzed 2-DE gels. Since some protein spots, which had a spot volume of $>200\,000$, appeared close to saturation levels (spots 62, 90, 121 marked in Fig. 1A), we also used a variation of method A, entitled method B, to check if such spot volumes would have an influence on the results. In this case, the volume of each spot was divided by the total volume of all spots, which had a volume of less than 200 000. Since these types of normalization method tend to produce extremely small values, the results were multiplied by a scaling factor. ImageMaster 2-DE suggests two options for the scaling factor to use. One option is to multiply by a constant factor. The other option, which we have used, is to multiply by the total area of all the evaluated spots in the gel. Gels with more spots would have a higher total spot volume, usually resulting in a lower normalized volume; this method compensates for the differences in spot density.

Normalized volumes of protein spots, which have shown circadian changes with an amplitude of variance ≥ 4 at the different LL time points within the four independent data sets, were further examined. With both normalization methods, two proteins (spots 138 and 163) were found, which met these criteria (Fig. 1B, C; Tables 1 and 2). The corresponding protein of spot 138 has an isoelectric point (IEP) of 9.15 and an average molecular weight of 36.2 kDa. Its amount is peaking in the subjective night (LL41). The lowest expression is present at LL29 in the middle of the subjective day. The other circadian expressed protein (spot 163) has an apparent IEP of 8.13 and an average molecular weight of 29.5 kDa. Its expression increases from the subjective day to the subjective night with a maximum of expression at LL41. Both proteins show a rhythm in their amount under conditions of constant dim light and temperature demonstrating that their expression is controlled by the circadian clock.

The corresponding proteins from spots 138 and 163 were cut out of the gels and tryptic digested. The resulting peptides were separated by reversed phase HPLC and analyzed by LC-ESI-MS.

3.2. Identification of a circadian expressed PDI by MS

For the protein from spot 163, the MS analysis identified two peptides with high score ($X_{corr}>2$ if charge $z=2$; Table 3). They could be found in eight overlapping ESTs and an EST assembly from *C. reinhardtii*. The consistent sequence from these ESTs fits completely into a definite part of the genome sequence belonging to scaffold 2212 and shows a high similarity to a putative PDI from *Arabidopsis thaliana* (Fig. 2). The conserved sequence ***WCGHCK*** is a consensus sequence for PDI-like proteins and the most characteristic part of a thioredoxin fold domain. It represents a redox active site with two vicinal thiols [18]. PDI is known as an oxidoreductase, which assists e.g. in the folding of newly synthesized proteins in the endoplasmic reticulum [19]. It typically catalyzes the formation, reduction, and isomerization of disulfide bonds during protein folding. Interestingly, a PDI, which is expressed in the chloroplast in *C. reinhardtii*, has been already described and was named RB60. It is a major component of the *psbA* mRNA binding protein complex, encoding the D1 protein of photosystem II and is implicated in the redox-responsive regulation of translation in *C. reinhardtii* chloroplasts in a light/dark-dependent manner [9–11].

Sequence comparison of PDI sequences from different organisms shows the highest similarity of the *C. reinhardtii* sequence with the putative PDI from *A. thaliana* and a known PDI from *Nicotiana tabacum* (Fig. 2). Thereby, two mammalian PDIs and the *C. reinhardtii* RB60 seem more closely related to each other than to the other plant PDIs.

3.3. Identification of a circadian expressed TPR protein by MS

For the protein from spot 138, the MS analysis identified several peptides with high score ($X_{corr}>2$ if charge $z=2$; Table 4). They could be found in two overlapping ESTs and an EST assembly. In the genome database, the MS analysis identified four additional peptides, which all are present in scaffold 18. The sequences of the ESTs fit completely into scaffold 18 (from bp 1772 to 2689). The four peptides, which are not present in the ESTs, are encoded downstream of the genomic DNA in scaffold 18 from bp 3936 to 4688.

The mentioned sequences show 61% similarity to the ami-

Table 2
Circadian changes of spot 138, based on normalization method A or B

Normalization method (see Section 2.5)	LL25	LL29	LL37	LL41
Method A	13(±7)%	8(±2)%	48(±11)%	74(±17)%
Method B	19(±6)%	8(±2)%	50(±12)%	74(±17)%

The values indicate the percentages of normalized spot volumes as well as the S.E.M.

Table 3
LC-ESI-MS and bioinformatic analysis of protein spot 163 against the EST and the genomic databases from *C. reinhardtii*

Identified peptides	z charge	MH+ peptide mass	X_{corr}	Sequence data of the peptides	EST clones and genome scaffolds
1	2	1657.5	4.6	(K) LENVHDLTPDNFDK*	av621757, av622593, av621931, av625285, av621911, av634721, av625181, av637653, EST assembly 20021010.8415.1 scaffold 2212*
2	2	1013.5	3.86	(K) LGELVAADPK	

*Not found in scaffold by MS since separated by an intron.

no-terminal TPR domain (amino acid residues 1–72) of protein F20D22.4 from *A. thaliana* (Fig. 3), which is considered as a TPR-containing protein. Examples for well-known TPR-containing proteins are the peroxisomal targeting import receptor (Pex5p) in yeast [20] or the splicing factor Clf1p (crooked neck-like factor 1 protein) [12]. In plants like soybean there are protein–protein interaction modules, called HOP (heat shock protein (HSP)70/HSP90 organizing protein), which contain TPRs that mediate the assembly of HSP70/HSP90 multi-chaperone complexes [21]. Sequence comparison of the translated *C. reinhardtii* EST and the TPR protein from *A. thaliana* with HOPs from three different organisms show a high degree of similarity in the N-terminal TPR domain (Fig. 3), which represents the binding site of HOP to HSP70, at least in humans [22].

TPR proteins have been proposed to interact preferably with WD-40 repeat (tandem repeats of about 40 residues, each containing a central Trp[W]-Asp[D] motif) proteins [23]. Interestingly, the molecular mechanisms of the circadian clocks in *Drosophila melanogaster* and *Neurospora crassa* show a direct involvement of WD-40 proteins in the regulation of the circadian oscillator loop [24,25]. Notably, there are also TPR proteins described in *C. reinhardtii* that are suggested to regulate translation/initiation of chloroplast messages and which are part of multiprotein complexes including RNA [26,27].

```
                  40                  50                  60
     ---------------+-------------------+-------------------+-
 34  (...)  V V A L V G - - - - - - - - - - - A L A L T A H A A G  C.reinhardtii 20021010.8415.1
102  (...)  K G S L E P Q K Y E G - P R N A E A L A E Y V N K E G  A.thaliana pPDI NP.182269
103  (...)  K G S L E P K K Y E G - G R T A E A L A E F V N S E G  N.tabacum PDIpr CAA72092
319  (...)  E E E M T K Y K P E S D E L T A E K I T Q F C H H F L  R.norvegicus NP.037130
318  (...)  E E E M T K Y K P E S E E L T A E R I T E F C H R F L  H.sapiens PDIpr P07237
345  (...)  M E K N K K F R M E G - E F T A D N V A K F A E S V V  C.reinhardtii cpPDI T07927

                  70                  80                  90
     ---------------+-------- ---------+-------------------+-
 53  G G G D D P T I K - - - - L E N - - - - - - V H D L T P D N F  C.reinhardtii 20021010.8415.1
131  G T N V K L A A V - - - - P Q N - - - - - - V V V L T P D N F  A.thaliana pPDI NP.182269
132  G T N V K I A S T - - - - P S S - - - - - - V V V L S P D N F  N.tabacum PDIpr CAA72092
349  E G K I K P H L M S Q E L P E D W D K Q P V K V L V G K N F  R.norvegicus NP.037130
348  E G K I K P H L M S Q E L P E D W D K Q P V K V L V G K N F  H.sapiens PDIpr P07237
374  D G T A Q A V L K S E A I P E D P Y E D G V Y K I V G K T V  C.reinhardtii cpPDI T07927

                 100                 110                 120
     ---------------+-------------------+-------------------+-
 74  D K I V N G - A K H V L I E F Y A P W C G H C K R M V P E Y  C.reinhardtii 20021010.8415.1
152  D E I V L D Q N K D V L V E F Y A P W C G H C K S L A P T Y  A.thaliana pPDI NP.182269
153  D E I V L D E T K D V L V E F Y A P W C G H C K S L A P I Y  N.tabacum PDIpr CAA72092
379  E E V A F D E K K N V F V E F Y A P W C G H C K Q L A P I W  R.norvegicus NP.037130
378  E D V A F D E K K N V F V E F Y A P W C G H C K Q L A P I W  H.sapiens PDIpr P07237
404  E S V V L D E T K D V L L E V Y A P W C G H C K K L E P I Y  C.reinhardtii cpPDI T07927

                 130                 140                 150
     ---------------+-------------------+-------------------+-
103  K K L G E L V A A D P K L K N R V V V A K V N A D A H R S L  C.reinhardtii 20021010.8415.1
182  E K - - - - V A T V F K Q E E G V V I A N L D A D A H K A L  A.thaliana pPDI NP.182269
183  E K - - - - V A T S F K Q E E D V V I A N L D A D K H R D L  N.tabacum PDIpr CAA72092
409  D K - - - - L G E T Y K D H E N I V I A K M D S T A N - - E  R.norvegicus NP.037130
408  D K - - - - L G E T Y K D H E N I V I A K M D S T A N - - E  H.sapiens PDIpr P07237
434  K K - - - - L A K R F K K V D S V I I A K M D G T E N - - E  C.reinhardtii cpPDI T07927

                 160                 170
     ---------------+-------------------+-
133  G D K F D V R G F P T I K F F P A G K P A T (...)      285  C.reinhardtii 20021010.8415.1
208  G E K Y G V S G F P T L K F F P K D N K A G (...)      361  A.thaliana pPDI NP.182269
209  G E K Y G V S G F P T L K F F P K G N K A G (...)      359  N.tabacum PDIpr CAA72092
433  V E A V K V H S F P T L K F F P A S A D R T (...)      508  R.norvegicus NP.037130
432  V E A V K V H S F P T L K F F P A S A D R T (...)      508  H.sapiens PDIpr P07237
458  H P E I E V K G F P T I L F Y P A G S D R T (...)      533  C.reinhardtii cpPDI T07927
```

Fig. 2. Alignment of the new identified *C. reinhardtii* PDI from EST assembly 20021010.8415.1 with different PDI sequences spanning the location of the second thioredoxin-like domain. The underlined sequences represent the peptides that were identified by LC-ESI-MS, the residues in bold are highly conserved among all PDIs. The sequence ***WCGHCK*** is a thioredoxin-like active site. pPDI: putative PDI; PDIpr: PDI precursor; cpPDI: chloroplast PDI (also named RB60).

Table 4
LC-ESI-MS and bioinformatic analysis of protein spot 138 against the EST and the genomic databases from *C. reinhardtii*

Identified peptides	z charge	MH+ peptide mass	X_{corr}	Sequence data of the peptides	EST clones and genome scaffolds
1	2	1348.9	3.27	(-) ADPENAVLYSNR*	av631788, av642168, EST assembly 20021010.379.3 scaffold 18*
2	2	1052.0	3.05	(-) AAALYTQAIK	
3	2	958.3	2.31	(-) LRPEWEK	
4	2	1792.8	4.99	(-) PTPVSAEEFGPLPPLIK	scaffold 18
5	2	1169.5	4.46	(-) MAAAAVLAIIPK	
6	2	1417.9	3.34	(-) HTWFMPIATEGK	
7	2	1280.6	2.91	(-) ATVAYPQTWSR	

*Not found in scaffold by MS since separated by an intron.

Acknowledgements: We thank the DFG for supporting our research. We are grateful for information supplied by the *C. reinhardtii* genome project of DOE, USA.

References

[1] Dunlap, J.C. (1999) Cell 96, 271–290.
[2] Harmer, S.L., Panda, S. and Kay, S.A. (2001) Annu. Rev. Cell Dev. Biol. 17, 215–253.
[3] Panda, S., Hogenesch, J.B. and Kay, S.A. (2002) Nature 417, 329–335.
[4] Mittag, M. and Wagner, V. (2003) Biol. Chem. 384, 689–695.
[5] Harris, E. (2001) Annu. Rev. Plant Physiol. Plant Mol. Biol. 52, 363–406.
[6] Hippler, M., Klein, J., Fink, A., Allinger, T. and Hoerth, P. (2001) Plant J. 28, 595–606.
[7] Lacombe, E., Doorselaere, J.V., Boerjan, W., Boudet, A.M. and Grima-Pettenati, J. (2000) Plant J. 23, 663–676.

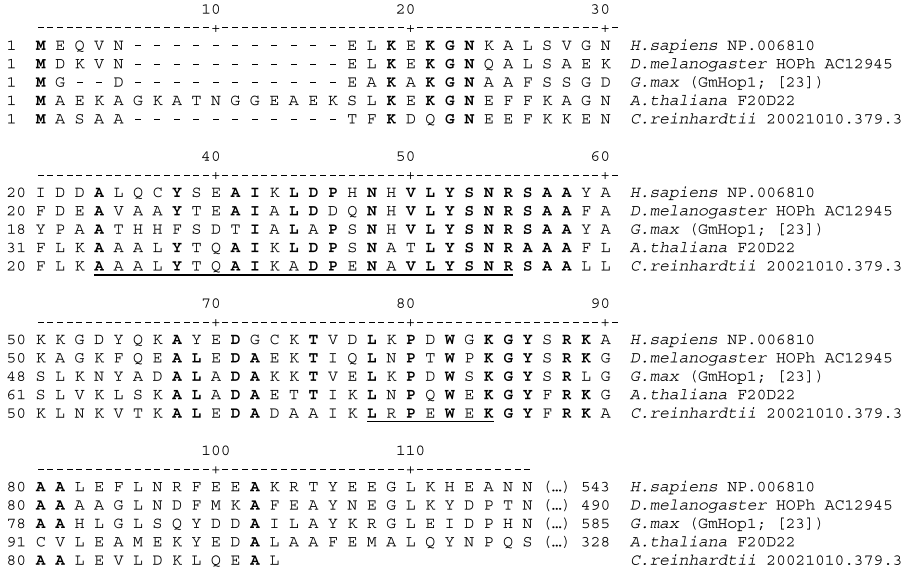

Fig. 3. Alignment of the translated *C. reinhardtii* EST assembly 20021010.379.3 with the N-terminal TPR domains of the protein F20D22.4 from *A. thaliana* and HOPs from different species. The underlined sequences represent the peptides that were identified by LC-ESI-MS. The residues in bold are highly conserved amino acids. HOPh: HOP homolog.

[8] Liu, J., Shen, X., Nguyen, V.A., Kunos, G. and Gao, B. (2000) J. Biol. Chem. 275, 11846–11851.
[9] Danon, A. and Mayfield, S.P. (1994) EMBO J. 13, 2227–2235.
[10] Kim, J. and Mayfield, S.P. (1997) Science 278, 1954–1957.
[11] Trebitsh, T., Meiri, E., Ostersetzer, O., Adam, Z. and Danon, A. (2001) J. Biol. Chem. 276, 4564–4569.
[12] Wang, Q., Hobbs, K., Lynn, B. and Rymond, B.C. (2003) J. Biol. Chem. 278, 7875–7883.
[13] Mittag, M. (1996) Proc. Natl. Acad. Sci. USA 93, 14401–14404.
[14] Görg, A., Obermaier, Ch., Boguth, G., Csordas, A., Diaz, J.-J. and Madjar, J.-J. (1997) Electrophoresis 18, 328–337.
[15] Görg, A., Boguth, G., Obermaier, Ch., Posch, A. and Weiss, W. (1995) Electrophoresis 16, 1079–1086.
[16] Mortz, E., Vorm, O., Mann, M. and Roepstorff, P. (1994) Biol. Mass Spectrom. 23, 249–261.
[17] Stauber, E.J., Fink, A., Markert, C., Kruse, O., Johanningmeier, U. and Hippler, M. (2003) Eukaryot. Cell 2, 978–994.
[18] Åslund, F. and Beckwith, J. (1999) J. Bacteriol. 181, 1375–1379.
[19] Freedman, R.B., Hirst, T.R. and Tuite, M.F. (1994) Trends Biochem. Sci. 19, 331–336.
[20] Gurvitz, A., Wabnegger, L., Langer, S., Hamilton, B., Ruis, H. and Hartig, A. (2001) Mol. Genet. Genomics 265, 276–286.
[21] Zhang, Z., Quick, M.K., Kanelakis, K.C., Gijzen, M. and Krishna, P. (2003) Plant Physiol. 131, 525–535.
[22] Chen, S., Prapapanich, V., Rimerman, R.A., Honoré, B. and Smith, D.F. (1996) Mol. Endocrinol. 10, 682–693.
[23] van der Voorn, L. and Ploegh, G.H. (1992) FEBS Lett. 307, 131–134.
[24] Ko, H.W., Jiang, J. and Edery, I. (2002) Nature 420, 673–678.
[25] He, Q., Cheng, R., Yang, Y., He, Q., Yu, H. and Liu, Y. (2003) EMBO J. 22, 4421–4430.
[26] Boudreau, E., Nickelsen, J., Lemaire, S.D., Ossenbuhl, F. and Rochaix, J.D. (2000) EMBO J. 19, 3366–3376.
[27] Vaistij, F.E., Boudreau, E., Lemaire, S.D., Goldschmidt-Clermont, M. and Rochaix, J.D. (2000) Proc. Natl. Acad. Sci. USA 97, 14813–14818.