DEPARTMENT OF CROP SCIENCES
DIVISION OF PLANT PROTECTION

MASTER THESIS

Recombinant Peptides as a novel plant protection strategy, own investigations

MATTHIAS WIEDEMANN, H0340707

SUPERVISORS

UNIV.PROF. DR.PHIL. AXEL SCHOPF AND ASSOC. PROF. DR. HOLGER BOHLMANN

VIENNA, JUNE 2012

ACKNOWLEDGMENT

I would like to thank Dr. Holger Bohlmann and Dr. Axel Schopf for supervising the thesis and for giving me their support. I would also like to thank DI Stephan Plattner, MSc. Muhammad Amjad Ali and MSc. Amjad Abbas, MPhil Kausar Hussain Shah as well as their other colleagues from Pakistan. Last but not least my thanks go to everybody in the plant protection team for being so supportive and friendly.

DEDICATED TO MY LITTLE NEPHEW MARKUS

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AMP – Antimicrobial Peptide

bp – base pairs

dd – double digested

DNA – deoxyribonucleic acid

E. coli – Escherichia coli

Etbr – Ethidiumbromide

kb – kilobases

LB – Luria-Bertani

MBP – Maltose Binding Protein

MR – Molecular Replacement

NMR – Nuclear Magnetic Resonance

ROS – Reactive Oxygen Species

SP – Signal Peptide

µl – microliter

PCR – Polymerase Chain Reaction

ABSTRACT

Thionins refer to a family of peptides mainly found in plants of the families *Poaceae*, *Brassicaceae* and *Santalaceae*. Typically, a thionin consists of 45-48 amino acid residues. 6-8 of these are cysteine residues, which form 3-4 disulphide bonds. Some thionins have cytotoxic activity and they are therefore of interest in the development of novel plant protection solutions.

The proteins are toxic to microbial cells, presumably attacking the cell membrane and rendering it permeable: this results in the inhibition of sugar uptake and allows potassium and phosphate ions, proteins, and nucleotides to leak from cells. Thionins are mainly found in seeds, where they act as a defense against consumption by animals. Barley (*Hordeum vulgare*) leafs contain thionin which is highly toxic to plant pathogens and is involved in the mechanism of plant defense against microbial infections.

Viscotoxins are thionins from the leaves and stems of the European mistletoe (*Viscum album*). There are several closely related isoforms that are toxic to different cell types. There are indications that viscotoxin A3 is not as poisonous as thionin 2.1 to *E. coli.* (Abbas, 2011)

The main task of this thesis was the cloning of Arabidopsis thionin THI2.1 and viscotoxin A3 into the pJOE49 vector with 6 different signal peptides: PelB, PhoA, MalE, OsmY, DsbA and OmpA. The expression of these toxins delivered by the signal peptides is subject to further study.

KURZZUSAMMENFASSUNG

Thionin bezieht sich auf eine Familie von Peptiden, die hauptsächlich in Weizen und europäischen Misteln gefunden wird. Typischerweise besteht ein Thionin aus 45-48 Aminosäureresten. 6-8 von diesen sind Cysteine und bilden 3-4 Disulfidbindungen. Einige Thionine haben antimikrobielle Aktivität und sind daher für die Entwicklung neuartiger Pflanzenschutzmittel interessant. Die Proteine sind giftig für mikrobielle Zellen. Vermutlich geht der Angriff auf die Zellmembran und macht sie durchlässig: dies führt zu einer Hemmung der Zucker-Aufnahme und ermöglicht Kalium und Phosphat-Ionen, Proteinen und Nukleotiden das Austreten aus der Zelle. Thionin aus Arabidopsis wird vor allem in Samen gefunden, wo es als Schutzmechanismus dient und die Keimlinge gegen pythopathogenen Mikroogrganismen verteidigt. Im Gerstenblatt *(Hordeum vulgare)* gibt es Thionin, das hochgiftig für Pflanzenpathogene ist und in den Mechanismus der pflanzlichen Abwehr gegen mikrobielle Infektionen eingebunden ist.

Viscotoxine sind kleine Proteine, die toxisch gegen eine Reihe von Zellen sind. Sie gehören zur Familie der Thionine und werden aus den Blättern und Stängeln der europäischen Mistel *(Viscum album)* hergestellt. Es gibt Hinweise darauf, dass diese Viscotoxine nicht so giftig sind wie Thionin 2.1, und besser in *E. coli* vermehrt werden können.

Die Hauptaufgabe dieser Arbeit war die Klonierung der beiden Toxine in den pJOE 49-Vektor mit sechs verschiedenen Signalpeptiden: PelB, PhoA, MalE, OsmY, DsbA and OmpA. In einer zukünftigen Studie könnte die Expression der beiden Toxine untersucht werden.

“Asking naive questions
is one of the most successful ways
of really moving ahead.”

Craig Venter

1 INTRODUCTION

Prokaryotic and eukaryotic cells as well have proteins that are able to entry into the secretory pathway under the precondition that they are with a specific targeting signal: a signal peptide (SP) This SP is in most cases a transient extension to the amino terminus of the protein and is cleaved off by a small class of enzymes known as signal peptidases after reaching its destination. (Heijne, 1990)

Eight families of antimicrobial peptides, with a size from 2 to 9 kD, were identified in plants until now. Thionins, defensins, hevein- and knottin- like peptides so-called lipid transfer proteins, MBP1 IbAMP, and the recently found snakins. There structure is compact, which is stabilized by 2-6 disulfide bridges. They are able to build up inducible or permanent defense barriers. (García-Olmedo, 1992)

1.1 ANTIMICROBIAL PEPTIDES

Proteins with a small molecular weight that have a broad defense array against bacteria, fungi and viruses are called antimicrobial peptides (AMPs). Many have a hydrophobic and hydrophilic side that was conserved throughout evolution. This constellation enables the molecules to be solvable in aquatic surroundings until entering the lipid-rich membranes. After accessing the cytoplasmatic membrane, the peptide starts to interact with the lipid bilayer and displacement of lipids by various mechanisms. (Yupping and Gallo, 2009) Defensins and cathelicidins are 2 groups of AMPs. Beside their very important function against microbes, there is evidence of AMPs being alerted in the immune response by receptor-dependent interactions. AMPs help in processes like wound healing or act in chemotaxis. Further investigations will lead to novel knowledge about this group of peptides. (Izadapanah, 2005)

1.2 THIONINS

Thionins are found in a variety of plants. They can be isolated from leafs, seeds or roots and are active against viruses, microbes and fungi. Thionins have a protein structure that throughout the whole family consists of a polypeptide chain of 45-48 amino acids and 3-4 internal disulphide bonds, have a similar three-dimensional structure and, with one exception called crambin, present a high amount of sequence homology. This involves a similarity of the dispersion of the hydrophobic and hydrophilic ends. Moreover, thionins demonstrate different toxic activity to fungi and bacteria, as well as animals and plant cells. This could indicate a role in plant defense against the action of bacteria, fungi, and insects. But until further research this stays uncertain.

Studies of the toxicity of thionin 2.1 show the need for an electrostatic binding of the positively charged proteins with the negatively charged phospholipids of the membrane, leading to huge turnarounds of the structure and integrity of the membrane. This has been observed in the cell plasma membranes and model membranes. (Caaveiro et al., 1997) (Huang et al., 1997) It is unclear why polypeptides from different plants with almost the same structures vary so much in function. The position of the amino acid and it' s charge must be related to their different effects on plasma and model membranes, but the molecular mechanism behind the toxicity of thionins is far from being understood. For example, it has been proposed that the toxicity of β-purothionin could be a result of its ability to generate ion channels in cell membranes, contrary to the toxicity of α-hordothionin and wheat α-thionin, which is binding to the membrane surface and misbalancing the surface. (Caaveiro et al., 1997) The investigation of the particular ways of intercommunication of thionins with plasma and model membranes is hence crucial in order to establish their function. Also, the study of the biological outcomes of these membrane-active proteins at a molecular stage is valuable not only from the biological point of view, but also from the medical one, as proteins and peptides would be a way to overcome multi resistance in bacteria like MRSA, which are not treatable otherwise anymore. Moreover, they might be a possible application against cancer, if differences in toxicity against tumor cells could result in new forms of treatment. (Giudici et al., 2003)

1.3 VISCOTOXINS

Thionins from mistletoes are called viscotoxins. They can be isolated from the leaves of mistletoes in relatively great amount. (Coulon, 2003) They consist of 46 amino acids, and are separated from the leaves and stems of European mistletoe (*Viscum album*). They show variants of toxic action against different tumor cells, which could be linked to the dissimilarity of the cytotoxicity. This dissimilarity may reflect structural differences in the three-dimensional structure. The plant defense role might be explained by the activity of viscotoxins, since their high expression rate results in greater resistance to pathogens. Six isoforms, namely A1, A2, A3, B, 1-PS, and U-PS, have been described. They are very strong when compared to viscotoxin B (VtB) that has weak toxic power. Viscotoxin A3 is building composites with negatively charged DNA, and it is currently being proposed that the helix-turn-helix motif of viscotoxin A3 could represent a new DNA-binding domain. The NMR structural determination of VtA3 has been announced only recently (Romagnoli et al., 2000), and the main form of VtA3 is very much alike to that found in the other members of the thionin family. There are differences in the amino acid sequence between VtA3 and VtB. The three-dimensional structure of VtA3 (Protein Data Bank code 1ED0) is the main form. There are changes in amino acid residues. One is localized in the antiparallel pair of α-helices and in

the segment that joins them. Moreover, all changes in charged amino acid residues are localized in one α-helix.

How viscotoxins A3 actually alter and modify membrane properties remains uncertain. The different phospholipid compositions of viscotoxin A3 VtA3 and VtB and the interaction with model membranes should possibly be investigated by infrared and fluorescence spectroscopies in an attempt to identify their mechanism of action. Since VtA3 is the strongest viscotoxin A3 whereas VtB is the weakest, the results could provide us with a new understanding of the molecular mechanism of action and of the linkage between the protein and the membrane, since they could possibly reveal that the interaction of thionins with the membrane might be more sophisticated than a simple electrostatic binding. (Giudici et al., 2003)

1.4 THIONIN GROUPS

Approximately 100 individual sequences of thionin, occur in more than 15 different plant species. The name of thionins is used for plant peptides of two different but well characterized groups. One is the α/β-thionins and one is the γ-defensins. Despite the common name and probably very distant common origin, they have quite differing three-dimensional structures.
Y-thionins should more appropriately be called plant defensins. Then they would have to be considered as a separate protein families, as many researchers are proposing it.
Traditionally the α/β-thionins were split up into five different classes (I, II, III, IV, and V) (Bohlmann and Apel, 1991):

Type I thionins are present in the endosperm of grains (the family Poaceae), that are highly basic, and consist of 45 amino acids. 8 of them are cysteines.
Type II thionins were extracted from leaves and nuts of the parasitic plant Pyrularia pubera and from the leaves of barley Hordeum vulgare. They are slightly less basic than typeI thionins (reduced positive charge from +10 to +7) and consist of 46–47 amino acids. Both type I and II thionins have four disulfide bonds. (Strec, 2006)
Type III thionins have been extracted from leaves and stems of mistletoe species, such as Dendrophthora clavata, Phoradendron tomentosum, Phoradendron liga, and Viscum album. As the class V thionins are not of importance for this work, they are not described further. (Strec, 2006)

1.5 THIONIN PROPROTEINS

Leaf thionins are encoded by a large gene family containing as many as 50 genes that are differtently regulated. (Bohlmann and Apel, 1987) Thionins get synthesized as much larger precursors with a molecular weight of about 15 kDa N-terminal to the thionin domain is a typ-

ical sequence that directs the proportion into the endoplasmic reticulum. The configuration of the proprotein is made of the actual thionin and a C-terminal extension. This C-terminal is called the acidic domain. It consists of a large number of manly acidic amino acids. The amino acids at most positions are variable, and only a few are highly conserved. Tyrosine at position 61 and a glycine at position 65 for example. The six cysteine residues are absolutely conserved. (Bohlmann, 1994)

1.6 DISULPHIDE BONDS

Maintenance of the tertiary structure of disulphide bonds is often supported by covalent bonds that are arranged in polypeptide chains of these proteins. Covalent bonds can either tie two amino acids in the same protein, or connect different polypeptide chains in a multi-subunit protein. The prevalent form is covalent sulfur – sulfur bonds that are cross linkages in proteins. They are formed when the cell is preparing newly synthesized proteins for transporting. Their form is catalyzed in the endoplasmic reticulum from an enzyme that couples two pairs of -SH groups with cysteine side chains. They are next to the folded protein. The confirmation of a protein is not changed by disulphide bonds but they act as an atomic lock that has the potential to put the protein back into the most desired confirmation. Lysozyme – which is to be found in tears and can dissolve bacterial cell walls – can keep its confirmation for a long time, as it is stabilized by cross – linkage of disulphide bonds. In the cytosol they normally fail to form because of the high concentration of reducing agents that put S-S bonds back into the cysteine -SH groups. But it is obvious that the proteins don't need the support of disulphide bonds in the mild environment of the cell. (Alberts et al., 2008)

1.7 SIGNAL PEPTIDES

Different classes of signal peptides are used to specify different cellular placement of proteins. But not all proteins possess signaling regions. Those which don't are maintained in the cytoplasm. The others are classified by their signal peptides. The significant factors determining interaction with the protein transport system are the properties of the amino acids that constitute the signal peptide region of a protein. Like this they determine the destination to which that protein is delivered.

Once its final destination has been reached, the amino acids comprising the signal peptide are often cleaved off the protein. The cleavage is catalyzed by enzymes known as signal peptidases. (Leversen, et al., 2009)

To estimate the probability of a singal peptide, which is called C-score, the program Signal P. was used. (Petersen et al, 2011) The graphical diagrams also contain a S-score which refers to the length of the signal peptide and the Y-score which supports the C-score to find the optimal cleavage site. (Petersen et al., 2011)

The normal configuration of signal peptides from various proteins can be described as a positively charged n-region. A neutral but polar c-region and a hydrophobic h-region are following. The (-3,-1)-rule states that the residues at positions -3 and -1 (relative to the cleavage site) must be neutral and small. So the cleaving can pass correctly. (Nielsen, et al., 1997)

1.7.1 PHOA

Alkaline phosphatase is encoded by the *phoA* gene in *E. coli*. The wild-type *phoA* gene has a signal sequence that allows the export of alkaline phosphatase into the periplasm where it is active. Due to the highly reducing environment, alkaline phosphatase is not active in the cytoplasm. Alkaline phosphatase activity can be detected on solid media containing X-P (a colorless compound cleaved to form a blue colored compound like Xgal). Gene fusions to *phoA* provide an assay for extra cytoplasmic proteins or domains of proteins. If *phoA* is fused to a domain of a protein that is in the periplasm, it will result in alkaline phosphatase activity (yielding an X-P + colony). If *phoA* is fused to a domain of a protein that is in the cytoplasm, it will NOT result in alkaline phosphatase activity (yielding an X-P – colony). (Shatkin and Soel, 1993)

1.7.2 PELB

The pelB gene encodes pectatelyase B, one of three pectatelyases identified in *Erwinia carotovora EC.* (Koebnik et al., 2000)

The signal peptide sequence consists of 22 N-terminal amino acids residues. Attachment to any other protein, at DNA level is possible. It is often used in genetic engineering, as it is transferring the fused proteins to the periplasm of Gram negative bacteria, like *E. coli* in our case. (Lei, 1987)

1.7.3 MalE

The maltose binding protein MalE (MBP) is part of the maltodextrin transport system in E. coli. The maltodextrins diffuse through the LamB pore, located in the outer membrane, into the periplasm. Here the sugar is bound to MalE which prevents it to diffuse back outside. MalE passes the sugar to a complex of MalF and MalG in the cytoplasmic membrane, where active transport is mediated by the ATPase MalK. (Ehrmann and Hofmann, 1998)

40622 Mr is the size of the MBP and it has a size of 30 x 40 x 65 Å. The two domains within the protein are very much alike regarding their secondary structure. The central pleated sheets are enclosed by helics. The domains are linked by 3 bridges, and in between is a channel where the sugar is bound. The hydroxyl groups of the sugar are cemented at side chains of the protein with hydrogen bonds. (Spurlino, 1992)

1.7.4 OsmY

OsmY stands for osmotically inducible gene of *E. coli*, and can be brought about to fold by a hyperosmotic stress of 8 to 10 and a 2- to 3-fold growth in a complex media. The gene outcome of the periplasmic protein has a molecular mass of 22 kDa on a SDS – PAGE gel. After 99.3 minutes a genetic fusion in the *E. coli* chromosome could be determined. Sequencing could be achieved after cloning and showed an open reading frame. The molecular mass of the precursor protein in the open reading frame is 21,090 with a mature protein of 18,150 after the signal peptide cleavage. The cleavage site Ala-Glu in the periplasmatic OsmY protein was validated in the open reading frame. (Yim and Villarejo, 1992)

1.7.5 DSBA

A family of periplasmic disulfide oxidoreductases catalyzes the correct disulfide bond formation in periplasmic and secreted proteins. Two possible pathways for the thiol-disulphide-oxidoreductase exist. One is called the isomerization pathway of DsbD, together with the periplasmatic proteins Dsbc and DsbG. The second is an oxidation pathway, belonging to the

inner membrane protein DsbB and the periplasmatic protein DsbA. The synergy of DsbA and reduced substrates results in catalyzing oxidation of the cysteine residues to disulphide bonds. (Bardwell and Hiniker 2004) After finishing the oxidation of the substrate, DsbA revolves to a reduced form again and goes back to an active oxidized state through awarding the electrons to DsbB. These electrons shift to the membrane-bound quinones, and relocate to the anaerobic electron acceptor molecular oxygen. (Bader et al., 1999)

DsbA acts monomeric, contrary to DsbC and DsbG that work in a dimming way. After oxidation enforced by DsbB, DsbC is transformed to a monomeric state and works as an oxidase. Through attachment of DsbA at the N-terminal domain of DsbC, a chimeric protein evolves and works like an isomerase. The isomerization pathways and the disulphide oxidation can replace each other. This could mean a similar substrate binding of these proteins, whereas DsbA is the predominant oxidant of disulphide in the periplasm. Several phenotypes of the DsbA strain can be determined. Examples are sensitivity to benzylopenicillin or loss of mobility that can all be referred to the calamity of disulphide bonds in the protein during the processes.

Modifications of members of the thioredoxin family can be complemented by DsbA, but the opposite is not possible. This indicates an important role in the folding. (Bardwell and Hiniker, 2004)

1.7.6 OMPA

Almost 50% of the outer part of the membrane of Gram-negative bacteria consists of proteins. A lot of prokaryotic proteins belong to the outer membrane A (OmpA) superfamily. Especially the OmpA of *E. coli* has been investigated in detail. It´s main function is to provide the physical linkage between the outer membrane and the peptidoglycan layer underneath. Another important function is the support of bacterial conjugation, as well as assisting as a receptor for some bacteriophages.

The structure has been confirmed by NMR, and contains 8 stranded beta barrels, that have been conserved by evolution. Crystallisation of the transmembrane domain is limited to 171 amino acids even though the total length would be 325 amino acids. The other 153 amino acids belonging to the C-terminal globular periplasmic domain could indicate a separate third domain. (Wang, 2002)

1.7 SIGNAL PEPTIDASE

In the 70´s the existence of a peptidase that is involved in the maturation of secreted proteins was first reported. Soon after, several membrane and secreted proteins that are typically synthesized as larger precursors and have an extra aminoterminal signal peptide were found.

They all share some common features. The N-terminus is basic, apolar in the middle and has small uncharged amino acid residues at the site of cleavage. The genetics revealed that the integrity of the hydrophobic stretch is essential for the stable transport of the proteins across the membrane frontier. (Dalbey et al., 1997)

2 EXPRESSION SYSTEMS

There are two main systems for the expression of recombinant proteins. From the time the cDNA got cloned, the researcher has to decide where he wants to amplify the proteins. It can be done in a prokaryotic (bacterial) or eukaryotic (Plant, yeast or mammalian cell) system. Which system will be chosen determines the vector that will be needed for cloning of the cDNA. As there are different promoters which function especially in *E. coli* and others that work best with yeast or mammalian systems. (Bernaudat et al., 2011)

Unfortunately, many AMPs are so toxic that they strongly interfere with the physiology of the expression systems. As a result, expression yields are dramatically reduced, and sometimes abolished. In fact, some AMPs are so toxic to the expression vector that it cannot be handled during the growth phase. (Fakhri, 2007)

2.1 ESCHERICHIA COLI

Escherichia Coli is one of the most widespread forms of expression system and despite the development of eukaryotic systems *E coli* remains the preferred host for recombinant protein expression. *E. coli* is easily transformable, grows quickly in simple media like LB, and requires inexpensive equipment for growth and storage. And in most cases, *E. coli* can be made to produce sufficient amounts of protein suitable for the intended application. (Gerstein, 2001)

2.2 YEAST

Yeasts integrate the easiness of genetic manipulation and fermentation of a microbial organism with the advantage to secrete and to modify proteins according to a general eukaryotic scheme. Yeasts are providing durable and effective platforms for the production of recombinant proteins. (Gelissen et al., 2005)

2.2.1 SACCHAROMYCES CEREVISIAE

Heterologous gene expression was first realized with the so called baker's yeast, *Saccharomyces cerevisiae*. For a long time this platform has been successfully used as vector system to produce various pharmaceuticals, especially insulin and HBsAg. However, this system has certain limitations and some disadvantages have to be taken into account. *S. cerevisiae* has a habit to hyperglycosylate recombinant proteins. Furthermore, N-linked carbohydrate chains are terminated by mannose attached to the chain via a α1, 3bond, which is considered to be

allergenic. Another problem is a consequence of the limited variety of carbon sources that can be used as a feedstock, putting boundaries to fermentation process design. (Gelissen et al., 2005) The amount of protein secretion has to be improved along with the stress tolerance as well. (Pérez-Torrado, 2010)

2.2.2 PICHIA PASTORIS

P. pastoris was initially destined for the production of single-cell proteins (SCP) for farm animals because of its fast growth in methanol-containing media, as this source of carbon from natural gas was very cheap in the late 60s. With increasing costs for methane during the oil price crisis in the 1970s, and the switch to soybean as the major animal feed, production of SCP decreased and its importance dwindled. Through genetic engineering a reliable way for gene expression was established and heterologous gene expression now became the main purpose of use. The amount of produced proteins exceeds *S. cerevisiae.* Elements that are derived from the methanol-metabolic pathway genes of *P. pastoris* can normally exist in the vegetative haploid state or the vegetative, multilateral budding way. The organism is considered to be homothallic. Diploid cells cultured in a standard vegetative medium remain diploid. When transferred to nitrogen-limited conditions, they undergo meiosis and produce haploid cells. Since it is most stable in its vegetative haploid state, easy isolation and characterization of mutants is possible. All *P. pastoris* expression strains are derived from strain NRRL-Y 11430. (Gelissen et al., 2005)

2.3 BACULOVIRUS-EXPRESSION SYSTEM

Baculovirus-mediated expression in insect cells has become well-established for the production of recombinant glycoproteins. Its frequent use arises from the relative ease and speed with which a heterologous protein can be expressed on the laboratory scale and the high chance of obtaining a biologically active protein. In addition to *Spodoptera frugiperda* Sf9 cells, which are probably the most widely used insect cell line, other mainly lepidopteran cell lines are exploited for protein expression. Recombinant baculovirus is the usual vector for the expression of foreign genes but stable transfection of – especially dipteran – insect cells presents an interesting alternative. Insect cells can be grown on serum free media which is an advantage in terms of costs as well as of biosafety. For large scale cultures, conditions have been developed which meet the special requirements of insect cells. With regard to protein folding and post-translational processing, insect cells are second only to mammalian cell lines. Evidence has been presented that many processing events known in mammalian systems do also occur in insects. One drawback however is the protein glycosylation and in par-

ticular the N-glycosylation, which in insects differs in many respects from that in mammals. (Altmann et al., 1999)

2.4 TRANSGENIC PLANTS

Using transgenic plants for human protein production has significant advantages compared to other expression systems. Transgenic plants are easy to handle. Gene transfer and recovery of transformants is simple. Proteins produced for the use in humans can be considered as extremely save. It is not possible for plants to transmit human pathogens. Through the use of transgenic plants the unplanned contamination of products with human pathogens like Lentivirus (AIDS) or Hepatitis C is avoided. (Ni, 1997)

2.4.1 TOBACCO

Mammalian cells often lack the ability of high level transgenic expression; whereas bacteria like *E. coli* miss the machinery for post-translational modification. Tobacco is considered as the optimal plant for genetically engineering. Cultivation costs are low and there is a high capacity of protein production. Plant systems have eukaryotic cell machinery for complex protein processing. They are much simpler for the adaption of large quantity production than reactors. (Ni, 1997)

2.4.2 ARABIDOPSIS THALIANA

Arabidopsis thaliana belongs to the mustard family, and it is native to Europe, North America and Asia. Various phenotypes are available for experiments. The Colombia and the Landsberg ecotypes are standardized and most widely used. The plants can be grown in petri dishes under fluorescence or in the greenhouse. They produce a high offspring and up to 5000 seeds in total. The genome size is relatively small with only 125 MB. It has only 5 chromosome pairs (2n = 10), for which detailed cards exist and is completely sequenced. The plant can be transformed very easily and the seeds can be stored for a long period of time (Meinke et al., 1998).

3 MATERIALS AND METHODS

3.1 VECTOR

The plasmid pJOE4905.1 is a rhamnose-inducible expression vector derived from pTST101 (Motejadded and Altenbucher, 2009) It was modified only by digestion, according to the steps shown in the diagram below:

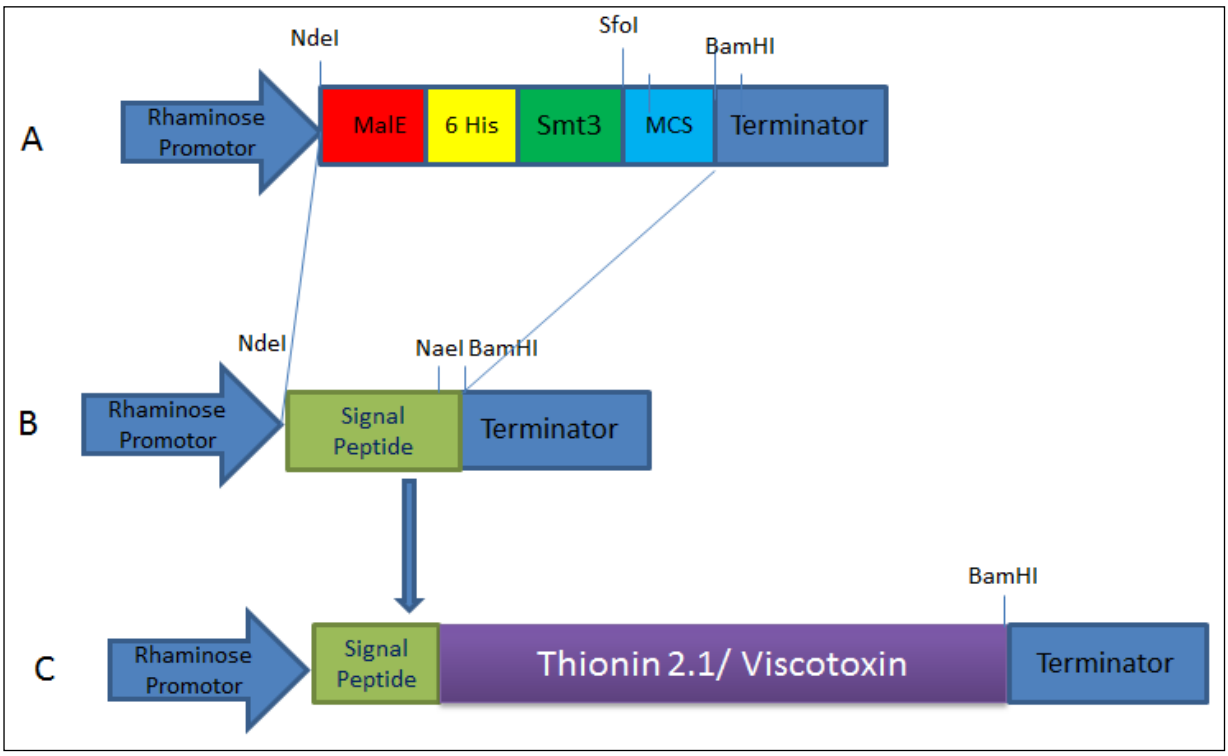

Figure 1: Vector pJOE 49 and cloning of signal peptides

A: The plasmid pJOE4905.1: The fusion partner maltose binding protein (MBP), an affinity His-tag and a Smt3 site. Cloning got carried out in the plasmid vector. The NdeI site N-terminal, the fusion partner and the BamHI site in the polylinker of the MCS were used for the integration of the different signal peptides.

B: PCR amplified signal peptide constructs were digested with NaeI and BamHI.

C: The Insert (thionin/viscotoxin) was digested with DraI and BamHI which produces blunt and sticky ends. Then they are compatible with the signal peptide and the plasmid.

3.2 PRIMERS USED FOR THIS WORK:

Enzyme recognition sites:

CA TATG= NdeI
GTAT AC

GCC GGC = NaeI / PdiI
CGG CCG

G GATCC = BamHI / BamHI-HF (High Fidelity)
CCTAG G

PhoA1: 5´-GAGCATATGAAACAGTCTACCTCGCGCTGGCGCTGCTACCGCTGCTfor

PhoA2: 5´-TATGGATCCTTAGCCGGCTTTGGTAACCGGGGTGAACAGCAGCGGCArev

OmpA1: 5´- GAGCATATGAAAAAAACCGCGATCGCGATCGCGGTTGCGCTGGCGGGfor

OmpA2: 5´-TATGGATCCTTAGCCGGCCTGCGCAACGGTGGCGAAACCCGCCAGCGrev

DsbAforNde: 5´-GAGCATATGAAAAAATCTGGCTGfor

DsbArevBam: 5´-TATGGATCCTTAGCCGGCAGACGAGAGAACGCrev

PelB1: 5´-GAGCATATGAAATACCTGCTGCCGACTGCGGCGGCGGGTCTGCTGCTGCfor

PelB2: 5´-TATGGATCCTTAGCCGGCCATCGCCGGTTGCGCCGCCAGCAGCAGCAGrev

OsmY1: 5´-GAGCATATGACCATGACCCGTCTGAAAATCTCTAAAACCCTGCTGGCGGTTATGCTGA

OsmY2: 5´-TATGGATCCTTAGCCGGCGTACGCAGAACCCGTCGCAACCGCAGAGGTCAGCATAACC

DsbA1: 5´- GAGCATATGAAAAAAATCTGGCTGGCGCTGGCGGGTCTGGTTCTGfor

DsbA2: 5´-TATGGATCCTTAGCCGGCAGACGCAGAGAACGCCAGAACCAGACrev

MalR2SPrevBam: 5´-TATGGATCCTTAGCCGGCCAGCGCAGACGCAGA

VaA3forDra: 5´-CAGTTTAAAAGCTGCTGCCCCAA

pJOEfor: 5`-CGAATTCAGGCGCTTTTTAG

pJOErev: 5`-CGCTTCTGCGTTCTGATTTA

3.3 RESTRICTION ENZYMES

As all DNA molecules consist of an almost equivalent mixture of the same four nucleotides, they cannot be instantly separated like e.g. proteins. Proteins can be separated on account of their different charges and binding properties. These problems were lessened with the discovery of restriction nucleases. These enzymes that are purified from bacteria cut the DNA double helix at particular sites that are defined by the local nucleotide sequence. A long double-stranded DNA molecule is cleaved into fragments of strictly defined sizes. Different enzymes have various sequence specifications, and it is quite easy to find enzymes that can create a DNA fragment that includes the gene of interest. Step by step purification of the gene from a mixture can be achieved by sorting of the DNA by size. Several species of bac-

teria produce dissimilar forms of restriction enzymes, which safeguard them from viruses by cutting off incoming viral DNA. A specific sequence of four to eight nucleotides in the DNA is recognized by each bacterial nuclease. Amino sequences that would be recognized by nucleases in the genome of the bacterium itself are protected from cleavage by methylation at an A or a C nucleotide. The sequences in foreign DNA are generally not methylated and so are cleaved by the restriction enzymes. Large numbers of restriction enzymes have been purified from various species of bacteria. Several hundred are commercially available. (Alberts et al., 2008) For this work NdeI, BamHI, NaeI (Fermentas) and NdeI, BamHI HF, PdiI and DraI (NEB) got used.

3.4 LIGATION

Ligation is the alignment of the ends of two (usually double-stranded) DNA molecules and the formation of a covalent linkage (phosphodiester bond) between them in one or both strands. A break in the sugar-phosphate backbone of a double-stranded DNA molecule that can be sealed simply by the formation of a phosphodiester bond is called a nick. If nucleotides are missing, then it is called a gap and cannot be sealed by ligation alone. (Howe, 2007) For this work, the signal peptide was always ligated with the open vector. After this, it was opened again and ligated with the thionin and viscotoxin inserts. Ligation was done at room temperature. At least for 4 hours or preferably overnight, with T4 DNA Ligase 5 u/ µl. (Fermentas) One control was without the insert. In the second control the enzyme and the insert.

Compound	**main**	**control 1**	**control 2**
Vector	5 µl	5 µl	5 µl
Insert	3 µl	0 µl	0 µl
T4 DNA Enzyme	1 µl	1 µl	0 µl
T4 DNA 10 x Buffer	1 µl	1 µl	1 µl
H2O	0 µl	3 µl	3 µl
Total	**10 µl**	**10 µl**	**10 µl**

Table 1: Master Mix for Ligation and Ligation Control

3.5 TRANSFORMATION

After ligation, the constructs were transformed into DH10 competent cells from Invitrogen. They have the ability to take up extracellular DNA. Their cell walls are made permeable by a heat shock or electric pulse. For this work, transformation through heat shock was chosen. The heat shock method is a

basic technique of molecular biology: The heat shock is initiated after the slow defrosting of DH10 competent cells on ice, which were used for cloning in this work. Through mutation this strain contains more plasmids per cell and is therefore suitable for high efficiency transformation of plasmids. After defrosting the ligation is added for five minutes. Meanwhile the shaker is set to 42 degrees for one minute to make the holes in the cell walls. After this, the cells are put on ice again for another five minutes for regeneration. The transformed competent cells then grow in antibiotic-free media for 45 minutes, before being put on a plate with agar media. One of the colonies from this plate that got verified by sequencing for containing the wanted constructs got picked by pipetting. It was transformed into BL 21 competent cells from New England Biolabs – NEB. The BL 21 cells contain a T7 promoter system for expression of toxic recombinant proteins. (Dumon-Seignovert et al, 2004)

3.6 PHOSPHATASE TREATMENT OF VECTOR

To minimize the possibility of self-ligated vectors after the transformation, the linearized vector was treated with Fermentas FastAP phosphatase to remove the 5' phosphate overhangs, necessary for ligation. This improves the percentage of colonies with inserts.

3.7 SINGLE AND DOUBLE DIGEST

Overnight cultures of the bacteria were made. Plasmid mini-preparations were purified (GeneJET Plasmid Miniprep Kit – Fermentas). For the double digestion of the plasmids the enzymes BamHI and NdeI were used. Depending on the availability, sets from New England Biolabs (NEB) or Fermentas were taken. Both companies offer adapted buffer solutions which were checked for compatibility with the online double digest finder. (NEB, 2012)The NEB buffer 4 has 100% activity with over 170 restriction enzymes. Fermentas offers the Tango Buffer 10 x for the enzymes used.

Compound	**Amount**
Insert	50 µl
Enzyme I(e.g. DraI)	1 µl
Enzyme II (e.g. BamHI)	1 µl
Tango Buffer 10 x	6 µl
dd H2O	2 µl
Total	60 µl

Table 2: Reference chart for the double digestion of plasmids and inserts.

3.8 GRADIENT PCR

To determine the optimum annealing temperature of primers for different templates, a gradient Mastercycler from Eppendorf got used. By gradually heating along a gradient the optimum temperature is established. The temperature of 57° could be determined for the pJOE primer pair.

3.9 COLONY PCR

Colony picking was done under sterile conditions. 16 colonies were picked with a needle and put 50µl LB medium which contained Ampicilin 100 mg/ml. Incubation time of the media at the shaker was 45 minutes at 37° and 180 rpm. A PCR master mix for the total amount of samples was prepared. The master mix was transferred to PCR strips with 8 containers and mixed with 1 µl of liquid from the standard Eppendorf reaction tubes (1,5 ml capacity). They got cycled in the PCR machine for 33 times with the pJOE primers at 57°.

Compound/ Samples	**8**	**16**	**30**	**40**	**50**
dNTP´s	4	8	15	20	25
Primer for.	4	8	15	20	25
Primer rev.	4	8	15	20	25
Buffer	20	40	75	100	125
TAQ Polymerase	1,6	3,2	6	8	10
H2O	159	318	600	800	1000
Total	200	400	750	1000	1250

Table 3: Master mix recipe for various amounts of samples.

4 RESULTS

In this paragraph, the work carried out for the thesis is described in chronological order, illustrated by pictures for a better understanding.

4.1 CLONING OF SIGNAL PEPTIDES

Initially, the inserts that would be the signal peptides, were only a bit longer than 100 bp. They were amplified with the specifically designed pairs of primers, and the suitable strain of *E. coli* as a template. After this, they were digested with the enzymes NdeI and BamHI.

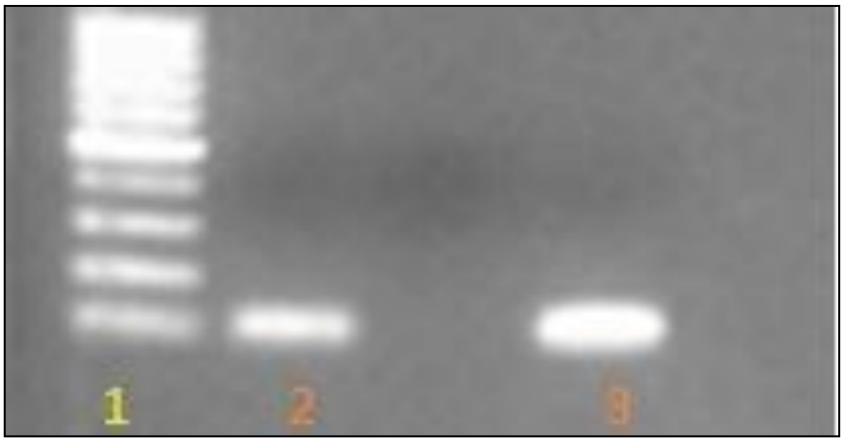

Figure 2: DsbA and OmpAdd insert.

4.2 DIGESTED PLASMIDS

Through the previous digestion of the inserts, it was possible to put them together with the vector pJOE49 that was double digested with NdeI and BamHI. In the next step, the vector and the insert were ligated as they were appropriate partners. Inserted into the vector pJOE 49, they looked longer than 100 bp. This was due to the fact that the pJOE primers started to amplify before. Overnight cultures can be digested directly without making mini-preparations to test whether the insert got cleaved. Then parts of the DNA get visible as dimmers as a side-effect. If the digestion with NaeI was successful, also BamHI was added to the mixture. After purification, the vector with the digested insert inside was ready for ligation with the two variations of thionins.

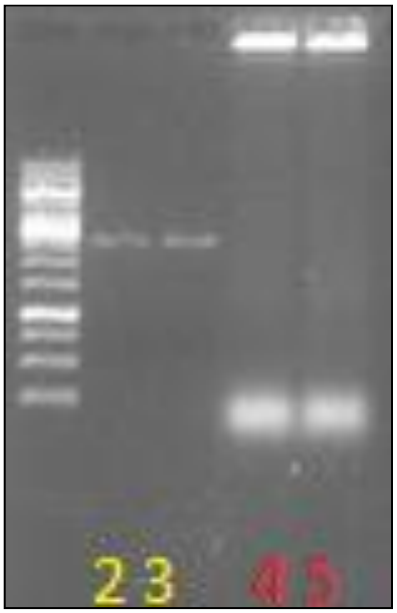

Figure 3: Digested plasmids.

Row 2 and 3: Plasmids OsmY signal peptides.

4.3 Amplified amino sequence of Thionin 2.1 with acidic domain and Strep-Tag

The thionin 2.1 was produced according to the same procedure as the signal peptides. After isolation from *Arabidopsis thaliana* and insertion into *E. coli,* it was prolonged with a Strep-Tag system that would simplify detection in an acrylamide gel. The Strep-Tag II is an eight-residue minimal peptide sequence (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) that exhibits intrinsic affinity toward streptavidin and can be fused to recombinant proteins in various fashions. (Schmidt and Skerra, 2007) It was cloned at the C-terminus of the THI2.1 thionin proprotein.

As it was inside the bacterial strain, it was possible to be amplified in the PCR machine by primers. To enable the ligation with the open vector, it was digested as well. NaeI is comparable with DraI, as they both have blunt ends. The back end can be cut with BamHI as it has a straight end. The amino sequence of thionin 2.1 with Strep-Tag and acidic domain is shown below:

KICCPSNQARNGYSVCRIRFSKGRCMQVSGCQNSDTCPRGWVNAILENSADAT-
NEHCKLGCETSVCGAMNTLQNSDASEIVNGASEQCAKGCSIFCTKSYV-
VPPGPPKLLGWSHPQFEK

4.4 AMPLIFIED AMINO SEQUENCE OF VISCOTOXIN A3

The principle described above is the same for the viscotoxin A3. However, this insert does not contain the strep tag as it could not be excluded that an additional tag would alter the activity of the thionin.

Thi2Va1.1 (Viscotoxin A3)

KSCCPNTTGRNIYNACRLTGAPRPTCAKLSGCKIISGSTCPSDYPK

4.5 LIGATION AND TRANSFORMATION

The digested signal peptides were ligated with the thionin 2.1 or viscotoxin A3 overnight. Then they were transformed into DH-10 competent cells by heat shock. The successful transformation can be seen by the numerousness of colonies. In the control plates almost no colonies are visible, as the vector could not re-ligate after the phosphatase treatment.

Figure 4: Successful transformation of PhoA with viscotoxin.

4.6 COLONY PCR

The colony PCR was the next step. Not all the colonies were transformed correctly. So a detailed check had to be carried out. Together with the signal peptide the amplified region of pJOE with thionin was around 600 bp long. This could easily be measured with the 100 bp ladder. In the picture below, one sequence containing only the signal peptide can be seen between the positive clones in row 2 and 3. The other positive clones (ladder 4 – 5) are shorter, as the viscotoxin A3 is missing the acidic domain.

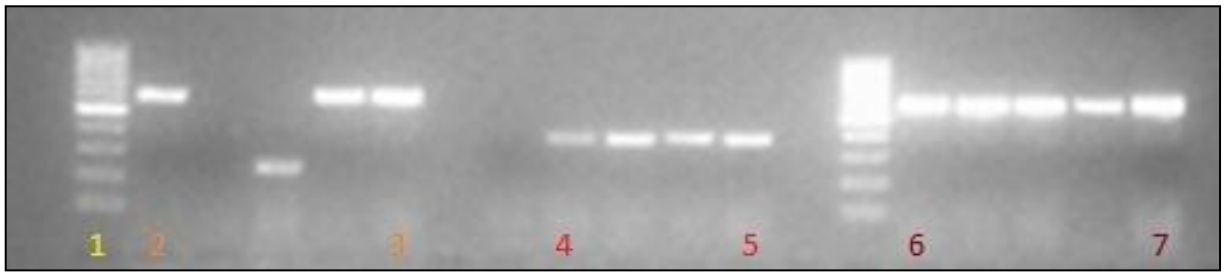

Figure 5: Colony PCR.

1st ladder: Colony 2 - 3 OsmY with THI 2.1 Strep-Tag/4 – 5 OsmYViscotoxin A3.

2ndladder: Colony 6 - 7 OmpA with THI 2.1 Strep-Tag.

4.7 PhoA

The whole process described in the previous paragraphs was repeated for each single construct, with no alterations. In a first step, we had the amino sequence for every signal peptide alone. The second construct then included the insertion of thionin 2.1 and the third amino sequence was used to illustrate the viscotoxin A3 constructs.

Enzyme Recognition Sites:

CA TATG= NdeI
GTAT AC

GCC GGC = NaeI / PdiI
CGG CCG

G GATCC = BamHI / BamHI-HF (High Fidelity)
CCTAG G

gtaatgaacaattcttaagaaggagatatacatatgaaacagtctaccatcgcgctggcg
V M N N S - E G D I H M K Q S T I A L A
ctgctgccgctgctgttcaccccggttaccaaagccggctaaggatcctgtacaagtaag
L L P L L F T P V T K A G - G S C T S K

Amino Sequence PhoAThionin 2.1 with Strep-Tag

HMKQSTIALALLPLLFTPVTKAKICCPSNQARNGYSVCRIRFSKGRCMQVSGCQNSDTCP
RGWVNAILENSADATNEHCKLGCETSVCGAMNTLQNSDASEIVNGASEQCAKGCSIFCTK
SYVVPPGPPKLLGWSHPQFEK

Amino Sequence PhoAViscotoxin A3

MKQSTIALALLPLLFTPVTKAKSCCPNTTGRNIYNACRLTGAPRPTCAKLSGCKIISGS
TCPSDYPK

4.8 PelB

gagcatatgaaatacctgctgccgactgcggcggcgggtctgctgctgctggcggcgcaa
E H M K Y L L P T A AA G L LLL A A Q
ccggcgatggccggctaaggatccat
P A M A G - G S

Amino SequencePelBThionin 2.1

MKYLLPTAAAGLLLLAAQPAMAKICCPSNQARNGYSVCRIRF-
SKGRCMQVSGCQNSDTCPRGWVNAILENSADATNEH
CKLGCETSVCGAMNTLQNSDASEIVNGASEQCAKGCSIFCTKSYVVPPGPPKLL-
GWSHPQFEKGSCTSKL

Aminio Sequence PelbViscotoxin A3

MKYLLPTAAAGLLLLAAQPAMAKSCCPNTTGRNIYNACRLTGAPRPTCAKLSGCKIIS-
GSTCPSDYPK

4.9 MalE

atacatatgaaaatcaaaacgggtgcgctgatcctggcgatgttctctgcgtctgcgctg
I H M K I K T G A L I L A M F S A S A L
gccggctaaggatccata
AG - G S I

Amino Sequence MalEThionin 2.1

MKIKTGALILAMFSASALAKICCPSNQARNGYSVCRIRFSK
GRCMQVSGCQNSDTCPRGWVNAILENSADATNEHCKLGCETSVCGAMNTLQNSDASEIVN
GASEQCAKGCSIFCTKSVVPPGPPKLLGWSTRSSKNKDPVQVSLAVW

Amino Sequence MalEViscotoxin A3

MKIKTGALILAMFSASALAKSCCPNTTGRNIYNACRLTGAPRPTCAKLSGCKIIS
GSTCPSDYPK

4.10 OsmY

gagcatatgaccatgacccgtctgaaaatctctaaaaccctgctggcggttatgctgacc
E H M T M T R L K I S K T L L A V M L T
tctgcggttgcgacgggttctgcgtacgccggctaaggatccata
S A V A T G S A Y A G - G S I

Amino Sequence OsmYThionin 2.1 with Strep-Tag

MTRLKISKTLLAVMLTSAVATGSAYAKICCPSNQARNGYSVCRIRFSKGRCMQVSGC
QNSDTCPRGWVNAILENSADATNEHCKLGCETSVCGAMNTLQNSDASEIVNGASEQCAKG
CSIFCTKSYVVPPGPPKLLGWSHPQFEK

Amino Sequence OsmYViscotoxin A3

MTMTRLKISKTLLAVMLTSAVATGSAYAKSCCPNTTGRNIYNACRLTGAPRPTCAKLSG
CKIISGSTCPSDYPK

4.11 DsbA

gagcatatgaaaaaaatctggctggcgctggcgggtctggttctggcgttctctgcgtct
E H M K K I W L A L A G L V L A F S A S
gccggctaaggatccata
AG - G S I

Amino Sequence DsbAThionin 2.1 with Strep-Tag

EGDIHMKKIWLALAGLVLAFSASAKICCPSNQARNGYSVCRIRF-
SKGRCMQVSGCQNSDTCPRGWVN
AILENSADATNEHCKLGCETSVCGAMNTLQNSDASEIVNGASEQCAKGCSIFCTKSYVVP
PGPPKLGWSHPQFEK

Amino Sequence DsbAViscotoxin A3

MKKIWLALAGLVLAFSASAKSCCPNTTGRNIYNACRLTGAPRPTCAKLSGCKIISGSTCPSDYPK

4.12 OmpA

gagcatatgaaaaaaaccgcgatcgcgatcgcggttgcgctggcgggtttcgccaccgtt
E H M K K T A I A I A V A L A G F A T V
gcgcaggccggctaaggatccata
A Q A G - G S I

Amino SequenceOmpAThionin 2.1 without Strep-Tag

MKKTAIAIAVALAGFATVAQAKICCPSNQARNGYSVCRIRFSKGRCMQVSGCQNSDTCPRG
WVNILGSCTSKLG

Amino SequenceOmpAThionin 2.1 with Strep-Tag

HMTMTRLKISKTLLAVMLTSAVATGSAYAKICCPSNQARNGYSVCRIRFSKGRCMQVSGC
QNSDTCPRGWVNAILENSADATNEHCKLGCETSVCGAMNTLQNSDASEIVNGASEQCAKG
CSIFCTKSYVVPPGPPKLLGWSHPQFEK

Amino SequenceOmpAViscotoxin A3

MKKTAIAIAVALAGFATVAQAKSCCPNTTGRNIYNACRLTGAPRPTCAKLSGCKIISGSTC
PSDYPKRILYK

5 DISCUSSION

To test if it would be possible to secret thionins with the help of one or several signal peptides together into surrounding media was the main basic approach of this thesis.

E. coli is a widely used expression system that is standardized for the large scale production of recombinant proteins. It is easy to establish and grows fast. *E. coli* itself and the vector that was chosen for this thesis had been in use before in the lab of Dr. Bohlmanns group.

The plasmid pJOE4905.1 allows the expression of fusion genes via an L-rhamnose-inducible promoter. This is favorable for the expression of toxic proteins like thionin 2.1 that can kill the bacteria and fungi. Cells can be grown until to a certain quantity before the Rhamnose gets added. So enough thionin can be produced before the media gets poisonous for the bacteria. To find positive clones was often problematic. Treatment with Fast AP phosphatase (Fermentas) improved the amount of positive clones. This together with the cheap way of production made *E. coli* the host of choice.

Pathways for extracellular secretion of proteins in *E. coli* are the Type I, II and V secretion systems.

Due to its simplicity the Type I secretion system (T1SS) is the most frequently used. The type I system is capable of transporting polypeptides of up to 800 kDa across the cell envelope in a few seconds. The system consists of only three subunits: the ABC protein, a Membrane Fusion Protein (MFP), and an outer membrane protein (OMP) (Holland et al., 2005).
The substrate haemolysin (HlyA), which includes toxins, lipases, haem-binding and S-layer proteins is transported through a channel formed by the interaction of the components of the system. Therefore, the substrate is transported directly forwarded to the medium without forming a periplasmic intermediate. The secretion signal needed for the recognition of the specific secretion machinery is located at the C-terminal end of HylA and comprises the last 46–60 amino acids of HlyA (Gentschev et al., 1996).

Extracellular secretion via the type II (T2SS) secretion system is a two-step process. Proteins following this pathway are synthesized as signal peptide-containing precursors and are transported across the inner membrane into the periplasmic space, where the signal peptide is cleaved of and the proteins get folded. In the second step, the proteins are transported outside the cell by a protein complex. It involves 12–16 protein components (Filloux, 2004).

The type V (T5SS) secretion system includes the autotransporter family, the two-partner system and the Oca family. Similar to the type II system, the secretion is a two-step process, involving first the translocation of the precursor through the inner membrane and then it's translocation through the outer membrane via a pore formed by a ß-barrel (Desvaux et al., 2004). The typeV secretion system is one of the simplest systems.

Signal peptides from *E. coli* and other sources of several outer membrane proteins that are transported by the general secretion pathway (GSP) have been used to achieve periplasmic or surface expression with various degrees of success. How this transfer is recognized, how it is fuelled, and how the signal sequence can assist in the process, remains to be seen. (Shokri et al., 2003)

The signal peptides from the genes PelB, PhoA, MalE, OmpA, DsbA and OsmY are secreting the proteins inside the bacteria to the periplasm. Oxidative conditions are predominant there. Through the outer membrane it is also possible for oxygen to diffuse. This is favorable for proteins with disulfide bonds like thionin 2.1 and viscotoxin, and a reason why it gets investigated as target place for expression.

The enzyme DsbA from the periplasmic disulfide oxidoreductases family catalyzes the correct disulfide bond formation in periplasmic and secreted proteins. Two possible pathways for the thiol-disulphide-oxidoreductase exist, and the isomerization pathway relies, just like, the oxidative pathyay on a protein relay.

PelB, OmpA, PhoA and MalE have been used for efficient secretory production of recombinant proteins in *E. coli*. There is no general rule for the selection of a proper signal sequence to guarantee the effective expression of a given recombinant protein. So the listed signal peptides had to be examined in a trial-and-error type procedure, together with several others. (Choi and Lee, 2004)
Also an extracellular expression of OsmY together with human target proteins could be reported. (Kotzsch et al., 2011)

The next step for this project will be the induction of the vector with the L-rhamnose at different temperatures and rpm speeds in the growing chamber. Different growing times will be adapted, as there is a possibility that the thionins get degraded by time. SDS phage would be used to determine if induction works and to find the place where the proteins got secreted.

E. coli has been used for both industrial- and laboratory-scale cytosolic production of recombinant proteins. Similar success with secretory and extracellular production of recombinant proteins will probably follow with more research efforts. (Coi and Lee, 2004)

6 CONCLUSION

With this thesis, the successful cloning of six signal peptides into the pJOE 49 vector in *E. coli* was demonstrated. It was shown that thionin 2.1 Strep-Tag and viscotoxin A3 can be attached to the constructs and remain stable inside the plasmids. Stock cultures were prepared after every step and were stored at -80°C in the freezer for further use.

Future research will focus on the expression of the toxins and should be performed at different temperatures in the growing chamber. Thionin 2.1 proteins will possibly be more visible on SDS – PAGE electrophoresis, whereas the viscotoxin A3 might be isolated by HPLC.

7 FIGURES

8 REFERENCES

Abbas MSc. A., 2011. personal communication. *Boku Vienna*, October 2011.

Alberts B., Johnson A., Lewis J., Raff M., Roberts K., Walter P., 2008. Molecular Biology of the cell . *New York: Garland Science*: pp. 423-425.

Altmann F., Staudacher E., Wilson I. B. H. and März L., 1999. Insect cells as hosts for the expression of recombinant glycoproteins. *Glycoconjugate Journal,* 16: pp. 109-123.

Bader M., Muse W., Ballou D. P., Gassner C., Bardwell J. C. A., 1999. Oxidative Protein Folding is Driven by the Elcetron Transport System. *Cell,* 98: pp. 217-227.

Bardwell J. C. A. and Hiniker A., 2004. In Vivo Substrate Specificity of Periplasmic Disulfide Oxidoreductases. *The Journal of Biological Chemestry, 279:* pp. 12967–73.

Bernaudat F., Frelet-Barrand A., Pochon N., Dementin S., Hivin P., Boutigny S., Rioux J. B., Salvi D., Seigneurin-Berny D., Richaud P., Joyard J., Pignol D., Sabaty M., Desnos. T., Pebay-Peyroula E., Darrouzet E., Vernet T., Rolland N., 2011 Heterologous Expression of Membrane Proteins: Choosing the Appropriate Host. *PloS one*, 6: p. 29191.

Bohlmann H. and Apel K., 1987. Isolation and characterization of cDNAs. *Bacterial disease resistance in plants*. pp. 446-454.

Bohlmann H. and Apel K., 1991. Thionins. *Annu. Rev. Plant Physiol., 42: pp. 227–240*

Bohlmann H., 1994. The role of thionins in plant protection. *Plant Science.*, 13: pp. 1–16.

Caaveiro J. M., Molina A., González-Mañas J. M., Rodríguez-Palenzuela P., García-Olmedo F., Goñi F. M., 1997. Differential effects of five types of antipathogenic plant peptides on model membranes. *FEBS Letters*, 410: pp. 338–342.

Choi J. H. and Lee S. Y., 2004. Secretory and extracellular production of recombinant proteins using *Escherichia coli. Applied Microbial Biotechnology*, 64: pp. 625–635.

Coulon A., Mosbah A., Lopez A., Sautereau A. M., Schaller G., Urech K.. Rougé P., Darbon H., 2003. Comparative membrane interaction study of viscotoxins A3, A2 and B from mistletoe (Viscum album) and connections with their structures. *Biochemical Journal,* 13: pp. 71-78.

Dalbey R. E., Lively M. O., Bron S., 1997. The chemistry and enzymology of the type I signal peptidase. *Protein Science*, 262: pp. 2–5.

Desvaux M., Parham N.J., Henderson I.R., 2004. The autotransporter secretion system. *Res Microbiol* 155: pp. 53–60.

Dumon-Seignovert L., Cariot G., Vuillard L., 2004. The toxicity of recombinant proteins in Escherichia coli: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression Purification*, 37: pp. 203-206.

Ehrmann M. and Hofmann E., 1998. The ABC maltose transporter. *Molecular Biology*, 29: pp. 685-694.

Fakhri S., 2007. Overview on the Expression of Toxic Gene Products in Escherichia coli. *Current Protocols*, 5,19: pp. 1-5.

Filloux A., 2004. The underlying mechanisms of type II protein secretion. *Biochim Biophys,* 1694: pp. 163–179.

García-Olmedo F., Molina A. and Allamilo J. A., 1992. *Plant defensine peptides.* Madrid: Laboratorio Bioquimica. pp: 1-7.

Gelissen G., Kunze G., Gaillardin C., Cregg J. M., Berardi E., Veenhuis M. and van der Klei I., 2005. New yeast expression platforms based on methylotrophic Hansenula polymorpha and Pichia pastoris and on dimorphic Arxula adeninivorans and Yarrowia lipolytica– A comparison. *FEMS Yeast Research,* 5: pp. 1079–1096.

Gentschev I., Mollenkopf H., Sokolovic Z., Hess J., Kaufmann S.H.R., Goebel W., 1996. Development of antigen-delivery systems based on the *Escherichia coli* hemolysin secretion pathway. *Gene,* 179: pp.133–140.

Gerstein S. A., 2001. *Molecular Biology Problem Solver: A Laboratory Guide.* s.l.: Wiley-Liss, Inc. pp. 462-465.

Giudici M., Pascual R., de la Canal L., Pfüller K., Pfüller U., Villalaín J., 2003. Interaction of viscotoxins A3 and B with membrane model systems: implications to their mechanism of action. *Biophysical Journal*, 85: pp. 971-981.

Heijne, Gunar von., 1990. Signal Peptide. *Journal of Membrane Biology*, 3: p. 115.

Holland I. B., Schmitt L., Young J., 2005. Type I protein secretion in bacteria, the ABC-transporter dependent pathway.*Molecular Membrane Biology* 22: pp. 29–39.

Howe C. J., 2007. *Gene Cloning and Manipulation.* The Edinburgh Building, Cambridge CB2 8RU, UK: Cambridge University Press. pp. 14-15.

Huang W., Vernon L. P., Hansen L. D., Bell J. D., 1997. Interactions of thionin from Pyrularia pubera with dipalmitoylphosphatidylglycerol large unilamellar vesicles. *Biochemistry*, 36: pp. 2860–2866.

Izadapanah A. MD., 2005. Antimicrobial Peptides. *Journal of the American Academy of Dermatology,* 52: pp. 381-390.

Kotzsch A., Vernet E., Hammarström M., Berthelsen J., Weigelt J., Gräslund S., Sundström M., 2011. A secretory system for bacterial production of high-profile protein targets, *Protein Science*, 20: pp. 597–609.

Lei S. P., Lin H. C., Wang S. S., Callaway J., Wilcox G., 1987. Characterization of the Erwinia carotovora pelB gene and its product pectate lyase. *Journal of Bacteriology*, 169: pp. 4379-4383.

Leversen N. A., de Souza G., Målen H., Prasad S., Jonassen I., Wiker H. G., 2009. Evaluation of signal peptide prediction algorithms for identification of mycobacterial signal peptides using sequence data from proteomic methods. *Microbiology; 155:* pp. 2375–2383.

Meinke D. W., Cherry J. M., Dean C., Rounsley S. d., Koorneef M., 1998. Arabidopsis thaliana: A Model Plant for Genome Analysis. *Science*, 282: pp. 262-268.

Motejadded H. and Altenbuchner J., 2009. Construction of a dual-tag system for gene expression, protein affinity purification and fusion protein processing. *Biotechnology Letters*, 31: pp. 543-9.

New England Biolabs., 2012. *www.neb.com.* [Online] 2012. [Cited: 06 07, 2012.] http://www.neb.com/nebecomm/DoubleDigestCalculator.asp

Ni H., 1997. *Expression of human protein c in transgenic tobacco.* Blacksburg, Virginia: pp. 43-45.

Nielsen H., Engelbrecht J., Brunak S., von Heijne G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng., 10:* pp.1-6.

Pérez-Torrado R., Panadero J., Hernández-López M., Prieto J. A., Randez-Gil F., 2010. Global expression studies in baker's yeast reveal target genes for the improvement of industrially-relevant traits: the cases of CAF16 and ORC2. *Microbial Cell Factories*, 9. pp. 2-9.

Romagnoli S., Ugolini R., Fogolari F., Schaller G., Urech K., Giannattasio M., Ragona L., Molinari H., 2000. NMR structural determination of viscotoxin A3 from *Viscum album* L. *Biochemical Journal*, 350: pp. 569–577.

Schmidt T. G. and Skerra A., 2007. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nature Protocols*, 2: pp. 1528-35.

Shatkin J. A. and Soel W., 1993. Membrane Topology Model of Escherichia coli oL-Ketoglutarate Permease by PhoA Fusion Analysis. *Journal of Bacteriology*, 175: pp. 565-567.

Petersen T. N., Brunak S., von Heijne G., Nielsen H., 2011. *SignalP 4.0: discriminating signal peptides from transmembrane regions.* Nature Methods, pp.785-786.

Shokri A., Sandén A. M., Larsson G., 2003. Cell and process design for targeting of recombinant protein into the culture medium of Escherichia coli. *Applied Microbial Biotechnology*, 60: pp. 654–664.

Spurlino J. C., 1992. The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *Journal of Biological Chemestry*, 266: pp. 5202-5219.

Strec B., 2006. Plant thionins - the structural perspective. *Cellular and Molecular Life Sciences*, 63: pp. 1370-1375.

Wang Y., 2002. The function of OmpA in Escherichia coli. *Biochem. Biophys. Res. Commun.*, 292: pp. 396–401.

Yim H. H. and Villarejo M., 1992. osmY, a new hyperosmotically inducible gene, encodes a periplasmic protein in Escherichia coli. *Journal of Bacteriology*, 174: pp. 3637–3644.

Yupping L. and Gallo R. L., 2009. AMPed Up immunity: how antimicrobial peptides have multiple roles in immune defense. *Trends Immunology*, 30: pp. 131-141.

9 ATTACHMENT

The C-score is the “cleavage site” score. For every position in a transmitted sequence, a C-score is reported, and it should only be outstanding at the distinct cleavage site. The cleavage position is referring to an individual number that highlights the residue at the starting point of the mature protein.

The Y-max value is a derivate of the C-score. Together with the S-core it ensures a better prediction of the cleavage site than the gross C-score alone. This is especially useful when several peaks occur and more precise information is sought.

The S-mean is supporting the calculation of the signal peptide length, starting from the N-terminal amino acid to the highest Y-max score.

The D-score gives a simplified average of the S-mean and the Y-max score. (Petersen et al., 2011)

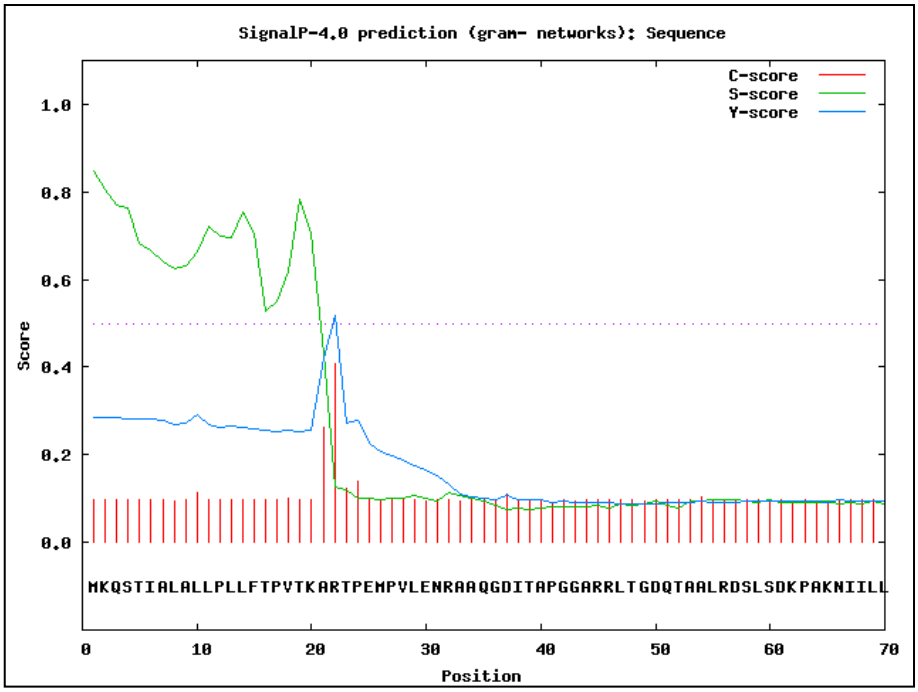

Figure 6: C-score PhoA.

Measure Position Value: Cut-off signal peptide?

max. C	22	0.408	
max. Y	22	0.517	
max. S	1	0.848	
mean S	1-21	0.681	
D	1-21	0.594	0.570 YES

Name=Sequence SP='YES' Cleavage site between pos. 21 and 22: TKA-RT D=0.594 D-cutoff=0.570 Networks=SignalP-noTM

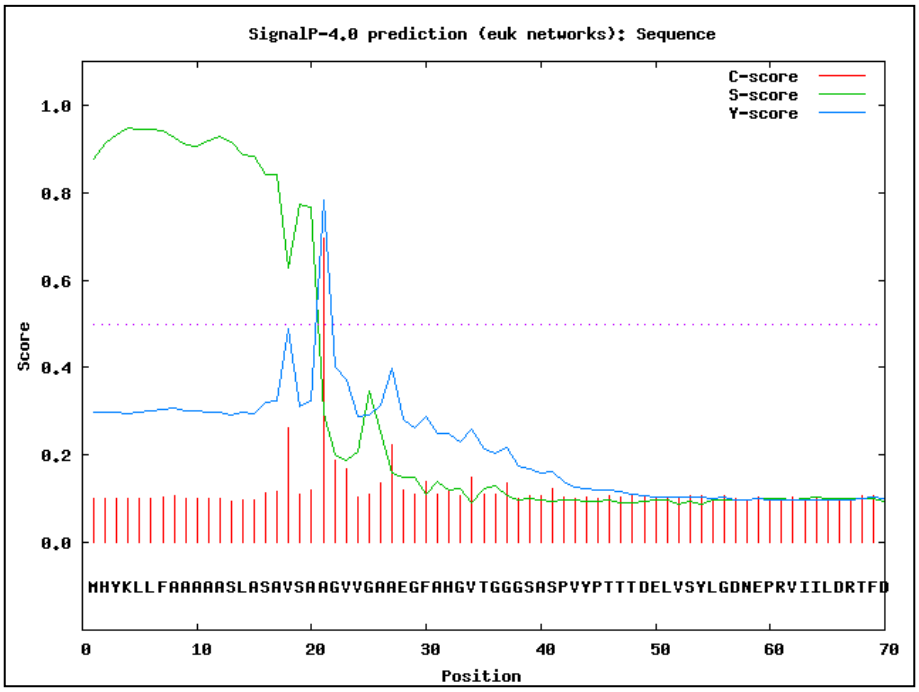

Figure 7: C-score PelB.

Measure Position Value: Cut-off signal peptide?

max. C	21	0.697	
max. Y	21	0.783	
max. S	4	0.947	
mean S	1-20	0.881	
D	1-20	0.836	0.450 YES

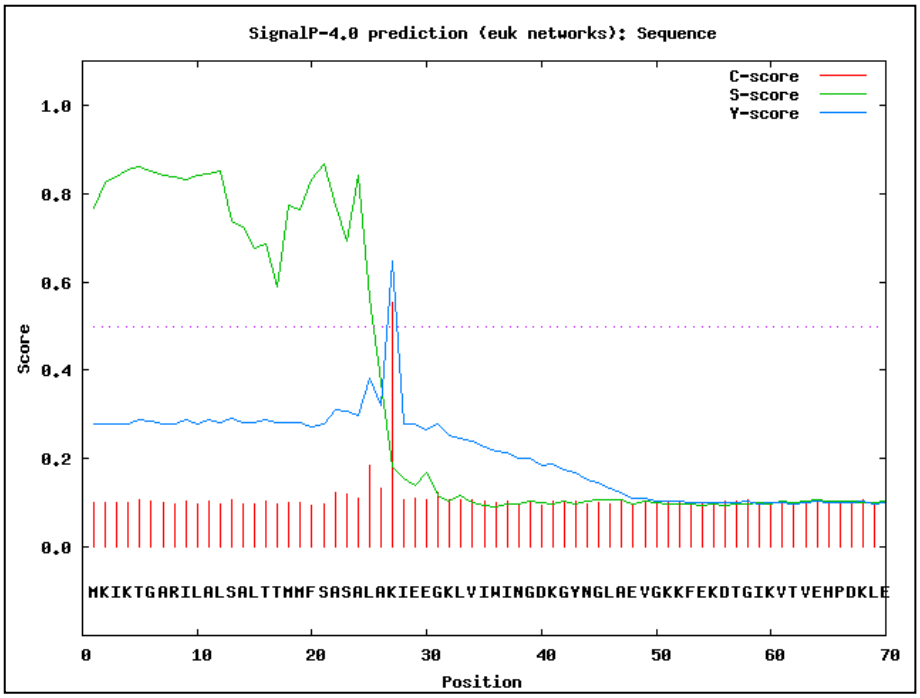

Figure 8: C-score MalE.

Measure Position Value: Cut-off signal peptide?
max. C 27 0.554
max. Y 27 0.648
max. S 21 0.867
mean S 1-26 0.767
D 1-26 0.712 0.450 YES

Name=Sequence SP='YES' Cleavage site between pos. 26 and 27: ALA-KI D=0.712 D-cutoff=0.450 Networks=SignalP-noTM

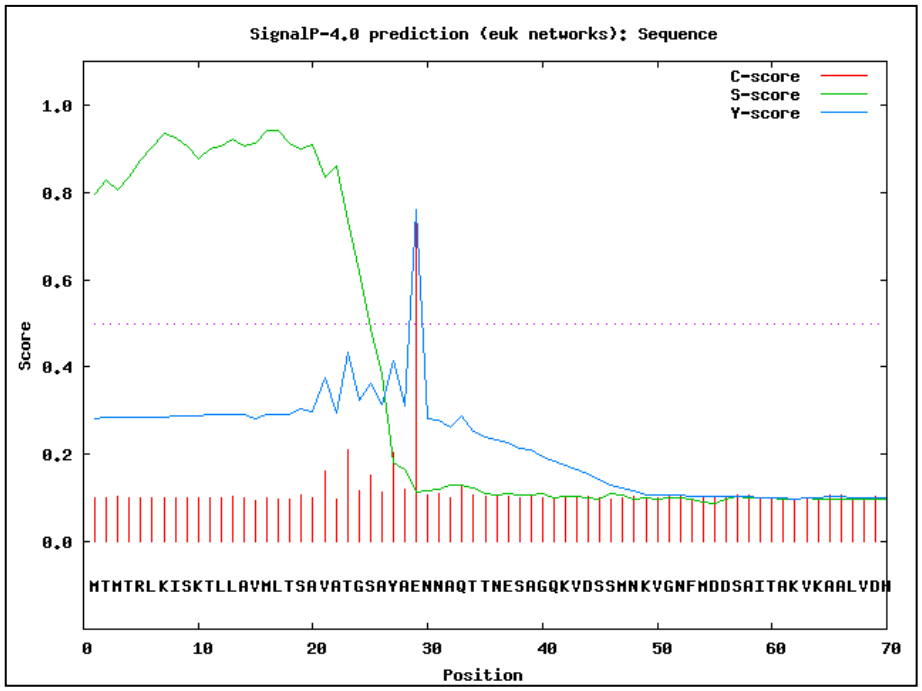

Figure 9: C-score OsmY.

Measure Position Value: Cut-off signal peptide?
max. C 29 0.747
max. Y 29 0.761
max. S 16 0.942
mean S 1-28 0.789
D 1-28 0.776 0.450 YES

Name=Sequence SP='YES' Cleavage site between pos. 28 and 29: AYA-EN D=0.776 D-cutoff=0.450 Networks=SignalP-noTM

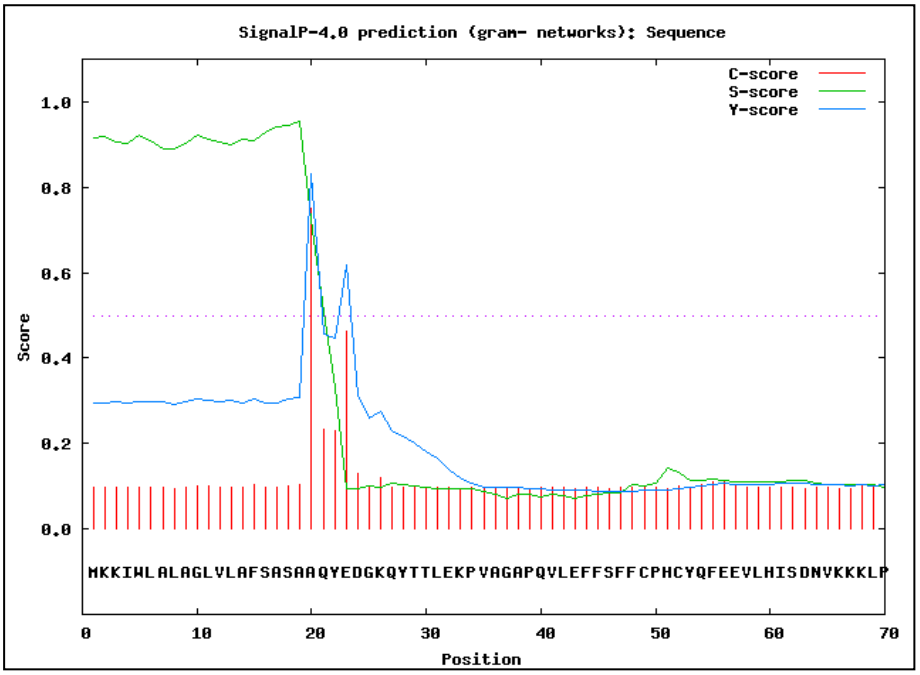

Figure 10: C-score DsbA.

Measure Position Value: Cut-off signal peptide?
max. C 20 0.750
max. Y 20 0.830
max. S 19 0.955
mean S 1-19 0.915
D 1-19 0.870 0.570 YES

Name=Sequence SP='YES' Cleavage site between pos. 19 and 20: ASA-AQ D=0.870 D-cutoff=0.570 Networks=SignalP-noTM

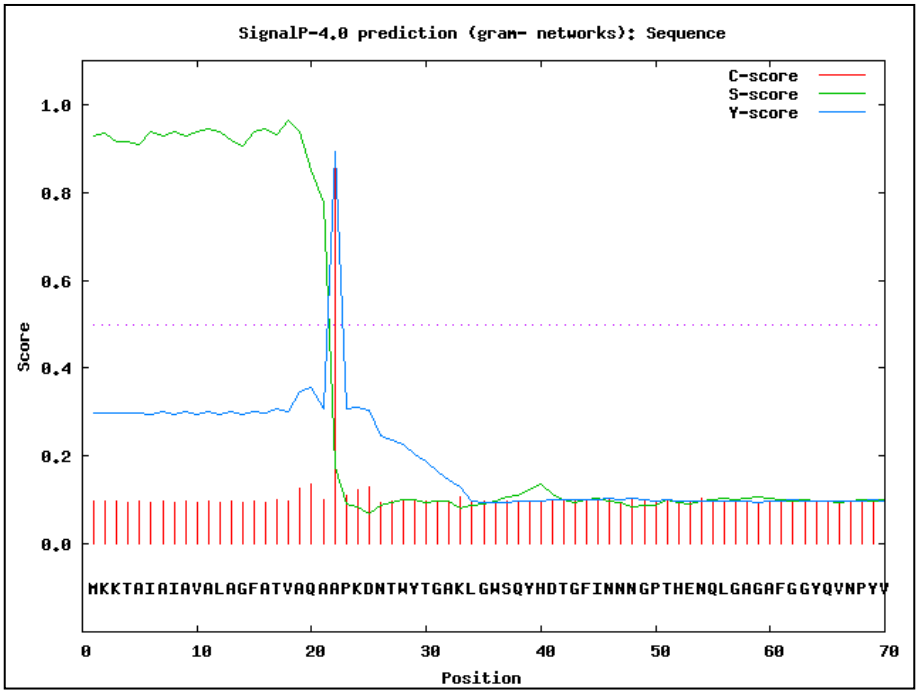

Figure 11: C-score OmpA.

Measure Position Value Cut-off signal peptide?
max. C 22 0.872
max. Y 22 0.894
max. S 18 0.963
mean S 1-21 0.920
D 1-21 0.906 0.570 YES

Name=Sequence SP='YES' Cleavage site between pos. 21 and 22: AQA-AP D=0.906 D-cutoff=0.570 Networks=SignalP-noTM
(Petersen, 2011)

SEQUENCES FROM AGOWA

PhoApJOE rev.

GCCAAGCTTACTTGTACAGGATCCTTAGCCGGCTTTGGTAACCGGGGTGAACAGCAGCGG
CAGCAGCGCCAGCGCGATGGTAGACTGTTTCATATGTATATCTCCTTCTTAAGAATTGTT
CATTACGACCAGTCTAAAAAGCGCCTGAATTCGCGACCTTCTCGTTACTGACAGGAAAAT
GGGCCATTGGCAACCAGGGAAAGATGAACGTGATGATGTTCACAATTTGCTGAATTGTGG
TGATCGCAATTGCCCTATAGTGAGTCGTATTAATTTCGTCGTATTAATTTCGATGCATGC
ATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGG
GGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGC
GCGATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATC
GCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTTACTGT
CTGCCCTCGTAGCATCGAGCATGCATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAAT
GGCAGCTTCAGTGGATTAAGTGGGGGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTAT
TCATACGGTTAAAATTTATCAGGCGCGATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCAT
TTTTACCTGTCTGCTGCCGTGATCGCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGG
GATTATGGTAAATCCACTTACTGTCTGCCCTCGTAGCCATCGAGCATGCACCATTCCTTG
CGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGC
ATAAGG

Reverse complement

CCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCG
CCGCCGCAAGGAATGGTGCATGCTCGATGGCTACGAGGGCAGACAGTAAGTGGATTTACC
ATAATCCCTTAATTGTACGCACCGCTAAAACGCGTTCAGCGCGATCACGGCAGCAGACAG
GTAAAAATGGCAACAAACCACCCGAAAAACTGCCGCGATCGCGCCTGATAAATTTTAACC
GTATGAATACCTATGCAACCAGAGGGTACAGGCCACATTACCCCCACTTAATCCACTGAA
GCTGCCATTTTTCATGGTTTCACCATCCCAGCGAAGGGCCATGCATGCTCGATGCTACGA
GGGCAGACAGTAAGTGGATTTACCATAATCCCTTAATTGTACGCACCGCTAAAACGCGTT
CAGCGCGATCACGGCAGCAGACAGGTAAAAATGGCAACAAACCACCCGAAAAACTGCCGC
GATCGCGCCTGATAAATTTTAACCGTATGAATACCTATGCAACCAGAGGGTACAGGCCAC
ATTACCCCCACTTAATCCACTGAAGCTGCCATTTTTCATGGTTTCACCATCCCAGCGAAG
GGCCATGCATGCATCGAAATTAATACGACGAAATTAATACGACTCACTATAGGGCAATTG
CGATCACCACAATTCAGCAAATTGTGAACATCATCACGTTCATCTTTCCCTGGTTGCCAA
TGGCCCATTTTCCTGTCAGTAACGAGAAGGTCGCGAATTCAGGCGCTTTTTAGACTGGTC
GTAATGAACAATTCTTAAGAAGGAGATATACATATGAAACAGTCTACCATCGCGCTGGCG
CTGCTGCCGCTGCTGTTCACCCCGGTTACCAAAGCCGGCTAAGGATCCTGTACAAGTAAG
CTTGGC

PhoAThionin 2.1pJOE for.

TACATATGAAACAGTCTACCATCGCGCTGGCGCTGCTGCCGCTGCTGTTCACCCCGGTTA
CCAAAGCCAAAATCTGCTGCCCTTCCAACCAAGCTAGAAATGGCTATAGTGTATGCCGTA
TAAGGTTTTCCAAGGGAAGGTGTATGCAAGTGAGTGGATGCCAAAACTCTGATACATGCC
CTCGAGGTTGGGTAAACGCCATTCTCGAAAACTCAGCTGATGCTACCAATGAGCACTGCA
AGTTAGGGTGTGAAACTTCTGTGTGTGGTGCCATGAACACTCTCCAGAACTCTGATGCAA
GTGAAATTGTGAATGGAGCGTCGGAACAATGTGCCAAGGGATGTTCTATTTTCTGTACCA
AGAGCTATGTAGTCCCACCTGGGCCGCCTAAACTGTTGGGTTGGTCTCACCCGCAGTTCG
AAAAATAAGGATCCTGTACAAGTAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCA
GCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCG
GCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCG
CCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAA
CGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCT
CTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGA
GGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCAT
CCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTGTTTATTTTTCTAAATACATT
CAATATGTATCCGCTCATGAGACATAACCCTGATAAATGCTTCAATAATATTGAAAAAGG
AGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTT

PhoAVisco A3 for.

ACATATGAAACAGTCTACCATCGCGCTGGCGCTGCTGCCGCTGCTGTTCACCCCGGTTAC
CAAAGCCAAAAGCTGCTGCCCCAACACCACCGGACGAAACATCTACAACGCCTGCCGCTT
AACCGGCGCTCCCCGCCCAACCTGTGCCAAACTGAGCGGTTGCAAAATCATCTCTGGATC
AACATGTCCATCCGATTATCCCAAATAGGATCCTGTACAAGTAAGCTTGGCTGTTTTGGC
GGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATA
AAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCA
GAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAAC
TGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTG
TTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGT
TGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCA
AATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTGTT
TATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGC
TTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTC
CCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA
AAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCG
GTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAA

PelbpJOE for.

TACCTGCTGCCGACTGCGGCGGCGGGTCTGCTGCTGCTGGCGGCGCAACCGGCGATGGCC
GGCTAAGGATCCTGTACAAGTAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGC
CTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGC
AGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCC
GATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACG
AAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCT
CCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGG
GTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCT
GACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTGTTTATTTTTCTAAATACATTCAA
ATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGA
AGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCC
TTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGG
GTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTC
GCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTTAAAGTTCTGCTATGTGGCGCGGTA

PelB Thionin2.1 pJOE rev.

GCCAAGCTTACTTGTACAGGATCCTTATTTTTCGAACTGCGGGTGAGACCAACCCAACAG
TTTAGGCGGCCCAGGTGGGACTACATAGCTCTTGGTACAGAAAATAGAACATCCCTTGGC
ACATTGTTCCGACGCTCCATTCACAATTTCACTTGCATCAGAGTTCTGGAGAGTGTTCAT
GGCACCACACACAGAAGTTTCACACCCTAACTTGCAGTGCTCATTGGTAGCATCAGCTGA
GTTTTCGAGAATGGCGTTTACCCAACCTCGAGGGCATGTATCAGAGTTTTGGCATCCACT
CACTTGCATACACCTTCCCTTGGAAAACCTTATACGGCATACACTATAGCCATTTCTAGC
TTGGTTGGAAGGGCAGCAGATTTTGGCCATCGCCGGTTGCGCCGCCAGCAGCAGCAGACC
CGCCGCCGCAGTCGGCAGCAGGTATTTCATATGTATATCTCCTTCTTAAGAATTGTTCAT
TACGACCAGTCTAAAAAGCGCCTGAATTCGCGACCTTCTCGTTACTGACAGGAAAATGGG
CCATTGGCAACCAGGGAAAGATGAACGTGATGATGTTCACAATTTGCTGAATTGTGGTGA
TCGCAATTGCCCTATAGTGAGTCGTATTAATTTCGTCGTATTAATTTCGATGCATGCATG
GCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGT
AATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGCGCG
ATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATCGCG
CTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTTACTGTCTG
CCCTCGTAGCATCGAGCATGCATGGCCCTTTCGCTGGG

Reverse complement

GCCAAGCTTACTTGTACAGGATCCTTATTTTTCGAACTGCGGGTGAGACCAACCCAACAG
TTTAGGCGGCCCAGGTGGGACTACATAGCTCTTGGTACAGAAAATAGAACATCCCTTGGC
ACATTGTTCCGACGCTCCATTCACAATTTCACTTGCATCAGAGTTCTGGAGAGTGTTCAT
GGCACCACACACAGAAGTTTCACACCCTAACTTGCAGTGCTCATTGGTAGCATCAGCTGA
GTTTTCGAGAATGGCGTTTACCCAACCTCGAGGGCATGTATCAGAGTTTTGGCATCCACT
CACTTGCATACACCTTCCCTTGGAAAACCTTATACGGCATACACTATAGCCATTTCTAGC
TTGGTTGGAAGGGCAGCAGATTTTGGCCATCGCCGGTTGCGCCGCCAGCAGCAGCAGACC
CGCCGCCGCAGTCGGCAGCAGGTATTTCATATGTATATCTCCTTCTTAAGAATTGTTCAT
TACGACCAGTCTAAAAAGCGCCTGAATTCGCGACCTTCTCGTTACTGACAGGAAAATGGG
CCATTGGCAACCAGGGAAAGATGAACGTGATGATGTTCACAATTTGCTGAATTGTGGTGA
TCGCAATTGCCCTATAGTGAGTCGTATTAATTTCGTCGTATTAATTTCGATGCATGCATG
GCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGT
AATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGCGCG
ATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATCGCG
CTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTTACTGTCTG
CCCTCGTAGCATCGAGCATGCATGGCCCTTTCGCTGGG

PelbVisco A3 for.

TGCTGCCGACTGCGGCGGCGGGTCTGCTGCTGCTGGCGGCGCAACCGGCGATGGCCAAAA
GCTGCTGCCCCAACACCACCGGACGAAACATCTACAACGCCTGCCGCTTAACCGGCGCTC
CCCGCCCAACCTGTGCCAAACTGAGCGGTTGCAAAATCATCTCTGGATCAACATGTCCAT
CCGATTATCCCAAATAGGATCCTGTACAAGTAAGCTTGGCTGTTTTGGCGGATGAGAGAA
GATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTT
GCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACG
CCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATC
AAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGG
TGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAAC
GGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGA
AGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTGTTTATTTTTCTAA
ATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATAT
TGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCG
GCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAA
GATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTT
GAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAA

MalEpJOE rev.

GCCAAGCTTACTTGTACAGGATCCTTAGCCGGCCAGCGCAGACGCAGAGAACATCGCCAG
GATCAGCGCGCCAGTTTTGATTTTCATATGTATATCTCCTTCTTAAGAATTGTTCATTAC
GACCAGTCTAAAAAGCGCCTGAATTCGCGACCTTCTCGTTACTGACAGGAAAATGGGCCA
TTGGCAACCAGGGAAAGATGAACGTGATGATGTTCACAATTTGCTGAATTGTGGTGATCG
CAATTGCCCTATAGTGAGTCGTATTAATTTCGTCGTATTAATTTCGATGCATGCATGGCC
CTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGTAAT
GTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGCGCGATC
GCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATCGCGCTG
AACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTTACTGTCTGCCC
TCGTAGCATCGAGCATGCATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGC
TTCAGTGGATTAAGTGGGGGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATAC
GGTTAAAATTTATCAGGCGCGATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTAC
CTGTCTGCTGCCGTGATCGCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTAT
GGTAAATCCACTTACTGTCTGCCCTCGTAGCCATCGAGCATGCACCATTCCTTGCGGCGG
CGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGG
GAGAGCGTCGACCGATGCC

Reverse Complement

GGCATCGGTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGG
TTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCTCGATGGCTACGAGGGCA
GACAGTAAGTGGATTTACCATAATCCCTTAATTGTACGCACCGCTAAAACGCGTTCAGCG
CGATCACGGCAGCAGACAGGTAAAAATGGCAACAAACCACCCGAAAAACTGCCGCGATCG
CGCCTGATAAATTTTAACCGTATGAATACCTATGCAACCAGAGGGTACAGGCCACATTAC
CCCCACTTAATCCACTGAAGCTGCCATTTTTCATGGTTTCACCATCCCAGCGAAGGGCCA
TGCATGCTCGATGCTACGAGGGCAGACAGTAAGTGGATTTACCATAATCCCTTAATTGTA
CGCACCGCTAAAACGCGTTCAGCGCGATCACGGCAGCAGACAGGTAAAAATGGCAACAAA
CCACCCGAAAAACTGCCGCGATCGCGCCTGATAAATTTTAACCGTATGAATACCTATGCA
ACCAGAGGGTACAGGCCACATTACCCCCACTTAATCCACTGAAGCTGCCATTTTTCATGG
TTTCACCATCCCAGCGAAGGGCCATGCATGCATCGAAATTAATACGACGAAATTAATACG
ACTCACTATAGGGCAATTGCGATCACCACAATTCAGCAAATTGTGAACATCATCACGTTC
ATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCAGTAACGAGAAGGTCGCGAATTCA
GGCGCTTTTTAGACTGGTCGTAATGAACAATTCTTAAGAAGGAGATATACATATGAAAAT
CAAAACTGGCGCGCTGATCCTGGCGATGTTCTCTGCGTCTGCGCTGGCCGGCTAAGGATC
CTGTACAAGTAAGCTTGGC

MalEThionin 2.1 pJOErev.

CGCCAAACAGCCAAGCTTACTTGTACAGGATCCTTATTTTTCGAACTGCGGGTAGACCAA
CCCAACAGTTTAGGCGGCCCAGGTGGGACTACCTAGCTCTTGGTACAGAAAATAGAACAT
CCCTTGGCACATTGTTCCGACGCTCCATTCACAATTTCACTTGCATCAGAGTTCTGGAGA
GTGTTCATGGCACCACACACAGAAGTTTCACACCCTAACTTGCAGTGCTCATTGGTAGCA
TCAGCTGAGTTTTCGAGAATGGCGTTTACCCAACCTCGAGGGCATGTATCAGAGTTTTGG
CATCCACTCACTTGCATACACCTTCCCTTGGAAAACCTTATACGGCATACACTATAGCCA
TTTCTAGCTTGGTTGGAAGGGCAGCAGATTTTGGCCAGCGCAGACGCAGAGAACATCGCC
AGGATCAGCGCGCCAGTTTTGATTTTCATATGTATATCTCCTTCTTAAGAATTGTTCATT
ACGACCAGTCTAAAAAGCGCCTGAATTCGCGACCTTCTCGTTACTGACAGGAAAATGGGC
CATTGGCAACCAGGGAAAGATGAACGTGATGATGTTCACAATTTGCTGAATTGTGGTGAT
CGCAATTGCCCTATAGTGAGTCGTATTAATTTCGTCGTATTAATTTCGATGCATGCATGG
CCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGTA
ATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAAGCGCGA
TCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATCGCGC
TGAACGCGTTTTAGCGGTGCGTACAATT

Reverse complement

AATTGTACGCACCGCTAAAACGCGTTCAGCGCGATCACGGCAGCAGACAGGTAAAAATGG
CAACAAACCACCCGAAAAACTGCCGCGATCGCGCTTGATAAATTTTAACCGTATGAATAC
CTATGCAACCAGAGGGTACAGGCCACATTACCCCCACTTAATCCACTGAAGCTGCCATTT
TTCATGGTTTCACCATCCCAGCGAAGGGCCATGCATGCATCGAAATTAATACGACGAAAT
TAATACGACTCACTATAGGGCAATTGCGATCACCACAATTCAGCAAATTGTGAACATCATC
ACGTTCATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCAGTAACGAGAAGGTCGCGA
ATTCAGGCGCTTTTTAGACTGGTCGTAATGAACAATTCTTAAGAAGGAGATATACATATGA
AAATCAAAACTGGCGCGCTGATCCTGGCGATGTTCTCTGCGTCTGCGCTGGCCAAAATCTG
CTGCCCTTCCAACCAAGCTAGAAATGGCTATAGTGTATGCCGTATAAGGTTTTCCAAGGGA
AGGTGTATGCAAGTGAGTGGATGCCAAAACTCTGATACATGCCCTCGAGGTTGGGTAAACG
CCATTCTCGAAAACTCAGCTGATGCTACCAATGAGCACTGCAAGTTAGGGTGTGAAACTTC
TGTGTGTGGTGCCATGAACACTCTCCAGAACTCTGATGCAAGTGAAATTGTGAATGGAGCG
TCGGAACAATGTGCCAAGGGATGTTCTATTTTCTGTACCAAGAGCTAGGTAGTCCCACCTG
GGCCGCCTAAACTGTTGGGTTGGTCTACCCGCAGTTCGAAAAATAAGGATCCTGTACAAGT
AAGCTTGGCTGTTTGGCG

MalEVisco A3pJOErev.

CCAAGCTTACTTGTACAGGATCCTATTTGGGATAATCGGATGGACATGTTGATCCAGAGA
TGATTTTGCAACCGCTCAGTTTGGCACAGGTTGGGCGGGGAGCGCCGGTTAAGCGGCAGG
CGTTGTAGATGTTTCGTCCGGTGGTGTTGGGGCAGCAGCTTTTGGCCAGCGCAGACGCAG
AGAACATCGCCAGGATCAGCGCGCCAGTTTTGATTTTCATATGTATATCTCCTTCTTAAG
AATTGTTCATTACGACCAGTCTAAAAAGCGCCTGAATTCGCGACCTTCTCGTTACTGACA
GGAAAATGGGCCATTGGCAACCAGGGAAAGATGAACGTGATGATGTTCACAATTTGCTGA
ATTGTGGTGATCGCAATTGCCCTATAGTGAGTCGTATTAATTTCGTCGTATTAATTTCGA
TGCATGCATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATT
AAGTGGGGGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTT
ATCAGGCGCGATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGC
CGTGATCGCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCAC
TTACTGTCTGCCCTCGTAGCATCGAGCATGCATGGCCCTTCGCTGGGATGGTGAAACCAT
GAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGTAATGTGGCCTGTACCCTCTGGTTGCA
TAGGTATTCATACGGTTAAAATTTATCAGGCGCGATCGCGGCAGTTTTTCGGGTGGTTTG
TTGCCATTTTTACCTGTCTGCTGCCGTGATCGCGCTGAACGCGTTTTAGCGGTGCGTACA
ATTAAGGGATTATGGTAAATCCACTTACTGTCTGCCCTCGTAGCCATCGAGCATG

Reverse Complement MalEVisco A3

CATGCTCGATGGCTACGAGGGCAGACAGTAAGTGGATTTACCATAATCCCTTAATTGTAC
GCACCGCTAAAACGCGTTCAGCGCGATCACGGCAGCAGACAGGTAAAAATGGCAACAAAC
CACCCGAAAAACTGCCGCGATCGCGCCTGATAAATTTTAACCGTATGAATACCTATGCAA
CCAGAGGGTACAGGCCACATTACCCCCACTTAATCCACTGAAGCTGCCATTTTTCATGGT
TTCACCATCCCAGCGAAGGGCCATGCATGCTCGATGCTACGAGGGCAGACAGTAAGTGGA
TTTACCATAATCCCTTAATTGTACGCACCGCTAAAACGCGTTCAGCGCGATCACGGCAGC
AGACAGGTAAAAATGGCAACAAACCACCCGAAAAACTGCCGCGATCGCGCCTGATAAATT
TTAACCGTATGAATACCTATGCAACCAGAGGGTACAGGCCACATTACCCCCACTTAATCC
ACTGAAGCTGCCATTTTTCATGGTTTCACCATCCCAGCGAAGGGCCATGCATGCATCGAA
ATTAATACGACGAAATTAATACGACTCACTATAGGGCAATTGCGATCACCACAATTCAGC
AAATTGTGAACATCATCACGTTCATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCA
GTAACGAGAAGGTCGCGAATTCAGGCGCTTTTTAGACTGGTCGTAATGAACAATTCTTAA
GAAGGAGATATACATATGAAAATCAAAACTGGCGCGCTGATCCTGGCGATGTTCTCTGCG
TCTGCGCTGGCCAAAAGCTGCTGCCCCAACACCACCGGACGAAACATCTACAACGCCTGC
CGCTTAACCGGCGCTCCCCGCCCAACCTGTGCCAAACTGAGCGGTTGCAAAATCATCTCT
GGATCAACATGTCCATCCGATTATCCCAAATAGGATCCTGTACAAGTAAGCTTGG

OsmYpJOErev.

GCCAAGCTTACTTGTACAGGATCCTTAGCCGGCGTACGCAGAACCCGTCGCAACCGCAGA
GGTCAGCATAACCGCCAGCAGGGTTTTAGAGATTTTCAGACGGGTCATGGTCATATGTAT
ATCTCCTTCTTAAGAATTGTTCATTACGACCAGTCTAAAAAGCGCCTGAATTCGCGACCT
TCTCGTTACTGACAGGAAAATGGGCCATTGGCAACCAGGGAAAGATGAACGTGATGATGT
TCACAATTTGCTGAATTGTGGTGATCGCAATTGCCCTATAGTGAGTCGTATTAATTTCGT
CGTATTAATTTCGATGCATGCATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGC
AGCTTCAGTGGATTAAGTGGGGGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCA
TACGGTTAAAATTTATCAGGCGCGATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTT
TACCTGTCTGCTGCCGTGATCGCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGAT
TATGGTAAATCCACTTACTGTCTGCCCTCGTAGCATCGAGCATGCATGGCCCTTCGCTGG
GATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGTAATGTGGCCTGT
ACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGCGCGATCGCGGCAGTT
TTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATCGCGCTGAACGCGTTT
TAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTTACTGTCTGCCCTCGTAGCCA
TCGAGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCT
GCTTCCTAATGCAGGA

Reverse complement OsmY

TCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGG
AATGGTGCATGCTCGATGGCTACGAGGGCAGACAGTAAGTGGATTTACCATAATCCCTTA
ATTGTACGCACCGCTAAAACGCGTTCAGCGCGATCACGGCAGCAGACAGGTAAAAATGGC
AACAAACCACCCGAAAAACTGCCGCGATCGCGCCTGATAAATTTTAACCGTATGAATACC
TATGCAACCAGAGGGTACAGGCCACATTACCCCCACTTAATCCACTGAAGCTGCCATTTT
TCATGGTTTCACCATCCCAGCGAAGGGCCATGCATGCTCGATGCTACGAGGGCAGACAGT
AAGTGGATTTACCATAATCCCTTAATTGTACGCACCGCTAAAACGCGTTCAGCGCGATCA
CGGCAGCAGACAGGTAAAAATGGCAACAAACCACCCGAAAAACTGCCGCGATCGCGCCTG
ATAAATTTTAACCGTATGAATACCTATGCAACCAGAGGGTACAGGCCACATTACCCCCAC
TTAATCCACTGAAGCTGCCATTTTTCATGGTTTCACCATCCCAGCGAAGGGCCATGCATG
CATCGAAATTAATACGACGAAATTAATACGACTCACTATAGGGCAATTGCGATCACCACA
ATTCAGCAAATTGTGAACATCATCACGTTCATCTTTCCCTGGTTGCCAATGGCCCATTTT
CCTGTCAGTAACGAGAAGGTCGCGAATTCAGGCGCTTTTTAGACTGGTCGTAATGAACAA
TTCTTAAGAAGGAGATATACATATGACCATGACCCGTCTGAAAATCTCTAAAACCCTGCT
GGCGGTTATGCTGACCTCTGCGGTTGCGACGGGTTCTGCGTACGCCGGCTAAGGATCCTG
TACAAGTAAGCTTGGC

OsmYThionin 2.1 pJOE for.

ACATATGACCATGACCCGTCTGAAAATCTCTAAAACCCTGCTGGCGGTTATGCTGACCTC
TGCGGTTGCGACGGGTTCTGCGTACGCAAAATCTGCTGCCCTTCCAACCAAGCTAGAAAT
GGCTATAGTGTATGCCGTATAAGGTTTTCCAAGGGAAGGTGTATGCAAGTGAGTGGATGC
CAAAACTCTGATACATGCCCTCGAGGTTGGGTAAACGCCATTCTCGAAAACTCAGCTGAT
GCTACCAATGAGCACTGCAAGTTAGGGTGTGAAACTTCTGTGTGTGGTGCCATGAACACT
CTCCAGAACTCTGATGCAAGTGAAATTGTGAATGGAGCGTCGGAACAATGTGCCAAGGGA
TGTTCTATTTTCTGTACCAAGAGCTATGTAGTCCCACCTGGGCCGCCTAAACTGTTGGGT
TGGTCTACCCGCAGTTCGAAAAATAAGGATCCTGTACAAGTAAGCTTGGCTGTTTTGGCG
GATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAA
AACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAG
AAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACT
GCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGT
TGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTT
GCGAAGCAACGGCCCGGAGGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCA
AATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAACTCTTTTGTTT
ATTTTTCTAAATACATTCAATATGTATCCGCTCATGAGAC

OsmYVisco A3 for.

ACATATGACCATGACCCGTCTGAAAATCTCTAAAACCCTGCTGGCGGTTATGCTGACCTC
TGCGGTTGCGACGGGTTCTGCGTACGCCAAAAGCTGCTGCCCCAACACCACCGGACGAAA
CATCTACAACGCCTGCCGCTTAACCGGCGCTCCCCGCCCAACCTGTGCCAAACTGAGCGG
TTGCAAAATCATCTCTGGATCAACATGTCCATCCGATTATCCCAAATAGGATCCTGTACA
AGTAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCA
GAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCA
CCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCT
CCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGA
CTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCC
GCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCC
GCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGC
GTTTCTACAAACTCTTTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGA
GACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAAC
ATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACC
CAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACA
TCGAACTGGATCTCA

DsbaThionin 2.1 pJOE rev.

AACAGTTTAGGCGGCCCAGGTGGGACTACATAGCTCTTGGTACAGAAAATAGAACATCCC
TTGGCACATTGTTCCGACGCTCCATTCACAATTTCACTTGCATCAGAGTTCTGGAGAGTG
TTCATGGCACCACACACAGAAGTTTCACACCCTAACTTGCAGTGCTCATTGGTAGCATCA
GCTGAGTTTTCGAGAATGGCGTTTACCCAACCTCGAGGGCATGTATCAGAGTTTTGGCAT
CCACTCACTTGCATACACCTTCCCTTGGAAAACCTTATACGGCATACACTATAGCCATTT
CTAGCTTGGTTGGAAGGGCAGCAGATTTTGGCAGACGCAGAGAACGCCAGAACCAGACCC
GCCAGCGCCAGCCAGATTTTTTTCATATGTATATCTCCTTCTTAAGAATTGTTCATTACG
ACCAGTCTAAAAAGCGCCTGAATTCGCGACCTTCTCGTTACTGACAGGAAAATGGGCCAT
TGGCAACCAGGGAAAGATGAACGTGATGATGTTCACAATTTGCTGAATTGTGGTGATCGC
AATTGCCCTATAGTGAGTCGTATTAATTTCGTCGTATTAATTTCGATGCATGCATGGCCC
TTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGTAATG
TGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGCGCGATCG
CGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATCGCGCTGA
ACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTTACTGTCTGCCCT
CGTAGCATCGAGCATGCATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCT
TCAGTGGATTAAGTGGGGGT

Reverse Complement

ACCCCCACTTAATCCACTGAAGCTGCCATTTTTCATGGTTTCACCATCCCAGCGAAGGGC
CATGCATGCTCGATGCTACGAGGGCAGACAGTAAGTGGATTTACCATAATCCCTTAATTG
TACGCACCGCTAAAACGCGTTCAGCGCGATCACGGCAGCAGACAGGTAAAAATGGCAACA
AACCACCCGAAAAACTGCCGCGATCGCGCCTGATAAATTTTAACCGTATGAATACCTATG
CAACCAGAGGGTACAGGCCACATTACCCCCACTTAATCCACTGAAGCTGCCATTTTTCAT
GGTTTCACCATCCCAGCGAAGGGCCATGCATGCATCGAAATTAATACGACGAAATTAATA
CGACTCACTATAGGGCAATTGCGATCACCACAATTCAGCAAATTGTGAACATCATCACGT
TCATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCAGTAACGAGAAGGTCGCGAATT
CAGGCGCTTTTTAGACTGGTCGTAATGAACAATTCTTAAGAAGGAGATATACATATGAAA
AAAATCTGGCTGGCGCTGGCGGGTCTGGTTCTGGCGTTCTCTGCGTCTGCCAAAATCTGC
TGCCCTTCCAACCAAGCTAGAAATGGCTATAGTGTATGCCGTATAAGGTTTTCCAAGGGA
AGGTGTATGCAAGTGAGTGGATGCCAAAACTCTGATACATGCCCTCGAGGTTGGGTAAAC
GCCATTCTCGAAAACTCAGCTGATGCTACCAATGAGCACTGCAAGTTAGGGTGTGAAACT
TCTGTGTGTGGTGCCATGAACACTCTCCAGAACTCTGATGCAAGTGAAATTGTGAATGGA
GCGTCGGAACAATGTGCCAAGGGATGTTCTATTTTCTGTACCAAGAGCTATGTAGTCCCA
CCTGGGCCGCCTAAACTGTT

DsbAVisco A3 pJOE rev.

TTGGGATAATCGGATGGACATGTTGATCCAGAGATGATTTTGCAACCGCTCAGTTTGGCA
CAGGTTGGGCGGGGAGCGCCGGTTAAGCGGCAGGCGTTGTAGATGTTTCGTCCGGTGGTG
TTGGGGCAGCAGCTTTTGGCAGACGCAGAGAACGCCAGAACCAGACCCGCCAGCGCCAGC
CAGATTTTTTTCATATGTATATCTCCTTCTTAAGAATTGTTCATTACGACCAGTCTAAAA
AGCGCCTGAATTCGCGACCTTCTCGTTACTGACAGGAAAATGGGCCATTGGCAACCAGGG
AAAGATGAACGTGATGATGTTCACAATTTGCTGAATTGTGGTGATCGCAATTGCCCTATA
GTGAGTCGTATTAATTTCGTCGTATTAATTTCGATGCATGCATGGCCCTTCGCTGGGATG
GTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGTAATGTGGCCTGTACCC
TCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGCGCGATCGCGGCAGTTTTTC
GGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATCGCGCTGAACGCGTTTTAGC
GGTGCGTACAATTAAGGGATTATGGTAAATCCACTTACTGTCTGCCCTCGTAGCATCGAG
CATGCATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAA
GTGGGGGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTAT
CAGGCGCGATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCG
TGATCGCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTT
ACTGTCTGCC

Reverse Complement

GGCAGACAGTAAGTGGATTTACCATAATCCCTTAATTGTACGCACCGCTAAAACGCGTTC
AGCGCGATCACGGCAGCAGACAGGTAAAAATGGCAACAAACCACCCGAAAAACTGCCGCG
ATCGCGCCTGATAAATTTTAACCGTATGAATACCTATGCAACCAGAGGGTACAGGCCACA
TTACCCCCACTTAATCCACTGAAGCTGCCATTTTTCATGGTTTCACCATCCCAGCGAAGG
GCCATGCATGCTCGATGCTACGAGGGCAGACAGTAAGTGGATTTACCATAATCCCTTAAT
TGTACGCACCGCTAAAACGCGTTCAGCGCGATCACGGCAGCAGACAGGTAAAAATGGCAA
CAAACCACCCGAAAAACTGCCGCGATCGCGCCTGATAAATTTTAACCGTATGAATACCTA
TGCAACCAGAGGGTACAGGCCACATTACCCCCACTTAATCCACTGAAGCTGCCATTTTTC
ATGGTTTCACCATCCCAGCGAAGGGCCATGCATGCATCGAAATTAATACGACGAAATTAA
TACGACTCACTATAGGGCAATTGCGATCACCACAATTCAGCAAATTGTGAACATCATCAC
GTTCATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCAGTAACGAGAAGGTCGCGAA
TTCAGGCGCTTTTTAGACTGGTCGTAATGAACAATTCTTAAGAAGGAGATATACATATGA
AAAAAATCTGGCTGGCGCTGGCGGGTCTGGTTCTGGCGTTCTCTGCGTCTGCCAAAAGCT
GCTGCCCCAACACCACCGGACGAAACATCTACAACGCCTGCCGCTTAACCGGCGCTCCCC
GCCCAACCTGTGCCAAACTGAGCGGTTGCAAAATCATCTCTGGATCAACATGTCCATCCG
ATTATCCCAA

OmpAThionin 2.1pJOE rev.

GCCAAGCTTACTTGTACAGGATCCTTAGCCGGCCTGCGCAACGGTGGCGAAACCCGCCAG
CGCAACCGCGATCGCGATCGCGGTTTTTTTCATATGTATATCTCCTTCTTAAGAATTGTT
CATTACGACCAGTCTAAAAAGCGCCTGAATTCGCGACCTTCTCGTTACTGACAGGAAAAT
GGGCCATTGGCAACCAGGGAAAGATGAACGTGATGATGTTCACAATTTGCTGAATTGTGG
TGATCGCAATTGCCCTATAGTGAGTCGTATTAATTTCGTCGTATTAATTTCGATGCATGC
ATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGG
GGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAAAATTTATCAGGC
GCGATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATC
GCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTTACTGT
CTGCCCTCGTAGCATCGAGCATGCATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAAT
GGCAGCTTCAGTGGATTAAGTGGGGGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTAT
TCATACGGTTAAAATTTATCAGGCGCGATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCAT
TTTTACCTGTCTGCTGCCGTGATCGCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGG
GATTATGGTAAATCCACTTACTGTCTGCCCTCGTAGCCATCGAGCATGCACCATTCCTTG
CGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGC
ATAAGGGAGAGCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGT
GGGCGCGGGCATGACTATCGTC

Reverse Complement

GACGATAGTCATGCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGC
ATCGGTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTG
AGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCTCGATGGCTACGAGGGCAGAC
AGTAAGTGGATTTACCATAATCCCTTAATTGTACGCACCGCTAAAACGCGTTCAGCGCGA
TCACGGCAGCAGACAGGTAAAAATGGCAACAAACCACCCGAAAAACTGCCGCGATCGCGC
CTGATAAATTTTAACCGTATGAATACCTATGCAACCAGAGGGTACAGGCCACATTACCCC
CACTTAATCCACTGAAGCTGCCATTTTTCATGGTTTCACCATCCCAGCGAAGGGCCATGC
ATGCTCGATGCTACGAGGGCAGACAGTAAGTGGATTTACCATAATCCCTTAATTGTACGC
ACCGCTAAAACGCGTTCAGCGCGATCACGGCAGCAGACAGGTAAAAATGGCAACAAACCA
CCCGAAAAACTGCCGCGATCGCGCCTGATAAATTTTAACCGTATGAATACCTATGCAACC
AGAGGGTACAGGCCACATTACCCCCACTTAATCCACTGAAGCTGCCATTTTTCATGGTTT
CACCATCCCAGCGAAGGGCCATGCATGCATCGAAATTAATACGACGAAATTAATACGACT
CACTATAGGGCAATTGCGATCACCACAATTCAGCAAATTGTGAACATCATCACGTTCATC
TTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCAGTAACGAGAAGGTCGCGAATTCAGGC
GCTTTTTAGACTGGTCGTAATGAACAATTCTTAAGAAGGAGATATACATATGAAAAAAAC
CGCGATCGCGATCGCGGTTGCGCTGGCGGGTTTCGCCACCGTTGCGCAGGCCGGCTAAGG
ATCCTGTACAAGTAAGCTTGGC

OmpA without Strep-TagpJOE rev

AGGGCAGACAGTAAGTGGATTTACCATAATCCCTTTAATTGTACGCACCGCTAAAACGCG
TTCAGCGCGATCACGGCAGCAGACAGGTAAAAATGGCAACAAACCACCCGAAAAACTGC
CGCGATCGCGCCTGATAAATTTTAACCGTATGAATACCTATGCAACCAGAGGGTACAGG
CCACATTACCCCCACTTAATCCACTGAAGCTGCCATTTTTCATGGTTTCACCATCCCAG
CGAAGGGCCATGCATGCTCGATGCTACGAGGGCAGACAGTAAGTGGATTTACCATAATC
CCTTAATTGTACGCACCGCTAAAACGCGTTCAGCGCGATCACGGCAGCAGACAGGTAAA
AATGGCAACAAACCACCCGAAAAACTGCCGCGATCGCGCCTGATAAATTTTAACCGTAT
GAATACCTATGCAACCAGAGGGTACAGGCCACATTACCCCCACTTAATCCACTGAAGCT
GCCATTTTTCATGGTTTCACCATCCCAGCGAAGGGCCATGCATGCATCGAAATTAATAC
GACGAAATTAATACGACTCACTATAGGGCAATTGCGATCACCACAATTCAGCAAATTGT
GAACATCATCACGTTCATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTGTCAGTAACG
AGAAGGTCGCGAATTCAGGCGCTTTTTAGACTGGTCGTAATGAACAATTCTTAAGAAGG
AGATATACATATGAAAAAAACCGCGATCGCGATCGCGGTTGCGCTGGCGGGTTTCGCCA
CCGTTGCGCAGGCCAAAATCTGCTGCCCTTCCAACCAAGCTAGAAATGGCTATAGTGTA
TGCCGTATAAGGTTTTCCAAGGGAAGGTGTATGCAAGTGAGTGGATGCCAAAACTCTGA
TACATGCCCTCGAGGTTGGGTAAACTAAATTCTCGGATCCTGTACAAGTAAGCTTGGC

Reverse complement

GCCAAGCTTACTTGTACAGGATCCGAGAATTTAGTTTACCCAACCTCGAGGGCATGTATC
AGAGTTTTGGCATCCACTCACTTGCATACACCTTCCCTTGGAAAACCTTATACGGCATAC
ACTATAGCCATTTCTAGCTTGGTTGGAAGGGCAGCAGATTTTGGCCTGCGCAACGGTGGC
GAAACCCGCCAGCGCAACCGCGATCGCGATCGCGGTTTTTTTCATATGTATATCTCCTTC
TTAAGAATTGTTCATTACGACCAGTCTAAAAAGCGCCTGAATTCGCGACCTTCTCGTTAC
TGACAGGAAAATGGGCCATTGGCAACCAGGGAAAGATGAACGTGATGATGTTCACAATTT
GCTGAATTGTGGTGATCGCAATTGCCCTATAGTGAGTCGTATTAATTTCGTCGTATTAAT
TTCGATGCATGCATGGCCCTTCGCTGGGATGGTGAAACCATGAAAAATGGCAGCTTCAGT
GGATTAAGTGGGGGTAATGTGGCCTGTACCCTCTGGTTGCATAGGTATTCATACGGTTAA
AATTTATCAGGCGCGATCGCGGCAGTTTTTCGGGTGGTTTGTTGCCATTTTTACCTGTCT
GCTGCCGTGATCGCGCTGAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAA
TCCACTTACTGTCTGCCCTCGTAGCATCGAGCATGCATGGCCCTTCGCTGGGATGGTGAA
ACCATGAAAAATGGCAGCTTCAGTGGATTAAGTGGGGGTAATGTGGCCTGTACCCTCTGG
TTGCATAGGTATTCATACGGTTAAAATTTATCAGGCGCGATCGCGGCAGTTTTTCGGGTG
GTTTGTTGCCATTTTTACCTGTCTGCTGCCGTGATCGCGCTGAACGCGTTTTAGCGGTGC
GTACAATTAAAGGGATTATGGTAAATCCACTTACTGTCTGCCCT

OmpAVisco A3pJOE for.

TATGAAAAAAACCGCGATCGCGATCGCGGTTGCGCTGGCGGGTTTCGCCACCGTTGCGCA
GGCCAAAAGCTGCTGCCCCAACACCACCGGACGAAACATCTACAACGCCTGCCGCTTAAC
CGGCGCTCCCCGCCCAACCTGTGCCAAACTGAGCGGTTGCAAAATCATCTCTGGATCAAC
ATGTCCATCCGATTATCCCAAATAGGATCCTGTACAAGTAAGCTTGGCTGTTTTGGCGGA
TGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAA
CAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAA
GTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGC
CAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTG
TTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGC
GAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAAT
TAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTGTTTAT
TTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTC
AATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCT
TTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAG
ATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTA
AGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAA

Reverse Complement OmpAVisco A3

TTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCG
CTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTA
CTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAAT
AAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT
TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAA
AGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAATTTGATGCC
TGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTCGCAACGTTCAA
ATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACAGATAAAA
CGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTTCCCTACT
CTCGCATGGGGAGACCCCACACTACCATCGGCGCTACGGCGTTTCACTTCTGAGTTCGGCA
TGGGGTCAGGTGGGACCACCGCGCTACTGCCGCCAGGCAAATTCTGTTTTATCAGACCGCT
TCTGCGTTCTGATTTAATCTGTATCAGGCTGAAAATCTTCTCTCATCCGCCAAAACAGCCA
AGCTTACTTGTACAGGATCCTATTTGGGATAATCGGATGGACATGTTGATCCAGAGATGAT
TTTGCAACCGCTCAGTTTGGCACAGGTTGGGCGGGGAGCGCCGGTTAAGCGGCAGGCGTTG
TAGATGTTTCGTCCGGTGGTGTTGGGGCAGCAGCTTTTGGCCTGCGCAACGGTGGCGAAAC
CCGCCAGCGCAACCGCGATCGCGATCGCGGTTTTTTTCATA

OmpAThionin 2.1Strep-TagpJOE rev.

CATATGACCATGACCCGTCTGAAAATCTCTAAAACCCTGCTGGCGGTTATGCTGACCTCT
GCGGTTGCGACGGGTTCTGCGTACGCAAAATCTGCTGCCCTTCCAACCAAGCTAGAAATG
GCTATAGTGTATGCCGTATAAGGTTTTCCAAGGGAAGGTGTATGCAAGTGAGTGGATGCC
AAAACTCTGATACATGCCCCCGAGGTTGGGTAAACGCCATTCTCGAAAACTCAGCTGATG
CTACCAATGAGCACTGCAAGTTAGGGTGTGAAACTTCTGTGTGTGGTGCCATGAACACTC
TCCAGAACTCTGATGCAAGTGAAATTGTGAATGGAGCGTCGGAACAATGTGCCAAGGGAT
GTTCTATTTTCTGTACCAAGAGCTATGTAGTCCCACCTGGGCCGCCTAAACTGTTGGGTT
GGTCTCACCCGCAGTTCGAAAAATAAGGATCCTGTACAAGTAAGCTTGGCTGTTTTGGCG
GATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAA
AACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAG
AAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACT
GCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGT
TGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTT
GCGAAGCAACGGCCCGGAGGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCA
AATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTGTT
TATTTTTCTAAATACATTC

Reverse Complement OmpAThionin 2.1

GAATGTATTTAGAAAAATAAACAAAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAG
GATGGCCTTCTGCTTAATTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCC
CTCCGGGCCGTTGCTTCGCAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGA
GAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTC
GTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCATGGGGAGACCCCACACTACCATCG
GCGCTACGGCGTTTCACTTCTGAGTTCGGCATGGGGTCAGGTGGGACCACCGCGCTACTG
CCGCCAGGCAAATTCTGTTTTATCAGACCGCTTCTGCGTTCTGATTTAATCTGTATCAGG
CTGAAAATCTTCTCTCATCCGCCAAAACAGCCAAGCTTACTTGTACAGGATCCTTATTTT
TCGAACTGCGGGTGAGACCAACCCAACAGTTTAGGCGGCCCAGGTGGGACTACATAGCTC
TTGGTACAGAAAATAGAACATCCCTTGGCACATTGTTCCGACGCTCCATTCACAATTTCA
CTTGCATCAGAGTTCTGGAGAGTGTTCATGGCACCACACACAGAAGTTTCACACCCTAAC
TTGCAGTGCTCATTGGTAGCATCAGCTGAGTTTTCGAGAATGGCGTTTACCCAACCTCGG
GGGCATGTATCAGAGTTTTGGCATCCACTCACTTGCATACACCTTCCCTTGGAAAACCTT
ATACGGCATACACTATAGCCATTTCTAGCTTGGTTGGAAGGGCAGCAGATTTTGCGTACG
CAGAACCCGTCGCAACCGCAGAGGTCAGCATAACCGCCAGCAGGGTTTTAGAGATTTTCA
GACGGGTCATGGTCATA