Spring 5-8-2018

Examining Manganese Oxidation in Bacteria through Bioinformatics and Environmental Sampling

T'onna Collins
collins_t@lynchburg.edu

Examining Manganese Oxidation in Bacteria through Bioinformatics

and Environmental Sampling

T'onna Collins

Seniors Honors Project

Submitted in partial fulfillment of the graduation requirements of the Westover Honors Program

Westover Honors Program

May, 2018

_____

Christine Terry, PhD

_____

Katherine Geszvain, PhD

_____

M. Zakaria Kurdi, PhD

_____

Nancy Cowden, PhD

Abstract

Manganese is an essential element for all living things. Manganese plays a role in geochemical cycles, like the nitrogen fixation cycle and the oxygen-evolving complex. The focus of this study is the manganese oxidation mechanism. Manganese oxidation plays a role in metabolic and antioxidant processes. *Pseudomonas putida* GB-1 is commonly used to study the process of manganese oxidation in bacteria. We used the previously identified Mn-oxidase genes, *mopA*, *mnxG* and *mcoA,* from *P. putida* GB-1 to look for conservation of manganese oxidation in other organisms. Basic Local Alignment Search Tool (BLAST) was used to find homologs of these three genes in other bacterial species. We focused our analysis on loci that were found in non-pseudomonad species. Sequence alignments were used to analyze similarities and differences on both an amino acid and DNA scales. To complement these bioinformatic studies, microbes were isolated from around the campus of Lynchburg College, Lynchburg, Virginia, and tested for the ability to oxidize manganese. Those that tested positive were identified through sequencing of the 16S rRNA gene. Additional studies of the manganese oxidizing capabilities of these strains were described. The goal was to see what genes are conserved from species to species in the manganese oxidation mechanism.

Introduction

Manganese (Mn) is an abundant, essential trace metal and is used by some enzymes that catalyze oxidation-reduction reactions. An oxidation-reduction, or redox, reaction is the transfer of electrons from one molecule to another. Oxidation is the loss of electrons and reduction is the gain of electrons. Redox reactions are integral to life as we know it. In the human body, for example, cellular respiration is a series of redox reactions where glucose is oxidized to carbon dioxide, oxygen is reduced to water, and energy is harvested as ATP. Manganese oxidation is important in antioxidant and metabolic functions in both eukaryotes and prokaryotes (Learman and Hansel 2014). The function of manganese changes based on its oxidation state. The oxidation state of an element represents the number of electrons lost or gained; Mn(IV) refers to the fact that this element has two fewer electrons than Mn(II). Some examples where Mn(II) is important include: (i)helping to split water during photosynthesis and (ii) aiding superoxide dismutase convert the highly toxic free radical, O_2, to hydrogen peroxide (Lohry 2007).

Six enzymes have been identified as putative manganese oxidases in bacteria: MoxA, MnxG, MofA, MopA, McoA, and CotA (Tang et al. 2014). Bacterial Mn (II)-oxidizing enzymes fall into two categories: multicopper oxidases (MCOs) and heme peroxidases (Geszvain et al. 2012). There are large MCOs like the MnxG enzymes from *Bacillus* and *Pseudomonas putida* GB-1. There are also small MCOs, including McoA, MoxA, MofA, and CotA. MopA is a member of the heme peroxidase family. A heme peroxidase contains heme, an iron-containing molecule, and uses hydrogen peroxide as an electron acceptor to catalyze oxidative reactions. It is difficult to study the genes responsible for manganese oxidation when there are distinctly different ways that manganese oxidation can occur.

Manganese oxidation is not essential for survival of some bacteria; *P. putida* GB-1 was able to live when manganese oxidation genes were deleted (Gezsvain, Smesrud, and Tebo 2016). With this discovery, the question arises regarding the purpose of manganese oxidation in microbes if it is not needed for survival. Manganese oxidation may not be crucial to the organisms performing the reaction but may be essential for other organisms. Manganese oxidation plays a role in geochemical cycles that are important to other organisms. For example, manganese plays a role in the nitrogen fixation cycle by helping plants grow and develop with the synthesis of polyamines (Weisany, Raei, and Allahverdipoor 2013). Manganese also plays a role in photosynthesis by oxidizing water to dioxygen in the oxygen-evolving complex of photosystem II. This is an important reaction because it is the source for almost all atmospheric oxygen (Pecoraro, Gelasco, and Baldwin 1996). This study provides an opportunity to learn about manganese oxidation in additional microbial species, therefore helping us to learn more about the genes and protein products involved.

Pseudomonas putida GB-1 is a bacterial model organism used to examine and analyze the enzymes linked to Mn(II) oxidation. *Pseudomonas putida* GB-1 contains genes for three different manganese oxidizing enzymes: (i) *mnxG* (PputGB1_2447), a MCO with near homology to *B. subtilis* MnxG, (ii) *mcoA* (PputGB1_2665), (iii) *mopA* (PputGB1_3553).

Pseudomonas putida GB-1 MnxG was directly identified as an ortholog of *Bacillus* MnxG via mass spectroscopy and was shown to be a Mn(II) oxidase through genetic analysis (Dick et al. 2007). Specific regions within a manganese oxidation enzyme bind to (or coordinate) copper ions that play a direct role in manganese oxidation. Within these copper-binding domains, histidine residues are necessary for the copper-binding function; *P. putida* GB-1 MnxG

(PputGB1_2447) exhibits a high level of homology to Bacillus *mnxG* for all but one of these copper-binding domains (Dick et al. 2007).

It is unusual for an organism to have two Mn(II) oxidase enzymes present (Dick et al. 2007). However, in *P. putida* FB-1, MnxG and McoA function as a partially redundant, complementary MCO pair (Geszvain, McCarthy, and Tebo 2013). A third Mn(II) oxidase enzyme, MopA, was discovered when *fleQ* was deleted from *P. putida* GB-1 (Geszvain et al. 2016). *FleQ* is a σ^{54}-dependent transcription factor that positively regulates MopA in *P. putida* GB-1. Deletion of *mopA* from a *fleQ*-positive strain of *P. putida* GB-1 did not result in a decrease of Mn(II) oxidation because pseudonomad MnxG and McoA can perform most of Mn(II) oxidation that was completed.

There are a variety of other bacteria besides *P. putida* and *Bacillus* that oxidize manganese. For example, *Citrobacter sp. FM-2* has multiple, multicopper oxidase genes that oxidize manganese (Tang et al. 2014). These genes have three conserved regions that are similar to MnxG, MofA, and MoxA (the multicopper oxidases that are found in *Pseudomonas* and *Bacillus* species). Manganese oxidation in *Citrobacter sp.* FM-2 has evolved independently based on the phylogenetic tree, with the closest similar genus being *Bacillus* (Tang et al. 2014).

There are some microorganisms in which Mn oxidation is activated in the presence of another bacterium. *Arthrobacter sp.* QXT-31 and *Sphingopyxis sp.* QXT-31 need each other to oxidize manganese (Liang et al. 2016). Neither strain was capable of independent Mn(II) oxidation.

In *Pseudomonas putida* GB-1 genes involved in Mn(II) oxidation are found in five, putative operons (Table 1). An operon is a set of genes grouped and regulated together on the

chromosome in order to coordinate their protein synthesis. For example, the *mnxG* operon in *P. putida* contains five other genes whose functions include electron transport and copper-chaperone proteins for the MCOs. These five operons are identifiable by their locus tags; locu*s* tags are gene identifiers used in the scientific community.

Table 1. Manganese-oxidizing operons found in *P. putida* GB-1.

Operon	Function of Proteins Expressed	Locus Tags	Reference
mnxG	Electron transport and copper chaperone proteins	PputGB1_2447-2452	Geszvain, McCarthy, and Tebo 2013
mcoA	Electron transport and copper chaperone proteins	PputGB1_2453-2458	Geszvain, McCarthy, and Tebo 2013
mnxR	Two component regulatory proteins involved in manganese oxidation	PputGB1_2519-2521	Geszvain and Tebo, 2010
	Proteins of unknown function that copurify with manganese oxidase activity	PputGB1_2550-2553	Buzzo, 2011
	Required for the export of oxidase activity	PputGB1_2665-2667	Geszvain, McCarthy, and Tebo 2013

Because *P. putida* GB-1 has multiple Mn(II) oxidizing enzymes, the question of why it has

so many arises. A possibility could be that MnxG and McoA are affected depending on the

environment. In liquid media, Δ*mcoA* was mildly defective, and Δ*mnxG* delayed oxidations.

Since Mn(II) oxidation occurs in the stationary phase, the oxidation could be delayed entry into

stationary phase. In solid media, Δ*mcoA* was more severely affected than Δ*mnxG*. Other factors

that can affect the two operons are O_2 levels and the presence of antibiotics or plasmids. In

conclusion, each enzyme may be optimal for a specific environmental condition in order to

increase manganese oxidation.

There are many commonalities between the manganese oxidases of species. However, how it occurs and how well it is conserved from one species to the other remains a mystery. What organisms are more likely to have this trait is still unknown. Do these organisms use only MCOs or heme-peroxidases to oxide manganese? That is a question that may be answered as scientists continue to observe and analyze manganese oxidation in microorganisms. From this particular research, there may be evidence of manganese oxidizing bacteria on the Lynchburg College campus in Lynchburg, Virginia. We selected soil and aquatic samples to investigate for the presence of manganese oxidation in microorganisms since manganese is usually found in soil, water, and rocks.

Materials and Methods

This research was carried out in two parts: bioinformatics and environmental/lab-based research.

Bioinformatic Analysis

Proteins with high degrees of similarity to McoA (PputGB1_2447), MnxG (Pputgb_2665), and MopA (Pptugb_3353) were found using the website Microbes Online (http://www.microbesonline.org). Six candidates from non-pseudomonad genera were selected and their amino acid sequences were aligned using Clustal Omgea (https://www.ebi.ac.uk/Tools/msa/clustalo/) to analyze similarities in the copper binding domains.

The operon structure of the genes coding for these proteins was examined using Microbes Online to determine if neighboring genes were also conserved.

Finally, each of the non-pseudomonad strains were examined to determine whether all three manganese oxidase genes might be present.

Sample Collection

The second part of the study was to examine if any sampled microbes on the Lynchburg College campus oxidized manganese. Five samples were collected on a partly cloudy day with rainfall from the previous night and the temperature was 24.4° C. Two samples were collected by the pipes near College Lake. One of those samples was collected from under the pipe in stagnant water (37.3988, -79.1869) and the other was obtained from running water in a streamlet from across the pipe (37.3989, -79.1886). One sample was collected from stagnant water next to the

dock at Beaver Point (37.3977, -79.1886). A sample was collected from dirt by mushrooms on a path near College Lake (37.3989, -79.1876). The last sample was obtained from dirt near a tree in front of Hobbs-Sigler Hall on campus (37.3989, -79.1869). Each sample was transferred to a LEPT plate with 100 g/ml of cycloheximide and phosphate buffer saline (PBS) for soil samples. Cycloheximide keeps mold from growing on the plates. The phosphate-buffered saline keeps cells isotonic. The sterile technique was used for microbial transfer. The samples were transferred differently based on their niche. For liquid samples, a clean, sterile dropper was used each time to obtain and transfer 100 µl to the plate. The samples were spread evenly across the surface of the plate. For the dirt samples, the tubes were brought to volume using PBS at the 7 ml mark on the tube and was left alone for the solid matter to settle. Afterwards, the directions for a liquid sample were followed. The plates were labeled, duplicated except for the Hobbs samples, incubated at room temperature for five to seven days and stored at 4° C.

Identification of Manganese Oxidizing Bacteria using the LBB test

Colonies were picked from the plates and streaked for single colonies on LEPT plates (Geszvain, McCarthy, and Tebo 2013). LEPT plates were made from 2.5 ml of 10% yeast extract, 250 µl of 1M $CaCl_2$, 415 µl of 1M $MgSO_4$, and 500 ml of H_2O. The mixture was then autoclaved and cooled to about 55° C. Then, 1 ml of 25% casamino acids (ACROS, 61204-1000 organics), 2.5 ml of 10% glucose, 5 ml of 1M Hepes at a pH 7.5, 50 µl of $MnCl_2$, 500 µl of 3.7mM $FeCl_3$, and 100 µl of 5x trace element mix.

After some time had passed for the bacteria to grow, the plates were examined for brown colonies since manganese oxides are brown in the LEPT media. To test the brown colonies, Leucoberbelin blue (LBB) was used to analyze the presence of manganese oxidation. The LBB test solution was made by combining 9.9 ml of water, 100 µl of 1% glacial acetic acid, 4.7 mg of

LBB solution I (Sigma-Aldrich, 432199) to create 0.4% LBB. The tube was stored by wrapping foil around the tube. Five µl of the LBB for each sample was placed on a parafilm strip and a brown colony from each sample was mixed with the drop of LBB solution by a pipette tip. If the solution changed from clear to blue, it indicated the presence of manganese oxidation.

Isolation of Pure Cultures of Manganese Oxidizing Strains

Once a brown colony was shown to oxidize manganese by the LBB test, another brown colony from the same sample was re-streaked to purify the sample. The purification of each sample happened by isolating single colonies on a new LEPT plate. Re-streaking for isolations occurred by sterilizing the area with a flame. An inoculating loop was sterilized by the flame. Once the loop was cooled, a manganese oxidizing sample was collected and re-streaked on a new plate by the quadrant streaking technique. The loop was sterilized again to kill the remaining bacteria, and the steps were repeated for the remaining samples.

Bacterial strains were named according to (i.) the site from which they were collected and (ii.) were given unique identifiers (e.g., A1, A2, etc.) to distinguish multiple isolates from the same site. For example, SBPA1 was collected from the stream by the pipe; BPDA and BPDB were collected from the Beaver Point dock. Other samples were collected from stagnant water by a pipe (STA), mulch by mushroom on path (MSH), and the mulch near Hobbs-Sigler Hall (HOB).

Identification of SBPA1 using PCR

After the successful isolation of SBPA1 strain, DNA was extracted for a polymerase chain reaction (PCR) and sent to Eton Bio Lab for DNA sequencing of each sample. To prepare a sample for PCR, the template was created by mixing 50 µl of NaOH and a colony of the selected sample and heating the mixture for 15 m at 95° C in a PCR tube. The sample was

collected by a sterile loop. The PCR mix contained 1.5 µl of the boiled sample, 0.25 µl of 0.05µM primer 8F, 0.25 µl of 0.05 µM primer 1492R (Turner et al. 1999), 25 µl of Hot StarTaq mastermix (Qiagen, Germantown, MD) and 23 µl of distilled H_2O to form 50 µl PCR solution in a PCR tube (Francis and Tebo 2001). The tubes were placed in the PCR machine for the heating cycle of 15 m at 95°C for the mastermix, 30 s at 95°C, 30 s at 51°C and 1 m at 72°C for 35 cycles, 10 m at 72°C, and 4°C indefinitely. Polymerase chain reaction products were visualized by gel electrophoresis. 0.40g of agarose and 20 ml of 1x Tris-borate-EDTA (TBE) was mixed together in a 125 ml flask and heated until it became a homogenous mixture. 2µl of Gel Green Nucleic Acid Stain staining dye (miniPCR, #RG-1550-01) was added, and the mixture was poured into a casting tray with a comb to create small, individual wells for the PCR products to be placed in. Once the gel solidified, it was transferred to the gel electrophoresis apparatus. The gel was then covered by 1x TBE. After the PCR reactions were complete, 2 µl of the product was placed in its labeled tube, 5 µl of the molecular marker, (1 kb DNA ladder), was placed in another tube, and 10 µl of loading dye was placed in all the tubes. The PCR products and the molecular weight marker were placed in each well, the apparatus was plugged in, and the gel ran for 30 m. The PCR products were cleaned up with Qiagen PCR cleanup kit (Qiagen, Cat No./ID: 28104). The PCR product was sent to Eton Bioscience (Research Triangle, NC) for sequencing. Six primers were used for good sequence coverage. Sequences were then compared to publicly available genome sequences using BlastN (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Afterward, a PCR reaction was created for *P. putida* GB-1 and the *Bacillus* strain to determine if *mopA* was present in the *Bacillus*. The pseudonomad was used as a positive control since it has been shown that *mopA* is present in this organism. The samples were prepared with a sterile loop to obtain one colony, and 30 µl of 0.2 M NaOH was introduced into each PCR tube.

The tubes were boiled at 95° C for 15 minutes. The primers mop_503F and mop_1586R (Tebo Lab, personal communication; Integrated DNA Technologies) were diluted 10x to transfer 2.5 µl of 10 µM primer concentration. A primer stock was made of 3 µl of each primer and 24 µl of distilled H_2O for a total of 30 µl. A PCR stock was produced with 17.5 µl of the HotstarTaq mastermix (Qiagen, Cat No./ID: 203443), 17.5 µl of the primer stock and 150 µl of distilled H_2O. 49 µl of the PCR mix was placed in the labeled PCR tube along with the corresponding PCR lysate. The PCR tubes ran in the PCR machine (BIORAD T100 ™ Thermal Cycler, Biorad headquarters should be listed here). The gel for the electrophoresis apparatus was made with 2% agarose, 20 ml of 1x TBE, and 2 µl of the staining dye. Once it was mixed, poured, and hardened, the gel was covered by the 1x TBE. Each well had 5 µl of distilled H_2O, 2 µl of 5x loading dye, and 3 µl of PCR product or the 1 kb DNA ladder molecular weight marker.

The next step was to investigate the presence of MopA in manganese-oxidizing strains; I did not look for other putative genes because MnxG and McoA are very similar to other MCOs, and it would be difficult to design primers specific enough to distinguish them. Examining the involvement of MCOs by looking at the effect of adding copper followed. Since MopA is calcium-binding, the stimulation of Mn oxidation using MopA by adding calcium to media was studied.

To keep the samples for later purposes, the identified sample was frozen (Geszvain, McCarthy, and Tebo 2013). The top of the tube was flamed for sterilization. One ml of 80% glycerol was transferred to the tube, 1.5 ml of the sample was then placed in the tube and frozen at -80° C. After 6 weeks, the identified SBPA1 and *P. putida* GB-1 were streaked from -80° C to nutrient agar plates.

.

Results

Bioinformatics

The six loci with highest similarity to the *mnxG* gene of *P. putida* GB-1 were found, using Microbes Online, in the following species: *Leptothrix cholodnii Sp-6*, *Stigmatella aurantiaca DW4/3-1*, *Pseudomonas fluorescens Pf0-1*, *Pseudomonas entomophila L48*, *Sorangium cellulosum ' So ce 56'*, and *Geobacter sulfurreducens PCA*. Alignment of the protein sequences for these loci are in Appendix A. The copper binding domains' amino acid sequences were aligned to analyze their similarity (Table 2). There was high similarity in each region. Seventy percent of the amino acids in region C were identical, in all six species; those in regions D, B, and A were 70%, 56%, and 61% identical, respectively. The genomic neighborhood surrounding each of these loci was examined to determine if they were found as part of an operon with similar structure (Table 3). The operon structure was conserved for these bacteria since each one had the genes found in the *P.* putida GB-1 *mnxG* operon.

The five loci with the highest similarity to the McoA of *P. putida* GB-1 were found in *Pseudomonas entomophila L48*, *Pseudomonas fluorescens Pf0-1*, *Leptothrix cholodnii SP-6*, and *Haliangium ochraceum DSM 14365*. *Leptothrix cholodnii SP-6* was found to contain two *mcoA* genes. The amino acid sequences of the copper-binding domains were aligned to examine the percentage of conservation (Table 4). The entire amino acid sequence alignment is available (Appendix B). The regions were highly conserved for *mcoA*. Seventy-six percent of region A was identical, region B was 92% identical, region C was 80% identical, and region D was 65% identical. The regions surrounding these loci were analyzed to determine possible conservation of operon structure (Table 5). The operon structure was completely conserved, except for the second *mcoA* operon in *L. cholodnii* SP-6, which was missing a *sco1* gene.

Four different organisms, *Methylobacterium chloromethanicum CM4*, *Rhodopseudomonas palustris BisA53*, *Roseobacter sp. MED 193r*, and *Roseobacter sp AzwK-3b*, contained loci most similar, as determined by Microbes Online, to MopA of *P. putida* GB-1 (Figure 6). One of the organisms, *Methylobacterium chloromethanicum CM4*, contain two distinct loci with similarity to MopA of *P. putida* GB-1. The similarity of MopA proteins were examined by aligning them to see the conservation. Approximately half of the amino acids in MopA were identical in the species. The full alignment is available as well (Appendix C).

Pseudomonas putida GB-1 contains two MCOs, (MnxG and McoA) and a heme peroxidase (MopA). The eleven species from this study that were found to contain genes with high similarity to loci of MnxG, McoA, or MopA, were examined to determine if they possessed similarities of the other genes as well (Table 7). Seven of the strains contained a MCO; of these, five strains contained both MnxG and McoA. Six strains contained heme peroxidase (MopA). In general, the bacteria were likely to have either a MCO or a heme peroxidases, but two strains contained all three genes. Some strains contained multiple loci of the genes; *L. cholodnii* had two *mcoA* loci and *M. chloromethanicum* had two *mopA* loci.

Identification of Cultures

Two of the five samples that were collected contained microorganisms able to oxidize manganese. The samples collected from stagnant water by a pipe (STA), mulch by mushroom on path (MSH), and the mulch near Hobbs-Sigler Hall (HOB) did not contain microorganisms that could oxidize manganese.

The sample taken near the dock at Beaver Point (BPD) had two distinct colonies on the original plate. They were isolated and are known as BPDA and BPDB. BPDA had three distinct colony types that were cream, brown, and orange (Figure 1). The brown colony (BPDA1) was

isolated onto a new LEPT plate and produced three distinct colonies as well. At this point, BPDA1 stopped oxidizing manganese. BPDB had two colony types, but only the dark brown colony (BPDB1) oxidized manganese. BPDB1 was isolated and grew three colonies, but only two of them had manganese oxidizers, the cream and yellow colonies (BPDB1.2 and BPDB1.3). Both colonies were isolated and BPDB1.3 had two distinct colonies based on size. Nonetheless, the colonies were isolated but could not oxidize individually. The sample from the stream by the pipe (SBP) contained a colony type that oxidizes manganese; this strain (SBPA1) was isolated for a pure culture. BPDA1 contained three colony types (cream, orange, and brown). None of these could be purified to isolation without losing the ability to purify manganese. A similar phenomenon was seen with BPDB1.

Polymerase Chain Reaction (PCR) and DNA Sequencing

The 16S rRNA gene of SBPA1 was amplified by PCR and sent to Eton Bioscience Lab for sequencing. This strain was identified as genus *Bacillus*. The species was inconclusive since the DNA sequence was found in many Bacillus species that are very closely related.

Table 2. The amino acid sequences of the copper binding domains of *mnxG* of *P. putida* GB-1 and those found to have high similarity. The colon represents strong relations. The period represents weak relations. An empty space shows no similarity between amino acid. The asterisks means the amino acids are identical. Lcho_3893 is *Leptothrix cholodnii Sp-6. Stigmatella aurantiaca DW4/3-1* is STIAU_6861. Pfl01_1975 is *Pseudomonas fluorescens Pf0-1*. Pputgb1_2447 is *Pseudomons putida Gb-1*. *Pseudomonas entomophila L48* is pseen2857. GSU1252 is *Geobacter sulfurreducens PCA*. *Sorangium cellulosum ' So ce 56'* is Sce2138.

Organisms' Locus Tag	Region C	Region D	Region B	Region A
Pputgb1_2447	**HIFHLHGHQW**	**VAGDAIYHCHFYPHFAQGMW**	**IGQHIHLPKWDLTAAD**	**FTHDHLGPSTHQQL**
Lcho_3893	HVFHLHNHQW	SAGDSIYHCHFYPHFAQGMW	IGQHIHLPKWDLTTTD	FTHDHYGPSTHQQI
Pfl01_1975	HIFHLHGHQW	VAGDAIYHCHFYPHFAQGMW	IGQHIHLPKWDLTAAD	FTHDHLGPSTHQQI
PSEEN2857	HIFHLHGHQW	VAGDAIYHCHFYPHFAQGMW	IGQHIHLPKWDLTAAD	FTHDHLGPSTHQQL
GSU1252	HVFHLHAQQW	IVGDTIFHCHFYPHFAQGMW	IGQHIHLVKFDVTASD	FTHDHFGPSTHQQT
STIAU_6861	HVFHLHAHQW	TPGDSIHHCHLYPHFAQGMW	VGQHIHLVKFDVMAAD	FTHDHFGPSSHQHH
Sce2138	HVFHLHGHQW	SVGDAIYHCHLYAHFAQGMW	IGQHIHLVKFDVTSSD	FTHDHFSASSHQHH
	*:**** :**	**:*.***:* *******	:****** *:*: ::*	***** . *:**:

Table 3. The comparison of the mnxG operon of *P. putida* GB-1 and its orthologs. The positive sign (+) indicates the gene is present in the bacteria. This operon is responsible for electron transport and copper chaperone proteins for manganese oxidation.

Organism	*mnxG*	propellor	sco1	sco1	livK	hypothetical protein
Pseudomonas putida GB-1	+	+	+	+	+	+
Stigmatella aurantiaca	+	+	+	+	+	+
Leptothrix cholodnii SP-6	+	+	+	+	+	+
Pseudomonas fluorenscens Pf0-1	+	+	+	+	+	+
Pseudomonas entomophila L48	+	+	+	+	+	+
Sorangium cellulosum 'So ce 56'	+	+	+	+	+	+
Geobacter sulfurreducens PCA *	+	+	+	+	+	+

Table 4. The amino acid alignment of the copper-binding domains of McoA in *P. putida* GB-1 and its orthologs. The semicolon represents strong relations. The period represents weak relations. An empty space shows no similarity between amino acid. The asterisks mean the amino acids are the same. Lcho_30649 and Lcho_2591 are *Leptothrix cholodnii Sp-6*. Pfl01_2001is *Pseudomonas fluorescens Pf0-1*. Pputgb1_2665is *Pseudomons putida Gb-1*. *Pseudomonas entomophila L48* is pseen2124. GSU1252 is *Geobacter sulfurreducens PCA*. Hoch_4297 is *Haliangium ochraceum DSM 14365*.

Organisms	Region A	Region B	Region C	Region D
Pputgb1_2665	**ISTHEHNGHSPAESDGF**	**WFHDHMLDFTAQNV**	**HPVHVHFEEG**	**FAGTYMEHCHNTQHEDSSML**
Hoch_4297	ISTHEHNGHNPAESDGY	WFHDHMLDFTAQNV	HPIHIHFEEG	FLGSFMEHCHNTQHEDHAML
Lcho_0649	ISTHMHNGHNPAESDGV	WFHDHMLDFTAQNV	HPVHVHFEEG	FAGTFVEHCHATQHEDNAML
Lcho_2591	ITTHEHNGHNPAESDGF	WFHDHMLDFTAQNV	HPIHIHFEEG	FAGTFVEHCHNTQHEDSSML
Pfl01_2001	ITTHEHNGHAPAESDGY	WFHDHMLDFTAQNV	HPVHVHFEEG	FAGTYMEHCHNTQHEDNSML
PSEEN2124	ISTHEHNGHSPAESDGF	WFHDHMMDFTAQNV	HPVHVHFEEG	FAGTYMEHCHNTQHEDTSML
	*:** **** ******	******:*******	**:*:*****	* *:::**** ***** :**

Table 5. The comparison of the *mcoA* operon to *P. putida* GB-1's *mcoA*. The positive sign (+) indicates the presence of the gene.

Organism	*mcoA*	sco1
Pseudomonas entomophila L48	+	+
Pseudomonas fluorescens Pf0-1	+	+
Leptothrix cholodnii SP-6	+	-
Leptothrix cholodnii SP-6	+	+
Haliangium ochracum DSM 14365	+	+

Table 6. The similarity (determined by Microbes Online) of *mopA* to *P. putida* GB-1's *mopA*.

Organism	Similarity (%)
Methylobacterium chloromethanicum CM4	42.37
Rhodopseudomonas palustris BisA53	46.09
Roseobacter sp. MED193	48.22
Roseobacter sp. AzwK-3b	48.42

Table 7. The inventory list of the organisms and their manganese oxidizing enzymes. The positive sign (+) indicates the presence of the gene. The negative sign (-) indicates the absence of the gene. The symbol, 2x, means there are two of that gene present.

Organism	*MnxG*	*mcoA*	*mopA*
Pseudomonas putida GB-1	pputgb1_2447	pputgb1_2665	pputgb1_3353
Leptothrix cholodnii Sp-6	+	2x	+
Stigmatella aurantiaca DW4/3-1	+	+	-
Pseudomonas fluorescens Pf0-1	+	+	+
Pseudomonas entomophila L48	+	+	-
Sorangium cellulosum ' So ce 56'	+	+	-
Geobacter sulfurreducens PCA	+	-	-
Haliangium ochraceum DSM 14365	-	+	-
Methylobacterium chloromethanicum CM4	-	-	2x
Rhodopseudomonas palustris BisA53	-	-	+
Roseobacter sp. MED 193r	-	-	+
Roseobacter sp AzwK-3b	-	-	+
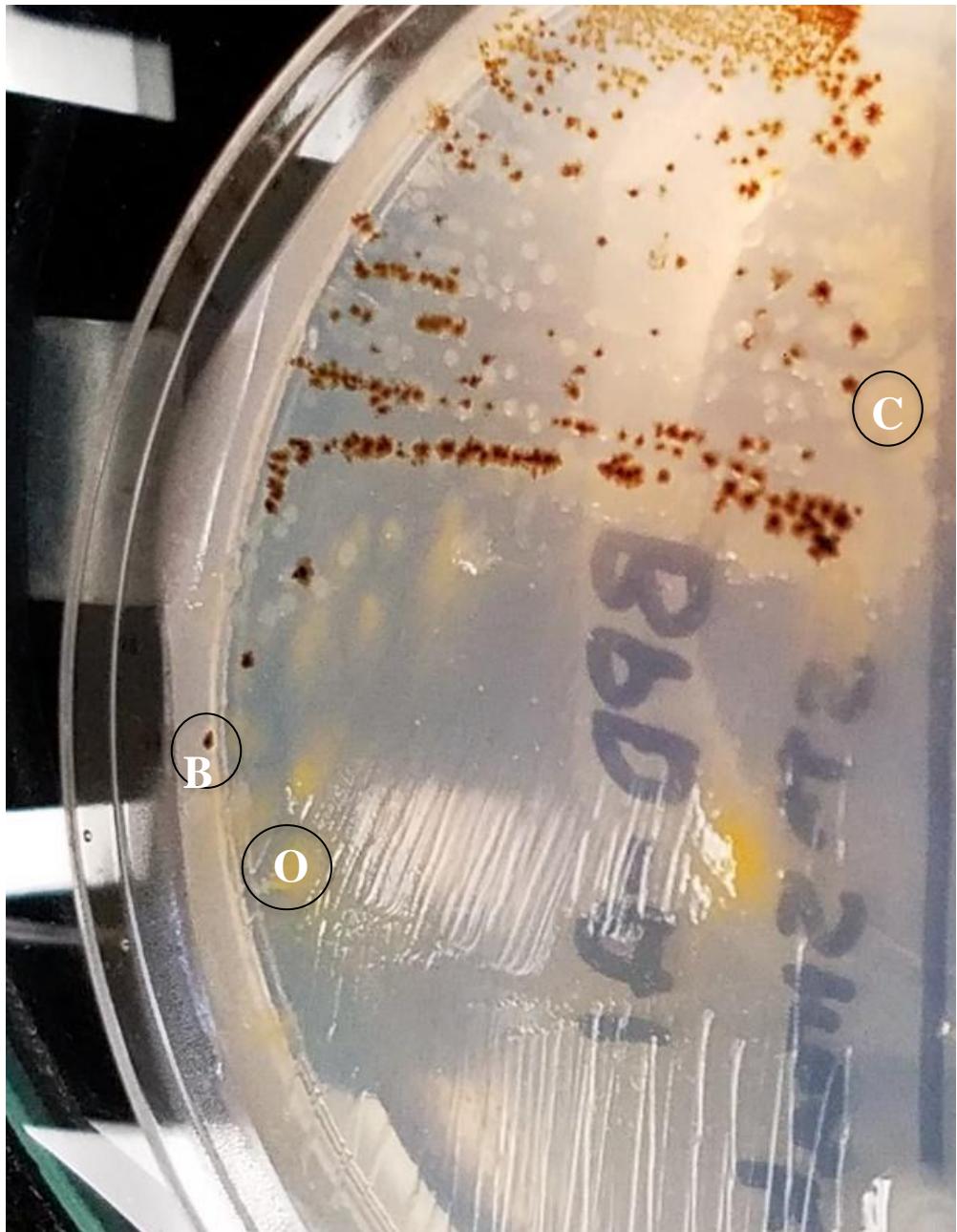
Figure 1. The three different types of colonies (cream (C), orange (O), and brown (B)) of BPDA1 on LEPT media.
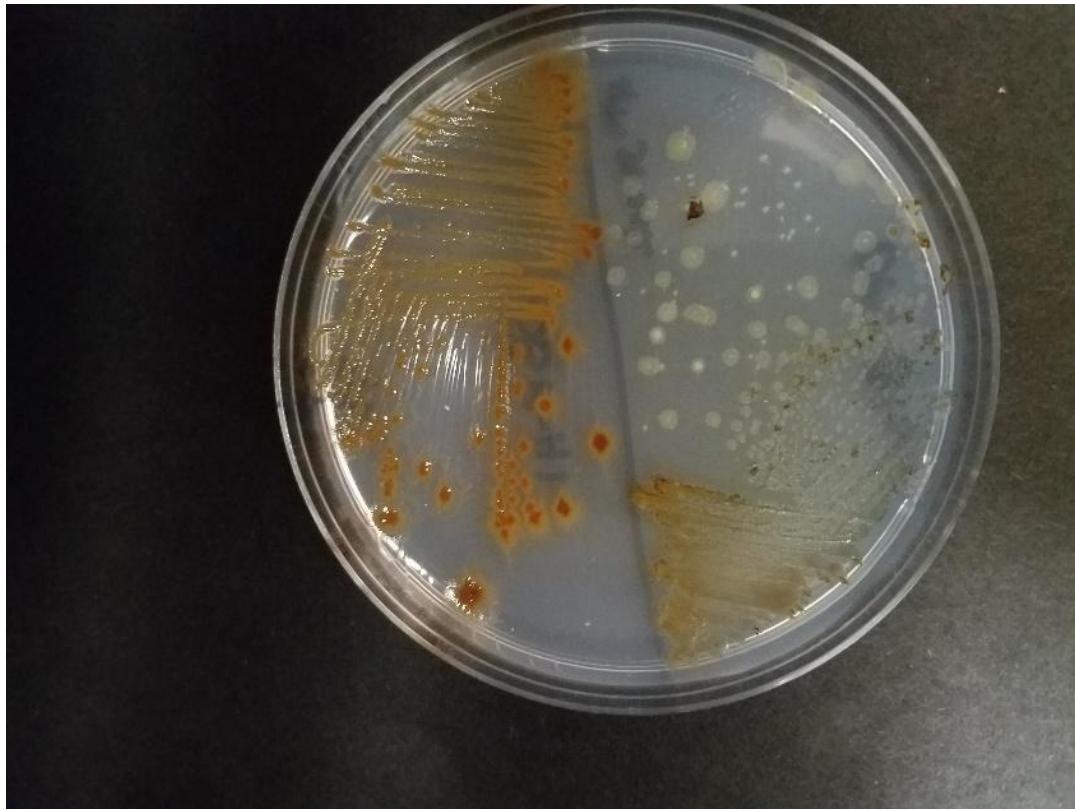
Figure 2. A visualization of SBPA1 (left) and BPDA1 (right) manganese oxidizing brown colonies on LEPT media.

Discussion

Within the bioinformatics component, we were able to find multiple bacteria with *mnxG*, *mcoA*, and *mopA* genes and their operons. The *mnxG* gene works with five other genes in an operon. The orthologs of *mnxG* in this study were found to have similar operon structure. The *mcoA* gene is part of a small operon with one *sco* gene. This operon structure was conserved in all of the organisms studied except for *Leptothriz cholodnii*. *Leptothriz cholodnii* has two *mcoA* genes, and only one of these genes is found in an operon with a *sco* gene.

The copper-domains of MnxG and McoA were highly conserved as well. Based on *P. putida* GB-1, the majority of the amino acids in these domains were conserved or strongly related. The factors that made the amino acids strongly related may have a similarity in the structure. Also, the amino acids may be different but have similar chemical properties like polarity.

Most bacteria in this study were found to have either MCOs or heme peroxidases. There were a couple of exceptions to this finding. *Leptothriz cholodnii* and *P. fluorescens* had all three enzymes, MnxG, McoA, and MopA. It was also common to see more than one MCO in an organism,. but some only had one MCO. There may be other manganese oxidizing enzymes present but undetected in our work since only mnxG, mcoA, and mopA were analyzed for this study.

For environmental sampling, there was success in the isolation of multiple manganese oxidizing bacteria. The SPBA1 strain was able to be isolated into a pure culture so it could be identified. It was identified as a *Bacillus* species. The BPDA1 strain was able to oxidize manganese, but the isolation of the oxidizing strain was not successful. It was difficult to isolate oxidizing colonies from the other colonies on the LEPT plate. Since isolation did not occur, a

pure culture could not be created for identification. BPDA1 may need the other bacteria on the medium to oxidize manganese. There are other studies that show organisms oxidizing manganese in the presence of certain, other organisms (Liang et al. 2016). Since we could not isolate the three colonies, BPDA1 may need the others to oxidize manganese.

In future work, we hope to identify the manganese oxidation genes in SBPA1. A bioinformatics approach can be used to detect manganese oxidization genes in the genome of our *Bacillus* species, then compare it to *P. putida* GB-1 manganese oxidizing bacteria.

We successfully isolated a manganese oxidizing bacterium from water samples on the Lynchburg College campus. Manganese is present in environmental soil and water samples (Swistock and Sharpe 2018). Therefore, our isolated *Bacillus* species may be aiding nearby organisms (i.e., plants) by providing them with manganese oxides. Therefore, although microorganisms may not need to oxidize manganese for their own survival, they may help plants, humans, other animals, and oxygen-evolving photosynthetic bacteria to survive.

Manganese levels are regulated in human drinking water for multiple reasons. From an aesthetic standpoint, it affects the color, taste and odor of water (Swistock and Sharpe 2018). Ingesting high levels of manganese can cause neurological problematic (Swistock and Sharpe 2018). Therefore, contributing to the collected knowledge of manganese oxidation could help to improve methods of water purification at waste water treatment plants.

References

Buzzo J. 2011. New insights in Mn(II) oxidation in *Pseudomonas putida* GB-1. M.S. thesis. Oregon Health & Science University, Portland, OR.

Dick GJ, Torpey JW, Beveridge, Terry J. Tebo, B. M. 2007. Direct identification of a bacterial manganese(II) oxidase, the multicopper oxidase MnxG, from spores of several different marine Bacillus species. Applied and Environmental Microbiology 74(4):1527-34.

Francis CA and Tebo BM. 2001. cumA multicopper oxidase genes from diverse Mn(II)-oxidizing and non-Mn(II)-oxidizing pseudomonas strains. Applied Environmental Microbiology 67(9):4272-8.

Gezsvain K, Butterfield C, Davis RE, Madison AS, Lee S, Parker DL, Soldatova A, Spiro TG, Luther III GW, Tebo BM. 2012. The molecular biogeochemistry of manganese (II) oxidation. Biochemical Society Transactions 40:1244-8.

Geszvain K, McCarthy JK, Tebo BM. 2013. Elimination of manganese(II,III) oxidation in *Pseudomonas putida* GB-1 by a double knockout of two putative multicopper oxidase genes. Applied and Environmental Microbiology 79(1):357-66.

Gezsvain K, Smesrud L, Tebo BM. 2016. Identification of a third Mn(II) oxidase enzyme in *Pseudomonas putida* sp;GB-1. Applied and Environmental Microbiology 82(13):3774-82.

Gesvain K and Tebo BM. 2010. Identification of a two-component regulatory pathway essential for Mn(II) oxidation in *Pseudomonas putida* GB-1. Applied and Environmental Microbiology 76(4):1224-31.

Learman DR and Hansel CM. 2014. Comparative proteomics of Mn(II)-oxidizing and non-oxidizing Roseobacter clade bacteria reveal an operative manganese transport system but minimal Mn(II)-induced expression of manganese oxidation and antioxidant enzymes. Environmental Microbiology Reports 6(15):501-9.

Liang J, Bai Y, Men Y, Qu J. 2016. Microbe–microbe interactions trigger Mn(II)-oxidizing gene expression. The ISME Journal (11):67-77.

Lohry R. 2017. Micronutrients: Functions, sources and application methods [thesis]. West: Lafayette: Purdue University. 1-15.

Pecoraro L, Gelasco A, Baldwin M. 1996. Reactivity and mechanism of manganese enzymes: A modeling approach. In: *Mechanistic bioinorganic chemistry*. Ann Arbor, MI: American Chemical Society. 265 p.

Swistock B and Sharpe W. 2018. Iron and manganese in private water systems [extension]. University Park, PA: The Pennsylvania State University. 1-3.

Tang W, Xia J, Zeng Z, Wu L, Ye G. 2014. Biological characteristics and oxidation mechanism of a new manganese-oxidizing bacteria FM-2; Bio-Medical Material and Engineering 24:703-9.

Turner S, Pryer KM, Miao VP, Palmer JD. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. Journal of Eukaryotic Microbiology 46(4):327-38.

Weisany W, Raei Y, Allahverdipoor KH. 2013. Role of some of mineral nutrient in biological nitrogen fixation. Bulletin of Environment, Pharmacology and Life Sciences 2(4):77-84.

Appendix A: The Amino Acid Sequence Alignment for *mnxG*.

The letters are the standard abbreviations for amino acids. The regions are highlighted and labeled. Lcho_3893 is *Leptothrix cholodnii Sp-6. Stigmatella aurantiaca DW4/3-1* is STIAU_6861. Pfl01_1975 is *Pseudomonas fluorescens Pf0-1*. Pputgb1_2447 is *Pseudomons putida Gb-1. Pseudomonas entomophila L48* is pseen2857. GSU1252 is *Geobacter sulfurreducens PCA. Sorangium cellulosum ' So ce 56'* is Sce2138.

```
Lcho_3893      -----MAG--AV----------------WAVAAWGGLAAGPAQAAITCARTVTANVVAFD
Pfl01_1975     MVGIHHAPPLLN----------------WLLLLVSMLSVELASAAVRCERNLVANVVAFD
pputgb1_2447   MTTPRK-GPLLS----------------SLVLA-LLGWESTSEAAVQCQRTLVANVVALD
pseen2857      MSTPRK-GPVLS----------------TLLLA-LLGWQASGEAAVQCQRTLTANVVALD
GSU1252        MRTRIKPGRIAGFLAQ------MM----------LLLAPTVSLAEVCANTIQADVVAID
STIAU_6861     ---------------------MPL---------VEPAPMAGTRVTCARTITANVVALD
Sce2138        MRLRWRPGWAMVSVAVLGCSGADAPPTEEESVLESQDELRYDFHAPASGRTVFADVVAID

Lcho_3893      KPLMYNRLGAGNVNGMMFALERDVIHASSQLPLTQGGAATPGQLDLRPDKRHRPLVLRVR
Pfl01_1975     QPLMFNRLGAQNANGMMYALRRDVVD-EHDVSLAYGGSAVPGKVSLRADKRPRPLVLRVA
pputgb1_2447   QPLMFNRLGAQNANGMLFALREDVVD-DKQVPLSKGGAAVPGKVTLRPDKRPRPIVLRVA
pseen2857      QPLMFNRLGAQNANGMMFALRQDVVD-EHQVPLNKGGAAVPGKVTLRPDKRPRPIVLRVA
GSU1252        QQIVHNRLGAFNPISMIFALRQDVVNAANGLTEAEGGILTPGNVRLRDDKRPRPLVLRVN
STIAU_6861     QVYTYNRFGSYNPSGMMYALKRDVEPIHG------SQPVGPGNARLRLDKRPRPLTLRAN
Sce2138        QMFVYDRLGAFNPGGMVYALRRDVVAADP------SQPIGPGNATLRPGKRPRPLVLRVN

Lcho_3893      QGDCLTVKLTNLLTAAPNPRHIPP------------------------------------
Pfl01_1975     AGDCLTINLQNLLDYQANPNKHFE------------------------------------
pputgb1_2447   AGDCLTVNLTNLLDYRANPNKHGI------------------------------------
pseen2857      AGDCLTVNLTNLLAYQANPNKHGI------------------------------------
GSU1252        EGDCLRISFQNLLAPTPLVLPPDAAGEF--------------------------------
STIAU_6861     VGDCLRIQFTNWLAPTRAAIPSPSQSQPTPSQVAQPTAGFAMVRKMLPTWLPTSSWDPLA
Sce2138        AGDRLSISFTNWLTPATGALPSTSAS----------------------------------

Lcho_3893      --------------NHVTPA----------------GQVFDVLIDEQIQDRAVSFHASG
Pfl01_1975     --------------GPEGEEGTV-----------NPGGADAFKVDEQVADRHVGFQVNG
pputgb1_2447   --------------EAEGPEGEEV---------ENEADEGFVADEQVAERMVGFQVNG
pseen2857      --------------HPEEPEGEEG----EEAEVPEVEDDENFVADQQVTDRHTGFQVNG
GSU1252        -------------------------------------AGQPATRHVGMAIMG
STIAU_6861     SLVTGRAVGGSWLGAGPEQNPGEEEEPDDEPGDDHGPGDADNAHKDNSPFTRRASVHVQG
Sce2138        -----------------------------------------TRDASIQVVG

Lcho_3893      LQVVDSIASDGSWVGANATDGHAPAGGGTRTYRLYAEKEGVFNVT-GGASFGSD-ANQGN
Pfl01_1975     LQAVNSIDDISSYTGRNANTL--VPPGASRSYVLYAEREGAFAVSSRGATFGGE-GAAGN
pputgb1_2447   MQAVNSIADISAYTGRNGNFF--VSPGSTRSYTLYAEREGAFAATSKAATFGGE-GTAGN
pseen2857      LQAVNSIADIAANTGRNGNFL--VAPGSTRSYTLYAEREGAFAVSSRGAPFGGE-GNAGN
GSU1252        VSFAD-MVSSGVNVGNNPVSG-YAAPGETKVYLVRGEHEGGYHINSIADNVSGE-GIQGQ
STIAU_6861     LQYLS-INDDGANVGKNPS-S-LVDPGNSTTYTLYADHEGVFFMYSMGATFGGE-GDGGS
Sce2138        LQIRN-IDALAGDVGKNES-S-LAAPGETRKYELFAGREGTFLMHSGGAMVGSEPGPGRR

Lcho_3893      ASNGLFGQVIVEPAGAKIYRSQVHEEELRLVADVNRNGVLEAAEKTATGQPKIVGYEAVY
Pfl01_1975     TANGLFGQVVVVPKSGRTYRNTLTEEEMRLATT----------GRTPAGHP-IVDYQARY
pputgb1_2447   VANGLFGQVVVVPKLGRTYRNTLTEEEMRLATT----------GRTATGQP-VIDYEARY
pseen2857      VANGLFGQVVVVPKGGRTYRNTLTEEEMRLATT----------GRTATGQP-VIDYEARY
GSU1252        TSFGLFGALNVQKKGAVWYRNQTTNAELLMATT----------GLTPSGHP-IIDYDAVY
STIAU_6861     TAQGLFGAVNVEPAGSKWFRSKVSAEVLQAATRKDTQGK---PLSNPDGTP-LIDYEARD
Sce2138        ISQTLYGAVTVEPPGAVAYRSQVTAAELSAAS------L---HPSNPDGTP-RIDYNAVD

Lcho_3893      PSDGVWLAEGKAGLPILNMMKCASATSCEIVHSEINAVVAGPNADGSFPAGTYPLERVGK
Pfl01_1975     PQREPWLREGKAGTPIISMVD-----GNEIISSESDAIVMGSNADGSFAPSTYPLESVGK
pputgb1_2447   PQVEPWIAEGKAGKPIIAMVN-----GNEIINSETDAIVMGPNPDGSFPKSTYPLESMGK
pseen2857      PQAEPWIAEGKAGKPIIAMID-----GNEILNSETDAIVMGPNPDGSFPKATYPLESVGK
GSU1252        PVG-----HRYAGQPILRILDPV---TNKIVHNEVNAIITGPNG-GDFPAGTY------Q
STIAU_6861     ----------SRDRPLLNILDTK---TYELWHGDLEAIITGYSR-TSVGT----------
Sce2138        ----------ASGQPILKMIDDA----GEIVHGDVNAVIAGFDE-TEVGS----------

Lcho_3893      RNPTVPNRLEPFRDFASIYHDETSNGQAFPGFYKTDPVFRYVLAGVKDAFMINYGSGGIG
Pfl01_1975     RNPAIPNRLEPFRDFASQFQDETAATQAFPAYW-ADPVMAHVLEPTRDSFMINYGSGGMG
pputgb1_2447   RNPALPNRLEAFRDFASQFADEVAGTQAFPGYW-ADPVMGHVLEPTRDSFMINYGSGGMG
```

```
pseen2857       RNPALPNRLEAFRDFAAQFADEVAGTQAFPGYW-ADPVMGHVLEPARDSFMINYAAGGMG
GSU1252         RNPTYPDRDRAFREFTVMFHDEVETVQPFPIFI--DPQFKFTLGGVKDGFMINYGSAAVG
STIAU_6861      ---FTSKDQGHFREITALYHDEIKAVQAFDEL-----EWNPTFHGVRDGFGVNYGVAGLG
Sce2138         ---PVSENMGYFREFTVLFHDILSVVQPVAEY-----QADPRYNGVRTGFGINYGAASLG


Lcho_3893       SEIIANRLGVGPMHDCLDCAYEEFFLTSFTVGDPALTVDVPANLGLEALLPGQTPPPGTQ
Pfl01_1975      AEVVANRLGVGPMHDCLSCAYEEFFLSSHTVGDVAMLVDVPANTGLENIAPGQTPSADQV
pputgb1_2447    AEVVANRLGVGPMHDCLSCAYEEFFLSAHTVGDIGTLVDVPANVGLEHIRPGEVPPASAV
pseen2857       AEVVANRLGVGPMHDCLSCAYEEFFLSSHTVGDIGTLVDIPANVGLERIRPGEVPPASAT
GSU1252         TEVVANRSGVGPTWDCAECKFEEFFLSSWAVSDPALLVDVPANTTD-------VNGNLIV
STIAU_6861      AELMANRAQIGPTKNCADCAFEEFFLESWAGGDPALNVEKDW------------------
Sce2138         SAVFANKQRIGPSKRCEECKFEEFFLSSWVNGDPALNIDRDD------------------


                                                            region C
Lcho_3893       GPKANYVIGAEDPSNVHHSYTGDFVKFRNSHIGK-EQHVFHLHNHQWLYNPNDDNSNYLD
Pfl01_1975      GVKATMALYPSEPSNVNHSYIGDFVKFRNTHNGH-EQHIFHLHGHQWLFNPNDDNSDYVD
pputgb1_2447    GVKASMALYPSEPANVHHSYIGDFTKFRNTHNGH-EQHIFHLHGHQWLFNPNDDNSDYID
pseen2857       GVKATMALYPAEPANVHHSYIGDFTKFRNTHNGH-EQHIFHLHGHQWLFNPNDDNSDYID
GSU1252         GPKATKALYPDDPSNVFHAYINDRAKIRNLHFGK-EFHVFHLHAQQWLFSPDDDGSNYLD
STIAU_6861      QGNAVRALYPDDPSNVHHSYLGDPVRIRNIHAGPAETHVFHLHAHQWKFSPGDDNSTYLD
Sce2138         AGNATEALYPDDPSNVSHGYLGDPVRIRNLHAGPSETHVFHLHGHQWLRSPGNEKSAYID


                                         Region D
Lcho_3893       AQGIGPGIGYTYEINFGGSGNRNKSAGDSIYHCHFYPHFAQGMWAHWRHHDTFEPGTILE
Pfl01_1975      AQGIGPGAGYTYEIANGGSGNRNRVAGDAIYHCHFYPHFAQGMWSMWRVHDVFEEGTRLE
pputgb1_2447    AQGIGAGVGYTYEIANGGSGNRNRVAGDAIYHCHFYPHFAQGMWAMWRVHDVFEEGTRLE
pseen2857       AQGIGPGVGYTYEIANGGSGNRNRVAGDAIYHCHFYPHFAQGMWAMWRVHDVFEPGTRLE
GSU1252         AQGIGPGGSYTYEIAYN-SGNRNKIVGDTIFHCHFYPHFAQGMWALYRLHDVFEAGTQLD
STIAU_6861      SQSIGPGAAFTYDINYGGSGNRNLTPGDSIHHCHLYPHFAQGMWALWRVHDVFEAGTP--
Sce2138         SQTIGPRAAYTYDITYGGGGNRNLSVGDAIYHCHLYAHFAQGMWGLWRTHDVFEAGTP--


Lcho_3893       ATPASYDGQGFASGTGYHDTLWGLAAGKPKAGSRAYPDSEIVAGTPIPALVPLPGKPMPV
Pfl01_1975      VSQQ--------GADGYHSEPYALRSGKPAAGARALPDGEIIAGTPIPAVVPLPGKAMAP
pputgb1_2447    VSGQ--------GENGFHSTPFALRSGKPAAGARALPDGEIVAGTPIPAIVPLPGKAMAP
pseen2857       VSGE--------GANGYHSTPFALRNGKPAAGARALPDGEIVAGTPIPAIVPLPGKAMAP
GSU1252         AN----------------------GLPIATARALPDGEIAAGTPIPAVVPIPTLPMAP
STIAU_6861      -----------------------------DRRLPDGEIAGGTPIPAVIPIPARAMPP
Sce2138         -----------------------------DRSLPDGELAAGTPTPAVVPIPGVALAP


Lcho_3893       MPGRVTTKVNP--DMAAA-----------------NPAKPVGSLTQVI---------DR
Pfl01_1975      MPGKVVVVPKIGETLVAGNDDDDEGDDDGEHH-GGNGESQAIGSLALVDRSEANRNADGS
pputgb1_2447    MPGKVVVVPKLSETLVAANDDDDDDEEGDDDHANQPPVRKAVGSLALVDRTEANRNADGT
pseen2857       MPGKVTVVPKLSETLVAEHDDD---EEEGDDHPDEPATPRAVGSLALVDRSETNRNADGS
GSU1252         IPGKVRIAQVPGYPGG-------------------------------------QVTYDEP
STIAU_6861      MPTYADTSVK-DANGH---------------------------------------Q---VLR
Sce2138         MPTYERTSVR-TADGK---------------------------------------T---VER


Lcho_3893       DVNPGYPFWIAGMEDVVGQRPPTPLLDMITKDEATALKASGDPLWSEIVPEQADGFDGGL
Pfl01_1975      LKNPGYPFWIGGMESSVGQRPPTPPLDMLDAATAQSLKASGKALWANLDPNQSGGWDGGL
pputgb1_2447    LKNPGYPFWIGGMESSVGNRPPTPPLDMLDPALARQLKDSGKALWANLDANQVDGWDGGL
pseen2857       LKNPGYPFWIGGMESSVGQRPPTPPLDMLDPALARQLKDSGKALWANLDPNQVDGWDGGL
GSU1252         DKNAGYPFSLGMK---AGHRPATPPLDLID------------------------DGGL
STIAU_6861      PAFPGYPFYVAGK---AGHRPPQPPLDLEE------------------------DGGL
Sce2138         DPMPGYPFYVAAV---AGHRPSQPPRDMEH------------------------DGGL


Lcho_3893       PRHALKGLAAGGVA--------HTVTSRLDFSKEVLKSNVAYYPESGTDLEKVAMAFHAQ
Pfl01_1975      QRHALDGVAAGGAA--------HTVTTSLDFSKEVTRAKPIYLPEEGTEVEQAAMAFHAK
pputgb1_2447    GRHALDGVSAGGEA--------ETTTTKLDFTKVVHKAKPIYLPEEGTEVEQAAMQFHAM
pseen2857       GRHALDGVSAGGEA--------VTTTTKLDFSKVVHKAKPIYLPEEGTEVEQAAMQFHAM
GSU1252         PRHIVRGTNNPLNPSTLDPAITHHEETTLSFDKVLVAAEAEQVPESGSRSERAAMDFHAR
STIAU_6861      PRHIVTSAIGPSTYGM---------PG-KRFYVDHHALNVKLLPQDGTLLEKKAMSFHGG
Sce2138         PRHYIANVPP-GGAAV---------GGRGLFDVQLSKADVQLLPKDGTPSEVAAANFHAG


Lcho_3893       RCV-----------------STSRPDGSPASCVADPVKGPVGGFILNGSKPVIGAPYHN
Pfl01_1975      KDH-----------------PSF--------ALLPGNQIVAKAFRTNGALPMAGAPYYE
pputgb1_2447    AEH-----------------PSY--------ALIPGSQPVAKAFRTNGALPTAGAPYYE
pseen2857       AEH-----------------PSF--------ALIPGSAPVAKAFRTNGALPTAGAPYYE
GSU1252         EFH-----------------PTFLPN----------GSPGQFVTNGLPPVPGAPYND
STIAU_6861      EFPGATSVQRPYYTADTVAAYPAYTAT-----------GTRGFFYVNGRKPVAGAPYAD
Sce2138         AFPGAVLAPERY--GFPVKGYPAFTAE-----------GKPALFLVNGQPPAAGAPFAD
```

```
Lcho_3893        PCIDDAGTVLNP-GVLGRFYSGELATATQAALNTRGASVFNSVSPRVYKGTNLQFD-AVF
Pfl01_1975       PCMDDRQKRLTSSAGSGEFASGDRL----DGMSFVGASTFTADRPRIYKAANIQFD-AVY
pputgb1_2447     PCMDDRGKRLTQSSGVGEFFSGESP----TGLNFRGASAFTADRPRIYKGANIQFD-AVY
pseen2857        PCMDDRGKRLTQSSGVGEFFSGENL----TGMSFRGSSTFTADRPRIYKAANIQFD-AVY
GSU1252          PCRGDKGGAVG-----------------------------VPRTYKGAVIQFDL-KM
STIAU_6861       PCPASA-----------------------------------PRRDYRAAYLQIDLQNV
Sce2138          PCPPNA-----------------------------------PVRTYRAAYVQTN-APV

Lcho_3893        NKAGYHYPQQRITALWQDVQGVIAKTQAPEPLVMRLNTFDCAVYHHSNLVPGYYEIDDYQ
Pfl01_1975       NKVGYHFPQARILALWEDAWPVITKQRPPEPLVMRMNTFDCVQYQQTNLVPATYEMDDYQ
pputgb1_2447     NKVGYHFPQARIIALWEDAWPVITKQRPPEPLVMRMNTFDCTMYQHTNLIPNIYEMDDYQ
pseen2857        NKVGYHFPQARILALWEDAWPVITKQRPPEPMVLRMNTFDCTQYVHTNLIPSFYEMDDYQ
GSU1252          NKVGHHFPQSRIITLWGDAQATVDGTRPPEPFFIRANSNDCVNFYHTNLVPSVYEQDDYQ
STIAU_6861       NREGWHDRQARIMVLNGDVADTQSGLRPPEPFFFRAESGECINFYATNLMPSELEADDFQ
Sce2138          NRAGWHDAQARMIVLEKDVEATLDGTRAPEPLFIRAQSGECVVFHATNLIPAALAADDFQ
```

Region A
```
Lcho_3893        VRTPTDIIGQHIHLPKWDLTTTDGAANGWNYEDGTMSPTTIRERIHAINCFNGHADECKG
Pfl01_1975       VRTPTDVIGQHIHLPKWDLTAADGSSNGWNYEDGVLSPGAVQERIHAIREFNQCAGS---
pputgb1_2447     VRTPTDVIGQHIHLPKWDLTAADGSANGWNYEDGILSPGSVVERVHAIRAYNGCTEG---
pseen2857        VRTPTDVIGQHIHLPKWDLTAADGSANGWNYEDGILSPGSVVERVHAIRAFNNCSEG---
GSU1252          VKTPTDVIGQHIHLVKFDVTASDGSANGFNYEDGTMSPDEVRERIHAFNLTGGLIQP---
STIAU_6861       IYTPTDTVGQHIHLVKFDVMAADGAGNGWNYEDGTFSFQEVLHRIDKITAAGGAFAA---
Sce2138          IFTPTDVIGQHIHLVKFDVTSSDGGANGFNYEDGALAMEEVVARIQAANALGGALAA---
```

```
Lcho_3893        GVLPGTGTGAR-LVAKDHPY-WGKVAANLGGAFPEQWRGARTTTQRWFTDPVVNTEGVDR
Pfl01_1975       --DPRDGTQAC-PKAKAHPF-FGQFG-------RSDWMGARTAMQRWFVDPVVNAKGVDR
pputgb1_2447     --DSRDGTAAC-PKAKQHPY-FGRFG-------RADWLGARTAMQRWFADPLVNVYNVDR
pseen2857        --DARDGTAAC-PKAKQHPY-FGRFG-------RADWLGARTAMQRWFADPLVNVHNVDR
GSU1252          -----DGVTKVALAAKPHPY-FGST------FNGRDITGARTTVQRWYIDNIRNNRGEDR
STIAU_6861       -----DGTVDSTGPRVPLSA---RP------HSRLGVMGAQTSTQRWWADPLTDSKGQER
Sce2138          -----GVRSGEPGARVRLGDPTRHP------RMTTPSLGAQTTTQRWFADPIMSEAGQDR
```

Region B
```
Lcho_3893        GLGIIFTHDHYGPSTHQQIGLYATVLAEPAGSRWAHNESGQQLGHDPVTGAPARTDTRLD
Pfl01_1975       GLGTIFTHDHLGPSTHQQIGLYATVLAEPAGSTWFHAETGEPLYS---------------
pputgb1_2447     GLGTIFTHDHLGPSTHQQLGLYATVLAEPAGSTWYHAETGEQLYN---------------
pseen2857        GLGTIFTHDHLGPSTHQQLGLYATVLAEPAGSTWYHAETGEQLYN---------------
GSU1252          TLGNVFTHDHFGPSTHQQTGLYASLLTEPQGSRWRDPQTGQFMGG---------------
STIAU_6861       PIETVFTHDHFGPSSHQHHGFYGALIVEPKGSKWRNPQTGVYYGT---------------
Sce2138          ALSTAFTHDHFSASSHQHHGLYGGLVVEPAGSIWRDPETGALFGD---------------
```

```
Lcho_3893        GTVFSDGGPTSWQAAILPQAQTGVYAGNTVKAGLCTSATQVNGVQCQKPFREFYLEFSDF
Pfl01_1975       G-ARQDGGPTSWQAVINTGDL-------------------DGDGKNDSFREFFLEYSDF
pputgb1_2447     PATRQDGGPTSWQAVIQTGDH-------------------DGDGKNDSYREFFLEYSDF
pseen2857        PATRQDGGPTSWQAVIQTGDH-------------------DGDGKNDSYREFFLEYSDF
GSU1252          --R-FDGGPTSWRADIITT----------------------NPAESYREFMVHVADF
STIAU_6861       --RPSDGGPTSWQADILTK----------------------DPANSYREFALAFADY
Sce2138          --R-SDGGPTSYRADILFPPG---------------------DHRSPFREFNLSIADY
```

```
Lcho_3893        QHAYEAGVYVGADPQGLPLNGTGAGESAVAFNAGNPAFASLGGQVDEAFRFAINPPAREQ
Pfl01_1975       QHAYEAGVYVGAGPNGVPNPQAF------------------PATADSFRYAINPPVRNN
pputgb1_2447     QHAYEAGVYVGAGPDGIPNAQSY------------------PATADSFRYAINPPVRGK
pseen2857        QHAYEAGVYVGAGPNGIPDGQSY------------------PATADSFRYAINPPVRGK
GSU1252          TLAYEDGACHTV-----PCVN--------------------PAKAIKPPGMEE
STIAU_6861       FPAYDACGQ-------------------------------------PVNAPNFKE
Sce2138          GPIYDECGR-------------------------------------PVNPSSQAP
```

```
Lcho_3893        IAPVFPDLVVEL-ANSVFCPSRPCPQAISVADPGMLVVNYRNEPVGLRIYDPAKVGPDGK
Pfl01_1975       ASTLLEGVLEVQGGQVPGCPSRPCPQAISVDDPGMFVVNYRNEPLALRVYDPNKVGPDGK
pputgb1_2447     ASNLLEAIVEERGGINPGCPSRPCPQAISVDDPGMFVVNYRNEPLALRVYDPNKVGPDGK
pseen2857        ASNLLEAIVETRGGLNPGCPSRPCPQAISVDDPGMFVVNYRNEPLALRVFDPNKVGPDGK
GSU1252          IGL--PF-LFRKPQICPNGTLPPCPEAISAEDPGTFLVNYRNEPVAERVRLPGT------
STIAU_6861       AAL--PW-AV-------NFQQTPMPEAISTSDPGAMIINYHNEPIPLRIGERTS---NCG
Sce2138          APL--PV-AI-------TH-GPSTPDAISWRDSGVGLINYRNEPIPLRIAERNC---STR
```

```
Lcho_3893        PGMQADGDKGDLAHALQSRTDRALSNMNVQPNA-------------------------
Pfl01_1975       RGMQADGLGGDLAYAMQSRTDRAIPAMNLAPNLVTAATGPTGGTTLFPPHINKGGSEPGD
pputgb1_2447     RGMQADGLAGDLSYALQTRTDRAIPAMNLSPSAITSAVGPTGGTTLFPPHINKAGAEPGD
```

```
pseen2857      RGMQADGLGGDLGYALQTRTDRAIPAMNLAPSAITQAVGPTGGTTLFPPHINKAGSEPGD
GSU1252        -NVQAPGLAGDLAYALSSRVQRANPLLNQQPAFY------------PPLT--PNVLPGD
STIAU_6861     NRVLRTDDRGDMANV--------------------------------F--RSDLHGD
Sce2138        AVVQKSGDAGAMHNV--------------------------------F--SSRVHGD


Lcho_3893      ------------------------------------------------------------
Pfl01_1975     PFTPMLRTYTGDNVRLRVHAGGHEEEHNVTLHGVKWLQSGSGFGNSSNSGWRASQMIGIS
pputgb1_2447   PFTPLLRTYSGDNVRLRMHAGGHEEEHNVTLHGVKWLQNGTGFGNSSNSGWKSSQMIGIS
pseen2857      PFTPMLRTYSGDNVRLRMHAGGHEEEHNVTLHGVKWLQNGSSGFGNSSNSGWKSSQMIGIS
GSU1252        PFTPLFQVYDNDRVNVRIQAGADEESHTASIHGVKWLQSYG----SPNSGFRNSQQLGIS
STIAU_6861     PYTPLLRGYEGDKVKIRLIQGSQEEQHSFSLHGAKWLHEGA----DPHSGYFNAQGIGIS
Sce2138        PATPLLRAYDGDRTLIRLTQGAQAEQHVFTVHGKKWLNEPT----DPDSGYANGQAIGVS


Lcho_3893      ------------------------------------------------------------
Pfl01_1975     EQMGFIAPVSML-SSSAATTGDYLYSMDASIEGYWSGIWGVMRNYTAKRNDLFAIPNNPN
pputgb1_2447   EQLGFMAPVSMI-SSSAATNGDYLYSLDAALEGYWNGIWGIMRNYTVQRADLFALPNNPQ
pseen2857      EQLGFMAPVSMI-SSSAATNGDYLYSLDAALEGYWNGIWGIMRNYTAQRSDLFPLPNNPQ
GSU1252        EQFQLRMPVIPD-RLQVGQTADYLYTLNASSDGYWSGIWGLMRSYAVRQPNLLPLPNNPI
STIAU_6861     EHFEFNLTNGLPALAGTYETADYMYQ-SASNGDLWNGMWGLLRTYTKRQTGVKTLGHLAL
Sce2138        EQADLALDAAPLFRKNAAGGADYLYA-GAPTDDLWEGRWGILRVHGERQTDLLPLSGASL


Lcho_3893      ------------------------------------------------------------
Pfl01_1975     PAGMRNTVAFEGSCPRI-----S--------------ANPNGIGTRPTVQRNYEVVAALA
pputgb1_2447   PVAMRNTVNFDGICPKT-----T--------------ANPNGIGTRTTVKRNYEIVAALA
pseen2857      PVAMRNTVNFDGICPKT-----T--------------ANPTGVGTRPTVKRSYEVVAALA
GSU1252        GTVPFTA-----------------------------------------------------
STIAU_6861     QAADLSTMDSAALSFREQLAESSWTPTPLLELPPESVLGRMDPDNEETIQRSYEVDDAGQ
Sce2138        P-----------------------------------------------------------


Lcho_3893      ------------------------------------------------------------
Pfl01_1975     NDILGNSLGLSIGDS----------------------AGLG-------------QHVG
pputgb1_2447   NDILENRNGVSINDP----------------------AGIG-------------QHVG
pseen2857      NDILENRQGVSISDP----------------------AGIG-------------QHVG
GSU1252        --------------------------------ANDASFSGPCPTSALVRSYDVTAVA
STIAU_6861     -QVRERKLTYAMTYPDKMRQLDARLVQEAESQGRFGQKAANVDACPRGSKVRLYRVSAID
Sce2138        ---------------------------------RGRRREVEVCPPEAQKRAYVVRAIT


Lcho_3893      ------------------------------------------------------------
Pfl01_1975     GPLNPAGGTLVLNSRTVSIPQVTVTDPEDGETITIGGQSGPLHDPTAILYVRKSDLDPVS
pputgb1_2447   GALKANGGTLVFNSRRTAIPLVSGVDPEDGEPFTIGGHSAPLHDPTAILYVRKADLDAST
pseen2857      GPLKANGGTLVFNSRRTSIPLVSGVDPEDGETFTIGGHSAPLHDPTAILYVRKADLDAST
GSU1252        ARTALPGGSLVYNSRIGP-----------------QGIGPLQDPDAVLYVRTGDLNPD-
STIAU_6861     AAHWLPGGKLVYNSKYG----------------------LYDPDAILFARNEYLTDL-
Sce2138        AKGNLPDDRLTYNATYG----------------------LHDPNAILFVLEEDLAAL-


Lcho_3893      ------------------------------------------------------------
Pfl01_1975     GKLKPGIPVEPLVLRAAAGDCINITLENRLPSVMPDLTQTAVMQGIVKRDRNSGLGSTTF
pputgb1_2447   GKLKAGVPVEPLVLRANAGDCISITLENRLPLVMPDLPSTAVMHNVVKRDRFGSEGSTAF
pseen2857      GKLKAGVPVEPLILRANAGDCISITLENRLPLVMPDLPSTAVMHNVVKRDRFDSEGATAF
GSU1252        GTLKAGVPVEPLVLRAAAGECIDVTLRNSLPAVLTESPGYSALHPIVE---------FF
STIAU_6861     --KSGKRQPEPLILRARAGECLQVELTNLLPDQLSKTPQWAYHTAITP----------YF
Sce2138        --RSGARPPEPLVLRAAAGECVEVTLVNDLPEDPKSPHWNYNAPIVD----------GF


Lcho_3893      ------------------------------------------------------------
Pfl01_1975     SNNLMRPSSHVGLHAQLLAYDITKSDGTNVGANPTQTVPPRVGSSGAYPTRTYQYYAGHL
pputgb1_2447   ANNLMRPSSHVGLHAQLLAYDITKSDGANVGLNPVQTVPPRAGTSGAYPTKVYQYYAGHL
pseen2857      ANNLMRPSSHVGLHAQLLAYDITKSDGVNVGLNPVQTVPPRSGSSGAWPTKVYQYYAGHL
GSU1252        NFNEVRTSSIVGLHPQLVEYDVTRSDGTVVGNNQDQTVPPGG-------VRQYRWYAGDV
STIAU_6861     NVNQVRPSNHVSLHPQLVNYDVNTDDGANVGLNALQTVPPGG-------KRVYRWYMGDY
Sce2138        NVNQVKPSNRVSLHAQLVGYDVTASDGANVGENPDQTVERGE-------QRTYTWYAGDI


Lcho_3893      ------------------------------------------------------------
Pfl01_1975     EREGKPVTQLGRSVDNINATAVEFGGLNFTP-ADVIKQPQKGLGGAMSILPIGATWVDDA
pputgb1_2447   EREGKPVSQLGRTVDNINTTAIEFGGLNLTP-SDFIKQPQKGLVGAMSILPQTATWTEDT
pseen2857      EREGKPTLQLGRTVDNINTTAIEFGGLNLTP-ADVIKQPQKGLVGGMSILPQAATWTEDT
GSU1252        KVVN-----NMR-----VATPIEFGASSLIS-TDLIKHASRGALGALIIEPLGSSWIEDY
STIAU_6861     RSSN-----NGSATRRGIYTPVEFGIVNLRNMADVVNHGMHGGIGALIVEPPDAVWTTDV
Sce2138        SVTT-----RGDV----RWQPVELGAVNLKDMADVVNHPMHGGIGALIVEPEGAVWYPSR


Lcho_3893      ------------------------------------------------------------
Pfl01_1975     -RKVNATVTAP-----------------------------GQTTYRDFAMVWHKALNTR
```

```
pputgb1_2447    ATRAQATVKVS----------------------------GQPDYRDFVTVWQRALNMR
pseen2857       ASRAQATVKVS----------------------------GQPDYRDFVTVWQRALNMR
GSU1252         PTPHPSVSAGQRP-----------------SRASATVIRANGTTFREQVLVIQDDVALR
STIAU_6861      GTDAQARVRYTVQSNQTDEDEDEDEDDEDEDEDDEDEDEDGKTETFRELVILYQDDLGLH
Sce2138         DSRAAAQVKY------------------------RDENGSEIWFREFVLLYQDEVPLQ


Lcho_3893       ------------------------------------------------------------
Pfl01_1975      WANGRPV-------EGIAAEGFGVPTDPQDNSSMAINYKTEPLWYRFGLAPDAPFGHADG
pputgb1_2447    WADGRPV-------EGINTEGNGAAGDPQDNGNMAVNYKTEPLWLRFGMAPDSPFGRANG
pseen2857       WADGRPV-------AGINTEGNGAAGDPQDNGNMAVNYKTEPLWLRFGVAPDAPFGRADG
GSU1252         F-----------GDNTPVPFVTGMEDALDTGMKAFNYRTEPLWFRLGFSPLNLPFQNQA
STIAU_6861      SNQSQFWDT-EGLSSGTALRNTRGIDDSQDTGQKAFNYHTEPLWARLGLPPQTDPGILND
Sce2138         SDQAAFRCLNPALGCGATLANVAGPPDPDFTGHKAFNYRAEPIWARLGVGPDAPLGTRPG


                                                           Region F
Lcho_3893       -----------------------------------------------==========--
Pfl01_1975      AGYGDMTNAHMAYSNALVGGDPQTPVLYAKPGQPFRTHILMPSGGSRGTTFQLDGHVWSL
pputgb1_2447    LGFGDVPNAHMAYANALVGGDPQTPVLYAKPGQPVRNHIVMPSGGSRGMVYQLDGHIWPL
pseen2857       LGFADIPNVHMAYSNALVGGDPQTPVLYAKPGQPVRNHIVMPSGGSRGMTYQLDGHVWPL
GSU1252         INFANVL------HNSTTGGDPETPVFTVNAGDETRVRLVQPAGHNRHHTFALYGHVWQR
STIAU_6861      QELSGL-------LSSTAYGDPATPVFTANAGEPLRLRVGHPSGHSRQHVFTLHGAEWQF
Sce2138         LDVSGV-------FSSAGRGDPATPIFTAGRYDRVRFRVLQASGHRRQHSFGLSNTEWAY

Lcho_3893       ------------------------------------------------------------
Pfl01_1975      NPFQAEKSDTGGYPMGTPGVGSVRFGYNPMSMYIGAHESILPAAHFSFMI-PSAGGSNAI
pputgb1_2447    HNYQAEKSDVDGYPMSLPGIGSVRFGNNPMAMYIGAQESVLPAAHFSFML-PSAGGANAV
pseen2857       HNFQAEKSDVDGYPMGLPGIGSVRFGYNPMAMFIGAQESVLPAAHFSFML-PSAGGANAV
GSU1252         EPHT-------------NN--STRLGLNPKSFWRGSQDVVGPASHWDFLLDHGAGGAFQA
STIAU_6861      NPWA-------------EGASSLRMGPNTTSSIVSSQGTMSVMQSWNFNPLYGAGGIAHA
Sce2138         NPWA-------------QGSGSRYQGRNSKSFAVSAHGGIGPMTAWTVTPYFRSGGMFDV

Lcho_3893       -----------------------------------
Pfl01_1975      PGDYLFRDYGAYGN-TSGLWGLLRVTNEPEPAPPAQ-
pputgb1_2447    AGDYLFRDYAAYGN-LSGLWGILRVTNEAPPATAPAQ
pseen2857       PGDYLFRDYAAYGN-TSGLWGILRVTNEAPPAVAPAQ
GSU1252         KGDYLYRDMLPIHF-LNGLWGIMRVQ-----------
STIAU_6861      PGDYLYRDQTSYQWSSGGLWGVLRVQ-----------
Sce2138         PGDRLYMDQSNA-MLTGGLWGIWRVTP----------
```

Appendix B: The Amino Acid Sequence Alignment for *mcoA*.

The letters are the standard abbreviations for amino acids. The regions are highlighted and labeled. Lcho_30649 and Lcho_2591 are *Leptothrix cholodnii Sp-6*. Pfl01_2001is *Pseudomonas fluorescens Pf0-1*. Pputgb1_2665is *Pseudomons putida Gb-1*. *Pseudomonas entomophila L48* is pseen2124. GSU1252 is *Geobacter sulfurreducens PCA*. Hoch_4297 is *Haliangium ochraceum DSM 14365*.

```
Hoch_4297       MKTEKPASRRARTAGMLLA-IAGAALTPIACTAEFDEAVDYVSSMVTGTVGPLPD-----
Lcho_0649       ----------MRRGAYILGAL-GAVSLSV-------------AVTAQVVGEVDIFGDPP
Lcho_2591       -----------------------------------------MAVAQSTGSLDFFGDPP
Pfl01_2001      --MMDRAMRDP-KKLFLLTPLSVFLMLALG-------------QMTGVRASPIDDENQPE
PSEEN2124       --MKKPAA--TSTHLFRLSS-VTALLLGLG-------------VAC-AVASGLDDVSQPP
Pputgb1_2665    --MKTPAVIHPNRYVFQLSS-VTALMLSLG------------LIT-AMASPLDDNSPPP

Hoch_4297       NLDPEHAYEPA----------------------LEGGTTGTAEDAAEEFNTGNLRES
Lcho_0649       PNDPSALIEPAIGDQAN----EDLKLLEHSNHSS--YEGTPNGEKADVYVEAAI--QPLQ
Lcho_2591       ATDPSAYVEPKIGADGEAYLL---AQANPANHGA--FDGGLTGTHVDPQAQTSV--PQAE
Pfl01_2001      ASDPSAYYDEPEDRAG---ALNAILTMPEANLDSFDLPDGVKGTRNTTRMENIL--PPTV
PSEEN2124       PTDPSAYSDPPADPVAAAAALEALKSMPEANEGSLELNDGAFGTRATVSTDNVL--PPAQ
Pputgb1_2665    PTDPSAYTDQPDDPTPA---LLNLSTLPEANEGSLELTDGVYGNRDTVRTDNVL--PPAL

Hoch_4297       GGGIDVPTGSKASPLYGAQPFTQKMLRFEEFGLQPLPTSTTAGTLSFPTPAD--------
Lcho_0649       QTSEKIPTGGKPSPDFGAKRYTTPLLMFEEFGPEKLDPSVQAGTNPLPRPKIGPLPAQDP
Lcho_2591       MTSANVPTNSKPSPLMGARMFDTPMLMFEEFGTDRLDPSVAAPTQPFPRPKIGAAPESDP
Pfl01_2001      QTSFNYPTNGKPSPLYGAQPFTQQLVLFEEFGPEKLDPTTPAAPLSFPPAAIGPAPAQDP
PSEEN2124       QTSSQYPTNGKPSPLFGAEPFTQQLLLFEEFGPEKLDPATPPYTLTFPLPMTGPAPAQDP
Pputgb1_2665    QTSDRYPTNGKPSPLFGAQPFTQQLLLFEEFGPEKLDPTLPPPDLTFPVPTLGAAPAQDP


Hoch_4297       ---AESSPTGAALDAFLAQ-DIFPYPTRLANVTDLNPWQTAIESFLGRTLDTPPAEGRPP
Lcho_0649       RSAARSAPDGTELEAFLAQPGMYPFPTHQSNTLDANPWRDDVAEFLGRS-VNSPAEGRPP
Lcho_2591       VSVARSAPDGVELDSFLAQRGIAPYPTRESNTLDQNPWAAEVADFLDRQLTGAPAEGRPP
Pfl01_2001      NNIARSAPPGTALDAFLRQPGLTPFPSQYANTVDRNPWQAQIELFLNRH-IGSSAEGRPP
PSEEN2124       NNVARSSPNGNALEAFLTQPGLTPFPSQYANVLDRNPWKAQIEVFLNRNSVGSPAEGRPP
Pputgb1_2665    NVVARSGPSGTALEAFLKQPGLYPFPTQFSNVLDRNPWKAQIEMFLNRHPVGSPAEGRPP

Hoch_4297       GEDWAHQRYTEFAPQAYFQVAQAGARDNLGARDSFQRHGYALGEFGPGGLYHNTTGLTGF
Lcho_0649       GEGWAHQRWNEMYPQAFFKTAQAGAGSRVNNGFRDQKQLHKYKVGEFGPGGLYHKPFLTSKA
Lcho_2591       GEGWAHQRWNEFSPQVFYKTAQTGARLNRGLRDSRQLHGFRVGEWAPGGLYHR-------
Pfl01_2001      GKGWAHQRWNEFYPQVAYKTVQTGARLNGGLRDSRQMHGYAVGEFGPGGLYHNVAGVPAT
PSEEN2124       GKGWSHQRWNEFYPQAAFKTAQAGARINQGLRDRKQLHGYSKGEFGPGGLYYQTSDIPTT
Pputgb1_2665    GKGWSHQRWNEFYPQAAFKTAQAGARINLGLRDRRQMHNYTVGEFAPGGLYYQTSDIPTT

Hoch_4297       DGTTDGIQVRFHPNFPVQNANAVWTFDGTMPPKLLMARYGEPLLMRHYNALPIDPAANHG
Lcho_0649       NGTTAGIDVRFHQNFPVQHQNSVWTFDGTMPPKLLMVRYGEPILMRHYNALPIDPAANKG
Lcho_2591       FGTNSGTKVQFHPGLPVQNHNAVWTFDGTMPPKLLMARYGQPLLMRHYNALPIDPAANKG
Pfl01_2001      DGTAKGVDPRFHPAMPVQDHNSVWTFDGTLPPKLLMVRYGQPVLMRHYNGLPIDPSANRG
PSEEN2124       LGTTKGIDTRFHPKMPLQGHKALWTFDGTFPVKLLMVRYGQPLLMRHYNALPIDPSANAG
Pputgb1_2665    LGTTKGIDTRFHPNFPLQNHKSLWTFDGTFPPKLLMVRYGQPILMRHYNALPIDPSANGG
```

```
                         Region A
Hoch_4297       FGLHTISTHEHNGHNPAESDGYTQAFYFPGQFYDYRWPMILAGHDSVNTGATDPRAGMPD
Lcho_0649       FGLHTISTHMHNGHNPAESDGVAQAFFFPGQFYDYRWPIQLAGYDTVNTGATDPRAGFPC
Lcho_2591       FGLHTITTHEHNGHNPAESDGFANAFFFPGQFYDYRWPMQLAGYDTINTTATDDRAAFPC
Pfl01_2001      FGLHTITTHEHNGHAPAESDGYANAFFFPGQYYDYRWPIQLAGYDSINTKAEDPRAAFPC
PSEEN2124       FGLHTISTHEHNGHSPAESDGFANAYFFPGQYYDYRWPIQLAGYDTINTRAQDPRAAFPC
Pputgb1_2665    FGLHTISTHEHNGHSPAESDGFANAYFFPGQYYDYRWPVQLAGYDTINTRAQDPRAAFPC
```

```
                                                                 Region B
Hoch_4297      GN-----------------GGITNIPGDWRETMSTHWFHDHMLDFTAQNVYKGNAAMMN
Lcho_0649      SPGETLFVNDASPGKKACPANGVIKIRGDWRETMSTQWFHDHMLDFTAQNVYKGNAVMMN
Lcho_2591      STGETLYVNDLSPGLKTCQ-NGRIRVRGDWRETMSTHWFHDHMLDFTAQNVYKGNAAMMN
Pfl01_2001     APGETLWVNDMQPAKKTCD-HGTIKIRGDWRETMSTHWFHDHMLDFTAQNVYKGNAAMMN
Pputgb1_2665   SPGETLFVNDGTPGLKTCQ-NGSIKIRGDWRETMSTHWFHDHMMDFTAQNVYKGNAV
PSEEN2124      SPGETLFVNDASPGLKTCE-NGSIKIRGDWRETMSTHWFHDHMLDFTAQNVYKGNAVMMNMMN

Hoch_4297      YYSALDRGNEAIN---------------DGVNLRLPSGTALDWGNRDYDVNLVIADKAWD
Lcho_0649      YYSAIDRGNEKIDCHYTDPNANDPSKVPTKVNLCLPSGSGMPWGNRDYDINLVIADKAWG
Lcho_2591      YYSALDRGNEVLN---------------DGINLRLPSGSALPWGNRDYDVNLLLADKAWD
Pfl01_2001     YYSALDRGNESVN---------------DGVNLRFPSGSALPWGNRDYDVNLVFADKAWD
PSEEN2124      YYSALDRGNEALQ---------------DGVNLRLPSGSAMPWGNRDYDVNLVVADKAWD
Pputgb1_2665   YYSALDRGNEALQ---------------DGVNLRFPSGSGMPWGNRDYDVNLVIADKAWD

Hoch_4297      QAGQLFFNIFNLDGFIGDQILVNWLWKPYLDVRARRYRFRILNGSVSRYFKIALVVERND
Lcho_0649      RDGQLWFNVFNKDGFLGDRILTNFVYHPYLDVRARRYRFRILPGSVSRYFALALVHERTD
Lcho_2591      KEGQLWFNVFNKDGFIGDRILTNWTYYPYLDVRARRYRFRILNGSVSRYFMVALVQQVSG
Pfl01_2001     QQGQLWFNPFNTDGFIGDQVLVNWQWKPTLDVRARSYRFRLLNGSVSRYFKLALVREVKG
PSEEN2124      ANGQLWFNPFNSDGFLGDQILVNWQYQPRLKVRARSYRFRILNGSVSRYFKFAVVREVAG
Pputgb1_2665   ANGQLWFNPFNTDGFLGDQILVNWQYQPKLKVRARSYRFRILNGSVSRYFKFAVVREIAG

Hoch_4297      NSGQFDGTE--NNTSYDLVPFYMIANDGNILEHSVHFNGSSLN----GYTNQKGRLPTQA
Lcho_0649      TKGEIKATVGGKKISYDRVPFHMVANDGNILEHAVPFDGTVDLDANGDKLEHKGTLPQMA
Lcho_2591      TAGEFQGPA-GSNVSYNRVPFHMVANDGNLLEHAVAFDGKVDLDRNGNKAEHKGMLPIQG
Pfl01_2001     TGGEFQGPR-NSGVSYSRVPFHMIANDGNIMEHSVPFDGSMDLDADGDKQNHNAILPTQG
PSEEN2124      NSGEFKGPS-GSNLSYNRVPFHLIANDGNLMEHAVPFDGSMDLNGDGDRQDHNAILPMQG
Pputgb1_2665   TSGEFKGPS-GSNVSYARVPFHMIANDGNIMEHAVPFDGTLDLNGDGNLQDNNGILPLQG

Hoch_4297      IAERYDIIVDFAQ--FAPGTKLYFVNMMEHDDGRGPKDIIPIRDILDGEYAPDLD-----
Lcho_0649      IAERYDIVVDFKK--FKAGDKLYFVNTQEHADPVKTNRRIPVADVLTGKYKPAVLEDKDG
Lcho_2591      IAERYDIVVDFAS--FAAGSKLYFVNVLEHTKGEVTGSKISLARILSESYKPL-ARDKDG
Pfl01_2001     IAERFDIIVNFAKNGIKPGDKLFFVNLQAQDDGKGPKEVIPLADVLSEKYLAV-IKQTS-
PSEEN2124      IAERYDIIVNFAKNGIKPGDKLYFVNLMEHQTGKGPNKAIALADVLSEKYKAV-IKQTS-
Pputgb1_2665   IAERYDIIINFAKNGIKAGDKLYFVNLEEHRTGKGPEGTISLADVLSEKYKAV-IKQTS-

Hoch_4297      ----GDRHRTDPTVGKFLEFRVHAYA--------GTDLSMDPAEYVAGNGLGPNGTDKTM
Lcho_0649      DGVADRWTDGDPGVGKFMELRIQSCKNAAGKAVACVDTSMNPADYVAGK--------KKM
Lcho_2591      DRVADRWENGDPGVGRFMELRVHPSS--------GTDKSMNPANFVAGK--------KKM
Pfl01_2001     --KGPQWDRGDPVVGKVLQLNVKAYT--------GQDLSMNPAAYEPAKPGKAE--GLVM
PSEEN2124      --KGPEWDKGDPVVGKFLQLLVQPYS--------GQDLSMDPALYEPAKPGKAA--GLVM
Pputgb1_2665   --KGPQWDKGDPAVGKFLQLLVQPYT--------GQDLSMDPVAYEPAKPGKAA--GLKM

Hoch_4297      IPLPG-----FTQTELDNAIHRTFEFGRSSGTDSAPWTVKTDGGSGLAMDPRRLSAAP--
Lcho_0649      VPLKLNRDDPVHHAKLQNATHREFKFGHSGGTDEQPWTIKTDGGPGYNADMRRITAAPQL
Lcho_2591      IPLTLDRDSAADYAKLSQARHRTFEFGRSGGTDAQPWTIRTDHDLAVGADPRRTSAAPQL
Pfl01_2001     IPLKIHRDNAADKALLAKALHRTFTFGRADGTDEAPWTIKTDSGFGFHMDPRRLNASTKL
PSEEN2124      IPLPLDPGSASDQAKLKVARHREFVFGRSDGTDSAPWTIKTDGGFGYSMDPRRINAAPQL
Pputgb1_2665   LPLPIDRNSATDLAKIKDARHREFIFGRSDGTDTTPWTIKTDGGFGYSMDPRRISAAPQL

                                       Region C
Hoch_4297      -----------EKGSVEIWHLENGGGGWAHPIHIHFEEGQIFRRDGGAPPAWERWARKDV
Lcho_0649      STGPTAGGSVQNESPYEVWYLELQ-GSWDHPVHVHFEEGIILRRDGKAPPEWEKWARKDI
Lcho_2591      TQGPTPAGFA-GNGTLEVWKLDLG-GGWSHPIHIHFEEGIVLRRGGKAPPDWEKWARKDV
Pfl01_2001     ASGPTDAGVN-GIGTLEVWNIRAGGTGWSHPVHVHFEEGIILSRGGKAPPEWEKWARKDV
PSEEN2124      ASEATDGGFS-GDGTLEVWKIKNGGNGWSHPVHVHFEEGVILSRGGKAPPEWEKWARKDV
Pputgb1_2665   ANQSTDGGFS-GDGTLEVWKIKNGGNGWSHPVHVHFEEGVILSRDGKAPPEWEKWARKDV

                                      Region D
Hoch_4297      YRVGPMPDSGDSVDIAIRFREFLGSFMEHCHNTQHEDHAMLMRWDLENPGQTVIMPTPMP
Lcho_0649      YRIGPDPHAGSSVEVALQFREFAGTFVEHCHATQHEDNAMLMRWDLERPGQTLAMPTPMP
Lcho_2591      FRLGPDSESGDSVEVALRFREFAGTFVEHCHNTQHEDSSMLLRWDVERPGQVLAMPTPMP
Pfl01_2001     YRIGSEADGLDNVEMAINFREFAGTYMEHCHNTQHEDNSMLLRWDLEKPGQLQLMPTPLP
```

```
PSEEN2124       YRIGPDSESTAEVEMAIRFREFAGTYMEHCHNTQHEDTSMLLRWDLEHPGQFQLMPTPLP
Pputgb1_2665    YRIGPDADSSEEVEMAIRFREFAGTYMEHCHNTQHEDSSMLLRWDIEHPGQFQVMPTPLP


Hoch_4297       SWDGVSYVPTFALDTFRTGDTDAAANASSNFGHLGN-----------------------
Lcho_0649       TWDGVHYTGSAALATFKSGAGTGTTYQLGQ-----------------------------
Lcho_2591       TWEGVDYVGQAALPTFRTGDGTGPSYAFGR-----------------------------
Pfl01_2001      SWDGVRYVNSAALPTFRTGDGFGPQVT-VKP----------------------------
PSEEN2124       GWDGVEYVDSAALPTFRTAASS-DDNSANKPPVAANDSAATSAGKPIVLNVLANDSDPDG
Pputgb1_2665    GWDGVQYMASVGLPTFRTKGHDDNDDPANKPPVAANDSAATTAGKPITLNVLANDTDPDG          .


Hoch_4297       ----------------------------------------------------------
Lcho_0649       ----------------------------------------------------------
Lcho_2591       ----------------------------------------------------------
Pfl01_2001      ----------------------------------------------------------
PSEEN2124       NLPLSVSSLAQPDSGQGTVSSNGTQVTYTPPATVASPFTASFTYAARDAKGLESLVPATV
Pputgb1_2665    NVPLTVTGLSQPDSGQGTVSTDGTTVTYTPPATVETPFTASFNYTARDAKGAESVTPATV


Hoch_4297       ----------------------------------------------------------
Lcho_0649       ----------------------------------------------------------
Lcho_2591       ----------------------------------------------------------
Pfl01_2001      ----------------------------------------------------------
PSEEN2124       SVAVTPALAMDEIQVSSATVQLRSNNRWTWELSGSTSVTANNSISVTANTTSGPLELGAA
Pputgb1_2665    SIAVSPAAAVDQVQVTSATVQVRSGNRYTWDVAGTTSVATGNSISVTAASTSGPLNLGNA


Hoch_4297       -------------------------------------------
Lcho_0649       -------------------------------------------
Lcho_2591       -------------------------------------------
Pfl01_2001      -------------------------------------------
PSEEN2124       TLTTSGTGARWRLSVTTVGSGPASPPTVTVRSALGQSVTAKISIK
Pputgb1_2665    TLTVATSGARWKLSVTTTGTGPATPATVTVKSALGQSVTVPISIK
```

Appendix C: The Amino Acid Sequence Alignment for *mopA*.

The letters are the standard abbreviations for amino acids. Pputgb1_3353 is *Pseudomonas putida Gb-1*. M*ethylobacterium chloromethanicum CM4* is Mchl_1979. *Rhodopseudomonas palustris BisA53* is RPE_1007. *Roseobacter sp. MED 193r* is MED193_02985. *Roseobacter sp AzwK-3b* is RAZWK3B_00405.

```
Mchl_1979       ------------------------------------------------------------
RPE_1007        ------------------------------------------------------------
Pputgb1_3353    ------------------------------------------------------------
RAZWK3B_00405   MVFVNTNDLKHILDQIRIAEQHAAGTPLTDLVANPLLPYGLRLVDGTLNNLTPGREEWGS
MED193_02985    MVFVNTNDLKHILDQIRIAEQHVAGTPLTELVQNPLLPYGLRLVDGTLNNLTPGREAWGS

Mchl_1979       ------------------------------------------------------------
RPE_1007        ------------------------------------------------------------
Pputgb1_3353    ------------------------------------------------------------
RAZWK3B_00405   ADQVMPRLLESTFLTQPDAAHLADPSPRGAPGAEPTSYLQTSGSVYDAEPRVISNLIADQ
MED193_02985    ADQVMPRLLDSTFLTQPDAAHLADPSPRGAPGAEPTSYLQTSGSVYDAEPRVISNLIADQ

Mchl_1979       ------------------------------------------------------------
RPE_1007        ------------------------------------------------------------
Pputgb1_3353    ------------------------------------------------------------
RAZWK3B_00405   TLSNPAVIAAALSHAGMTGQAMLTTANEIVQAYQRVIDAQAAAGNVDQALELQRQELQTA
MED193_02985    TLSNPAVVAAALSHSGLTGQAMLTAANEIVQAYQRVIDTQAATGNVDQALELQRQQLQSA

Mchl_1979       ------------------------------------------------------------
RPE_1007        ------------------------------------------------------------
Pputgb1_3353    ------------------------------------------------------------
RAZWK3B_00405   LDTATAELTAAQGDVAAKTTAKNEADQAVTQAQDTLDNAAATMVTLQGQGQVAEMQVALN
MED193_02985    LDTASAELAAAQNNVGAMTSDKADADQTVTDAQTTLDAAAATMSLLQSSGQVGDAQAALS

Mchl_1979       ------------------------------------------------------------
RPE_1007        ------------------------------------------------------------
Pputgb1_3353    ------------------------------------------------------------
RAZWK3B_00405   AAVAAHTQAEADLKAAQDELLQAESAAASMLTMHNAKQTTVTNLQGQKATADQELATAEG
MED193_02985    LAQTALIQSQNALNAAQGELDTATTSAASMLSMRDAKQTEVTNLQNQKTTADQELATTEA

Mchl_1979       ------------------------------------------------------------
RPE_1007        ------------------------------------------------------------
Pputgb1_3353    ------------------------------------------------------------
RAZWK3B_00405   LLSDARAALALETDAAGAVTEAQTAVENATTAVSNAQTRVNELDAQLTQAEADLADLNTQ
MED193_02985    LLSDAQDALLLETNGADAVTVAQTAVDDATTAVNAAQTHVNELDGQLQQAQSELDGLNAQ

Mchl_1979       ------------------------------------------------------------
RPE_1007        ------------------------------------------------------------
Pputgb1_3353    ------------------------------------------------------------
RAZWK3B_00405   LQAAEQTRDAAQTRVTAAQEAAAATQADREDAQALFDTEVSELATAQAEMLSALTNYITG
MED193_02985    LQTAQDAVTDTQARVTAAEQAADETQTARDSAQALFDTEVTEQASAQVEMLSALNVYLSG

Mchl_1979       ------------------------------------------------------------
RPE_1007        ------------------------------------------------------------
Pputgb1_3353    ------------------------------------------------------------
RAZWK3B_00405   GSLADFSVASAAYSAAREEKLAAETALNDAIDADDEADQAVIDETAALTTAQSRVSTLTT
MED193_02985    GSLADFSVASAAYSAAREEKLAAETALADAISADDAADQLAATENAALSTAQSQVTTLTG

Mchl_1979       ------------------------------------------------------------
RPE_1007        ------------------------------------------------------------
Pputgb1_3353    ------------------------------------------------------------
RAZWK3B_00405   QRDNDMTTRDGLSTQLDEANGLLTSAQNTLTAAEQTLQTAQDDNQAYVDAQNAVDLAQAD
MED193_02985    QRDAAITTRDGLNTQLGAANTVLSTAQGNLSTVQQDLLTAQQNNQAYIDAQEAVTQAQAD
```

```
Mchl_1979        ----------MAVKLNLQDLTFILKQIKIAEAHASGIKLTELRLDAAGTLLTDRGLYDAT
RPE_1007         ---------MTAFTLNQADLAFILRQIKVAEAHAAGTALTSIYVDAAGNVVAQ-------
Pputgb1_3353     -----------MANFSKSDLEFILKQIFIAEAHADGASLIDL-----------------
RAZWK3B_00405    VDAAQGIVDGLVAQLGVAEGELATLNTNLGNADAEVTRLEGV-----------------
MED193_02985     VDAAQVVIDGLVSQLTTVEGELTTLNINLTNANNEVTRLEGI-----------------

Mchl_1979        GNWLGDAAAPKAIADPHVPYGLRTVDG----TYNNLVPGRETWGSSGQPMPQLFEPTYLN
RPE_1007         ----GTPGAVLAIADPKVPAGLRTVNG----EDNNIVPGREEWGAADQSMPRLLTSSYTT
Pputgb1_3353     -----------LPNSQVPFGLRTVDG----SYNNLVAGQSEFGAADNSFLRLLDASYRA
RAZWK3B_00405    -----------VQTET--TDLETAAGGVTTAQQNLAAAQ----MADAAILAQLDVIADA
MED193_02985     -----------VQTET--TDLNTAQGGVTTAQQALVAAQ----TADAAILAQLDVIADA

Mchl_1979        DADGDTMALGPGAPVITNTNYGLPGSVADADPRIISNLVVDATLDNPAAIAAALRIAGSE
RPE_1007         GAGTLDLNGPAPGGAVTGGNYAGPGTIVDTAPRTVSNLIVDMSLNNPAAIIAALTFAGSG
Pputgb1_3353     NYV-----------------GTGNVVDSQPRTISNLIVDQTANNPAAVEANGG-----
RAZWK3B_00405    TVTLT--------EATATAAQAALDLTAAEGVEATKQVAVDTATSDLNALTAPGA-----
MED193_02985     TIALN--------SATTAAEQAATELSAAQDIETTKQSAFDTATTNLDALNAPGA-----

Mchl_1979        NVIADQRAITAAHEALKAAQAANPAGDHAVLQSNLDALLEQTGVTVTNGSIDVLNVSPDE
RPE_1007         DVLGDQGEITAAFLALKAARDADPLGDHATLQQALDDVLEQKGVTVTNGSIEVPNVAPDE
Pputgb1_3353     -----------AA---------------PVMSPGIDGV---FGTADDKPVFFIPNVSPDV
RAZWK3B_00405    -----------AEAALAAAQA-----TAAEAQTTLDTLLTTHNITMDGNNVLLPDVTPDE
MED193_02985     -----------ADAALAAAQA-----IATEAQTTLDTLLSTHAITMDGNNVMLPDVTPDE

Mchl_1979        GLSKPFNAWMTFFGQFFDHGLDLISKGGNGTVYVPLAADDPLV-LGQDGLAGTADDLAPH
RPE_1007         GLSAPFNAWMTFFGQFFDHGLDLISKGGAGTIYVPLAADDPLRTHGPDGVAGTGDEVSEQ
Pputgb1_3353     GLTAGFNAWMTFFGQFFDHGLDLVTKSSTDIVFIPLRPDDPLYNA------------SSP
RAZWK3B_00405    GLSAPYNSWMTLFGQFFDHGLDLVGKGGSGTVYIPLQPDDPLYDA------------NSP
MED193_02985     GLSAPYNSWMTLFGQFFDHGLDLVGKGGSGTVYIPLQPDDPLYDA------------TSP

Mchl_1979        LRFMTLTRAAQ-V-----------EGSQRNVTTPFVDQNQTYTSNASHQVFLREYAL-VD
RPE_1007         MAFMALTRATPAS-----------DGSQTNTTTPFVDQNQTYTSHASHQVFLREYSM-AS
Pputgb1_3353     TNFMVLSRAVRTAGADGVVGTADDGQ--PNTTSPFVDQSQTYSSHPSHQVFLREYMLDAA
RAZWK3B_00405    TNFMVLTRATNQPGPDGILGTADDIREHFNKTTPWVDQNQTYTSHPSHQVFLREYDIDAN
MED193_02985     TNFMVLTRATNQPGPDGILGTADDVREHFNKTTPWVDQNQTYTSHPSHQVFLREYDLDAN

Mchl_1979        GRPVATGRLLGG----------------ADGGLATWADVKFQARTILGIELTDADVSAV
RPE_1007         GVPMATGKLLGG----------------ADGGLATWADVKAQARDALGIELSDLDVFNV
Pputgb1_3353     GDPVATGRLITNRDLGADGKFGTADDGNSENGGMATWAVVKAQARDLLGINLTDADVHSV
RAZWK3B_00405    GNPVSNGYLLHG----------------QTGGMATWGDVKAQAATKLGIQLNDSDVLDG
MED193_02985     GSPVSNGYLLHG----------------QSGGMSTWGDVKAQAAAKLGIQLNDSDVLDG

Mchl_1979        PQLLVDAYGEFVRSANGLPQVMVGVGPGGQAVYAS----GSLAEPLKLSAIQLPVGTVLV
RPE_1007         PLLRTDPYGEFIRDANGFPQVIVGIGADGIPNTAD----DIVASGTNLAPFDLA------
Pputgb1_3353     PLLATDAYGNFLRGPNGMPQVVMRVNNGADGIAGTADDVTTLVEGNRAAP----------
RAZWK3B_00405    PLLATDPYGNFIPGANGLPQLVVPNPA---YVEGGTEPLNILIEGDLANP----------
MED193_02985     PLLATDAYGNFIPGANGLPQLVVANPA---YVEGGTEPLNILIEGDLANP----------

Mchl_1979        GPNGAQNVIEAGETVAAARTFNAFLDDIAHNAVPVAV----------NGVLRPDADALTG
RPE_1007         ---------SLNGGLGPVRTSHAFLDDIAHNAAPVVV----------GGVLAPDADSATG
Pputgb1_3353     -----------ISLANAVSTGHGFLDDIAHNAEPVKV---------GGVLQADADSAVG
RAZWK3B_00405    -----------VDASQALRNGHAFLEDISHTAFPKGMIDHDRNPMTPEIEVLPDADTDTG
MED193_02985     -----------VDASQAVRNGHAFLEDIAHNAVPGTYVVDRFTGE--TATKQADADTDTG

Mchl_1979        NAVQMNPQTGRNLEYDNELLDRHFVTGDGRGNENIGLTAVHHIFHSEHNRQIDAHKLTIL
RPE_1007         NDVAINPLTGQRLEYDNELLDRHYITGDGRGNENIGLTAVHHIFHSEHNRQVDSQKLTIL
Pputgb1_3353     NVQPVGP-GGNNLTYDNELLDAHYIAGDGRVNENIGLTAVHHVFHSEHNRLVQQTKDTLL
RAZWK3B_00405    NAIMPNI-FGMNETYDNELLDRHFIAGDGRGNENFGLTAVHHVFHTEHNRQAAEMKQTIL
MED193_02985     NAIIPNQ-FGQNETYDNELLDRHFIAGDGRGNENFGLTAVHHVFHSEHNRQTTEMKQTIL

Mchl_1979        QSGDLAFINDWLATDIAALPGNFAQMTPLGQLAYANTLSWDGERLFQAARFATEMQYQHL
RPE_1007         RSGDTAFINEWLATDIGGLPSGFASLSGLDQLAYANSLNWDGERLFQAARFATEMQYQHL
Pputgb1_3353     AAGDLAFLNEWLIDDVIAIPTTPA---------GIAALVWDGERLFQAAKFGTEMQYQHL
```

```
RAZWK3B_00405      DSGELAFINEWLATPITEADLATA---------TVDTVEWDGGRLFQAAKFTTEMQYQHL
MED193_02985       DSGELAFINEWLATPINEDELSTA---------AIDTLTWDGGRLFQAAKFTTEMQYQHL


Mchl_1979          VFEEFARKIQPLVDPFVFNPV--TEIDPSIFAEFANTVYRFGHSMLTENMPRLGPDGQ--
RPE_1007           VFEEFARKIQPAIDPFVFNSV--TDINPAIFSEFANTVYRFGHSMLTEGMPRLDGAGN--
Pputgb1_3353       VFEEFARTIQPQIDEFLAPNGYDTSINPAILAEFAHVVYRFGHSMLTETVDRFDPSFNPV
RAZWK3B_00405      AFEEFGRTVQPQIAAFGVNGS--AEIDGAVMAEFAHVVYRFGHSMLTENVHTMDPN----
MED193_02985       AFEEFGRTVQPQIAAFMVNAS--AEVDASIMAEFAHVVYRFGHSMLTENVQTMDPN----


Mchl_1979          -----ALDADLGLIDAFLNPLAFDNDGGLSHDESAAAIMRGMTIERGSEIDEFVVGALRN
RPE_1007           -----SMDSDLGLVEAFLNPVLFDNDGAISHDAGAAAIVRGMTIERGNEIDEFVVDALRN
Pputgb1_3353       SGDPANPDQQLGLIAAFLNPLAFAGS-GATADEAAGAIIRGVTRQLGNEIDEFVTEALRN
RAZWK3B_00405      -----GVNTSNGLIEAFLNPVAFDLDQTLTSDQAAGAVARGMSRETGANIDEFITSALRD
MED193_02985       -----GVNTSTGLIEAFLNPVAFDLDQTLTSDQAAGAVARGMSRETGANIDEFITSALRD


Mchl_1979          NLLGLPLDLAAINIARGRDTGTPTLNEARAQLYAATGSTFLTPYTSWVEMAANLKNPLSV
RPE_1007           NLLGLPLDLAAINIARGRDTGMPSLNDARTQLYAASGSTFLKPYDHWVDFAANLKNPASI
Pputgb1_3353       NLLGLPLDLPALNIARGRDTGIPSLNEARREFYGATGDSQLKAYISWADFADHLKHPASL
RAZWK3B_00405      NLVGLPLDLAALNIARGRETGVPSLNAAREQFYSATGSEFLKPYEGWSEFAANLKNPASI
MED193_02985       NLVGLPLDLAALNITRGRDTGVPSLNAAREQFYAATGSEFLKPYEGWSDYAANLKNPASI


Mchl_1979          VNFIAAYGTHGTVVAATTLAAKRDAAMALVFGGDG-----APTDRLDYLNSRGSWAGRE-
RPE_1007           VNFVAAYGTHSTIAGATTLEAKRLAAMELVFGVDQDGDATVAADRTAFLTGTGAWAGVE-
Pputgb1_3353       INFIAAYGTHSSITGATTEAAKRAAAVALVLGGAG-----APADRLDFLNSTGAWANVTL
RAZWK3B_00405      INFIASYGTHETIVNATTVVEKRAAATDLVLGGDN-----APADRLDFVNGTGAWADAE-
MED193_02985       INFIAAYGTHDTIANATTVVQKRAAATDLVLGGDT-----APADRLDFVNGTGAWATIE-


Mchl_1979          -----------------TGFGAVDLWIGGLAEKQMPFGGMLGSTFNAIFEAQMENLQDA
RPE_1007           -----------------TGLNRIDLWIGGLAEKKMPFGGMLGSTFNAIFELQLENLQDG
Pputgb1_3353       AGKDGIAGTADDIAGVTVSGVDAIDFWVGGLAEKKMPFGGMLGSSFNFVFETQLEALQNG
RAZWK3B_00405      -----------------TGINDVEFWIGGLAEDIMPFGGMLGSSFGFVFQQQMEALQNG
MED193_02985       -----------------TGINAVEYWIGGLAEAIMPFGGMLGSSFGFAFQQQMEALQNG


Mchl_1979          DRFYYLSRVQGQNFLNELEQNSFSKIMLANSSLSLPGPDGIRGTADDIVPRHIGVDAFAD
RPE_1007           DRFYYLTRTQGQNFLNMLEQNSFAKMIMANTDLAQPGPDGIRGTADDIVPRHIGVDSFAD
Pputgb1_3353       DRFYYLSRTAGMNFGTELENNSFAKLIMLNSDVT----------------HLSNTVFLT
RAZWK3B_00405      DRFYYLARTAGMNMIAELENNSFASMIVRNTDIKDGG------------AHIPANIFSS
MED193_02985       DRFYYLSRTNGMDMLGGLENNSFASMIMRNTDIADGG------------AHIPANIFSS


Mchl_1979          YDFELEVNAANQLDQNGAAPGRDPTGNDPVLEAMGLGKVVRDDPGTAADEGASGFHASVN
RPE_1007           YDYVLEVDEANQEDYNGAAAGKDPHGADPFMEALGLGKVIRDDPGTAGP-----------
Pputgb1_3353       PTFTLEVNQANQFTGLGADGKADPTGGI---EINGVEIVPL--V----------------
RAZWK3B_00405      MDVILEVDQAVQSMPDPVSTDVDPFLA-----AMGTTMVERATA----------------
MED193_02985       MEYILEVDQSVQAMADPVSTELDPFLA-----AMGTTLVERETG----------------


Mchl_1979          ALVRRFGADGSPTGALVDGSEDGVGGAGSPVTWADLKANAAKLGIALTQADMLDAPVLRI
RPE_1007           ------------------------------------------------------------
Pputgb1_3353       ------------------------------------------------------------
RAZWK3B_00405      ------------------------------------------------------------
MED193_02985       ------------------------------------------------------------


Mchl_1979          GADGRLAFAPSSSVPEAVAVANGSFEGLALVAGQEGVILDGNGNYTTTSPAGWTIAGGVG
RPE_1007           ------------------------------------------------------------
Pputgb1_3353       ------------------------------------------------------------
RAZWK3B_00405      ------------------------------------------------------------
MED193_02985       ------------------------------------------------------------


Mchl_1979          GLFAPADAVVDPAGRDGANVVWLRGGATLSQEDGTTLQAGVGYTYSFKVGDRTDFTWPGA
RPE_1007           ------------------------------------------------------------
Pputgb1_3353       ------------------------------------------------------------
RAZWK3B_00405      ------------------------------------------------------------
MED193_02985       ------------------------------------------------------------
```

```
Mchl_1979          EARLVAVGGANPVTLGTLTLTEPADGQWGTFTLATGVVPSALVGLQLRLEIRNTGSGDAQ
RPE_1007           ------------------------------------------------------------
Pputgb1_3353       --------------------------------------------------------I----
RAZWK3B_00405      --------------------------------------------------------T----
MED193_02985       --------------------------------------------------------T----

Mchl_1979          ILVDDIELVRTAPAYRSDLTPAQTPGYDPAADPFLRDGAGNVLRTGQSIASPAADLDATV
RPE_1007           ------------------------------------------------------------
Pputgb1_3353       ------------------------------------------------------------
RAZWK3B_00405      ------------------------------------------------------------
MED193_02985       ------------------------------------------------------------

Mchl_1979          VDPAALNLPFATGHYLRFTGGEHVLVGGTDGNDTIITDFGDDGIWGDAGDDRIEAGAGVD
RPE_1007           -----------DANYIRFSGGEHIVVGGTSGNDTIITDFGDDGIWGDDGDDRIESGAGVD
Pputgb1_3353       RD--NPDTVGPDSNYLHYTGEDHVVLGGTSGNDIIISGDGDDTVYGDAGDDVLEGGAGND
RAZWK3B_00405      VDAPLIDGVREYDNFLKFNGGEHAVLGGTDQRDILNGGLGDDALWGGAGDDLLIGDAGVN
MED193_02985       GDAPLVDGAREYDNLLKFNGGEHVVLGGTDQRDILVGGLGDDALWGGAGDDLLIGDSGVN

Mchl_1979          LVNGGAGDDIITDSGDTGDFLKGEDGNDVIANSNGI---DILMGGRGKDAIFVGVD-ATE
RPE_1007           LVNGGAGNDIITDSGDTGDFIKGDEGDDVIANSNGI---DILMGGSGKDVVFVGVD-DTE
Pputgb1_3353       AVLGGAGDDIITDSF-GDNRLEGNAGNDVIVAGSMLAAGNLILGGDGQDFII-TTEDIST
RAZWK3B_00405      TLRGGAGNDILKDGD-DVSFLHGEDGDDVISAGGG--IGELMFGGKGNDAILMGSDDAAT
MED193_02985       TFRGGDGNDIIKDGD-DISFLHGEDGDDVISAGGG--AAELIFGGRGNDAILMGRDDAKH

Mchl_1979          VFAGEGDDFVIGGDDADLLMGNEGDDWMEGGGGFDTTAGDNSELFFNSAIKGHDVMFAGG
RPE_1007           VFAGEGDDFVLGGDGVDFLLGNEGDDWMEAGGGFDTTAGDNSELFFNSAIKGHDVMFAGS
Pputgb1_3353       TFGGQGDDFILGAKTNLPPTGNEGDDWIEKGTQ-DGAPGDNFAPLLGDEVVGNDIFVGGG
RAZWK3B_00405      AFGMDGNDFIYGGAGGEFMFGGEGDDWIEGGDGFDMVMGDNGDPFGGSRVVGHDVLDGGA
MED193_02985       TFAGMGNDFVFGGSGADIVDGGEGDDWIEGGDGFDFLVGDNNDNLGGSRVIGHDVLMGGA

Mchl_1979          DEHDFDGESGDDIMVQGESVMRNEGMFGFDWAIYKGNQIAANADMRIPIFT---------
RPE_1007           EEHDFDAESGDDIMVQGESVMRNEGMFGFDWAIFKGMSLNGYADMNIPIFT---------
Pputgb1_3353       GFDEMIGEGGDDIFVGSDAQDKMDGMSGFDWITNKNDKVGVTVDLSLAALAQPHGNAPNQ
RAZWK3B_00405      NDNDLLGEGGDDILIQREGVHVNEGDLGFDWVAHKDAMTGADIDLSRRLET---------
MED193_02985       NDNDLHGESGDDILFQGEGTHVNLAELGYDWIAHKGLTTHAEVDLTERVTT---------

Mchl_1979          --------TEEADILRNRFDKTEGLSGWRLNDTLIGDDRTAAANADAEAPAGAPIAAANE
RPE_1007           --------TEQADILRNRFDRVEALSGWDNDDTLIGDSRVFGDIAP------GDTAATTE
Pputgb1_3353       NAGVFNPVGASPASILDRFAEVEGVSGSNYADVLKGDDVDAVTILN------HGGAT---
RAZWK3B_00405      --------NETQAALRDRYDLVEAASGTQFDDFLYGDDREGAALPD------GIDLLAND
MED193_02985       --------EQ-TEFFRDRYIDAEAVSGTQNDDLIWGDNRLGNEAPA------DGAILDNE

Mchl_1979          GVFFNDGLDAAGIARIAGLDQIVSLAS---GQQFFEAGNILLGGAGSDTLQGNGGNDILD
RPE_1007           GIFFNDGLDQAGIDRIAGLDQIVQVGT----TGFFEAGNVLLGGAGSDLIEGNGGDDVLD
Pputgb1_3353       ----GSA--LTNVALIRGLQQFLADA--GLPTTGFATGNIMLGGNGSDLIEGRGGDDLID
RAZWK3B_00405      ATLFANELTQAGVDRIAGLRELLADMMGVDPADTFTGGNILLGGGGSDVIQGRGGDDVID
MED193_02985       VTLFGNELTQEGVNRIEGLRELLGNLMGDAPGDLFTGGNILLGGGGSDMIQGRGGDDVID

Mchl_1979          GDRWLNVRISIRNPAD---AGQEMATADSLKHVFDDSAANQARGWAGKSLFELMIDRVIS
RPE_1007           GDRWLNVRIRITGTGEANTADNQIATVDSLKHVFTAADA-ADPSWVGKSLFELMIARTIV
Pputgb1_3353       GDKWLNVRIAVYAPGDVNHTGPEIASFDSMVDM----------------IPFMLDRTYN
RAZWK3B_00405      GDKALNVRISVRDPQDP---TVELRSINSLFEI----------------SNELLDGSIS
MED193_02985       GDKWLDARISIRDQTDP---AIEIRSIDSLEEI----------------IPQLLDGSIS

Mchl_1979          PTQLAIVREVITTGAT-AADVDTAVFNDIRANYTITRAANGTLTVTHTTLT---N-----
RPE_1007           PNQLHIVREIIEDD-G-AGDTDIAVFNDIRANYAITGTADGSIRVEHVTVS---AVIDPA
Pputgb1_3353       PGQLKAVREILPGTSTGGAAFDTAIFSGVQSEYVVSQNTKGTADVADDVW----------
RAZWK3B_00405      PGQLQIVREILDAGQEG--DVDTAVYFDVRANYNIRTNADGSVTVDHIAPPIEGAFIDPI
MED193_02985       PSQLQITREIVDRGQEG--DVDTAVYFDVRANYSIRANADGSVTVDHINVTEGPNPIDPV

Mchl_1979          ---P--AVDDGTDTLRNIEKLRFADGTADVALVLNQPFDSLTIRPFDADGDDSSTLVATL
RPE_1007           NGRN--LVSDGVDTLRNFEVLRFADGDFALTPP------ELRLNGFDAGVFADNFDNAAL
Pputgb1_3353       TVTDTVAGRDGVDTLLHIERLQFADSQRVLVTGV--------------------------
```

```
RAZWK3B_00405      NGQPNVEGGDGTDRLTNVEILRFADMEIDLRTDP------------------------
MED193_02985       NGDPNPILGDTDRLTGIEILRFADMEIDLRTNP-------------------------

Mchl_1979          VNRVNATTRPVTLQWQVLADNGQWRNVTGADGQVTNGGTFFTPTGATGVEIRVVANWTST
RPE_1007           N--NSTGSTPWTSAWVESNDSGGV---------TTGQ------------IRIDAG----
Pputgb1_3353       ----------------------------------------------------------
RAZWK3B_00405      ----------------------------------------------------------
MED193_02985       ----------------------------------------------------------

Mchl_1979          VAGNTGLQQTASIQSAFVGTAAAEDITGSASPNVILGRDGDDDIEGEVGNDAIYAGSGDD
RPE_1007           ----------NSNTLQFIGNLPANGSNGAE---ITRGI----NLAGATQATLAYSA-NPD
Pputgb1_3353       ----------------------------NA-----QPTGS---------PTITDGNGAA
RAZWK3B_00405      ---------------------------ASNGLIQGTDGDDVLNGSNGRDAIFAGAGND
MED193_02985       ---------------------------ANNGLIQGTGGNDVLDGSNGPDAIFAGAGDD

Mchl_1979          RVDGGEGDDTLLGNDGADTLIGRTGNNTLDGGNDDDQLSGGHGNDRLIGGAGTDTAIYSG
RPE_1007           NLDPGESVTVQFSADGVSFV----NLNTIT---------GNGGNQP-------FTHTLAG
Pputgb1_3353       ITVGDLLTVSVAGVRDAD-------------------------------------NISAGN
RAZWK3B_00405      VINGLAGNDVIFGEGGSD-------------------------------------TITWNA
MED193_02985       AINGLAGNDVIFGEGGSD-------------------------------------TITWNA

Mchl_1979          PIAAYSFERNA-------QGEVVVSDNLGAEGD-GVDTLT-TIEQ-IQMGNDLTPYA---
RPE_1007           PFTANAAVRFVATAITGNANEVVTINDLSVSFTRPASPVTTDVEASFTEDGAAVPIASLP
Pputgb1_3353       PLGTLA----------------------------------------------------
RAZWK3B_00405      PNGGYDVVVGHG----------------------------------------------
MED193_02985       PLGGYDVVVGHG----------------------------------------------
-
Mchl_1979          LVANGT---AAVDIVVGTAGNNTLSGGAGNDLVFAGAGNDNVLWRTGDGRDFVDGGAGTD
RPE_1007           SIADDSATMASARIVLT-------NASAGDQLVIGGLPAGI----AGAIDTTVA---GQI
Pputgb1_3353       ----------------------------------------DRS---------V-----S--
RAZWK3B_00405      --------------------IVNDGGTDMLVINGDGTEDG---------F-----TMY
MED193_02985       --------------------IVNDGGTDMLVINGDGTEDG---------F-----TMY

Mchl_1979          SFRIMNGTGPVQQLTLAQARAQFANLSFRDDTQTVVV---------RNGIVIAELKNVEQ
RPE_1007           T-VTLTGTAS--RA---TYQEAIQAVAFRNTSQAPSVDDRIIHVTVNDGLVDSNIATS-T
Pputgb1_3353       --YYWQ--FEADPG-----SGVFQD--------------II-----------LLP-AGD
RAZWK3B_00405      TVAAWTGQAPVDPS-----SEIIIT-----R-TVAGVETAIM-----------EVRGIEE
MED193_02985       TVAAWAGQAPFDPL-----SEIIIT-----R-TVAGVETPIM-----------EVRGIEE

Mchl_1979          VAVNTVATGAPVVTDPTPTNGLVSPTEGQPLGALVAAIQDADGLGAFSLRWQQS------
RPE_1007           IHVTAVNDAPNAVNDRIVTNNVNAAAFQVPDWALTA--NDTDADGPFVVTGINGANNLTA
Pputgb1_3353       LAF------------------------------QSADG-TSFK------------
RAZWK3B_00405      VAI--------------------------------NGVQGIETIN------------
MED193_02985       VTI--------------------------------NGVQGVETIN------------

Mchl_1979          ---GDNGQTWTDIAGNAAGTLNYTPGQAQVGDVLRLRVSFTDGAGNPEEL-FSAPTGIVG
RPE_1007           THPAGGPTTITDTGFTNGGSFNYTVSDGLATDAANVQ-----------------V--
Pputgb1_3353       ----------------------VSPDLAGLSLRVKAIYQDAHGTTEVL-FSAPT----
RAZWK3B_00405      ----------------------AVGDFAGTSLAPNTIYVNGTAGREVIDFSGFTSTQ-
MED193_02985       ----------------------AVGDFAGTSLAPNTIYVNGTAGNEVIDFSGFTSTQ-

Mchl_1979          DSFTGTALNRTFNGTAGDDIANGADTALFGIQPNDTMNGGAGNDILNGRGGND-------
RPE_1007           ----VRDVTGAINGNNGTDI-------LVGDGAANTLEGAGGNDLIFAGGGDDSINWTAT
Pputgb1_3353       ------A---V----------------VQPGA--PVVPT-------------------
RAZWK3B_00405      -----RL---V---------------VDAGDGRDVITGGAGNDILNAGAGDDEITWSVG
MED193_02985       -----RI---V---------------VNAGDGSDIITGGAGNDILNAGAGRDEITWSVG

Mchl_1979          --TFIQVSTDGRDRVDGGAGT-DTYQLNGAAGPETFRIYSRSAWLQVAGNTEAQLAASTE
RPE_1007           SIGAVDIANDGRDFADGGAGTLDRFVVNGSGSAETFVVYAAADALAA---GFTGFKPGTE
Pputgb1_3353       ---------PATPVVDATAGG-EGLHMV--RSDLNFIL------------------AQ
RAZWK3B_00405      ---------GGSDVVDGGAGE-DTYTINGDASDETFRVYAADSWTG------APLATGTD
MED193_02985       ----------GSDVVDGGAGD-DTYTINGDASDETFRVYAADSWTG------APLATGTD
```

```
Mchl_1979        IVITRNGTGAAAIVAELDNVEEI-RVNTLQVTSPGGQNGGANGGDTIQVIGSFTGTSLNF
RPE_1007         IVITRNG----AVIAELDNVEEI-TINT-----------GAGNDSVTAVGNFNPTSLAF
Pputgb1_3353     IKIAEADAAGQDILSLIPNIRAPLGLRAV----------DGSNNNLMNLNGHN------
RAZWK3B_00405    IVITRDAGAGEQVIGQLQNIEEI-QINTA-----------GGVD-NVEVIGDFNRQLSLH
MED193_02985     IVITREADAGEQVIGQLQNIEEI-QINTA----------GGTN-DVQVIGDFNPTALNF

Mchl_1979        NTITIDGSSGDDTVDMAALTSAHRIVFRSNGGHDTIVGTLRPQDVIELPAGSNRADYVAA
RPE_1007         NTITINGDGGNDTVDVSTLQSAHRILFRSNGGNDTIVGNLRPQDVVELAPGTDLASYTLA
Pputgb1_3353     --NTQFGAAD--N--VFPRL-TDPVF-------------------NPAEGAPAGFFGP
RAZWK3B_00405    TSR---------------------------------------------------------
MED193_02985     NTISVMGEVA--NVDVTGRLSTHRVFVETESRTTTVTGTLTEGDVIVVPVDIDPATLT--

Mchl_1979        AGAN--G-MTTLSNGSHTITFSA------------------AGGMPRITVDAGADGGDG
RPE_1007         DNGN--G-TKTLSNGTHSITFTG------------------A--VPPQFENTPPPAGND
Pputgb1_3353     GSPAIPGSSYQQTSGPV-------------------------------------------
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     --ETIDGTTVTRTSGDFTLTYTVDDPTTFPIIVQSGSGLANALRGVATINEETGAPQGRG

Mchl_1979        EGVTGAFA-------YTAADIDGLEALVRGQR------PDNA-----GDDDVPTGYRELS
RPE_1007         EGVSGAFE-------YTPSDLSGLLALVNGQQ------PANG-----DDDGGPTGVRELS
Pputgb1_3353     ---FDSQPRTISNLIVDQTSNNPAAYAT---------AYDPGADGV-----LNFGAPGND
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     TVALGTLPRG--NVLLTQDDIAALKYMVTGQGEAPILPPEPGARAENEPADLIVGTRDLE

Mchl_1979        GHGNNLDHPT-WGSADQAFIRLTQARYGEADANGNRAINPIFDGLDARTISNILGTQEAG
RPE_1007         GHNNNQANPN-WGAADQPFIRITNAHYGDEDANGNRAINPVFDGLDPRTISNILGTQEAN
Pputgb1_3353     D----V----LKDGVRIVASPGMDGQFG-----------TTDDHDVY----LFENTAAD
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     GLTNNLLNPEISGGATLPFSRVTEARYAGIGEDGAGIVNPVFDDLDARAISNALGAQDAD

Mchl_1979        LPKAGNDANIFFMAMGQYIDHGLDFLPKGGNGSIVIGAPGGG----------APGSNNPA
RPE_1007         LPHAGNDANIFFMAMGQYIDHGLDFLGKGGNGSIQIGAAGGG----------APGSDNPA
Pputgb1_3353     AGLS-APFNAWMTFFGQFFDHGLDLVTKGGSGTIYIPLQPDDPLYVEGGFTNFMVVTRAT
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     AAKA-ASANMFMMSFGQYFDHGLTFIPKGGHDPITIGGADMG----------RPSGDNPA

Mchl_1979        DLTRGTVMAVDANGVPQHKNQTSPYIDQNQAYGSNALVGQFLRE----------------
RPE_1007         DLTRGSVAGY-ENGVPQHVNRTSPYVDQNQAYGSNDLVGQFLRE----------------
Pputgb1_3353     NLPGPDGILGNADDIHEHTNTTTPFVDQNQTYSSHPSHQVFLRAYVMTDDGPVATGRLIT
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     DLTRATATINPETGEIEHTNITSPVVDQNQVYGSSNLVGQLLRE----------------

Mchl_1979        ---SDGAQGVGMRLLAGAPDPSNPAFNLL----PTLRELVNHHWQADTIFAG---PDGPI
RPE_1007         ---GDGNGGVGAHLFAGGPDPSNPQFSLL----PTLRELIEEHWSNNTVFHSESLPDGAV
Pputgb1_3353     NRDLGADGKFGTAD-----DTEIGGMATWKVVKAQARDLLG------INLTDADVDNVPL
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     ---SGSNGGFGAHVLMGQEDPSASGFQLM----ATLRELLDHHTQAGTVFTDTDKGDVTL

Mchl_1979        SFRTYYTNFALSEGVTGTL-----------------------------------------
RPE_1007         AFRDYFPGLVQ-------------------------------------------------
Pputgb1_3353     LATDAYGNFIKGPNGYPMVVMKGVDGIAGTADDQQVEGNPLAPISLTNAVRTGHQFLADI
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     --EGYYPDLFNEDGTY--------------------------------------------

Mchl_1979        FNTETGAFDPQVLTKLVGNFMGSGHPLLLDTNP-----FISVLDHFVAGDGRANENFALT
RPE_1007         ----GGVINTAMLPGMISNFMGTSHALVLDANP-----FINVLDHYVAGDGRANENFALT
Pputgb1_3353     AHNAVPVFSGGVLAPDADNAVGNAVPVNPQTGANLAYDNELLDAHYIAGDGRVNENIGLT
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     --------NAATIKDLSDDFMGEGWPLLIDTNP-----FMNLLDHFVGGDGRANENVGLT

Mchl_1979        SIHTVWARNHNYHVEKLLESGFEG-------------------------TPEQVFQ
RPE_1007         SIHTIWARNHNHHVEGLEAAGFQG-------------------------TAEELFQ
Pputgb1_3353     AVHAIFHAEHNRLVAQTMDTVLDSHDLAFLNEWLLNPVTALPVTPAEIDALVWNGERLFQ
```

```
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      SMHTVWARNHNYHVDQLLASGYDAD---------------------------TPEELFQ

Mchl_1979         AAKMVNEAEYQRVVFDEYLETLIGGLRSDGTHGFEAYDPNVDVAISHEFAAAVFRFGHSL
RPE_1007          AAKMINEAEYQRVVFDEYLETLLGGLRSQGTHGFEEYDPGADAGISHEFAAAVFRFGHSL
Pputgb1_3353      AAKFGTEMQYQHLVFEEFARTVQPRVDL-FFAPTQVYDVDLDASIVAEFAHTVYRFGHSM
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      AARILNIGEYQQVVFNDFADSLLGGLQGSGTHGHDKYDPTTDARISHEFAAAAYRFGHSQ

Mchl_1979         IGQTLNVKGADGE----------TVPVSLFDAFLNPSNDPSVFTAPLP----PGYVPQPG
RPE_1007          IGQTMTVLDADGN----------PTQVNLFDAFLNPSNDPSVFPSPLP----PGYTPQPG
Pputgb1_3353      LTETVDRFDIDFNVIKDPASANPDQQLGLIAAFLNP--------------LAYAASGV
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      IGQSMTLKDVDADGN--PF----TVEVPLFDIFLNPTNDPDAFTADFGTLEQYGYKPQSG

Mchl_1979         YAQYGVGGIIGGTIEQAAEDVDFNIVDAVRNDLVRIRADLFAFNVARGWDVGLGTLNQVR
RPE_1007          YAQHGVNAIIGGTVSQPAEDVDFNIVDAVRNDLVRINADLFAFNVARGWDVGLGTLNQVR
Pputgb1_3353      TPEDATSAIVRGVTRQGGNEIDEFVTEALRNNLLGLPLDLPAINIARGRDVGIPSLNAVR
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      YAQYGVDNILGGLVEQPSEEVDLQVVDAVRNDLVRVSADLFAFNVARGRDVGLGTLNQVK

Mchl_1979         ADLAASTNPYIRDAVGFAGGDLSPYASWEDFQARNGLSDAVIAQFRQAYPDLVLAA----
RPE_1007          QDLAASTNPYVSEAVGFAGGDLSPYTSWEDFQQRNGLSQAVIEQFKQAYPDLQLAA----
Pputgb1_3353      RDVYGQ----------TGDTQLKPYTSWVDLVQHLKHPESLI-NFIAAYGTHSSITGATT
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ADLAASDNRYISEAIDLSDMSMTPYTDWEDFQARNGLSDEMIAKFQTAYPAL--------

Mchl_1979         ------------------ADIAAFRAI-----------NGDIAIAMQADGTGVVKGIDR
RPE_1007          ------------------ADVAAFQAI-----------NPDIDIAMNDDGTGIVKGIDR
Pputgb1_3353      LLEKRAAAMALVFGGEGAPADRLDFLNSSGAWANVTLPGKDGVLGTADDLKAVTITGVDA
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      -----------VLTVDTEQYD--AFVEA-----------NPDIALIDNGDGTMTVKGIDR

Mchl_1979         LDLWVGGLAEKHI-NNGVVGQTFWVVLHEQFDRLQDGDRFYYLERFDNFDFYENLVDGQG
RPE_1007          VDLWVGGLAEQHI-NGGLVGQTFWVVLSEQFERLQDGDRFYYISRFDNFDFYENFIDGQE
Pputgb1_3353      IDLWIGGLAEEKTPFGGMLGSTFNFVFENQMEKLQDGDRFYYLERTSGLSMNAELES-NS
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      VDLWVGGLAEQHI-QDGVVGHTFWVLIHEQLDRLQEGDRFYYVDQIGDLPVYNNFISNLT

Mchl_1979         FSDIVARNTGLTVLPEHIFELSDEDGPGTE-PGDDDDDDDGVTDPVGGDPDEDEDGPTDP
RPE_1007          FADIIARNTGMTGLPEHMFRTDPIDDENNQNPDDNEGDDDGT--PVG-------------
Pputgb1_3353      FAKLIMANTSAAHLPGLVFSDPGLYLELDQSK----------------QYNDGLGHADP
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      FGDIVTRNTGMTDLPQDVFSYTGDEIVEDNGTADQQTQQ---------PPVTDDAGNANQ

Mchl_1979         VGGGGDAGGDEDEDDGVTDPVGGGDDPGDD-----EDDG-------QGDGDGTTDPVGGG
RPE_1007          --NGDDETDDEDDD----DTASGGDDEDDD-----ETAG----------GDDDDDEITGG
Pputgb1_3353      L-------------------GE--NGEQVVF-----------------RDSPLTA-
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ADTGQDDTGQDDNGQGQ--TAQGDDTQGEMTVDNSTQTAQGGASADQEPVSNDDDTIDGG

Mchl_1979         DGDED-G--PGTGP-GTNPPVNHAPGVIA-GGANGDVLNGTAGADTILGLDGDDNILAGG
RPE_1007          DDDDDDE--TVGGD-DDVTPPPSTPGVIRTGTPQPDVLVGSAGDDNIVAFADADVIIADA
Pputgb1_3353      -------------------GPDSNYIRYAGAEHIVLGGTNGDDILVSSEGDDTVW---
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      DSSDADTETPASDPSTEQSPNGSDEGLVKSGTDLGDALVGAAGDDILSGHEGNDMLV---

Mchl_1979         GADVVRAGAGNDFVDAGEGRDVVFAGDGDDDVLSGGGADMVYGDGGNDRILAGAGNDLVT
RPE_1007          GADAISAGDGADFVNAGAGRDVVFAGAGDDQVFGGDHADTIYGDAGNDRIFGDQGNDMIN
Pputgb1_3353      --------------GDAGNDRIEGGDGNDQLRGGAGDDIISDMGGDDNIQGGDGNDVLH
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      --------------GNAGDDMLFGGDGRDDMLGGDGADMLMGGAGSDNIIGGDGDDMII
```

```
Mchl_1979        AGAGRDTVIGGEGDDLFVAETGDGDDTYWGDEMGGGLGSDTLDMSAITANIAVNLGTGLA
RPE_1007         AGAGDDTVFGGAGSDLLVAEIGDGNDVYFGDDSEGGSGIDTLDISAATVAVTVNLGAGAL
Pputgb1_3353     GGNGVNLIIGGFGNDFIVTGEDAS--EAIG-----GQGNDFILGSKA---NEQDMGNEGD
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     GGSGSDILMGNAGDDTFIGTDGDCNDIYYG-----GEGSDTIDMASVTSNLTVRLGNAGT

Mchl_1979        GRGSA---------TSTQSGRDVLWGVENVVTGSGNDDITASDAANVMD----------G
RPE_1007         LQGSV---------VSSQTGNDTIWGIENVNTGSGNDTITASNAVNVIN----------G
Pputgb1_3353     DWIEKGTSDGAPGDNFDPLGNDPIIGNDVFIGGNENDKFNGEGGDDIMVGSLGFGDRYIG
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     DRGSV---------TTEEGGRDTIWSVENFVGGSGDDTIFASDAANVLD----------G

Mchl_1979        GEGSDTYRFGSAAAANG--------------------------------DTIEGFRPGDK
RPE_1007         GAGDDTFRFLTASAADG--------------------------------DTILGFEPGDR
Pputgb1_3353     GSGYDWATFKGLAQGVTIDYSDRFFDVPPVPGSGASALVRFDIMEGLSGSAHGDFLRGDN
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     GDGNNTFVFETA-----------------------------A----NAQGDH-----

Mchl_1979        IDLSAIDADA---------GLAGNQAFTLATGAAFTGVGQLLVTQETRD-----------
RPE_1007         LDLSGMDANL---------VSGGDQSFVLING-AFNAAGQLAVSFESRE-----------
Pputgb1_3353     EDAASLPTNGATGSVLTNISLINGLSSLLAAGATFYDGGNIILGGSGSDLIEGRGGDDIL
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     --------------------INGL-----------SAGDLLQFGTGSDMLELTWGDADM

Mchl_1979        -DG-----DYVVVQGNTAG---------DASPEFKLAIKGNTAPTAADFTL---------
RPE_1007         -DG-----DFTIVQGNIDG---------NAEGDFKIEIAGHHNLTASNVGL---------
Pputgb1_3353     DGDKWLNVRISVRANH-DGTGPEIASFDSMEPMVPFMLNGTYNPGQLVIVREILNGTDNY
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     QD------SFSFSEGDEAGTTKISGSI--EDEQFELTVSGRFD---------------Y

Mchl_1979        ------------------------------------------------------------
RPE_1007         ------------------------------------------------------------
Pputgb1_3353     DTAVFSGVATDYSVVVDGNAVIVTDLVAGRDGVDHLTGIERLQFSDRSQASGVGTAVNAG
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     DNVG-------------------------------------------------------
-
Mchl_1979        ------------------------------------------------------------
RPE_1007         ------------------------------------------------------------
Pputgb1_3353     PTGHLVILDDATGVRDDTPVTGQLLRVTPLAVHDADNVSAANLTGAVNGPVAYYWQVENL
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     ------------------------------------------------------------

Mchl_1979        ------------------------------------------------------------
RPE_1007         ------------------------------------------------------------
Pputgb1_3353     PGSGVYEDITFVAAGEVSRAIGTTYRVTDDVAGLSIRVRAVYQDAKGTLEIVDSSPNSAP
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     ------------------------------------------------------------

Mchl_1979        ------------------------------------------------------------
RPE_1007         ------------------------------------------------------------
Pputgb1_3353     SAGPAVTGLLVQNQTLTANPATIVDADGLSNPQFTFQWQSNRGVGWVNIAGAISSTFVLT
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     ------------------------------------------------------------

Mchl_1979        ------------------------------------------------------------
RPE_1007         ------------------------------------------------------------
Pputgb1_3353     QAQVGQNMRVVVSYVDDFGVHESVASDILDPVANVNDAPTGAVLISDTTPDQGQTLTALT
RAZWK3B_00405    ------------------------------------------------------------
MED193_02985     ------------------------------------------------------------

Mchl_1979        ------------------------------------------------------------
RPE_1007         ------------------------------------------------------------
```

```
Pputgb1_3353      GGIADLDGLGAFSFQWQQGTGGTFANIAGATAATFTPGFAQGNQQLRVIVRYTDGFGTLE
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ------------------------------------------------------------


Mchl_1979         ------------------------------------------------------------
RPE_1007          ------------------------------------------------------------
Pputgb1_3353      SLTSAATAAVTVPPGVVLTGTNAANTLTGGAGNDVISGLGGNDILNGLAGADQLFGGVGN
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ------------------------------------------------------------
-
Mchl_1979         ------------------------------------------------------------
RPE_1007          ------------------------------------------------------------
Pputgb1_3353      DTLNGGDDADLLNGEDGNDILNGGLGADVMNGGAGNDTFVVDNVGDTVTEALNGGTDLVQ
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ------------------------------------------------------------


Mchl_1979         ------------------------------------------------------------
RPE_1007          ------------------------------------------------------------
Pputgb1_3353      TSLASYTLGANVENLTYTGSSAFTGTGNALANTITGGAGNDLLNGGAGADRLVGGVGNDT
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ------------------------------------------------------------


Mchl_1979         ------------------------------------------------------------
RPE_1007          ------------------------------------------------------------
Pputgb1_3353      YVVDNAGDVVVEATGAGTDLVRTTLASYTMAANVENLTYTGVGNFSGTGNGLANIINGAA
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ------------------------------------------------------------


Mchl_1979         ------------------------------------------------------------
RPE_1007          ------------------------------------------------------------
Pputgb1_3353      GNDTLAGDGGNDILNGNAGNDTLNGDAGNDQLFGGLGADRLNGGGGDDSLDGGDGNDTLL
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ------------------------------------------------------------


Mchl_1979         ------------------------------------------------------------
RPE_1007          ------------------------------------------------------------
Pputgb1_3353      GDAGNDTLLGGAGDDSLDGGNGNDSLQGGDGNDTLFGDVGTDTLIGGAGNDFLNGAGGND
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ------------------------------------------------------------


Mchl_1979         ------------------------------------------------------------
RPE_1007          ------------------------------------------------------------
Pputgb1_3353      TVVGGAGNDTMMATDGNDVFQFAAGFGNDLIINFDAIAAGGQDRLDITALNITAATFAAS
RAZWK3B_00405     ------------------------------------------------------------
MED193_02985      ------------------------------------------------------------


Mchl_1979         ----------------------------------------
RPE_1007          ----------------------------------------
Pputgb1_3353      VTIADVGADTLVSIGAADSIRLVGVADATTVTAADFILAG
RAZWK3B_00405     ----------------------------------------
MED193_02985      ----------------------------------------
```