RESEARCH ARTICLE

iTRAQ-based quantitative proteomic analysis of a toxigenic dinoflagellate *Alexandrium catenella* and its non-toxic mutant

Shu-Fei Zhang, Yong Zhang, Zhang-Xian Xie, Hao Zhang, Lin Lin and Da-Zhi Wang

State Key Laboratory of Marine Environmental Science/College of the Environment and Ecology, Xiamen University, Xiamen, P. R. China

Received: May 11, 2015
Revised: September 1, 2015
Accepted: September 22, 2015

Paralytic shellfish toxins (PSTs) are a group of potent neurotoxic alkaloids produced by cyanobacteria and dinoflagellates. The PST biosynthesis gene cluster and several toxin-related proteins have been unveiled in cyanobacteria, yet little is known about dinoflagellates. Here, we compared the protein profiles of a toxin-producing dinoflagellate *Alexandrium catenella* (ACHK-T) and its non-toxic mutant (ACHK-NT), and characterized differentially displayed proteins using a combination of the iTRAQ-based proteomic approach and the transcriptomic database. Totally 3488 proteins were identified from *A. catenella*, and proteins involved in carbohydrate, amino acid and energy metabolism were the most abundant. Among them, 185 proteins were differentially displayed: proteins involved in amino acid biosynthesis, protein and carbohydrate metabolism and bioluminescence were more abundant in ACHK-T, while proteins participating in photosynthesis, fatty acid biosynthesis, and the processes occurring in peroxisome displayed higher abundances in ACHK-NT. Seven toxin-related proteins were identified but they varied insignificantly between the two strains. Different carbon and energy utilization strategies were potentially related to the toxin producing ability, and the regulation mechanism of PST biosynthesis was more complex in dinoflagellates. Our study provides the first comprehensive dataset on the dinoflagellate proteome and lays the groundwork for future proteomic study.

Keywords:
Alexandrium catenella / Dinoflagellate / iTRAQ / Microbiology / Paralytic shellfish toxin / Quantitative proteomics

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Paralytic shellfish toxins (PSTs) are a group of potent neurotoxic alkaloids responsible for paralytic shellfish poisonings (PSPs) around the world. In the past few decades, PSPs have attracted more and more concern due to the global increase in the frequency, intensity and geographic distribution of PST-producing algal blooms [1].

PSTs are synthesized by two different kingdoms of life, cyanobacteria and dinoflagellates. Much effort has been devoted to unveil the biosynthesis mechanism of PSTs in both kingdoms. The gene cluster encoding 26 toxin-related

Correspondence: Dr. Da-Zhi Wang, State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen 361005, P. R. China
E-mail: dzwang@xmu.edu.cn
Fax: +86-592-2186055

Abbreviations: **ACHK-NT**, non-toxic mutant of *Alexandrium catenella*; **ACHK-T**, toxin-producing dinoflagellate *A. catenella*; **KEGG**, Kyoto Encyclopedia of Genes and Genomes; **NCBInr**, NCBI non-redundant protein database; **PSP**, paralytic shellfish poisoning; **PST**, paralytic shellfish toxin

Colour online: See the article online to view Figs. 1–4 in colour.

Significance of the study

PSTs are a group of potent neurotoxic alkaloids produced by cyanobacteria and dinoflagellates. However, in contrast with cyanobacteria, little is known about the PST biosynthesis of dinoflagellates. This study, for the first time, applied the iTRAQ-based proteomic approach combining with the transcriptomic database to compare the protein profiles of a toxin-producing dinoflagellate *Alexandrium catenella* (ACHK-T) and its non-toxic mutant (ACHK-NT). The results indicated that processes involved in synthesis of various precursors or substrates for toxin production were strengthened in ACHK-T while more carbon and energy were shifted to fatty acid biosynthesis instead of toxin production in ACHK-NT. Different carbon and energy utilization strategies were potentially related to the toxin producing ability, and PST biosynthesis was more complex in dinoflagellates. Our study provides new insight into the PST biosynthesis in dinoflagellates and lays the groundwork for future proteomic study.

proteins has been characterized and several toxin-related proteins have been identified in cyanobacteria [2, 3]. However, PST biosynthesis in the dinoflagellates remains obscure owing to their massive genome and unique chromosomal characteristics [4]. So far, some putative homologs of cyanobacterial PST synthesis genes have been identified in dinoflagellates [5], however, only *sxtA* and *sxtG* participating in the first two PST-biosynthetic steps are characterized [6, 7].

Many genes in dinoflagellates are post-transcriptionally regulated, and proteins rather than genes are the actual executant of cellular processes [8]. This suggested that the study of proteins might help to uncover the toxin biosynthesis mechanism in dinoflagellates. Global techniques such as proteomics have been applied to mine the proteins involved in toxin biosynthesis [9–11]. However, the inherent drawbacks of the traditional gel-based proteomic approach and the severe lack of gene and protein information of the dinoflagellates limit large-scale protein identification [12]. Only a few proteins are putatively identified and the regulation mechanism of PST biosynthesis in the dinoflagellates remains unclear [13]. Recently an alternative, non-gel based quantitative proteomic method, iTRAQ, was developed, which quantifies proteins on the basis of peptide labeling and allows large-scale identification and accurate quantification of proteins from multiple samples within broad dynamic ranges of protein abundance [14]. This approach is applied to mine toxin-related proteins in a cyanobacterium [3].

In our previous studies, a non-toxic mutant of the toxigenic *Alexandrium catenella* (ACHK-NT) is obtained [10]. Transcriptomes of ACHK-NT and its parental toxic strain (ACHK-T) are compared and 66 812 unigenes are annotated [15], which partly makes up for the lack of the dinoflagellate genome and provides a database for protein identification. Here, we compared the protein profiles of ACHK-T and ACHK-NT, and identified differentially displayed proteins using a combination of the iTRAQ approach and the transcriptomic database. The purpose of this study was to mine toxin-related proteins through large-scale protein identification and quantitative comparison, and unveil those biological processes potentially related to PST biosynthesis.

2 Materials and methods

2.1 Organisms and culture conditions

Cultures of ACHK-T and ACHK-NT were grown in K medium [16] at 20°C. An irradiance of approximately 100 $\mu E/m^2 \cdot s$ was provided using cool white fluorescent bulbs under a 14: 10 h light: dark photoperiod. Cells of ACHK-T and ACHK-NT in the mid-exponential growth phase were harvested for proteomic analysis as previously described [15].

2.2 Protein preparation

Protein extraction was conducted following a previously reported protocol [10]. Briefly, approximately 1×10^6 cells of each strain were harvested in the mid-exponential growth phase, and protein was extracted using TRIzol, chloroform and ethanol and precipitated using isopropanol. The protein obtained was briefly washed with ethanol and air dried. After being rehydrated and alkylated, the protein was dissolved and centrifuged, and the supernatant was used for protein quantification and subsequent iTRAQ labeling.

2.3 Peptide labeling

For 4-plex iTRAQ labeling, ACHK-T and ACHK-NT were compared using two biological replicates for each. A total of 100 μg of protein from each sample was digested using Trypsin Gold (Promega, Madison, WI, USA) with the ratio of protein: trypsin = 30: 1 at 37°C for 16 h. Digested samples were dried using a vacuum centrifuge and reconstituted in 0.5 M tetraethyl-ammonium bromide. Each digest sample was labeled following the manufacturer's instructions: Tag_{113}, ACHK-NT1; Tag_{115}, ACHK-NT2; Tag_{117}, ACHK-T1 and Tag_{119}, ACHK-T2. After 2 h incubation, labeled samples were pooled and dried using vacuum centrifugation.

2.4 Cation exchange fractionation

The dried samples were reconstituted with 4 mL buffer A (25 mM NaH_2PO_4 in 25% ACN, pH 2.7) and were used for cation exchange fractionation using an LC-20AB HPLC pump system (Shimadzu, Kyoto, Japan). The peptides were eluted at a flow rate of 1 mL/min with a gradient of solvent A for 10 min, 5–35% solvent B (25 mM NaH_2PO_4, 1 M KCl in 25% ACN, pH 2.7) for 30 min, 35–80% buffer B for 1 min. Elution was monitored by measuring the absorbance at 214 nm, and the eluted peptides were pooled into 12 fractions, desalted with a Strata X C18 column (Phenomenex) and vacuum-dried, and then reconstituted in solvent C (5% ACN, 0.1% formic acid).

2.5 LC-MS/MS analysis

All peptide samples were separated on a nanoACQuity (Waters) system and analyzed on an AB SCIEX TripleTOF 5600 System. After sample injection, peptides were separated at an analytical flow rate of 300 nL/min with a segmented gradient of solvent D (95% ACN, 0.1% formic acid) at 5% for 1 min, 5 to 35% for 40 min, 35 to 80% for 5 min, at 80% for 5 min. Initial chromatographic conditions were restored in 2 min.

The MS was operated with an RP of ≥30 000 FWHM for TOF MS scans. For IDA, survey scans were acquired in 250 ms and as many as 30 product ion scans were collected if exceeding a threshold of 120 counts per second and with a 2+ to 5+ charge-state. Total cycle time was fixed to 3.3 s. Q2 transmission window was 100 Da for 100%.

2.6 Bioinformatics analysis

In our previous study, transcriptomes of ACHK-T and ACHK-NT are sequenced and the corresponding unigenes are generated [15]. In the present study, the amino acid sequences translated from the CDS of unigenes were used as the protein database.

MASCOT genetic format files converted from raw data files using Proteome Discoverer 1.2 (PD 1.2, Thermo) were searched and protein identification was performed using the MASCOT search engine (Matrix Science; version 2.3.02). The search parameters were as follows: fragment mass tolerance at 0.1 Da; peptide mass tolerance at 0.05 Da; trypsin as the enzyme, allowing up to one missed cleavage; and peptide charges of 2+ and 3+. iTRAQ labeling and carbamidomethylation were defined as fixed modifications. Gln→pyro-Glu (N-term Q), Oxidation (M), Deamidated (NQ) were the potential variable modifications. Specifically, an automatic decoy database search was performed to estimate the FDR. Peptides at the 95% confidence interval were counted as identified, and each identified protein involved at least one unique peptide.

Functional annotations of the proteins were conducted using the Blast2GO program against the KEGG and NCBInr.

For protein quantitation, the quantitative ratios were weighted and normalized using the median ratio in Mascot. A protein was considered to be differentially displayed if it contained at least two unique peptides, had similar ratios and at least one significant ratio of ≥1.2 (higher abundance in ACHK-T) or ≤0.83 (higher abundance in ACHK-NT) in ACHK-T1/ACHK-NT1 and ACHK-T2/ACHK-NT2.

2.7 Identification of toxin-related proteins

To identify toxin-related proteins, BLAST was performed with toxin-related protein sequences of the cyanobacterium *Cylindrospermopsis raciborskii* T3 [2] against the identified proteins. All hits with an e-value ≤ $1E^{-5}$ were retrieved as confident identifications. The long and short isoforms of *sxtA* are identified as comp66169_c0 and comp20666_c0, and *sxtG* as comp63408_c0, in our previous transcriptomic study [15], so we directly searched the three gene accessions in our protein database.

3 Results

3.1 Proteome overview

In total, 13 961 of the whole output 1 61 251 spectra were matched to 7662 peptides with an approximately 8.7% spectrum utilizing rate. Using the Mascot search engine, 3488 proteins were identified from 7476 unique peptides that collectively matched 13 431 unique spectra with a FDR of 1.5%. The molecular weight of most proteins was distributed between 20 and 70 kDa (Fig. 1A) and 59% of them possessed sequence coverage higher than 5% (Fig. 1B). Reproducibility assessment showed that 50% variation covered 91% of the proteins in ACHK-NT2/ACHK-NT1 and 93% in ACHK-T2/ACHK-T1 (Fig. 1C), which was similar to previous studies using the iTRAQ-based proteomic approach [17, 18].

3.2 Protein functional annotation

Of the 3488 proteins identified, photosynthesis related proteins, such as peridinin chlorophyll-a binding protein, light-harvesting protein and glyceraldehyde-3-phosphate dehydrogenase were the most abundant based on spectrum number (Supporting Information data 1).

Based on the KEGG categories, 1648 proteins were annotated into 17 groups. "Carbohydrate metabolism", "Amino acid metabolism" and "Energy metabolism" ranked as the top three abundant categories in "Metabolism", while "Translation" and "Folding sorting and degradation" dominated the "Genetic Information Processing" (Fig. 2A). Further

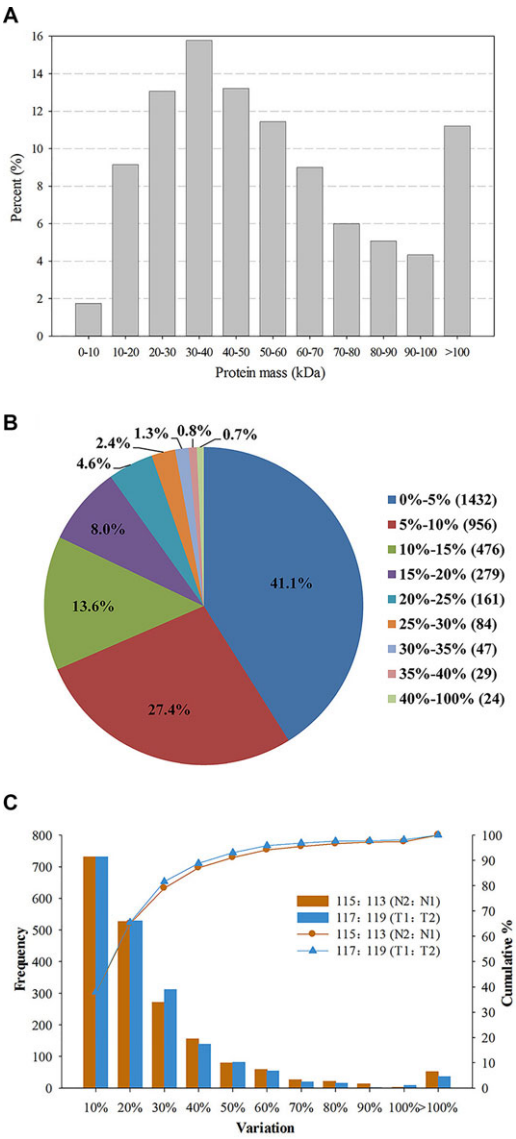

Figure 1. Overview of all proteins identified in *A. catenella*. (A) Molecular weight distribution; (B) coverage of proteins by the identified peptides; (C) frequency distribution (bars) and cumulative percentage (lines) of proteins. The former is from both ACHK-T and ACHK-NT two replicates across different ranges of variations, the latter is defined as the cumulative number of proteins falling within the defined variation range against the total number of proteins.

classification of the third pathway hierarchy showed that "Ribosome", "Protein processing in endoplasmic reticulum" and "Spliceosome" were the most frequently detected pathways (Fig. 2B).

3.3 Differentially displayed proteins

A total of 185 differentially displayed proteins were identified. Among them, 106 were more abundant in ACHK-T while 79 displayed higher abundances in ACHK-NT. Functional classifications of these proteins based on KEGG and NCBInr annotations are presented in Supporting Information Table 2 and Fig. 3. It should be pointed out that a large number of proteins were annotated as unknown functions.

The proteins with higher abundances in ACHK-T fell into 10 major biological process groups (Fig. 3A). Among them, 11 were involved in amino acid metabolism and three in purine biosynthesis. Seven proteins were assigned to carbohydrate metabolism, especially the glycolysis process. Moreover, two luciferase (LCF) isoforms and one luciferin-binding protein (LBP) involved in bioluminescence were more abundant in ACHK-T.

The proteins with higher abundances in ACHK-NT were classified into eight major biological process groups (Fig. 3B). Six proteins were involved in photosynthesis, including chlorophyll a/b binding proteins and photosystem II proteins. Interestingly, quite a number of proteins participated in the processes occurring in peroxisome, including carbon metabolism of photorespiration, fatty acid metabolism, glyoxylate cycle and reactive oxygen species (ROS) metabolism. Moreover, 13 ribosome subunit proteins and two isoforms of rRNA biogenesis protein RRP5 also displayed higher abundances.

3.4 Toxin-related proteins

With a cutoff e-value $\leq 1E^{-5}$, a total of 21 homologs of seven toxin-related proteins were obtained, sxtG, sxtH, sxtT, sxtU, sxtO, sxtZ and ompR (Table 1). As for toxin-related genes found in dinoflagellates, only the product of *sxtG* (comp63408_c0) annotated as an amidinotransferase was detected. However, none of them varied significantly between the toxic and non-toxic strains.

4 Discussion

This study compared the protein profiles of ACHK-T and ACHK-NT, and identified 185 differentially displayed proteins participating in various biological processes (Fig. 4). Variations of these processes revealed different carbon and energy utilization strategies, which might be related to the different toxin-producing abilities between these two strains.

4.1 Enhanced biological processes in ACHK-T

4.1.1 Amino acid and purine biosynthesis

Previous studies show that S-adenosylmethionine (SAM), arginine and acetyl-CoA are the substrates essential for PST biosynthesis [2, 19]. In toxic cyanobacteria, SAM provides

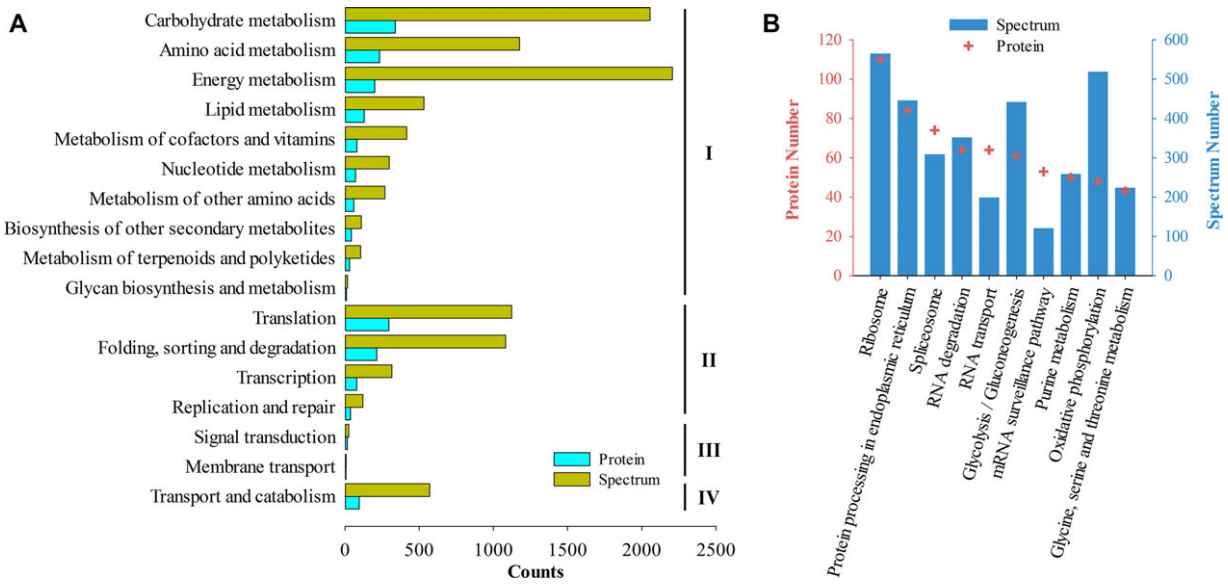

Figure 2. (A) KEGG classification of the annotated proteins and corresponding spectra based on secondary pathway hierarchy: (I) metabolism; (II) genetic information and processing; (III) environmental information and processing; (IV) cellular processes; (B) top ten abundant KEGG pathways of the annotated proteins.

a methyl-group to acetyl-CoA, converting it to propionyl-CoA [2]. In our study, SAM synthetase, an enzyme involved in SAM synthesis, displayed higher abundance in ACHK-T with a fold change of 1.278, indicating the active SAM biosynthesis occurring in ACHK-T, which needs more methionine as substrate. Thus, as expected, vitamin-B12 independent methionine synthase and 5-methyltetrahydrofolate-homocysteine methyltransferase were also more abundant in ACHK-T (fold change >1.3). They catalyze the final step of methionine biosynthesis with or without cobalamin as a cofactor [20]. Moreover, a bifunctional protein involved in arginine biosynthesis, glutamate N-acetyltransferase/amino-acid N-acetyltransferase and another seven proteins participating in the biosynthesis of serine family amino acids, aromatic amino acids and lysine were also more abundant in ACHK-T [20], indicating an enhanced amino acid biosynthesis in ACHK-T. Overall, higher abundances of these proteins in ACHK-T might provide more substrates for toxin synthesis.

Along with this, a more active purine biosynthesis might also occur for three enzymes: phosphorribosylformylglycinamidine synthase (purL), phosphoribosylformylglycinamidine cyclo-ligase (purM) and adenylosuccinate synthase (purA) were more abundant in ACHK-T (fold change >1.2). PurL and purM catalyze the fourth and fifth step of inosine monophosphate (IMP) synthesis while purA catalyzes one of the two steps from IMP to adenosine monophosphate, producing adenylosuccinate (SAMP) [21]. It is noteworthy that some intermediates of purine biosynthesis such as SAMP, share a similar structural formula with the PSTs. Thus, we could not exclude the possibility that an enhanced purine biosynthesis might provide more intermediates for PST biosynthesis in ACHK-T.

4.1.2 Carbohydrate metabolism

Carbohydrate metabolism not only generates energy for cells, but also provides various intermediates for the biosynthesis of other cellular compounds. In this study, four enzymes involved in the glycolysis displayed higher abundances in ACHK-T (fold change >1.2): glucose-6-phosphate 1-epimerase, triosephosphate isomerase, enolase and pyruvate kinase. These enzymes catalyze the production of glyceraldehyde-3-phosphate, phosphoenolpyruvate and pyruvate [22] which are vital substrates for the biosynthesis of serine, aromatic family amino acids, pyruvate family amino acids and nucleic acids [20]. In addition, the 2-methylcitrate dehydratase family domain containing protein, flavoprotein subunit and iron-sulfur subunit of succinate dehydrogenase were also more abundant in ACHK-T. These proteins participate in the biosynthesis of numerous compounds including amino acids and purine [22]. Therefore, higher abundances of these proteins in ACHK-T might provide more substrates for amino acid and purine biosynthesis, and subsequently for toxin biosynthesis.

4.1.3 Protein metabolism

In this study, quite a number of the higher abundance proteins in ACHK-T were related to protein metabolism processes, including transcription, translation, PTM and protein degradation. Among them, nucleolin, rRNA biogenesis protein RRP5, splicing factors or subunits and polyadenylate-binding protein participate mainly in the synthesis and maturation of ribosomes and mRNA involved in the protein

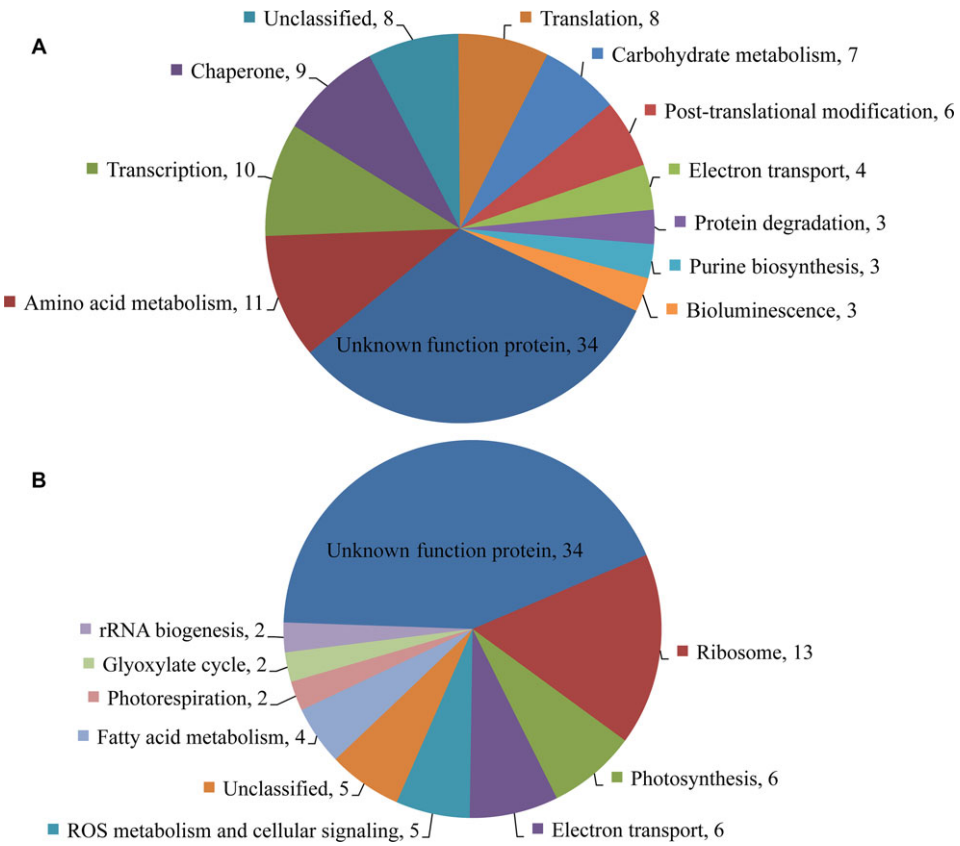

Figure 3. Functional distribution of differentially displayed proteins between ACHK-T and ACHK-NT. (A) Proteins with higher abundances in ACHK-T; (B) proteins with higher abundances in ACHK-NT.

translation [23], while some factors, such as translation initiation factor 5B, elongation factor Tu and 1-gamma, peptide chain release factor, cysteinyl-tRNA synthetase and ribosome recycling factor are essential for protein translation [24], indicating an enhanced protein translation in ACHK-T. Proteins involved in protein processes in endoplasmic reticulum such as PTM and protein degradation also displayed higher abundances, suggesting that PTM and degradation of protein were more active in ACHK-T, which coincided with the strengthened protein translation. In addition, several molecular chaperones such as FK506-binding protein 4/5, DnaJ homolog subfamily proteins, molecular chaperone GrpE, HSP and chaperonin GroEL displayed higher abundances in ACHK-T (fold change >1.3). These chaperones regulate various biological processes especially the protein folding processes [25]. Overall, higher abundances of these proteins suggested more active protein processes in ACHK-T, which might be an adaptive response to active toxin biosynthesis, since PST accumulation and protein biosynthesis present a direct correlation in toxic *A. fundyense* [26].

Table 1. Blast analysis of toxin-related proteins in *A. catenella*

Protein	Homologs	Top hit protein	Top score/ e-value	ACHK-T1/ ACHK-NT1	ACHK-T2/ ACHK-NT2	Putative function
sxtG	1	comp63408_c0_orf1	82.8/3.00E^{-18}	–	–	Amidinotransferase
sxtH	4	comp55174_c0_orf1	83.6/1.00E^{-18}	1.029	0.882	Phenylpropionate dioxygenase
sxtO	1	comp43217_c0_orf1	181/3.00E^{-58}	0.751	1.096	Adenylylsulfate kinase
sxtT	4	comp55174_c0_orf1	84.7/7.00E^{-19}	1.029	0.882	Phenylpropionate dioxygenase
sxtU	13	comp100612_c0_orf1	133/1.00E^{-38}	–	–	Short-chain alcohol dehydrogenase
sxtZ	1	comp61883_c0_orf1	55.1/6.00E^{-09}	0.903	1.055*	Signal transduction
ompR	1	comp17794_c0_orf1	71.6/7.00E^{-15}	0.894	0.786	Signal transduction

ACHK-T/ACHK-NT: protein abundance ratio between two samples; -: unquantifiable; *: p-value≤0.05.

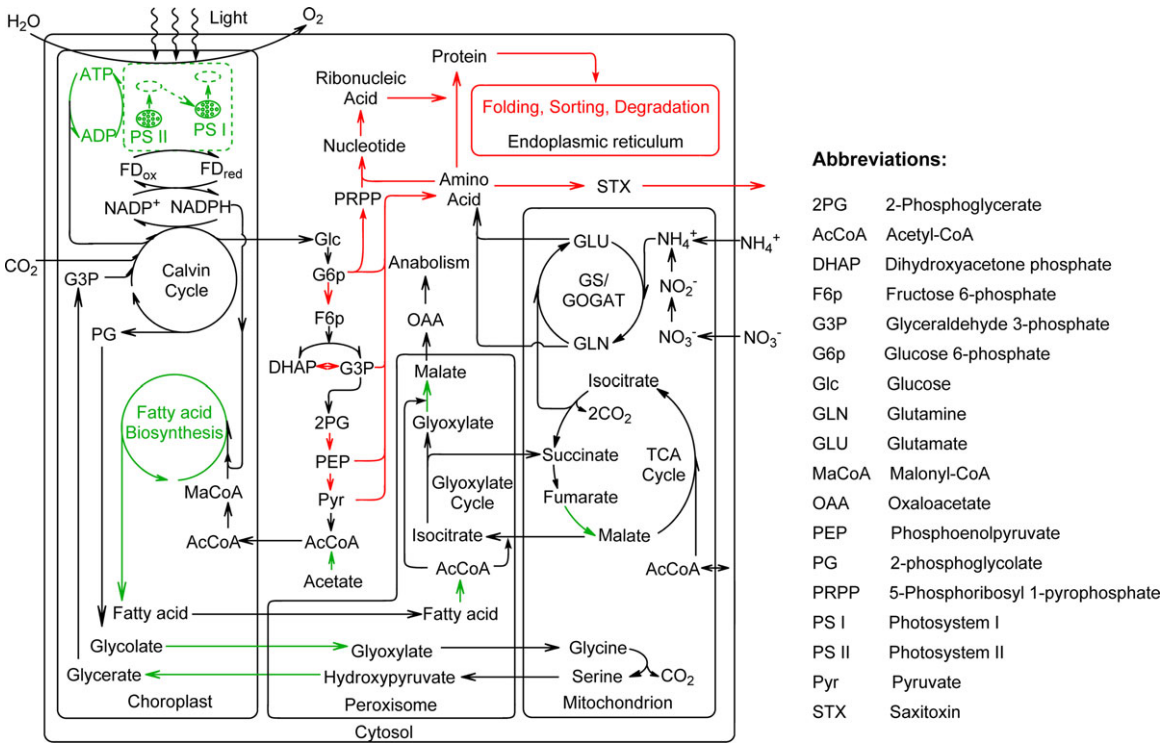

Figure 4. The proposed scheme illustrating the enhanced processes in ACHK-T (red) and in ACHK-NT (green). Adopted and modified from KEGG (map00010, map00630, map00910, map00195) and [40].

4.1.4 Bioluminescence

Dinoflagellates are the most common sources of bioluminescence in the ocean, and the luciferin- LCF reaction is responsible for this. In the present study, two LCF isoforms and one LBP altered significantly in abundance with an increasing fold change from 2.282 to 3.683, which was consistent with our previous study [10]. Usually, luciferin is bound with LBP at a relatively high pH value in case of being oxidized by LCF. When the proton flows to the scintillon resulting in a drop in pH, the luciferin is released and enzymatically oxidized with concomitant light emission [27]. It should be pointed out that luciferin is a nitrogen-enrichment tetrapyrrole [28] and the oxidation of luciferin will produce nitrogen-rich intermediates which possess a similar structure to PSTs, suggesting a potential relationship between bioluminescence and toxin biosynthesis. Higher abundances of LCF and LBP in ACHK-T might increase the production of nitrogen-enrichment intermediates, which provides more precursors for toxin biosynthesis.

4.2 Enhanced biological processes in ACHK-NT

4.2.1 Photosynthesis

Six photosynthesis related proteins, including light-harvesting complex I chlorophyll a/b binding proteins (LHCA1 and LHCA4), photosystem II proteins (psbA, psbB and psbC) and porphobilinogen synthase (ALAD) were more abundant in ACHK-NT (fold change >1.3). LHCA1 and LHCA4 are two antenna proteins in the light-harvesting complex, and they play a coordinator role among the antenna pigments of the photosystem and balance the excitation energy between photosystem I and II [29]. PsbA, psbB and psbC are three subunits assembled in photosystem II with psbA acting as the reaction center protein, and psbB and psbC responsible for light-harvesting [30]. ALAD participates in the synthesis of porphobilinogen which is a common precursor for all natural tetrapyrroles including chlorophylls. Higher abundances of these proteins indicated enhanced photosynthetic carbon fixation and energy production in ACHK-NT which may contribute to its higher growth rate [10].

4.2.2 Fatty acid biosynthesis

3-hydroxyacyl-[acyl-carrier-protein] dehydratase (fabZ) and enoyl-[acyl-carrier protein] reductase I (fabI) are two enzymes catalyzing the last two steps in the cycle of fatty acid elongation [31], and they displayed higher abundances in ACHK-NT (fold change >1.2). FabZ catalyzes the dehydration of β-hydroxyacyl-ACPs and then the product is reduced by fabI to generate a two carbons longer acyl-ACP. All the carbons for fatty acid elongation are derived from the pool of acetyl-CoA present in the plastid. Considering the fact that the

acetyl-CoA concentration in chloroplasts is far from enough for the high acetyl-CoA-consumptional fatty acid biosynthesis, it was therefore not surprising that acetyl-CoA synthetase functioning as acetyl-CoA synthesis was also more abundant in ACHK-NT (fold change, 1.311). This pathway is one of the major sources for the acetyl-CoA pool from which the substrates of fatty acid come [31]. Higher expressions of these enzymes implied that more acetyl-CoA or carbon might be utilized for fatty acid biosynthesis in ACHK-NT, instead of toxin biosynthesis.

4.2.3 Processes occurring in peroxisome

The peroxisome is observed in dinoflagellates but little is known about its physiological function in these algae [32]. Many studies based on photosynthetic plants reveal that the peroxisome contains enzymes which participate in the oxidative photosynthetic carbon cycle of photorespiration, the glyoxylate cycle, fatty acid β-oxidation and ROS metabolism [33]. In our study, proteins involved in these processes were more abundant in ACHK-NT (fold change >1.2). (S)-2-hydroxy-acid oxidase and hydroxypyruvate reductase 2 are two important enzymes involved in photorespiration in all photosynthetic organisms [34]; enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase is an enzyme participating in the β-oxidation of fatty acids; and malate synthase (MLS) and fumarate hydratase are enzymes involved in the glyoxylate cycle. Notably, MLS is one of two hallmark enzymes of the glyoxylate cycle in peroxisome [35]. In addition, several proteins involved in cellular defense against ROS [36–38], were also more abundant in ACHK-NT, including poly [ADP-ribose] polymerase, glutathione S-transferase, dynamin GTPase and monothiol glutaredoxin. Higher expressions of these proteins implied that processes in the peroxisome were strengthened in ACHK-NT, which might be a cellular response to the homeostasis imbalance caused by the loss of toxicity.

4.3 Toxin-related proteins

Several toxin-related proteins are putatively identified from the dinoflagellates using proteomic approaches [10,11]. However, these proteins are identified based on homology searching and the relationships between these proteins and toxin genes remain to be ascertained. Furthermore, the low-throughput protein identification of gel-based proteomic approaches also impedes exploration of toxin-related proteins. In this study, a combination of the iTRAQ-based proteomic approach and transcriptomic database was applied. SxtG and 21 homologs of seven toxin-related proteins were identified: sxtG, sxtH/T and sxtU directly involved in PST biosynthesis; sxtO related to the conversion of PST analogs; and sxtZ and ompR related to transcriptional regulation of PST synthesis [2]. The expressions of sxtZ and ompR in *A. catenella* further suggested that toxin production in dinoflagellates might be regulated in a similar manner as in cyanobacteria. However, compared with our previous study, several toxin-related proteins encoded by *sxtA*, *sxtB*, *sxtD*, *sxtI* and *sxtF/M* were not detected in this study [15]. A recent study also identifies only seven toxin-related proteins in cyanobacteria in spite of their clear biosynthesis mechanism [3]. We have no clear explanations for this inconsistency. Low abundances and specific cellular localizations of these proteins [3, 6], PTM [39] and a limited transcriptome database might be responsible for the poor protein identification. Moreover, a large number of differentially displayed proteins were not annotated with clear functions. These proteins might be involved in toxin biosynthesis, which needs further study.

Quantitative analysis showed no significant abundance variations of toxin-related proteins between ACHK-T and ACHK-NT. A similar situation is also observed in our previous transcriptomic study of these two strains [15]. These results indicated that toxin-related genes or proteins might not be unique to toxin biosynthesis; they might also be involved in other biological processes.

4.4 Concluding remarks

This study, for the first time, compared the protein profiles of ACHK-T and ACHK-NT using a combination of iTRAQ-based proteomic approach and a transcriptomic database. Proteins involved in carbohydrate, energy and amino acid metabolism dominated those identified. Quantitative comparison indicated different carbon and energy utilization strategies between ACHK-T and ACHK-NT: processes involved in synthesis of various precursors or substrates for toxin production were strengthened in ACHK-T while more carbon and energy were shifted to fatty acid biosynthesis instead of toxin production in ACHK-NT (Fig. 4). Twenty-one homologs of seven cyanobacterial toxin-related proteins and sxtG of dinoflagellates were identified as candidates involved in toxin biosynthesis; however, their abundance displayed insignificant differences between ACHK-T and ACHK-NT, suggesting the complex regulation mechanism of PST biosynthesis in dinoflagellates. Overall, this study provided a new insight into PST biosynthesis in dinoflagellates and lays the groundwork for future proteomic study. The iTRAQ-based proteomic approach could be utilized more appropriately in future work as increasing genomic information on the dinoflagellates appears.

This work was partially supported by research grants from the National Natural Science Foundation of China (Project No. 41176094 and 41230961). D.-Z. Wang was also supported by the 'Ten Thousand Talents Program' for leading talents in science and technological innovation. We thank Professor John

Hodgkiss from The University of Hong Kong for his help with English.

The authors have declared no conflict of interest.

5 References

[1] Hallegraeff, G. M., A review of harmful algal blooms and their apparent global increase. *Phycologia* 1993, *32*, 79–99.

[2] Kellmann, R., Mihali, T. K., Jeon, Y. J., Pickford, R. et al., Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. *Appl. Environ. Microbiol.* 2008, *74*, 4044–4053.

[3] D'Agostino, P. M., Song, X. M., Neilan, B. A., Moffitt, M. C., Comparative proteomics reveals that a saxitoxin-producing and a nontoxic strain of *Anabaena circinalis* are two different ecotypes. *J. Proteome Res.* 2014, *13*, 1474–1484.

[4] Lin, S. J., Genomic understanding of dinoflagellates. *Res. Microbiol.* 2011, *162*, 551–569.

[5] Hackett, J. D., Wisecaver, J. H., Brosnahan, M. L., Kulis, D. M. et al., Evolution of saxitoxin synthesis in cyanobacteria and dinoflagellates. *Mol. Biol. Evol.* 2013, *30*, 70–78.

[6] Stüken, A., Orr, R. J. S., Kellmann, R., Murray, S. A. et al., Discovery of nuclear-encoded genes for the neurotoxin saxitoxin in dinoflagellates. *PLoS One* 2011, *6*, e20096.

[7] Orr, R. J. S., Stüken, A., Murray, S. A., Jakobsen, K. S., Evolutionary acquisition and loss of saxitoxin biosynthesis in dinoflagellates: the second "core" gene, *sxtG*. *Appl. Environ. Microbiol.* 2013, *79*, 2128–2136.

[8] Monroe, E. A., Van Dolah, F. M., The toxic dinoflagellate *Karenia brevis* encodes novel type I-like polyketide synthases containing discrete catalytic domains. *Protist* 2008, *159*, 471–482.

[9] Chan, L. L., Sit, W. H., Lam, P. K. S., Hsieh, D. P. H. et al., Identification and characterization of a "biomarker of toxicity" from the proteome of the paralytic shellfish toxin-producing dinoflagellate *Alexandrium tamarense* (Dinophyceae). *Proteomics* 2006, *6*, 654–666.

[10] Wang, D. Z., Li, C., Zhang, Y., Wang, Y. Y. et al., Quantitative proteomic analysis of differentially expressed proteins in the toxicity-lost mutant of *Alexandrium catenella* (Dinophyceae) in the exponential phase. *J. Proteomics* 2012, *75*, 5564–5577.

[11] Wang, D. Z., Gao, Y., Lin, L., Hong, H. S., Comparative proteomic analysis reveals proteins putatively involved in toxin biosynthesis in the marine dinoflagellate *Alexandrium catenella*. *Marine Drugs* 2013, *11*, 213–232.

[12] Lin, D., Tabb, D. L., Yates, J. R., Large-scale protein identification using mass spectrometry. *Biochim. Biophys. Acta (BBA)-Proteins Proteomics* 2003, *1646*, 1–10.

[13] Wang, D. Z., Zhang, H., Zhang, Y., Zhang, S. F., Marine dinoflagellate proteomics: Current status and future perspectives. *J. Proteomics* 2014, *105*, 121–132.

[14] Fuller, H. R., Morris, G. E., in: Leung, H. C. E. (Ed.), *Integrative Proteomics* InTech, Croatia 2012, pp. 347–362.

[15] Zhang, Y., Zhang, S. F., Lin, L., Wang, D. Z., Comparative transcriptome analysis of a toxin-producing dinoflagellate *Alexandrium catenella* and its non-toxic mutant. *Marine Drugs* 2014, *12*, 5698–5718.

[16] Keller, M. D., Selvin, R. C., Claus, W., Guillard, R. R. L., Media for the culture of oceanic ultraphytoplankton 1, 2. *J. Phycology* 1987, *23*, 633–638.

[17] Glen, A., Gan, C. S., Hamdy, F. C., Eaton, C. L. et al., iTRAQ-facilitated proteomic analysis of human prostate cancer cells identifies proteins associated with progression. *J. Proteome Res.* 2008, *7*, 897–907.

[18] Du, C., Liang, J. R., Chen, D. D., Xu, B. et al., iTRAQ-based proteomic analysis of the metabolism mechanism associated with silicon response in the marine diatom *Thalassiosira pseudonana*. *J. Proteome Res.* 2014, *13*, 720–734.

[19] Shimizu, Y., Microalgal metabolites. *Chem. Rev.* 1993, *93*, 1685–1698.

[20] Umbarger, H. E., Amino acid biosynthesis and its regulation. *Annu. Rev. Biochem.* 1978, *47*, 533–606.

[21] Boldt, R., Zrenner, R., Purine and pyrimidine biosynthesis in higher plants. *Physiologia Plantarum* 2003, *117*, 297–304.

[22] Plaxton, W. C., The organization and regulation of plant glycolysis. *Annu. Rev. Plant Biol.* 1996, *47*, 185–214.

[23] Ginisty, H., Sicard, H., Roger, B., Bouvet, P., Structure and functions of nucleolin. *J. Cell Sci.* 1999, *112*, 761–772.

[24] Kapp, L. D., Lorsch, J. R., The molecular mechanics of eukaryotic translation. *Annual Review of Biochemistry* 2004, *73*, 657–704.

[25] Fink, A. L., Chaperone-mediated protein folding. *Physiol. Rev.* 1999, *79*, 425–449.

[26] Taroncher-Oldenburg, G., Kulis, D. M., Anderson, D. M., Coupling of saxitoxin biosythesis to the G_1 phase of the cell cycle in the dinoflagellate *Alexandrium fundyense*: temperature and nutrient effects. *Nat. Toxins* 1999, *7*, 207–219.

[27] Nicolas, M. T., Nicolas, G., Johnson, C. H., Bassot, J.-M., Hastings, J. W., Characterization of the bioluminescent organelles in *Gonyaulax polyedra* (dinoflagellates) after fast-freeze fixation and antiluciferase immunogold staining. *J. Cell Biol.* 1987, *105*, 723–735.

[28] Nakamura, H., Kishi, Y., Shimomura, O., Morse, D., Hastings, J. W., Structure of dinoflagellate luciferin and its enzymic and nonenzymic air-oxidation products. *J. Am. Chem. Soc.* 1989, *111*, 7607–7611.

[29] Jansson, S., The light-harvesting chlorophyll *a*/*b*-binding proteins. *Biochim. Biophys. Acta (BBA)-Bioenergetics* 1994, *1184*, 1–19.

[30] Barber, J., Photosystem II: a multisubunit membrane protein that oxidises water. *Curr. Opin. Struct. Biol.* 2002, *12*, 523–530.

[31] Ohlrogge, J., Browse, J., Lipid biosynthesis. *Plant Cell* 1995, *7*, 957–970.

[32] Bibby, B. T., Dodge, J. D., The ultrastructure and cytochemistry of microbodies in dinoflagellates. *Planta* 1973, *112*, 7–16.

[33] Palma, J. M., Corpas, F. J., del Río, L. A., Proteome of plant peroxisomes: new perspectives on the role of these organelles in cell biology. *Proteomics* 2009, *9*, 2301–2312.

[34] Bauwe, H., Hagemann, M., Fernie, A. R., Photorespiration: players, partners and origin. *Trends in Plant Sci.* 2010, *15*, 330–336.

[35] Pracharoenwattana, I., Smith, S. M., When is a peroxisome not a peroxisome? *Trends Plant Sci.* 2008, *13*, 522–525.

[36] Reumann, S., Babujee, L., Ma, C. L., Wienkoop, S. et al., Proteome analysis of *Arabidopsis* leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* 2007, *19*, 3170–3193.

[37] Handy, D. E., Loscalzo, J., Redox regulation of mitochondrial function. *Antioxid. Redox Signal.* 2012, *16*, 1323–1367.

[38] Strosznajder, R. P., Jesko, H., Zambrzycka, A., Poly (ADP-ribose) polymerase. *Mol. Neurobiol.* 2005, *31*, 149–167.

[39] Kellmann, R., Neilan, B. A., Biochemical characterization of paralytic shellfish toxin biosynthesis in vitro. *J. Phycology* 2007, *43*, 497–508.

[40] Guerra, L. T., Levitan, O., Frada, M. J., Sun, J. S. et al., Regulatory branch points affecting protein and lipid biosynthesis in the diatom *Phaeodactylum tricornutum*. *Biomass Bioenergy* 2013, *59*, 306–315.

