npg

ORIGINAL ARTICLE

The central role of selenium in the biochemistry and ecology of the harmful pelagophyte, *Aureococcus anophagefferens*

Christopher J Gobler[1], Alexei V Lobanov[2], Ying-Zhong Tang[1], Anton A Turanov[2],
Yan Zhang[2,3], Martina Doblin[4], Gordon T Taylor[1], Sergio A Sañudo-Wilhelmy[5],
Igor V Grigoriev[6] and Vadim N Gladyshev[2]

[1]*School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA;* [2]*Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA;* [3]*Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China;* [4]*Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, Sydney, New South Wales, Australia;* [5]*Department of Biological Sciences and Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA and* [6]*US Department of Energy, Joint Genome Institute, Walnut Creek, CA, USA*

The trace element selenium (Se) is required for the biosynthesis of selenocysteine (Sec), the 21st amino acid in the genetic code, but its role in the ecology of harmful algal blooms (HABs) is unknown. Here, we examined the role of Se in the biology and ecology of the harmful pelagophyte, *Aureococcus anophagefferens*, through cell culture, genomic analyses, and ecosystem studies. This organism has the largest and the most diverse selenoproteome identified to date that consists of at least 59 selenoproteins, including known eukaryotic selenoproteins, selenoproteins previously only detected in bacteria, and novel selenoproteins. The *A. anophagefferens* selenoproteome was dominated by the thioredoxin fold proteins and oxidoreductase functions were assigned to the majority of detected selenoproteins. Insertion of Sec in these proteins was supported by a unique Sec insertion sequence. Se was required for the growth of *A. anophagefferens* as cultures grew maximally at nanomolar Se concentrations. In a coastal ecosystem, dissolved Se concentrations were elevated before and after *A. anophagefferens* blooms, but were reduced by >95% during the peak of blooms to 0.05 nM. Consistent with this pattern, enrichment of seawater with selenite before and after a bloom did not affect the growth of *A. anophagefferens*, but enrichment during the peak of the bloom significantly increased population growth rates. These findings demonstrate that Se inventories, which can be anthropogenically enriched, can support proliferation of HABs, such as *A. anophagefferens* through its synthesis of a large arsenal of Se-dependent oxidoreductases that fine-tune cellular redox homeostasis.

Introduction

Selenium (Se) is an essential element that is required for the biosynthesis of selenocysteine (Sec), the 21st naturally occurring amino acid in the genetic code (Stadtman, 1996). As a trace

element, Se is the 66th most abundant element in the earth's crust, but can be anthropogenically enriched in aquatic environments by numerous agricultural, industrial and urban activities including fossil fuel (coal) combustion, refinery activity, irrigation of Se-rich soils, fertilizer use and sewage discharge (Bodek *et al.*, 1988; Cappon, 1991; Cutter and Cutter, 2004; US EPA, 2009). The essential nature of Se is related to its occurrence in proteins in the form of Sec residues. Sec is inserted co-translationally in response to the UGA codon, which is recoded from its normal stop codon function if a control element (called SECIS element) is encountered in the 3′-UTR of eukaryotic and archaeal

mRNAs, or immediately downstream of UGA in the coding regions of bacterial selenoprotein genes (Berry *et al.,* 1991; Hatfield and Gladyshev, 2002). Selenoproteins can be a thousand times more effective as catalysts than their sulfur-based cysteine homologs (Kim *et al.,* 2006), a fact that likely accounts for the cellular investment in Se-dependent pathways and the specialized machinery used for Se insertion into proteins (Lobanov *et al.,* 2009). Eukaryotic selenoproteomes vary significantly in size but are generally larger in aquatic organisms compared with terrestrial organisms and are particularly large in microalgae, which may have a more consistent supply of bioavailable Se in seawater compared with terrestrial organisms (Lobanov *et al.,* 2009).

Se has been shown to be required by many phytoplankton with examples of absolute Se requirements reported within Bacillariophyceae, Chlorophyceae, Phaeophyceae, Raphidophyceae, Rhodophyceae, Prymnesiophyeae and Prasinophyceae (Fries, 1982; Wheeler *et al.,* 1982; Keller *et al.,* 1984; Harrison *et al.,* 1988; Imai *et al.,* 1996). Se is also required by many dinoflagellates (Dinophyceae), particularly those known to form harmful algal blooms (HABs) including *Scrippsiella trochoidea* (Harrison *et al.,* 1988), *Katodinium rotundatum* (Harrison *et al.,* 1988), *Gymnodinium catenatum* (Doblin *et al.,*1999, 2000; Band-Schmidt *et al.,* 2004), *Alexandrium minutum* (Doblin *et al.,*1999), *Protoceratium reticulatum* (Mitrovic *et al.,* 2004), *Karenia selliformis* (Mountfort *et al.,* 2006), *Cochlodinium polykrikoides* (Lee, 2008), *Prorocentrum minimum*Baines and Fisher, 2001*), Chattonella verruculosa* (Imai *et al.,* 1996), and *Peridinium gatunense* (Lindstrom, 1991). The pelagophyte *Aureococcus anophagefferens* forms HABs (referred to as 'brown tides') that have destroyed shellfisheries and coastal habitats in the United States, South Africa and China (Gobler and Sunda, 2012) and recent sequencing and annotation of its genome (Gobler *et al.,* 2011) revealed an unusually large number of predicted Sec-containing proteins, even when compared with large selenoproteomes in other protistan taxa (Lobanov *et al.,* 2007; Palenik *et al.,* 2007). Despite this recent finding, little is known regarding the organization, composition and properties specific to the selenoproteome as well as the importance of Se in the ecology of *A. anophagefferens*. Moreover, although the role of macro-nutrients (nitrogen, phosphorus) in the occurrence of HABs has been well-studied (Anderson *et al.,* 2008; Heisler *et al.,* 2008), the importance of trace elements, including Se, for these events has rarely been considered, despite the widespread requirement of Se among harmful algae (Harrison *et al.,* 1988; Lindstrom, 1991; Doblin *et al.,*1999, 2000; Baines and Fisher, 2001; Band-Schmidt *et al.,* 2004; Mitrovic *et al.,* 2004; Mountfort *et al.,* 2006; Lee, 2008; Gobler *et al.,* 2011).

Here, we examined the role of Se and selenoproteins in the biology and ecology of the harmful alga, *A. anophagefferens*. The Se requirements of *A. anophagefferens* were established in culture. The selenoproteome, its components and SECIS (Sec insertion sequence) elements of this species were characterized via informatics and metabolic labeling. The dynamics of naturally occurring blooms of this species and dissolved Se concentrations were quantified and, finally, Se-enrichment experiments were performed with field populations.

Materials and methods

Culture experiments
Se concentrations required to achieve maximal growth rates in axenic *A. anophagefferens* (clone CCMP1984) were determined using G-medium made from artificial seawater (Doblin *et al.,*1999) supplemented with differing concentrations of Se added as sodium selenite. Cultures were grown at 21 °C in an incubator with a 14:10 h light:dark cycle, illuminated by a bank of fluorescent lights that provided a light intensity of $\sim 100\,\mu$mol photons per m² per sec to cultures. These conditions approximated temperature and light exposures found in Long Island estuaries during summer months when *A. anophagefferens* blooms (Gobler and Sunda, 2012). Cultures were maintained for a minimum of four transfers at each concentration before the collection of final growth rate data to ensure that cells were fully acclimated to treatment conditions and that the carryover of Se from the initial, full strength media was eliminated. Cellular growth rates were calculated for cultures based on cell densities (determined microscopically) in exponential growth phases, using the formula $\mu = [\ln (B_t/B_0)]/t$, where B_0 and B_t are the initial and final abundances, and t is the incubation duration in days. Growth rates were averaged over the entire exponential phase, which typically persisted for 3–6 days, depending on the concentration of Se in the media. Se-limitation of cultures was confirmed by the stimulation of growth of cultures in stationary phase at concentrations below 5 nM following the addition of 10 nM Se. The Monod kinetic terms μ_{max} (maximum growth rate) and K_S (half-saturation constant) were derived using Michaelis–Menten curve fitting functions within Kaleidagraph version 4.1.1 (Synergy Software, Reading, PA, USA).

Identification of genes coding for selenoproteins in A. anophagefferens
Identification of selenoprotein genes in the *Aureococcus* genome was performed using SECISearch (Novoselov *et al.,* 2002; Fomenko *et al.,* 2007), which analyzes primary sequences and secondary structures and then calculates the free energy for various parts of potential SECIS elements. The search was carried out using both default and loose patterns of SECISearch to accommodate identification of unusual SECIS structures. The searches also were then

extended to identify organism-specific structures by a modified SECISearch (Novoselov *et al.*, 2002; Fomenko *et al.*, 2007). For identified candidate SECIS elements, open reading frames (ORFs) were predicted in the regions upstream of the SECIS elements. The presence of at least one homologous protein in the NCBI non-redundant database was used as an additional requirement. At the final step, a manual sequence and homology analysis of predicted selenoprotein ORFs located upstream of candidate SECIS elements was carried out.

In addition to the SECIS-based search procedure, the genome was analyzed with TBLAST against all known selenoprotein sequences to identify homologs of previously described selenoprotein genes. A third procedure included a search for Sec/Cys pairs in homologous sequences (Novoselov *et al.*, 2002; Fomenko *et al.*, 2007). We extracted all ORFs containing in-frame UGA codons, and homologs containing Cys in place of Sec were found by identifying the Sec/Cys pairs flanked by homologous sequences. For this procedure, we used TBLASTX to examine all potential ORFs with in-frame UGA codons against NCBI non-redundant protein database. All candidate ORFs were then tested with SECISearch for the presence of SECIS elements. Finally, PSI-BLAST was used to identify additional more distant homologs. All data sets obtained with the three independent methods were combined and the proteins were classified as homologs of previously known selenoproteins, novel selenoproteins and candidate selenoprotein genes.

Metabolic labeling of the A. anophagefferens selenoproteome

To experimentally verify the occurrence of selenoproteins, *A. anophagefferens* cells were subjected to metabolic labeling with ^{75}Se. The green alga, *Chlamydomonas reinhardtii,* was labeled in parallel and used as a control. The metabolic labeling was performed as described previously (Novoselov *et al.*, 2002). Briefly, 100 ml of each culture was metabolically labeled with 100 μCi of ^{75}Se ([^{75}Se]selenous acid, specific activity 1000 Ci mmol^{-1}, Research Reactor Facility, University of Missouri, Columbia, MO, USA) for 48 h. Cells were collected, resuspended in PBS and sonicated. Approximately 20 and 50 μg of total soluble protein from *A. anophagefferens* and *C. reinhardtii*, respectively, were resolved by SDS–PAGE and transferred onto a PVDF membrane (Invitrogen, Carlsbad, CA, USA). Selenoprotein patterns were visualized with a PhosphorImager.

Ecosystem study

During the spring and summer of 1998, seawater samples were collected twice weekly to fortnightly from West Neck Bay, Stony Brook, NY, USA, a small (1 km^2), shallow (2–4 m), enclosed embayment on Shelter Island, within the Peconic Estuary of eastern Long Island. The absence of tributaries and point source anthropogenic inputs make this bay a relatively simple system, where chemical inputs can be largely attributed to groundwater seepage or internal recycling (Gobler and Sañudo-Wilhelmy, 2001). All sampling material used in this study was prepared using trace metal clean techniques (Gobler *et al.*, 2002). Water samples were collected with a peristaltic pump connected to acid-washed Teflon tubing on a bamboo pole extended 4 m upwind and lowered to a depth of 1 m. Dissolved seawater samples were obtained by filtration through trace metal clean, polypropylene capsule filters (0.2 μm). All analyses and sample manipulation performed after collection were conducted within vertical flow clean benches supplied with HEPA-filtered air. Triplicate chlorophyll *a* samples were collected on GF/F glass fiber filters (nominal pore size = 0.7 μm) and analyzed by standard fluorometric methods (Parsons *et al.*, 1984). *A. anophagefferens* densities were determined on field samples fixed with glutaraldehyde (1% final concentration). Samples were enumerated by direct count methods employing a Zeiss model D-7082 epi-fluorescent microscope (Jena, Germany) and an immunofluorescent label as described by Anderson *et al.*, 1989. The amount of chlorophyll *a* in West Neck Bay attributable to *A. anophagefferens* was estimated by assuming a constant chl *a* per cell value (0.035 ± 0.003 pg per cell for nutrient replete cultures (Gobler and Sañudo-Wilhelmy, 2001)) and multiplying this value by *A. anophagefferens* densities. Such approximations have been used successfully in the past to compare *A. anophagefferens* biomass with that of the total algal community (Gobler and Sañudo-Wilhelmy, 2001; Caron *et al.*, 2004).

Groundwater entering West Neck Bay through the intertidal zone was sampled from 1 m, Teflon-lined, PVC piezometers with 2.5 cm horizontal screened slits along the lower 25 cm at six locations across the Bay on multiple occasions. High groundwater seepage rates at West Neck Bay allowed piezometers to fill with fresh groundwater (salinity < 0.1 ppt) when sampled during low tide. Groundwater was sampled using a peristaltic pump equipped with acid-washed Teflon tubing. To ensure representative groundwater was sampled, piezometers were purged at < 100 ml min^{-1} and samples were not obtained until the conductivity, dissolved oxygen, temperature and pH of the pumped groundwater stabilized (Puls and Paul, 1995). Filtered groundwater or seawater samples were acidified to pH 1.6 with HCl, stored under dim light at room temperature and then analyzed for total dissolved Se as described by Cutter, 1982, 1983. Briefly, total dissolved Se was determined by boiling a 4 M HCl acidified sample for 1 h with the addition of potassium persulfate. Next, within a glass stripping vessel, selenite was quantitatively converted to

hydrogen selenide by adding sodium borohydride to samples along with sulfanilamide (to eliminate interference due to nitrite). The evolved hydrogen selenide was stripped from solution using helium and trapped in a borosilicate U-tube packed with silanized glass wool and immersed in liquid nitrogen. After the trap was removed from the LN2, an atomic absorption spectrometer fitted with an open quartz tube furnace burning an air–hydrogen flame was used to detect the hydride. Instrument response (as peak area) was recorded on a chromatographic integrator. To ensure accuracy, all determinations utilized the standard additions method of calibration, and all samples were analyzed in triplicate to quantify precision (found to be 4%). Detection limits for dissolved Se were 0.02 nM.

Field experiments were conducted on 4 June, 22 June, 8 July and 21 July 1998 using water from West Neck Bay. Within 2 h of collection, 50 ml of seawater was transferred to trace metal clean, 60 ml polycarbonate flasks in a HEPA laminar flow hood to prevent trace metal contamination. Triplicate flasks were amended with sodium selenite (10 nM) or were left unamended as a control treatment. The selenite solution was cleaned of trace metals with Chelex-100 ion exchange resin (Bruland, 1980), filter sterilized (0.2 μm) and frozen before use. Amended and unamended flasks were incubated at the same temperature found in West Neck Bay during water collection under 125 μmol photons per m² per sec of light on a light:dark cycle which mimicked summer conditions in West Neck Bay (14 h: 10 h) and the light found at ∼1.4 m in the bay (Gobler and Sañudo-Wilhelmy, 2001). After 48 h, samples from each flask were preserved to a final concentration of 1% glutaraldehyde for enumeration of *A. anophagefferen*s as described above. Net specific growth rates of *A. anophagefferens* were determined as described for culture experiments. Differences between control and treatment flasks were resolved via Student's *t*-test.

Results and discussion

Requirement of Se for growth of A. anophagefferens
A. anophagefferens (CCMP9184) cultures displayed a sigmoidal increase in growth rate in response to increasing selenite concentrations (Figure 1) and fit the Michaelis–Menten equation ($R = 0.97$). Cultures did not grow below 10 pM selenite, but displayed a gradual increase in growth above 10 pM and near maximal growth rates (∼0.6 per day) at concentrations of selenite at or above 5 nM (Figure 1). *A. anophagefferens* had a half-saturation constant (K_s) of 0.27 ± 0.08 nM for selenite and a μ_{max} of 0.60 ± 0.03 per day.

The A. anophagefferens selenoproteome
Computational searches for selenoprotein genes using a combined SECISearch, SECISearch adapted

for *A. anophagefferens*, and the Sec/Cys approaches extended the number of detected selenoproteins from the originally predicted 56 selenoprotein genes (26) to a total of 59. This represents the largest currently known selenoproteome of any organism, being two times larger than the one present in *Ostreococcus lucimarinus*, previously considered the largest eukaryotic selenoproteome (27,28), and four-fold larger than those of diatoms that compete with *A. anophagefferens* and green algae (Figure 2). The *A. anophagefferens* selenoproteome is characterized by the occurrence of almost all known eukaryotic selenoproteins, for example, 20 out of 25 human selenoproteins have selenoprotein homologs in *A. anophagefferens* and 12 out of 12 *Chlamydomonas* selenoproteins. Three selenoproteins have previously been detected only in bacteria (outside of *A. anophagefferens*; Table 1) and most selenoproteins had closest homologs from completely different groups of organisms. In addition, several selenoproteins occurred in multiple copies
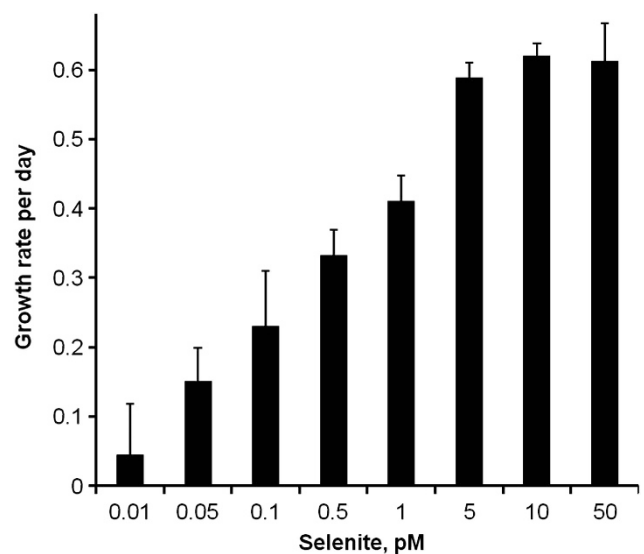

Figure 1 Growth rates of *A. anophagefferens* in media with differing concentrations of selenite.
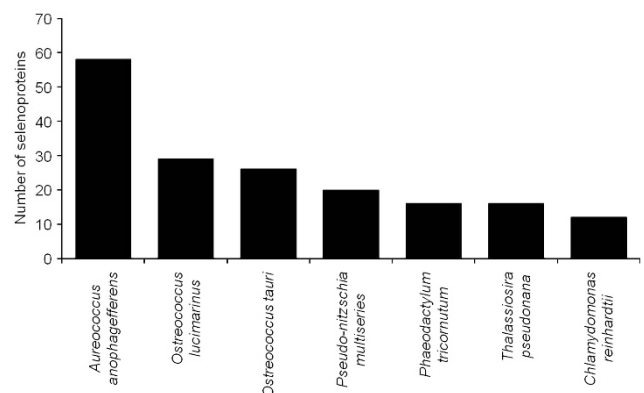

Figure 2 Size of selenoproteomes of seven eukaryotic phytoplankton for which complete genomes are available.

Table 1 Selenocysteine-containing proteins encoded in the *A. anophagefferens* genome

Protein name	PID	Chlamydomonas	Human	Organism with closest homolog
Thioredoxin domain containing protein	77962			*Micromonas sp.*
Thioredoxin domain containing protein	78111			*Brassica napus*
Thioredoxin domain containing protein	78110			*Micromonas sp.*
Thioredoxin domain containing protein	78109			*Thalassiosira pseudonana*
Thioredoxin domain containing protein	78108			*Emiliania huxleyi*
Glutathione peroxidase	77994	GPx4a, GPx4b	GPx1, GPx2, GPx3, GPx4, GPx6	*Mizuhopecten yessoensis*
Glutathione peroxidase	77995			*Micromonas pusilla*
Glutathione peroxidase	77996			*M. pusilla*
Glutathione peroxidase	77997			*Mizuhopecten yessoensis*
Glutathione peroxidase	78003			*Phaeodactylum tricornutum*
Methionine sulfoxide reductase A	77998	MsrA		*Volvox carteri*
Methionine sulfoxide reductase A	77999			*V. carteri*
Methionine sulfoxide reductase A	78000			*Chlamydomonas reinhardtii*
Methionine sulfoxide reductase A	78001			*C. reinhardtii*
Methionine sulfoxide reductase B	78002		MsrB1	*Micromonas sp.*
Methionine sulfoxide reductase B	77940			*Monosiga brevicollis*
Methionine sulfoxide reductase B	78004			*Micromonas sp.*
Methionine sulfoxide reductase B	77942			*Micromonas sp.*
Glutaredoxin	77970			*Chlorella variabilis*
Glutaredoxin	77969			*C. variabilis*
Glutaredoxin	77966			*Ectocarpus siliculosus*
Peroxiredoxin	78116			*M. pusilla*
Peroxiredoxin	78117			*M. pusilla*
Peroxiredoxin	77974			*Emiliania huxleyi*
Selenoprotein T	77992	SelT	SelT	*Arabidopsis thaliana*
Selenoprotein W	77920	SelW1, SelW2	SelW, SelV	*Danio rerio*
Selenoprotein W	77991			*D. rerio*
Selenoprotein W	78120			*E. siliculosus*
Selenoprotein W	78121			*E. siliculosus*
Selenoprotein W	78119			*E. siliculosus*
Selenoprotein O	77993		SelO	*Vibrio caribbenthicus*
Selenoprotein U	77989	SelU		*Lepeophtheirus salmonis*
Selenoprotein U	77988			*Physcomitrella patens*
Selenoprotein M	77987	SelM	SelM	*E. siliculosus*
Selenoprotein M	78118			*E. siliculosus*
Fe-S oxidoreductase	77968			*E. siliculosus*
Selenoprotein Sep15	77986		Sep15	*Picea sitchensis*
Selenoprotein H	77985	SelH	SelH	*V. carteri*
Iodothyronine deiodinase	77984		DI1, DI2, DI3	*Branchiostoma floridae*
Thioredoxin reductase	77983	TR1	TR1, TR2, TR3	*Thalassiosira pseudonana*
Selenoprotein K	77979	SelK	SelK	*Strongylocentrotus purpuratus*
Methyltransferase	78115			*Paenibacillus lactis* (bacterium)
Methyltransferase	77963			*Lyngbya majuscule* (bacterium)
Thiol:disulfide interchange protein	77965			*Chlorella sp.*
Fe-S reductase	77972			*Karenia brevis*
UGSC-containing protein	77971			*Pseudochattonella farcimen*
Membrane Selenoprotein	77964	MSP		*V. carteri*
GILT superfamily protein	77961			*Populus trichocarpa*
Rhodanase	78106			*Patulibacter sp. I11* (bacterium)
Protein disulfide isomerase	78005			*Alexandrium catenella*
Protein disulfide isomerase	77982			*Micromonas pusilla*
Protein disulfide isomerase	77977			*Phaseolus acutifolius*
Hypothetical protein	78107			*E. siliculosus*
Hypothetical protein	78112			*Thalassiosira pseudonana*
Hypothetical protein	78113			*E. siliculosus*
Hypothetical protein	77980			*Micromonas pusilla*
Hypothetical protein	77978			*V. carteri*
Hypothetical protein	77975			*C. reinhardtii*
Hypothetical protein	77967			*E. siliculosus*

PID are protein identification numbers as assigned on the US Department of Energy's Joint Genome Institute genome browser. *Chlamydomonas* and Human columns show selenoprotein homologs to *Aureococcus* selenoproteins in these organisms.

(Table 1, Supplementary Table 1). For example, methionine sulfoxide reductases were represented by four MsrA isozymes and four MsrB isozymes (Table 1). Several novel selenoproteins with unknown functions were also detected (Table 1). Further analysis of the selenoproteome revealed dominance of the thioredoxin fold, present in more than half of detected selenoproteins. In all of these

proteins, Sec was located in place of the catalytic Cys found in thioredoxin. This observation and the additional presence of many other oxidoreductases (MsrAs, MsrBs, GILT, TRs and so on) point to the redox function of Sec in selenoproteins and the role of Se in regulating redox homeostasis in *A. anophagefferens*.

A. anophagefferens selenoproteins have a Sec residue encoded by UGA codon, which is inserted with the help of unusual SECIS elements (Figure 3, Supplementary Table 2), which are RNA structures present in the 3′-UTRs of selenoprotein genes. One unique feature of *A. anophagefferens* SECIS elements is their unusually high diversity (Figure 3). Although several SECIS elements found in the *A. anophagefferens* genome correspond to the canonical SECIS model, a surprisingly large number of SECIS elements did not (Figure 3). A unique feature of *A. anophagefferens* SECIS elements is its unusually long stem. A typical length of the stem (distance from the SECIS core to the apical loop) is 10–13 nucleotides, while in *Aureococcus* SECIS elements it can be nearly twice as long (Figure 3). Another characteristic feature is the extremely small, virtually absent apical loop. The size of the apical loop in the canonical model varies from 7–27 nucleotides, while many *Aureococcus* SECIS elements have apical loops consisting of only 3 nucleotides and often lacking a conserved unpaired AA sequence (Figure 3). It remains to be seen if there are any extra constrains that disrupt nucleotide pairing near the end of the stem to increase the size of the apical loop.

Owing to the highly unique and diverse nature of SECIS elements in *A. anophagefferens* our informatics search (SECISearch) to identify selenoprotein genes had to be adjusted to recognize the *A. anophagefferens* consensus structure developed based on the analyses of SECIS elements in known selenoprotein genes, allowing the selenoproteome to be fully characterized. Although the genome of *A. anophagefferens* encodes the largest selenoproteome, the set of selenoproteins reported here represents a conservative estimate: the true number of selenoproteins is likely larger. The 59 predicted selenoprotein genes did not include less reliable predictions that could not be unambiguously confirmed owing to lack of homologs in other organisms, identifiable SECIS elements in the 3′-UTR (mostly due to insufficient length, or unusual form of the potential SECIS elements) and possibly incorrect gene models. In addition, proteins that were highly similar to each other were not always reported as separate entities, as additional verification was required to avoid over reporting caused by mis-assembly, splicing variants and sequencing errors. We detected components of Sec insertion machinery, including Sec-specific elongation factor EFsec, a SECIS-binding protein and Sec tRNA (Figure 3d). The latter was not detected by tRNAscan using standard parameters, but could be identified in a maximum sensitivity
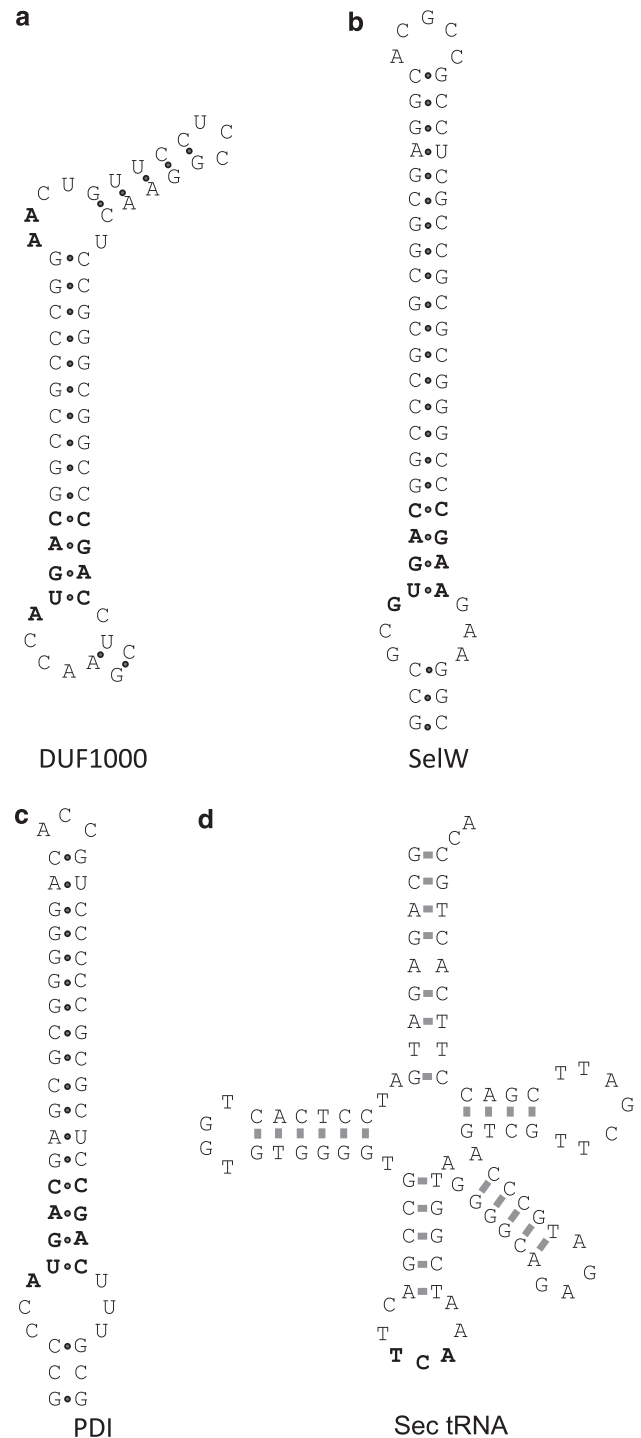

Figure 3 (**a**) A typical eukaryotic SECIS element that is present in the 3′-UTR of DUF1000. Such structure occurred in only some *Aureococcus* selenoprotein genes. (**b**, **c**) Typical *Aureococcus* SECIS structures, in 3′-UTRs of SelW and PDI, respectively, featuring a long stem and a small apical loop that lacks a conserved AA sequence and also lacks a mini-stem. An SBP2-binding site at the bottom of the main step is shown in bold. (**d**) *A. anophagefferens* Sec tRNA. The anticodon is shown in bold.

mode of this program. Overall, it is very likely, that the number of selenoproteins encoded in the *A. anophagefferens* genome exceeds 60.

Metabolic labeling of A. anophagefferens with ^{75}Se
Metabolic labeling of *A. anophagefferens* along with another microalga, *Chlamydomonas reinhardtii*, which has 12 selenoprotein genes (Novoselov *et al.*, 2007; Palenik *et al*, 2007; Figure 4), revealed both a large number of expressed selenoproteins in *A. anophagefferens* and a significantly greater abundance of these proteins compared with *C. reinhardtii*. Abundant selenoprotein bands in the 8–28 kDa region corresponded to the predicted masses of the majority of *Aureococcus* selenoproteins. These observations provide further evidence of the enhanced use of Sec and selenoproteins by *A. anophagefferens* compared with other organisms.

Promotion of A. anophagefferens brown tides by Se
During a coastal ecosystem study, dissolved Se concentrations in West Neck Bay, averaged 0.71 ± 0.37 nM (Figure 5). *A. anophagefferens* densities steadily increased from 1.0×10^4 cells per ml in May to $>1.0 \times 10^5$ cells per ml in June (Figure 5), representing 10–25% of the total algal community. *A. anophagefferens* reached peak densities in early July, as cell densities grew to 5.6×10^5 cells per ml (Figure 6) representing $>95\%$ of total algal biomass. During the bloom peak dissolved Se levels decreased to their lowest levels of the study (0.05 nM; Figure 5). Following the bloom's peak, dissolved Se concentrations rose to >1 nM and remained elevated as *A. anophagefferens* densities steadily declined to nearly zero (Figure 5). Groundwater entering West Neck Bay contained, on average, 3.4 ± 0.3 nM dissolved Se. During incubation

experiments conducted in June and late July, the addition of selenite did not significantly alter the growth of *A. anophagefferens* (Figure 7). During the peak of the *A. anophagefferens* bloom (8 July), however, the addition of selenite significantly increased *A. anophagefferens* growth rates compared with unamended controls (Figure 7; $P < 0.001$; *t*-test). Clearly, Se availability was not the primary driver of this *A. anophagefferens* blooms, but its importance in supporting brown tide proliferation is evident from its drawdown during the bloom and its stimulatory effect under Se-depleted conditions.

A. anophagefferens possess the largest selenoproteome of any living organism, consisting of at least 59 selenoproteins, and this study demonstrates that its ability to form HABs can be dependent on the bioavailability of Se. *A. anophagefferens*' SECIS elements are diverse and, in some cases, unique. Organism-specific features of SECIS elements are known for several species. For example, the majority of *Ostreococcus* SECIS elements possess an extra-long mini-stem (Lobanov *et al.*, 2007). The SECIS elements of evolutionarily unrelated proteins in *Dictyostelium discoideum* are highly conserved with five nucleotides preceding and two nucleotides following the SECIS core being identical to all identified SECIS elements (Novoselov *et al.*, 2002). This example substantiates convergent evolution in SECIS structure as well as constraints associated with SECIS-binding proteins. In addition, SECIS elements in *Neospora* and *Toxoplasma* show a non-canonical GGGA sequence in the SECIS core (Novoselov *et al.*, 2002; Fomenko *et al.*, 2007). Thus, deviations from the eukaryotic SECIS model, while rare, may be expected, and *A. anophagefferens* joins the group of organisms with lineage-specific SECIS functions. The diversity of SECIS elements found in the *A. anophagefferens* genome also suggests a higher-than-normal flexibility in the function of Sec insertion machinery. Once additional genomes with similar properties
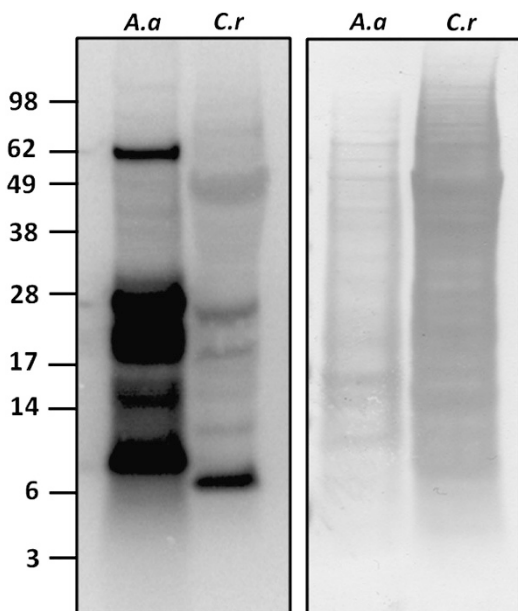

Figure 4 Metabolic labeling of *Aureococcus anophagefferens* (A.a.) and *Chlamydomonas reinhardtii* (C.r.) with ^{75}Se using 20 μg of A.a. and 50 μg of C.r. proteins separated by SDS–PAGE (left panel) and Coomassie blue staining to show protein loading (right panel). Migration of size markers (in kDa) is shown on the left.
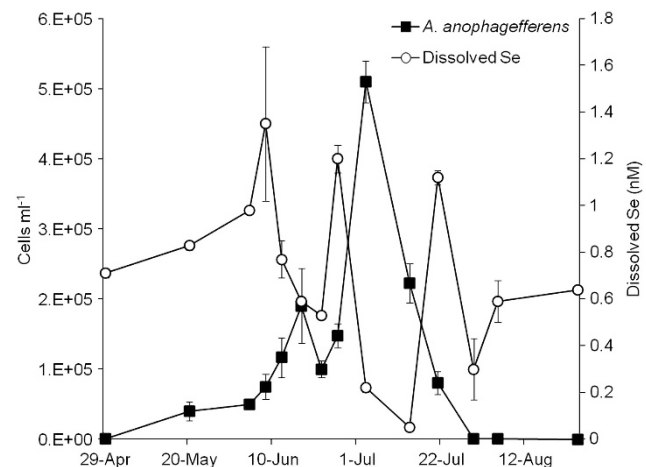

Figure 5 Temporal dynamics of dissolved Se and *A. anophagefferens* abundances during a brown tide bloom in West Neck Bay, NY, USA.

Figure 6 Nucleotide alignment of *Aureococcus* SECIS elements. Conserved nucleotides are highlighted. The SBP2-binding site is shown in red above the sequences.
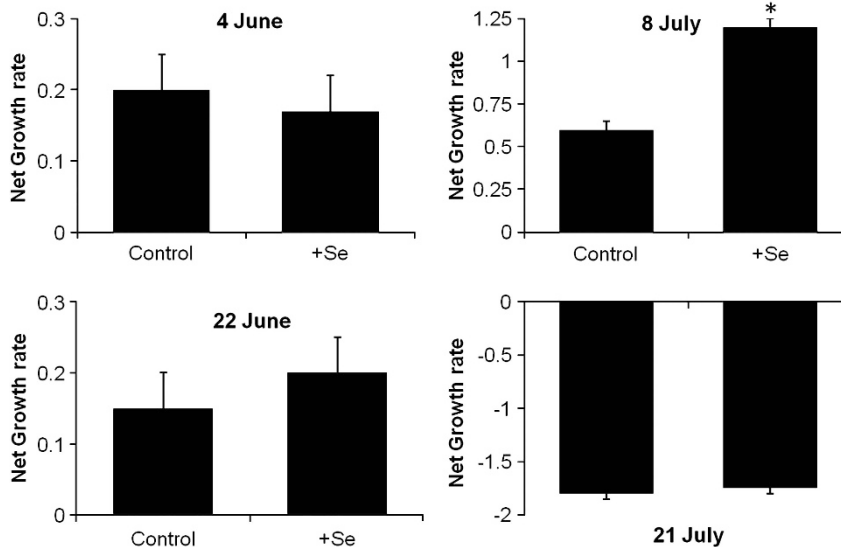

Figure 7 Net growth rates of *A. anophagefferens* during incubation experiments conducted using water from in West Neck Bay, NY, USA, during a brown tide event. The asterisk denotes that the addition of Se significantly enhanced the net growth rates of *A. anophagefferens* during the 8 July experiment ($P<0.001$).

become available, it will be of interest to identify the elements responsible for recognition of such variable SECIS element structures.

The *A. anophagefferens* selenoproteome contains both the majority of eukaryotic selenoproteins (for example, 80% of human and 100% of *Chlamydomonas* selenoprotein families) and several selenoproteins that have previously been detected only in bacteria, suggesting they were acquired through lateral gene transfer from bacteria. In addition, the closest homologs of *A. anophagefferens* selenoproteins are scattered across diverse

eukaryotes, again pointing to lateral transfers of some of these genes. Finally, many completely new selenoproteins were discovered, mostly with unknown functions. Location of Sec in these proteins, however, identifies the catalytic Cys residues in homologs of these proteins in other organisms (Fomenko *et al.*, 2007; Novoselov *et al.*, 2007).

Having the largest and most diverse selenoproteome identified to date implies these selenoproteins have an important role in the fitness and functioning of *A. anophagefferens*. More than half of *A. anophagefferens* selenoproteins possess a thioredoxin fold, and many additional proteins are homologs of known thiol oxidoreductases. In addition, *A. anophagefferens* selenoproteins represent nearly all known redox regulatory systems (thioredoxin, glutaredoxin, glutathione peroxidase, peroxiredoxin, methionine sulfoxide reductase and so on) as well as systems involved in disulfide bond formation and isomerization (protein disulfide isomerase, GILT and so on). The functions of several selenoproteins are unknown, but it is clear that at least some (and possibly most) of them are oxidoreductases. These findings indicate that *A. anophagefferens* uses Se to fine-tune its cellular redox homeostasis (Stadtman, 1996; Kim *et al.*, 2006; Lobanov *et al.*, 2009), an ability that may impart a competitive advantage to this species over its competitors during algal blooms. The shallow nature and rapid vertical mixing times (~ 1 h; Milligan and Cosper, 1997) of estuaries where *A. anophagefferens* resides chronically expose phytoplankton to extremes in light and temperature, which in turn are likely to create elevated levels of free radicals and intracellular oxidation. For example, exposure to excessive light causes algae to produce reactive oxygen species that result in high levels of oxidative stress and damage unless quickly detoxified (Waring *et al.*, 2010). The large number and multiple redundancies of selenoproteins in *A. anophagefferens* are likely to help protect these cells against such oxidation through the removal of hydroperoxides and the repair of oxidatively damaged proteins (Stadtman, 1996). Hence, the large arsenal of selenoproteins in *A. anophagefferens* that prevent cellular oxidation likely infers a competitive advantage to this alga, particularly during summer when it forms HABs and when vertical gradients in temperature and irradiance are at their annual extreme (Cosper *et al.*, 1987).

A highly enhanced use of Sec in *A. anophagefferens* compared with other organisms that co-occur with it should make this organism particularly competitive during HABs when other, non-Se resources such as nitrogen and light can be scarce (Gobler *et al.*, 2011). Sec-containing proteins can be a thousand times more effective as catalysts than their Cys homologs (Kim *et al.*, 2006) and, therefore, selenoproteins may help *A. anophagefferens* to more efficiently organize and execute its metabolism. Increased reliance on Se to support

biosynthesis of proteins also suggests that *A. anophagefferens* can efficiently sequester Se from its environment, a hypothesis confirmed by direct metabolic labeling of *Aureococcus* with ^{75}Se as well as by the decline of dissolved Se in estuaries experiencing brown tides.

Although maximal growth rates for *A. anophagefferens* occurred at Se at or above 5 nM, its half-saturation constant for growth was 0.27 nM. In the environment, dissolved Se concentrations ranged between 0.05 and 1.35 nM, above and below the half-saturation constant, but below the levels that elicited maximal growth of cultures. Importantly, multiple species of Se are present in marine environments (Cutter and Bruland, 1984) and a substantial portion of the total dissolved Se pool may be comprised of forms that are not readily accessible to phytoplankton (Baines and Fisher, 2001; Cutter and Cutter, 2004). As such, the bioavailable pool of Se for *A. anophagefferens* during the HAB in West Neck Bay was likely a fraction of the total dissolved pool (as low as 5% in some estuaries; Cutter and Bruland, 1984; Cutter and Cutter, 2004). In marine ecosystems, nutrient concentrations below half-saturation constants are often considered limiting to algal growth (Caperon and Meyer, 1972; Fisher *et al.*, 1992). The lowest concentrations of Se (0.05–0.22 nM) observed during the peak of the *A. anophagefferens* bloom were below its half-saturation constant and levels required for maximal growth, likely due to the large Se demand created by the bloom. These low Se concentrations likely restricted the growth of *A. anophagefferens* as the addition of 10 nM Se as bioavailable selenite during this period significantly enhanced *A. anophagefferens* growth rates. Hence, while a lack of Se may have contributed to the demise of the *A. anophagefferens* bloom in West Neck Bay, these findings further demonstrate that high levels of Se loading can promote the proliferation of HABs caused by this alga.

A. anophagefferens blooms in shallow, enclosed estuaries such as West Neck Bay where Se concentrations averaged ~ 1 nM, but never blooms in deeper estuaries or continental shelf regions that are characterized by lower Se concentrations (Cutter and Bruland, 1984; Measures et al., 1984; Cutter and Cutter, 1995, 2001). The formation of blooms exclusively in shallow estuaries ensures that *A. anophagefferens* has access to a rich supply of Se required to synthesize these ecologically important and catalytically superior enzymes (Stadtman, 1996; Hatfield and Gladyshev, 2002; Kim *et al.*, 2006). As *A. anophagefferens* relies on selenoproteins for growth and as a scarcity of Se may prohibit bloom formation in off-shore waters, Se is likely to have a key role in shaping the niche space and bloom occurrences of this species. Moreover, as some phytoplankton do not require Se (Harrison *et al.*, 1988), Se availability is likely to shape the succession and composition of phytoplankton communities in general.

Se can be anthropogenically enriched in aquatic environments (Bodek *et al.*, 1988; Cappon, 1991; Cutter and Cutter, 2004; US EPA, 2009). Consistent with this hypothesis, the watershed of West Neck Bay is densely populated with dwellings that are not connected to a sewage treatment plant but rather have cesspools that leach into the groundwater that enters West Neck Bay. Levels of Se measured in groundwater entering the Bay were four-fold greater than concentrations in the bay and other marine surface waters (Cutter and Bruland, 1984; Measures et al., 1984; Cutter and Cutter 1995, 2001, 2004), likely due to wastewater contamination from cesspools and/or fertilizer use on the watershed (Cappon, 1991; McBride and Spiers, 2001; US EPA, 2009). As groundwater is nearly the exclusive source of freshwater to this (Schubert, 1998) and other estuaries which host brown tides (LaRoche *et al.*, 1997), the anthropogenic loading of Se into this and perhaps other coastal ecosystems may lead to the intensification of HABs caused by *A. anophagefferens*. As Se is required by many other phytoplankton that form HABs, including *C. verruculosa* (Imai *et al.*, 1996), *S. trochoidea* (Harrison *et al.*, 1988), *K. rotundatum* (Harrison *et al.*, 1988), *G. catenatum* (Doblin *et al.*,1999, 2000; Band-Schmidt *et al.*, 2004), *A. minutum* (Doblin *et al.*,1999), *Protoceratium reticulatum* (Mitrovic *et al.*, 2004), *K. selliformis* (Mountfort *et al.*, 2006), *C. polykrikoides* (Lee 2008), *P. minimum* (Baines and Fisher, 2001) and *P. gatunense* (Lindstrom, 1991). Anthropogenic loading of this element may contribute to the risk of these events globally.

Acknowledgements

References

Anderson DM, Burkholder JM, Cochlan WP, Glibert PM, Gobler CJ, Heil CA *et al.* (2008). Harmful algal blooms and eutrophication: examples of linkages from selected coastal regions of the United States. *Harmful Algae* **8**: 39–53.

Anderson DM, Kulis DM, Cosper EM. (1989). Immunofluorescent detection of the 'brown tide' organism, *Aureococcus anophagefferens*. In: Cosper EM, Bricelj VM, Carpenter EJ (edn) *Novel phytoplankton blooms: Causes and impacts of recurrent Brown Tides and other unusual blooms* Lecture notes on coastal and estuarine studies. Springer-Verlag, p 213–228.

Baines SB, Fisher NS. (2001). Interspecific differences in the bioconcentration of selenite by phytoplankton and their ecological implications. *Mar Ecol Prog Ser* **213**: 1–12.

Band-Schmidt CJ, Morquecho L, Lechuga-Deveze CH, Anderson DM. (2004). Effects of growth medium, temperature, salinity and seawater source on the growth of *Gymnodinium catenatum* (Dinophyceae) from Bahia Concepcion, Gulf of California, Mexico. *J Plankton Res* **26**: 1459–1470.

Berry MJ, Banu L, Chen YY, Mandel SJ, Kieffer JD, Harney JW *et al.* (1991). Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3′ untranslated region. *Nature* **353**: 273–276.

Bodek I, Lyman WJ, Reehl WF, Rosenblatt DH. (1988). *Environmental inorganic chemistry: properties, processes, and estimation methods*. Pergamon Press: New York, NY, USA.

Bruland KW. (1980). Oceanographic distribution of cadmium, nickel, and copper in the North Pacific. *Earth Planet Sci Lett* **47**: 176–198.

Caperon J, Meyer J. (1972). Nitrogen-limited growth of marine phytoplankton. II Uptake kinetic and their role in nutrient growth of phytoplankton. *Deep-Sea Res* **19**: 619–632.

Cappon CJ. (1991). Sewage sludge as a source of environmental selenium. *Sci Total Environ* **100**: 177–205.

Caron DA, Gobler CJ, Buck NJ, Lonsdale DJ, Cerrato RM, Schaffner RA *et al.* (2004). Microbial herbivory on the brown tide algal, *Aureococcus anophagefferens*: results from natural ecosystems, mesocosms and laboratory experiments. *Harmful Algae* **3**: 439–457.

Cosper EM, Dennison WC, Carpenter EJ, Bricelj VM, Mitchell JG, Kuenstner SH *et al.* (1987). Recurrent and persistent 'brown tide' blooms perturb coastal marine ecosystem. *Estuaries* **10**: 284–290.

Cutter GA. (1982). Selenium in Reducing Water. *Science* **217**: 829–831.

Cutter GA. (1983). Elimination of nitrite interference in the determination of selenium by hydride generation. *Anal Chim Acta* **149**: 391–394.

Cutter GA, Bruland KW. (1984). The marine biogeochemistry of selenium: a re-evaluation. *Limnol Oceanograph* **29**: 1179–1192.

Cutter GA, Cutter LS. (1995). Behavior of dissolved antimony, arsenic, and selenium in the Atlantic Ocean. *Mar Chem* **49**: 295–306.

Cutter GA, Cutter LS. (2001). Sources and cycling of selenium in the western and equatorial Atlantic Ocean. *Deep-Sea Res II* **48**: 2917–2931.

Cutter GA, Cutter LS. (2004). Selenium biogeochemistry in the San Francisco Bay: changes in water column behaviour. *Estuarine Coast Shelf Sci* **61**: 463–476.

Doblin MA, Blackburn SI, Hallegraeff GM. (1999). Comparative study of selenium requirements of three phytoplankton species: *Gymnodinium catenatum, Alexandrium minutum* (Dinophyta) and *Chaetoceros cf. tenuissimus* (Bacillariophyta). *J Plankton Res* **21**: 1153–1169.

Doblin MA, Blackburn SI, Hallegraeff GM. (2000). Intraspecific variation in the selenium requirement of different geographic strains of the toxic dinoflagellate *Gymnodinium catenatum*. *J Plankton Res* **22**: 421–432.

Fisher TR, Peel ER, Ammerman JW, Harding LW. (1992). Nutrient limitation of phytoplankton in Chesapeake Bay. *Mar Ecol Prog Ser* **82**: 51–63.

Fomenko DE, Xing W, Adair BM, Thomas DJ, Gladyshev VN. (2007). High-throughput identification of catalytic redox-active cysteine residues. *Science* **315**: 387–389.

Fries L. (1982). Selenium stimulates growth of marine macroalgae in axenic culture. *J Phycol* **18**: 328–331.

Gobler CJ, Berry DL, Dyhrman ST, Wilhelm SW, Salamov A, Lobanov AV *et al.* (2011). Niche of harmful alga *Aureococcus anophagefferens* revealed through ecogenomics. *Proc Natl Acad Sci USA* **108**: 4352–4357.

Gobler CJ, Donat JR, Consolovo JA, Sañudo-Wilhelmy SA. (2002). Physicochemical speciation of iron during coastal algal blooms. *Mar Chem* **77**: 71–89.

Gobler CJ, Sañudo-Wilhelmy SA. (2001). Temporal variability of groundwater seepage and Brown Tide bloom in a Long Island embayment. *Mar Ecol Prog Ser* **217**: 299–309.

Gobler CJ, Sunda WG. (2012). Ecosystem disruptive algal blooms of the brown tide species, *Aureococcus anophagefferens* and *Aureoumbra lagunensis*. *Harmful Algae* **14**: 36–45.

Harrison PJ, Yu PW, Thompson PA, Price NM, Phillips DJ. (1988). Survey of selenium requirements in marine Phytoplankton. *Mar Ecol Prog Ser* **47**: 89–96.

Hatfield DL, Gladyshev VN. (2002). How selenium has altered our understanding of the genetic code. *Mol Cell Biol* **22**: 3565–3576.

Heisler J, Glibert PM, Burkholder JM, Anderson DM, Cochlan W, Dennison WC *et al.* (2008). Eutrophication and harmful algal blooms: a scientific consensus. *Harmful Algae* **8**: 3–13.

Imai I, Itakura S, Matsuyama Y. (1996). Selenium requirement for growth of a novel red tide flagellate *Chattonella verruculosa* (Raphidophyceae) in culture. *Fisheries Sci* **62**: 834–835.

Keller M, Guillard RRL, Provasoli L, Pintner IJ. (1984). Nutrition of some marine ultraplankton clones from the Sargasso Sea. *EOS* **65**: 898.

Kim HY, Fomenko DE, Yoon YE, Gladyshev VN. (2006). Catalytic advantages provided by selenocysteine in methionine-s-sulfoxide reductases. *Biochem* **45**: 13697–13704.

LaRoche J, Nuzzi R, Waters R, Wyman K, Falkowski PG, Wallace DWR. (1997). Brown tide blooms in Long Island's coastal waters linked to variability in groundwater flow. *Global Change Biol* **3**: 397–410.

Lee YS. (2008). Utilization of various nitrogen, phosphorus, and selenium compounds by *Cochlodinium polykrikoides*. *J Environ Biol* **29**: 799–804.

Lindstrom K. (1991). Nutrient requirements of the dinoflagellate *Peridinium gatunense*. *J Phycol* **27**: 207–219.

Lobanov AV, Fomenko DE, Zhang Y, Sengupta A, Hatfield DL, Gladyshev VN. (2007). Evolutionary dynamics of eukaryotic selenoproteomes: large selenoproteomes may associate with aquatic and small with terrestrial life. *Genome Biol* **8**: R198.

Lobanov AV, Hatfield DL, Gladyshev VN. (2009). Eukaryotic Selenoproteins and selenoproteomes. *Biochim Biophys Acta* **1790**: 1424–1428.

McBride MB, Spiers G. (2001). Trace element content of selected fertilizers and dairy manures as determined by ICP-MS. *Comm Soil Sci Plant Anal* **32**: 139–156.

Measures CI, Grant B, Khadem M, Edmond JM. (1984). Distribution of Be, Al, Se and Bi in the surface waters of the western North Atlantic and Caribbean. *Earth Planet.Sci Lett* **71**: 1–12.

Milligan AJ, Cosper EM. (1997). Growth and photosynthesis of the 'brown tide' microalga *Aureococcus anophagefferens* in subsaturating constant and fluctuating irradiance. *Mar Ecol Prog Ser* **153**: 67–75.

Mitrovic SM, Amandi MF, McKenzie L, Furey A, James KJ. (2004). Effects of selenium, iron and cobalt addition to growth and yessotoxin production of the toxic marine dinoflagellate *Protoceratium reticulatum* in culture. *J Exp Mar Biol Ecol* **313**: 337–351.

Mountfort D, Beuzenberg V, MacKenzie L, Rhodes L. (2006). Enhancement of growth and gymnodimine production by the marine dinoflagellate, *Karenia selliformis*. *Harmful Algae* **5**: 658–664.

Novoselov SV, Lobanov AV, Hua D, Kasaikina MV, Hatfield DL, Gladyshev VN. (2007). A highly efficient form of the selenocysteine insertion sequence element in protozoan parasites and its use in mammalian cells. *Proc Natl Acad Sci USA* **104**: 7857–7862.

Novoselov SV, Rao M, Onoshko NV, Zhi H, Kryukov GV, Xiang Y *et al.* (2002). Selenoproteins and selenocysteine insertion system in the model plant cell system. *Chlamydomonas reinhardtii. EMBO J* **21**: 3681–3693.

Palenik B, Grimwood J, Aerts A, Rouzé P, Salamov A, Putnam N *et al.* (2007). The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc Natl Acad Sci USA* **104**: 7705–7710.

Parsons TR, Maita Y, Lalli CM. (1984). *A manual of chemical and biological methods for seawater analysis.* Pergamon.

Puls RW, Paul CJ. (1995). Low-flow purging and sampling of ground water monitoring wells with dedicated systems. *Ground Water Monitoring Rev* **15**: 116–123.

Schubert CE. (1998). *Areas contributing ground water to the Peconic Estuary, and ground-water budgets for the North and South Forks and Shelter Island, Eastern Suffolk County.* Geological Survey Water Resources Investigations Report: New York, NY, USA 97–4136.

Stadtman TC. (1996). Selenocysteine. *Ann Rev Biochem* **65**: 83–100.

US EPA (2009). *Targeted National Sewage Sludge Survey Report—Overview.* EPA, 822-R-08-014.

Waring J, Klenell M, Bechtold U, Underwood GJC, Baker NR. (2010). Light-induced responses of oxygen photoreduction, reactive oxygen species production and scavenging in two diatom species. *J Phycol* **46**: 1206–1217.

Wheeler AE, Zingaro RA, Irgolic K, Bottino NR. (1982). The effect of selenate, selenite and sulfate on the growth of six unicellular marine algae. *J Exp Mar Biol Ecol* **57**: 181–194.