Research Article

Transcription Regulation of Plastid Genes Involved in Sulfate Transport in Viridiplantae

Vassily A. Lyubetsky, Alexander V. Seliverstov, and Oleg A. Zverkov

Institute for Information Transmission Problems (Kharkevich Institute), The Russian Academy of Sciences, Moscow 127994, Russia

Correspondence should be addressed to Vassily A. Lyubetsky; lyubetsk@iitp.ru

Received 1 April 2013; Accepted 12 June 2013

Academic Editor: William Piel

Copyright © 2013 Vassily A. Lyubetsky et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study considers transcription regulation of plastid genes involved in sulfate transport in the parasites of invertebrate (Helicosporidium sp.) and other species of the Viridiplantae. A one-box conserved motif with the consensus TAAWATGATT is found near promoters upstream the cysT and cysA genes in many species. In certain cases, the motif is repeated two or three times.

1. Introduction

This study focuses on selected species of the Viridiplantae, particularly, the genus Helicosporidium sp. (class Trebouxiophyceae), which comprises green algae parasitizing flies of the species Simulium jonesi [1–3]. Plastids of these parasites are a good target for antibiotic treatment, as earlier was shown for apicomplexan parasites of vertebrates (Toxoplasma gondii and Plasmodium spp. [4]).

The plastome of Helicosporidium sp. is relatively small, about 37 kb. Most of the plastome genes encode tRNA, rRNA, ribosomal proteins, and subunits of the bacterial-type RNA polymerase. One of two nonhousekeeping proteins is the CysT subunit of a sulfate ABC transporter.

Sulfate ABC transporters in cyanobacteria and proteobacteria consist of two identical ATP-binding CysA proteins, two transmembrane proteins (CysT and CysW), and a sulfate-binding protein SbpA. In the cyanobacteria Synechocystis sp. PCC 6803 [5], genes encoding the sulfate transporter subunits are arranged in a single operon sbpA-ssr2439-cysT-cysW-cysA. In cyanobacteria, no data on expression is available for this operon; however, in Escherichia coli and some other proteobacteria, genes of the sulfate transporter subunits are known to be regulated in the single operon cysPTWAM (further details are given in Discussion).

Plastomes of vascular plants lack genes of the sulfate transport system except for rare instances of cysT and cysA. However, the green alga Helicosporidium sp. retains cysT. Plastomes of the rhodophyte Cyanidium caldarium and Cyanidioschyzon merolae and the cyanelle genome of Cyanophora paradoxa lack cysT homologues but possess distant homologues of cysA presumably involved in the transport of zinc or manganese (further details are given in Results). Similar proteins are involved in the transport of molybdenum, zinc, and manganese and belong to a large family of transporters of ions, sugars, peptides, and more complex organic molecules. For example, the transcription regulation of the ziaA gene (encoding a polypeptide similar to a P-type ATPases involved in transporting heavy metals) is described in the cyanobacterium Synechocystis PCC 6803 [6].

The sulfate transport in plastids is necessary for the synthesis of many sulfur-containing compounds. For example, in Spinacia oleracea, the lack of sulfates leads to considerable changes in the expression of cysteine synthesis genes [7]. Also, plastids of many algae synthesize thiamine and other sulfur-containing compounds. For example, the lipoic acid is synthesized in plastids of apicomplexan parasites [8].

In this paper, we consider the expression regulation of cysT and cysA in Viridiplantae, in particular, Helicosporidium sp. and Pycnococcus provasolii, where cysT is present and cysA is absent.

In proteobacteria, the regulation mechanism of transcription initiation of cysA and cysT is known. The CysB protein is a transcription factor of the LysR family and acts as a tetramer. This protein binds DNA upstream the −35 box of a promoter and activates transcription initiation of
Table 1: CysA and CysT proteins encoded in plastids of the Viridiplantae and in the cyanobacterium *Synechocystis* sp. PCC 6803.

| Species                             | DNA locus    | GC, % | CysA          | CysT          |
|-------------------------------------|--------------|-------|---------------|---------------|
| *Synechocystis* sp. PCC 6803        | NC_017277.1  | 49    | YP_005652904.1 | YP_005652902.1 |
| Nephrolepis olivacea                | NC_000927.1  | 42    | NP_050872.1   | NP_050928.1   |
| Pycnococcus provasolii              | NC_012097.1  | 40    | None          | YP_00260832.1 |
| *Bryopsis* hypnoides                | NC_013359.1  | 33    | YP_003227066.1 | YP_003227041.1 |
| *Chlorella* variabilis              | NC_015359.1  | 34    | YP_004347745.1 | YP_004347762.1 |
| *Chlorella* vulgaris                | NC_001865.1  | 32    | NP_045832.1   | NP_045890.1   |
| *Coccomyxa subellipsoidea*          | NC_015084.1  | 51    | YP_004222028.1 | YP_004221988.1 |
| *Helicosporidium* sp.               | NC_008100.1  | 27    | None          | YP_635918.1   |
| *Leptosira* terestris               | NC_009681.1  | 27    | YP_001382174.1 | YP_001382135.1 |
| *Parachlorella* kessleri            | NC_012978.1  | 30    | YP_003058285.1 | YP_003058340.1 |
| *Mesostigma viride*                 | NC_002186.1  | 30    | NP_038429.1   | NP_038441.1   |
| *Chlorokybus* atmoophyicus          | NC_008822.1  | 36    | YP_001019096.1 | YP_001019170.1 |
| *Zygnema* circumcarinatum           | NC_008171.1  | 31    | YP_636486.1   | YP_636569.1   |
| *Anthoceros* formosae               | NC_004543.1  | 33    | NP_777407.1   | NP_777462.1   |
| *Aneura* mirabilis                  | NC_010359.1  | 41    | Pseudogene    | Pseudogene    |
| *Marchantia* polymorpha             | NC_001319.1  | 29    | NP_039293.1   | NP_039346.1   |
| *Ptilidium* pulcherrimum             | NC_015402.1  | 33    | Pseudogene    | Pseudogene    |

many operons. For proteobacteria *Salmonella typhimurium*, *Escherichia coli* [9], and *Klebsiella aerogenes* [10], the CysB activation of *cysPTWAM*, *cysK*, *cysIH*, *cysDNC*, *sbp*, and L-cysteine transport genes, as well as CysB autorepression is described in detail. Binding of CysB to DNA is not directly dependent on sulfate concentration but requires high concentrations of acetylserine. Also, proteobacteria lack a distinct binding motif for CysB.

2. Materials and Methods

The list of species is given in Table 1. Genomes were obtained from GenBank, NCBI. Clustering of proteins was performed using the method described in [11, 12].

An original algorithm from [13, 14] was employed to search for bacterial-type promoters. Relevant promoter sequences and the evolutionary impact of DNA point mutations on polymerase binding affinity are described in [15, 16]; experimental evidence was obtained using the *psbA* promoter of *Sinapis alba* [17].

A novel method based on clique search in a multipartite graph [18] was used to identify conserved motifs. In current modification of the method, the nucleotide similarity was estimated accounting for the GC content in plastid DNA. Namely, if the average GC-rate was \( p \), the additive contribution for a mismatch at any position in the calculation of the distance between two words of equal lengths was \((1 - p) \) for a A-T pair, \( p \) for a C-G pair, and \( 1/2 \) for a S-W pair, where \( S = \{C,G\} \) and \( W = \{A,T\} \). A large-scale search for binding sites was performed using formulas from [19]. Protein alignments were constructed with ClustalW v. 2.0.3 [20].

3. Results

3.1. Analysis of the Domain Structure and Phylogenetic Classification of Proteins. The *cysT* gene is present in the following species of Viridiplantae: Chlorophyta (*Bryopsis hypnoides*, *Nephrolepis olivacea*, *Pycnococcus provasolii*, *Chlorella variabilis*, *Chlorella vulgaris*, *Coccomyxa subellipsoidea* C-169, *Helicosporidium* sp. ex *Simulium jonesii*, *Leptosira* terestris, and *Parachlorella* kessleri), Streptophyta (*Chlorokybus atmoophyicus*, *Mesostigma viride*, and *Zygnema circumcarinatum*), Bryophyta (*Anthoceros formosae*, *Marchantia polymorpha*, and *Aneura mirabilis*), and *Ptilidium pulcherrimum* (*A. m.* and *Pp.* are pseudogenes) [21].

CysT proteins are conserved across green algae, cyanobacteria, and proteobacteria and function as the transmembrane domain of the ABC transporter (Figure 1). Short N-terminal regions of these proteins can vary. Another exception is CysT in *Bryopsis hypnoides* and *Leptosira terestris* that have truncated C-termini.

Cyanobacteria possess strong potential orthologs of *cysT* in the Viridiplantae. Cyanobacterial CysT is considered to be ancestral.

As CysA and CysT functions are linked, *cysA* and *cysT* normally either coexist or are absent in plastids of all Viridiplantae with the exception of *Helicosporidium* sp. and *Pycnococcus provasolii*. (Note that the *cysA* ortholog in *Marchantia polymorpha* is named *mbp.X*.) Viridiplantae species lacking *cysA* and *cysT* are closely related to the species that contain both genes [11, 22]. Unexpectedly, *cysA* and *cysT* are present in Bryophyta while they are absent in many highly organized algae close to land plants (*Chara vulgaris*, *Staurastrum punctulatum*). They are also absent in plastomes of *Physcomitrella patens* and all vascular plants. In green algae, these genes are mainly present in the class Trebouxiophyceae (genera *Chlorella*, *Coccomyxa*, *Helicosporidium*, *Leptosira*, and *Parachlorella*).

Sequences of plastid CysA and their orthologs from cyanobacterium are well aligned (Figure 2). CysA in all Viridiplantae has a highly conservative N-terminus domain which is typical for the ATP-binding cassette.
of ABC transporters. In all studied Chlorophyta, except for *Nephroselmis olivacea*, this protein possesses a short C-terminus. Conversely, in the Streptophyta, *Nephroselmis olivacea*, and cyanobacteria, C-termini are long and conservative. In *Mesostigma viride* and *Chlorokybus atomophyticus*, this domain is homologous to the TOBE domain involved in sulfate binding [23]. According to the Pfam 26.0 database [24], e-value for this domain is 0.0017 in *M. viride* and 0.00007 in *Ch. atomophyticus*. Other plastid-encoded proteins, although also being well conserved, have a lower score for this domain. There is no sulfate-binding CysP (SbpA) subunit in plastids which could indicate that cysP is located in the nucleus.

Outside of the Viridiplantae group, CysA orthologs are encoded in plastids of *Cyanidium caldarium* (NP_045139.1), *Cyanidoschyzon merolae* (NP_848950.1), and *Cyanophora paradoxa* (NP_043273.1). Interestingly, cyanobacteria CysA orthologs differ from plastid CysA orthologs. The NP043273.1 protein of *C. paradoxa* is an ortholog for the substrate-binding subunit of a zinc or manganese transporter.

3.2. Analysis of the Genomic Context. Genes upstream and downstream cysT and cysA are listed in Table 2. The rpl32 gene located upstream cysT encodes the ribosomal protein L32 and in most cases belongs to the same DNA strand as cysT; refer to Table 2. *Pycnococcus provasolii*, *Bryopsis hypnoides*,

---

**Figure 1:** A multiple alignment of CysT orthologs from the cyanobacterium *Synechocystis* sp. PCC 6803 and plastids of the Viridiplantae. Conservativity is shown above and below the columns.

**Figure 2:** A multiple alignment of CysA orthologs from the cyanobacterium *Synechocystis* sp. PCC 6803 and plastids of the Viridiplantae.
Helicosporidium sp., Leptosira terrestris, and Zygmena circumbinatum have different gene configurations of this loci; refer to Table 2. Only in Chlorokybus atmosphyticus, rpl32 is both upstream of cysT and belongs to a different strand than cysT. In most cases, the intergenic region upstream the cysT gene is quite long. The ycf1 gene is present downstream cysT in most algae. A few other genes are found downstream cysT: tRNA in Bryophyta and the alga Leptosira terrestris; rpl21 (L21 protein) in the alga Zygmena circumbinatum; rpoA (alpha subunit of bacterial-type RNA polymerase) in Bryopsis hypnoides; refer to Table 2.

The accD gene is located upstream cysA and belongs to the same DNA strand in Trebouxiophyceae algae, however, for Leptosira terrestris. In Nephroselmis olivacea, however, a tRNA gene upstream cysA is on the complementary strand. In Bryopsis hypnoides, ccsA is upstream cysA, and the intergenic region is very short. In Streptophyta, cysA is surrounded by tRNA genes, and they often belong to the complementary strand which suggests the presence of a promoter directly in the upstream region of cysA.

3.3. Searching for Bacterial Type Promoters. Only two candidate bacterial-type promoters are found in 5'-leader regions of the considered genes; refer to Table 2. The exception is the cysA gene in Anthoceros formosae, for which we detect three potential promoters of similar quality. In Chlorella vulgaris, the single promoter candidate is located upstream cysA and has the unusual ~35 box, AAGAAA, which was the reason for its rejection. However, in Ch. variabilis, a good potential promoter is detected in the upstream region of this gene with a TG-extension of the ~10 box. Promoters were not found in the upstream regions of either cysA or cysT in Nephroselmis olivacea, Pycnococcus provasolii, Bryopsis hypnoides, Leptosira terrestris, Anceura mirabilis, and Pitilidium pulcherrimum. Promoters were not found in the upstream region of cysA in Chlorokybus atmosphyticus and in the upstream region of cysT in Zygmena circumbinatum. We speculate that in these cases these genes are transcribed as a part of an operon or by an RNA polymerase of the phage type.

3.4. Searching for the Conservative Motif. Transcription regulation of plastid genes involved in the sulfate transport in the parasites of invertebrate (Helicosporidium sp.) and in other species of the Viridiplantae is considered. A one-box conserved motif with the consensus TAAWAGATT is found near the promoters in the upstream regions of the cysT and cysA genes in many species. In some cases, the motif is repeated two or three times. In the upstream region of the cysA promoter in alga C. subellipsoidea C-169, however, the entire motif is repeated twice and is supplemented with its partial repeat at the 5'-terminus. The motif is not present near the promoters in Chlorokybus atmosphyticus and Marchantia polymorpha. Deviations from the motif consensus are often the same in the same taxonomic lineage, which may reflect the variability of the transcription factor between lineages. The consensus was obtained from multiple alignments of 28 regions upstream two genes in 9 species (Coccomyxa subellipsoidea, Chlorella variabilis, Chlorella vulgaris, Helicosporidium, Parachlorella kessleri, Mesostigma viride, Chlorokybus atmosphyticus, Zygmena circumbinatum, and Anthoceros formosae). The LOGO profile of this motif is shown in Figure 3.

In most species, the motif is found upstream the ~35 box or is overlapping the promoter. In the cysA upstream region in Zygmena circumbinatum and Anthoceros formosae, the motif is detected between ~35 and ~10 boxes or is overlapping the ~10 box of the promoter.

4. Discussion

We believe that the found motif represents binding sites of a transcription factor because of its positional linking with
the promoter. The variable distance between the motif and the promoters suggests a repressor role of a putative transcription factor. Repeating of the motif is typically associated with a cooperative factor binding. This cooperativity can compensate for the motif variability, which is the case of *Coccomyxa subellipsoidea*.

The motif sequence confirms that *Helicosporidium* sp. belongs to the class Trebouxiophyceae. Its conservativity emphasizes the importance of *cysT* in plastids of parasites and suggests its key importance in understanding the role of the plastids in virulence. Indeed, plastids often synthesize many chemicals, which are usually provided by mitochondria [7, 25].

In *Leptosira terestris*, *cysA* and *cysT* are not predicted to have the regulatory sites, unlike other Trebouxiophyceae and their close relatives, which suggests a shift in the transporter domain. But the *CysW* protein is absent in plastids. This observation conforms with considerable changes in the CysA sequence in *Leptosira*. The lack of regulatory sites in *Nephrospum olivacea*, *Pycnococcus provasolii*, *Bryopsis hypnoides*, and *Marchantia polymorpha* may suggest a demising role of the protein, which is consistent with the loss or pseudogene nature of *cysA* and *cysT* in *Aneura mirabilis* and *Atlidium pulcherrimum*. The absence of bacterial-type promoters upstream of *cysA* and *cysT* is often associated with changes in genes order on the chromosome. This effect may be explained by the de novo formation of phage-type promoters (possibly activated by another factor), or the inclusion of *cysA* or *cysT* in another operon. In general, the sulfate transport can be regulated by changing the expression level of a nuclear encoded sulfate-binding domain *CysP* (SbpA).

CysB binding sites in proteobacteria [9, 10] differ considerably from binding sites of a putative factor that we predicted for *cysT* and *cysA*. However, the two most conserved motif positions in plastids coincide with the two conserved positions in experimentally characterized sites upstream *cysPTWAM*, *cysK*, *cysJH*, *cysLNC*, *sbp*, and *cysB* in proteobacteria. This evidence is however insufficient to establish the identity of our predicted motif and the CysB binding sites in proteobacteria. In *E. coli*, both proteins CysT and CysW consist of transmembrane domains that are very close to each other. Their genes belong to the sulfate transport operon *cysPTWAM*. But the CysW protein is absent in plastids. We hypothesize that CysW is replaced by the second CysT copy in plastids. The *cysT* and *cysA* genes do not form an operon in plastids, so we assume the CysT protein’s double abundance over the CysA protein. It conforms to our hypothesis.

In *E. coli*, CysT and CysW are transmembrane domains with similar structure. Their genes belong to the same operon with the sulfate transport subunit *cysPTWAM*; however, *cysW* is absent in plastids. We hypothesize that in plastids the CysW subunit is functionally replaced by another copy of CysT, and the *cysT* mRNA concentration is twice as high compared to *cysA* mRNA. This hypothesis is indirectly supported by the fact that *cysT* and *cysA* are not included in one operon in plastids, and thus their mRNA expression levels may differ considerably.

**Acknowledgments**

The authors are deeply grateful to L. Yu. Rusin, A. V. Lyubetskaya and S. A. Korolev for a valuable discussion of the results and methods, helping in their preparation. This work was partly supported by the Ministry of Education and Science of the Russian Federation (State Contract 14.740.11.1053 and Agreement 8481).

**References**

[1] A. Tartar, D. G. Boucias, B. J. Adams, and J. J. Becnel, “Phylogenetic analysis identifies the invertebrate pathogen *Helicosporidium* sp. as a green alga (Chlorophyta),” *International Journal of Systematic and Evolutionary Microbiology*, vol. 52, no. 1, pp. 273–279, 2002.

[2] A. P. de Koning and P. J. Keeling, “The complete plastid genome sequence of the parasitic green alga Helicosporidium sp. is highly reduced and structured,” *BMC Biology*, vol. 4, article 12, 2006.

[3] J. F. Pombert and P. J. Keeling, “The mitochondrial genome of the entomoparasitic green alga Helicosporidium,” *PloS ONE*, vol. 5, no. 1, Article ID e8954, 2010.

[4] T. A. Sadovskaya and A. V. Seliverstov, “Analysis of the 5'-Leader regions of several plastid genes in protozoa of the phylum apicomplexa and red algae,” *Molecular Biology*, vol. 43, no. 4, pp. 552–556, 2009.

[5] N. Tajima, S. Sato, F. Maruyama et al., “Genomic structure of the cyanobacterium *Synechocystis* sp. PCC 6803 strain GT-S,” *DNA Research*, vol. 18, no. 5, pp. 393–399, 2011.

[6] C. Thelwell, N. J. Robinson, and J. S. Turner-Cavet, “An SmtB-like repressor from *Synechocystis* PCC 6803 regulates a zinc exporter;,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 18, pp. 10728–10733, 1998.

[7] H. Takahashi and K. Saito, “Subcellular localization of spinach cysteine synthase isoforms and regulation of their gene expression by nitrogen and sulfur,” *Plant Physiology*, vol. 112, no. 1, pp. 273–280, 1996.

[8] N. Thomsen-Zieger, J. Schachtner, and F. Seeber, “Apicomplexan parasites contain a single lipoic acid synthase located in the plastid,” *FEBS Letters*, vol. 547, no. 1–3, pp. 80–86, 2003.

[9] N. M. Kredich, “The molecular basis for positive regulation of cys promoters in *Salmonella typhimurium* and *Escherichia coli*,” *Molecular Microbiology*, vol. 6, no. 19, pp. 2747–2753, 1992.

[10] A. S. Lynch, R. Tyrrell, S. J. Smerdon, G. S. Briggs, and A. J. Wilkinson, “Characterization of the CysB protein of *Klebsiella aerogenes*: direct evidence that N-acetylserine rather than O-acetylserine serves as the inducer of the cysteine regulon,” *Biochemical Journal*, vol. 299, no. 1, pp. 129–136, 1994.

[11] O. A. Zverkow, L. Rusin Yu, A. V. Seliverstov, and V. A. Lyubetsky, “Study of direct repeats in micro evolution of plant
mitochondria and plasts based on protein clustering,” *Moscow University Biological Sciences Bulletin*, vol. 68, no. 2, pp. 58–62, 2013.

[12] O. A. Zverkov, A. V. Seliverstov, and V. A. Lyubetsky, “Plastid-encoded protein families specific for narrow taxonomic groups of algae and protozoa,” *Molecular Biology*, vol. 46, no. 5, pp. 717–726, 2012.

[13] A. V. Seliverstov, E. A. Lysenko, and V. A. Lyubetsky, “Rapid evolution of promoters for the plastome gene ndhF in flowering plants,” *Russian Journal of Plant Physiology*, vol. 56, no. 6, pp. 838–845, 2009.

[14] V. A. Lyubetsky, L. I. Rubanov, and A. V. Seliverstov, “Lack of conservation of bacterial type promoters in plasts of Streptophyta,” *Biology Direct*, vol. 5, article 34, 2010.

[15] V. A. Lyubetsky, O. A. Zverkov, L. I. Rubanov, and A. V. Seliverstov, “Modeling RNA polymerase competition: the effect of σ-subunit knockout and heat shock on gene transcription level,” *Biology Direct*, vol. 6, article 3, 2011.

[16] V. A. Lyubetsky, O. A. Zverkov, S. A. Pirogov, L. I. Rubanov, and A. V. Seliverstov, “Modeling RNA polymerase interaction in mitochondria of chordates,” *Biology Direct*, vol. 7, article 26, 2012.

[17] A. Homann and G. Link, “DNA-binding and transcription characteristics of three cloned sigma factors from mustard (*Sinapis alba* L.) suggest overlapping and distinct roles in plastid gene expression,” *European Journal of Biochemistry*, vol. 270, no. 6, pp. 1288–1300, 2003.

[18] V. A. Lyubetsky and A. V. Seliverstov, “Some algorithms related to finite groups,” *Information Processes*, vol. 3, no. 1, pp. 39–46, 2003 (Russian).

[19] Z. Su, V. Olman, F. Mao, and Y. Xu, “Comparative genomics analysis of NtcA regulons in cyanobacteria: regulation of nitrogen assimilation and its coupling to photosynthesis,” *Nucleic Acids Research*, vol. 33, no. 16, pp. D290–D301, 2012.

[20] J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins, “The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools,” *Nucleic Acids Research*, vol. 25, no. 24, pp. 4876–4882, 1997.

[21] N. J. Wickett, L. L. Forrest, J. M. Budke, B. Shaw, and B. Goffinet, “Frequent pseudogenization and loss of the plastid-encoded sulfate-transport gene cysA throughout the evolution of liverworts,” *American Journal of Botany*, vol. 98, no. 8, pp. 1263–1275, 2011.

[22] V. A. Lyubetsky, A. V. Seliverstov, and O. A. Zverkov, “Elaboration of the homologous plastid-encoded protein families that separate paralogs in magnoliophytes,” *Mathematical Biology and Bioinformatics*, vol. 8, no. 1, pp. 225–233, 2013 (Russian).

[23] E. V. Koonin, Y. I. Wolf, and L. Aravind, “Protein fold recognition using sequence profiles and its application in structural genomics,” *Advances in Protein Chemistry*, vol. 54, pp. 245–275, 2000.

[24] M. Punta, P. C. Coggill, R. Y. Eberhardt et al., “The Pfam protein families database,” *Nucleic Acids Research*, Database Issue 40, pp. D290–D301, 2012.

[25] R. J. M. Wilson, K. Rangachari, J. W. Saldanha et al., “Parasite plastsids: maintenance and functions,” *Philosophical Transactions of the Royal Society B*, vol. 358, no. 1429, pp. 155–164, 2003.