DUAL-TARGETING OF NADP$^+$-ISOCITRATE DEHYDROGENASE

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in the Department of Plant Sciences

University of Saskatchewan

Saskatoon

By

John David McKinnon

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree from the University of Saskatchewan. I agree that the Libraries of this University may make it freely available for inspection. Moreover, I agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the department or the dean of the college in which this thesis work was done. It is understood that any copying or publication or use of this thesis or parts there of for financial gain without approval by the University of Saskatchewan and the author's written permission is prohibited. It is also understood that due recognition shall be given to the author and to the University of Saskatchewan in any scholarly use which may be made of any material used in this thesis.

Requests for permission to copy or to make either use of the material presented in this thesis in while or part should be addressed to:

Head of the Department of Plant Science

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5A8

ABSTRACT

Many mitochondrial and chloroplast proteins are encoded in the nucleus and subsequently imported into the organelles via active protein transport systems. While usually highly specific, some proteins are dual-targeted to both organelles. In tobacco (*Nicotiana tabacum* L.), the cDNA encoding the mitochondrial isoform of $NADP^+$-dependent isocitrate dehydrogenase ($NADP^+$-ICDH) contains two translational ATG start sites, indicating the possibility of two tandem targeting signals. In this work the putative mitochondrial and chloroplastic targeting signals from $NADP^+$-ICDH were fused to a yellow fluorescent protein (YFP) to generate a series of constructs and introduced into tobacco leaves by *Agrobacterium*-mediated transient transfection. The subsequent sub-cellular locations of the ICDH:YFP fusion proteins were then examined under the confocal microscope. Constructs predicted to be targeted to the chlroplast all localized to the chloroplast. However, this was not the case for constructs that were predicted to be mitochondrial targeted. While some constructs localized to mitochondria, others appeared to be chloroplast localized. This was attributed to an additional 50 amino acid residues of the mature $NADP^+$-ICDH protein which was present in those constructs. In addition, during the process of generating these constructs our sequence analysis indicated a stop codon present at amino acid position 161 of the mature $NADP^+$-ICDH protein from both Xanthi and Petit Havana cultivars of tobacco. This was confirmed by multiple sequencing reactions and created discrepancies with the reported sequence present in the database. The results of this study raise interesting questions with regard to the targeting and processing of $NADP^+$-ICDH.

ACKNOWLEDGMENTS

I would like to express my deep gratitude and appreciation to my research supervisor Dr. Gordon Gray, for the immense help advice and guidance I have received during my graduate program. Sincerer thanks to the members of my advisory committee, Dr. Bruce Coulman, Dr. Kirstin Bett and Dr. Troy Harkness for their valuable guidance during my graduate program and in preparation of this dissertation. Many thanks to Dr. Ken Wilson for all his help with cloning and the use of the confocal microscope. Dr. Peta Bonham-Smith and Dr. Frederica Brandizzi for supplying me with plasmid vectors used during this work.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

LIST OF ABBREVIATIONS

Abbreviation	
ddH_2O	deionized distilled water
GFP	green fluorescent protein
ICDH	isocitrate dehydrogenase
LB	Luria-Bertani medium
$NADP^+$-ICDH	$NADP^+$-dependent isocitrate dehydrogenase
TIM	translocase of the inner mitochondria membrane
TOM	translocase of the outer mitochondria membrane
TIC	translocon of the inner chloroplast membrane
TOC	translocon of the outer chloroplast membrane
YFP	yellow fluorescent protein

1.0 INTRODUCTION

Plant cells contain multiple cellular organelles and in order for proteins to be imported into these organelles they must contain targeting signals (or targeting peptides). Targeting signals are peptides which range from 20-60 amino acid residues in length and are usually found at the N- or C-terminal ends of the protein. While targeting signals for different organelles are usually quite distinct in primary sequence, secondary structure and location in the precursor protein, mitochondria and chloroplasts are two cellular organelles which use similar targeting signals. More than 90% of mitochondrial and chloroplastic proteins are encoded in the nucleus, synthesized on cytosolic ribosomes with a cleavable amino-terminal targeting peptide and then imported post-translationally into the organelle (Bhusham et al., 2006). After import, targeting peptides are cleaved off by peptidases in the respective organelles. Classical mitochondrial and chloroplastic targeting peptides are N-terminal, almost devoid of acidic amino acids, rich in arginine, leucine, alanine and hydroxylated residues and often show amphiphilic structures (Bhusham et al., 2006).

The majority of proteins imported into organelles have only one destination. However, a growing number of proteins have been shown to be dual-targeted, and although encoded by a single nuclear gene, are translated in the cytosol and targeted post-translationally to both mitochondria and chloroplasts (Peeters and Small, 2001). There are two proposed mechanisms by which dual targeting can be achieved: the first is through a dual targeting signal. Dual targeting signals result in two or more proteins with distinct targeting peptides. This can occur as a result of alternative transcription or translation initiation, alternative splicing, or post-translational modification. The second mechanism employs an ambiguous targeting signal.

Ambiguous signals are single targeting polypeptides that can be recognized and transported by the import machinery of more than a one organelle (Danpure, 1995; Small et al., 1998).

$NADP^+$-dependent isocitrate dehygrogenase (ICDH; EC 1.1.1.42) belongs to a multi-enzyme family with isoforms existing in several subcelluar locations including the cytosol, mitochondria, chloroplast, and peroxisome. This enzyme catalyzes the oxidative decarboxylation of isocitrate, producing 2-oxoglutarate (α-ketoglutarate) and CO_2 while converting $NADP^+$ to NADPH (Horton et al., 2002). Previously, work examining the mitochondrial isoform in tobacco (*Nicotiana tabacum* L. cv Xanthi) showed that the cDNA contained two in-frame translational start sites (Gálvez et al., 1998). Furthermore, this study concluded that the enzyme with its full-length targeting signal was localized predominantly to the mitochondrion and weakly targeted to the chloroplast (Gálvez et al., 1998). While this was not a detailed study, it did provide preliminary evidence to suggest that the protein derived from this gene was a good candidate for dual-targeting to both mitochondria and chloroplasts.

1.1 Thesis Objectives

In this work I will examine the localization of mitochondrial $NADP^+$-ICDH and test the hypothesis that this isoform is dual-targeted to both the mitochondrion and chloroplast. This will be accomplished by generating truncated constructs of $NADP^+$-ICDH at the two putative targeting signals and fusing them with a yellow fluorescent protein (YFP). Tobacco (*Nicotiana tabacum* L. cv Petit Havana) leaves will then be transiently transfected with *Agrobacterium tumefaciens* containing the ICDH:YFP fusion proteins and their sub-cellular location examined under the confocal microscope.

2.0 LITERATURE REVIEW

Eukaryotic cells are subdivided into various membrane-bounded compartments called cellular organelles. The endoplasmic reticulum, the Golgi apparatus, lysosomes and peroxisomes possess one boundary membrane. In contrast to these organelles, mitochondria and chloroplasts are bordered by two membranes. Because of these two membranes a complex protein import system has evolved. Based on structural/functional similarities it was suggested that mitochondria and chloroplasts were derived from bacteria which were incorporated into eukaryotic cells by a process called endosymbiosis (Margulis et al., 1981). As a result of this process, a genetic transfer of nucleotide information has occurred where genetic information has been transferred from the incorporated organelle to the nucleus. During this evolution cellular organelles lost most of their genome. Today, the majority of organelle proteins are encoded by nuclear genes, synthesized on cytosolic ribosomes and have to be imported into cellular organelles through an import system from the cytosol (Lang et al*.,* 1999). More than 90 % of mitochondrial and chloroplast proteins are encoded in the nucleus (Pfanner and Geissler, 2001).

The vast majority of proteins destined for mitochondria and chloroplasts are encoded by genes located in the nucleus and synthesized in the cytosol on ribosomes. Most of these proteins carry a cleavable N-terminal signal, or targeting peptide. This targeting signal carries the information required for targeting to the correct organelle. Chloroplast transit peptides are on average longer than mitochondrial pre-sequences. However, both are remarkably similar with respect to the amino acid composition (Bhushan et al., 2006). They exhibit a high abundance of hydroxylated, hydrophobic and positively charged amino acids and very low abundance of acidic amino acid residues (Zhang and Glaser, 2002). It was found that plant

mitochondrial targeting sequences were rich in serine when compared to targeting sequences from non-plant sources. These sequences form amphipathic α-helices, which are important for import (Emanuelsson et al., 2000), whereas chloroplast transit peptides are generally unstructured. Most of the proteins destined for these organelles are imported from the cytosol in a specific manner. Protein import involves multiple steps, from the interaction of the transported protein with cytosolic factors to the activity of specific receptors at the membrane envelope (Cunillera et al., 1997). Some proteins have evolved to be dual-targeted, meaning they are transported to two or more cellular compartments. The majority of reports on single gene products targeted to more than one sub-cellular compartment are commonly related to mitochondria and chloroplasts.

2.1 Import Targeting Signals

Targeting signals for different compartments are usually quite distinct in primary sequence, secondary structure and location in the precursor protein. The location of the targeting signal is frequently found at the N-terminal end of proteins destined for the mitochondrion or chloroplast (Mackenzie, 2005). The import machinery and targeting signals for these organelles appear to be similar. Classical mitochondrial and chloroplast targeting signals are N-terminal, almost devoid of acidic amino acids but rich in arginine, leucine, alanine and hydroxylated residues often showing amphiphilic structures (Small et al., 1998). These sequences are composed of approximately 20-60 amino acid residues (Millar et al., 2006). A common component in targeting sequences is the abundant amount of positively charged, hydroxylated and hydrophobic amino acid residues and the absence of negatively charged residues. Mitochondrial targeting sequences have the potential to form an amphipathic α-helix with a positively-charged face on one side and a hydrophobic surface on the other. The

amphipathic structure of the pre-sequences is thought to be important for recognition by the mitochondrial import machinery of the many cellular organelles (Abe et al., 2000).

It has been demonstrated that protein targeting is signal dependent. The small subunit of ribulose-1,5-bisphosphate carboxylase (Rubisco) is encoded in the nucleus, translated on cytosolic ribosomes and transported into the chloroplast (Ellis et al., 1981). The transit peptide embedded in this subunit contains the information for chloroplastic import as was proven by its ability to direct passenger proteins into the chloroplast. Schreier et al. (1985) created a chimeric gene construct consisting of the promoter, first exon and intron, as well as part of the second exon of the small subunit Rubisco gene fused to the amino-terminal end of the neomycin phosphotransferase II gene. The fusion protein was found to be processed and located within the chloroplasts of the transformed plants (Schreier et al., 1985). In addition, when attached to non-mitochondrial proteins, mitochondrial targeting sequences can specifically direct the passenger protein through the import complexes into the matrix showing access to the organelle is targeting signal dependent (Hurt et al., 1984; Abe et al., 2000).

2.1.1 Ambiguous Targeting Signals

Targeting signals for different organelles are usually quite distinct in some manner. However, the ambiguous targeting signal arises from a gene encoding a single precursor protein which carries a targeting signal that is recognized by the import apparatus of both organelles (Small et al., 1998). Peeters and Small (2001) defined these sequences (ambiguous targeting signals) as being recognized by both the mitochondrial and chloroplast import machinery without further modification. There are more then twenty proteins that are targeted to both these organelles by ambiguous signals. These signals are more hydrophobic and usually

contain more leucine and phenylalanine, but fewer alanine residues than specific mitochondrial and chloroplast targeting signals.

It has been shown that the dual-targeting capability of pea (*Pisum sativum* L.) glutathione reductase to both mitochondria and chloroplasts is dependent on a single targeting peptide. The precursor protein of pea glutathione reductase is dual-targeted to mitochondria and chloroplasts by means of an N-terminal targeting signal of 60 amino acid residues, which has been described as an ambiguous targeting signal (Creissen et al., 1995). This signal has a high serine content (17%), an abundance of aliphatic amino acid residues, an overall positive charge and an uncharged N-terminus which is characteristic of many ambiguous targeting signals.

2.1.2 Mitochondrial Targeting Signals

Newly synthesized mitochondrial pre-proteins contain specific targeting signals, or mitochondrial targeting sequences, that are usually bound by factors which maintain the translocation conformationb of the pre-protein. These factors include chaperones of the heat-shock protein 70 family (Hsp70), as well as specific factors like mitochondrial import stimulation factor that recognize mitochondrial targeting signals (Komiya et al., 1996; Mihara et al., 1996). The majority of mitochondrial pre-proteins are imported post-translationally (Neupert, 1997). The translocase of the outer mitochondrial membrane (TOM) mediates the entry of all nuclear encoded mitochondrial proteins into the mitochondria. This complex functions as a receptor for mitochondrial proteins and provides a protein conducting channel (Lister et al., 2005). Proteins are threaded in an unfolded conformation though the TOM complex (Lister et al., 2005). After crossing the outer membrane through the TOM complex, imported pre-proteins are directed to one of two translocases of the inner membrane (the TIM

complexes; Bauer et al., 2000). The action of the TIM/TOM apparatus requires the hydrolysis of ATP and GTP at different levels.

The TOM complex, is a multi-subunit complex approximately 450 kDa in size, composed of seven subunits (the TOM holo complex); Tom70, Tom40, Tom22, Tom20, Tom7, Tom6 and Tom5 (Neupert, 1997; Pfanner and Geissler 2001; Fig. 2.1). The TOM complex mediates the translocation across and insertion into the outer membrane of all nuclear encoded mitochondrial pre-proteins. Pre-proteins are recognized on the outer surface of the mitochondria by the receptor subunits Tom20, and Tom70 (Hines and Schatz, 1993). They are then transferred into the general import pore (GIP) of the TOM complex and translocated through the outer membrane into the intermembrane space. The GIP is part of the TOM core complex, which is composed of Tom40, Tom22, Tom7, Tom6, and Tom5, but does not contain the receptor subunits (Neupert, 1997). The two receptor proteins Tom20 and Tom70 show different protein specificities. Deletion of both the Tom20 and Tom70 is considered a lethal mutation (Ramage et al., 1993). Both Tom70 and Tom20 contain an amino-terminal membrane anchor and a hydrophilic C-terminal cytosolic domain of 65 kDa and 17 kDa (Söllner et al., 1989; Lister et al., 2005). Both the Tom20 and Tom70 are used for binding of targeting sequences and transfer of proteins into the GIP.

All pre-sequence carrying pre-proteins are directed to the TIM23 complex which consists of the essential membrane proteins Tim17, Tim23 and Tim50 (Bauer et al., 2000). These three proteins associate with the membrane bound Tim44 and the mitochondrial matrix heat-shock protein mtHsp70 (Yamamoto et al., 2002). Pre-proteins can also be directed to the TIM22 complex, consisting of essential membrane proteins Tim22, Tim54 and Tim18 (Koechler et al., 2000; Fig. 2.1). After import into the matrix, the targeting signals of the imported proteins are cleaved off by mitochondrial processing peptidase (MPP; Nunnari et al.,

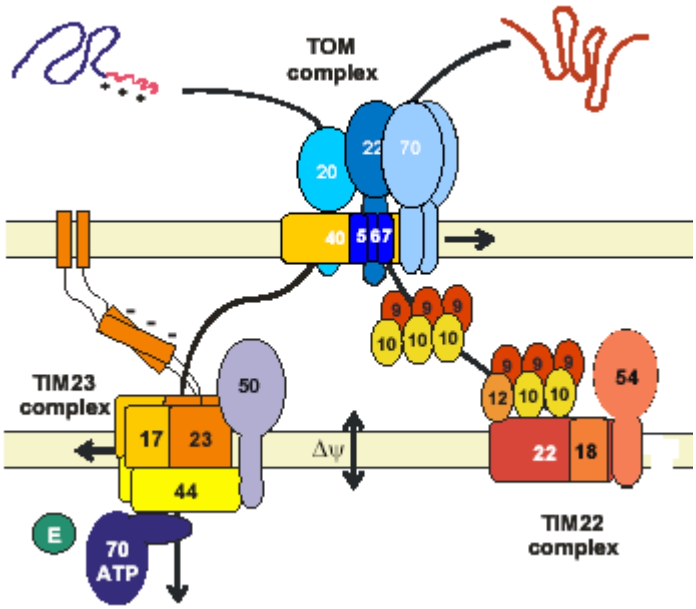

Figure 2.1 The general import pathway into mitochondria. Pre-proteins first bind to specialized import receptors of the TOM complex at the outer membrane and then are transferred to the general insertion pore (GIP). For further translocation, the TOM complex cooperates with the TIM23 and TIM22 complexes in the inner membrane (modified from Bauer et al., 2000).

1993; Neupert, 1997). The majority of the time complete removal of the pre-sequences is followed by folding of the protein into the correct conformation and is essential for obtaining the functionality of imported proteins. There are several matrix proteins including rhodanese, 3-oxo-acyl-CoA-thiolase and chaperonin 10 that are synthesized with a non-cleavable N-terminal targeting signals which have characteristics similar to those of the cleavable signals (Hammen et al., 1996). One matrix protein, the DNA helicase Hmi1, has been found to contain a pre-sequence-like targeting signal at its carboxy terminus (Lee et al., 1999).

2.1.3 Chloroplastic Targeting Signals

The vast majority of chloroplast proteins are synthesized on cytosolic ribosome's and post translationally imported into the chloroplast. The majority of chloroplast precursor proteins have cleavable, N-terminal targeting signals (Soll and Schleiff, 2004; Jarvis 2008). Targeting signals direct precursor proteins to the chloroplast in an organelle-specific way. The majority of pre-proteins are translocated across the organelle membrane using a multicomponent translocon at the outer and inner chloroplast envelope membranes, known as the TOC/TIC complex (Peeters and Small, 2001; Jarvis 2008). The pre-protein for chloroplasts contain a stromal import sequence or both a stromal and thylakoid targeting sequence. In the stroma, the import sequence is cleaved off and intra-chloroplast sorting and folding continues. Pre-proteins that contain a cleavable transit peptide are recognized in a GTP-regulated manner by receptors of the outer-envelope translocon, the TOC complex. The pre-proteins cross the outer envelope through an aqueous pore and are then transferred to the translocon in the inner envelope, the TIC complex (Becker et al., 2004). The TOC and TIC translocons function together during the translocation process. The action of the TOC/TIC apparatus requires the hydrolysis of ATP and GTP at different levels, indicating energetic requirements and

regulatory properties of the import process (Becker et al., 2004). The stromal processing peptidase then cleaves the transit sequence to produce the mature form of the protein, which can fold into its native form. The main subunits of the TOC and TIC complexes have been identified and characterized *in vivo*, *in organello* and *in vitro* (Jarvis and Soll, 2001).

Once a precursor protein arrives at the surface of the chloroplast, a highly specific recognition process is initiated to ensure the protein is indeed destined for this organelle. After membrane recognition, pre-proteins are then threaded into the import complex in extended conformation from N-terminus to C-terminus (Becker et al., 2004; Smith 2006). During import, the TOC and TIC complexes come together at contact sites and the precursor protein passes through both membranes.

The four pea TOC components that have been described are called Toc159, Toc34, Toc75 and Toc64, according to their molecular masses (Fig. 2.2). Toc159 (also called Toc160) appears to be a major point of contact for precursor proteins arriving at the complex (Becker et al., 2004). Toc34 belongs to the same unique class of GTP-binding proteins as Toc159, since it shares significant homology with Toc159 outside of the conserved GTP-binding site motifs (Jarvis and Soll, 2001). Like Toc159, it is an integral membrane protein and is attached to the outer envelope membrane by a C-terminal membrane anchor. Toc75 plays a central role in the import mechanism as it stably associates with Toc159 and Toc34 (Jarvis and Soll, 2001).

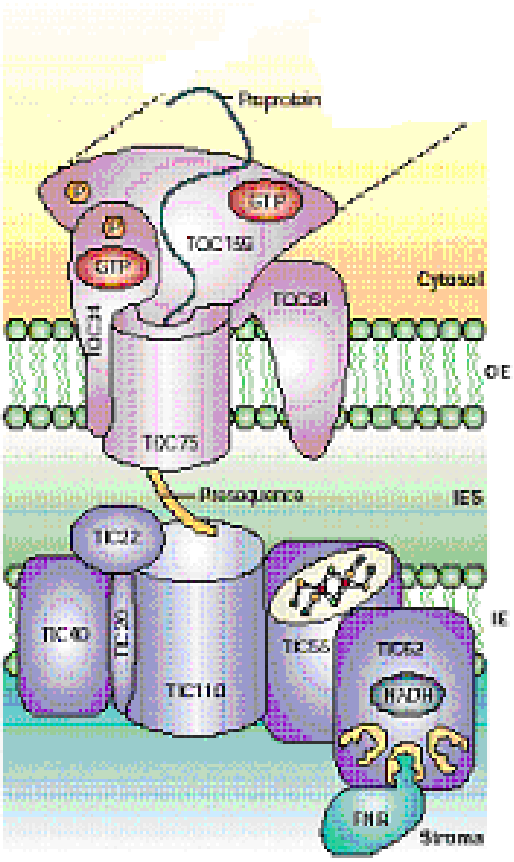

Figure 2.2 The chloroplast translocon complexes. Pre-proteins first bind to specialized import receptors of the TOC complex at the outer membrane. The Tic proteins comprising the TIC complex assist in further translocation (adapted from Soll and Schleiff, 2004).

The Tic proteins that have been identified to date are Tic110, Tic55, Tic40, Tic22 and Tic20 (Fig. 2.2). Tic110 is proposed to have a role in recruiting chaperones to the stromal face of the Tic complex (Hörmann et al., 2007). Tic55 plays a regulatory role during import by responding to changes in redox status within the chloroplast as it contains a Rieske-type iron sulfur cluster and a mononuclear iron-binding site (Schnell, 2000). Many genes expressed in the chloroplast are subject to redox regulation. Since proteins can affect the redox state in the organelle it is not surprising to find that chloroplast protein import is effect by redox regulation (Schnell, 2000). Tic40 during shares similarities to hsp70-interacting protein at its C-terminal end. It has been speculated that Tic40 regulates the chaperones responsible for driving chloroplast protein import (Jarvis and Soll, 2001). Tic22 is associated with the outer surface of the inner envelope membrane, suggesting that it might act as a receptor for pre-proteins as they emerge from the TOC complex, or mediate the association of TOC and TIC complexes (Kouranov et al., 1998). Tic20 is deeply embedded within the inner envelope membrane suggesting a role for protein conductance through the chloroplastic inner membrane (Kouranov et al., 1998). The structural organization of the protein suggests a channel-forming role, and has led to the idea that it is primarily involved in protein translocation (Soll and Schleiff, 2004).

2.2 Dual-Targeting Mechanisms

Many genes have been found to contain multiple in-frame 5' transcriptional/translational start sequences (ATG/AUG). It has been shown in several cases that two protein products are able to arise from the same gene. If the sequence between the start sites contains a cellular targeting signal then the protein initiated from the upstream start site will contain targeting information that the protein initiated from the downstream start site will

lack. This will lead to the protein being targeted to multiple destinations. There are three ways this is thought to occur which are illustrated in Fig. 2.3 and described below.

2.2.1 Alternative Transcription Starts

Many genes have multiple ATG sites meaning they also can potentially have multiple transcription starts sites. If one or more of these sites occurs downstream of the first ATG the encoded N-terminal sequence following the downstream ATG will not contain the information that the protein generated from the upstream ATG will possess (Small et al., 1998). This mechanism results in two transcripts being produced, the longer transcript possibly encoding a targeting sequence which is absent from the shorter transcript (Fig. 2.3A). This is the most common way in which multiple protein products are generated from a single gene in most organisms. Many house-keeping genes have multiple transcription starts as these enzymes are usually required in several cellular compartments. In plants, the monodehydroascorbate reductase (MDAR) is an enzymatic component of the ascorbate-glutathione cycle that is one of the major antioxidant systems of plant cells for the protection against the damage produced by reactive oxygen species (ROS). The MDAR activity has been described in several cellular compartments, such as chloroplasts, cytosol, mitochondria, glyoxysomes and leaf peroxisomes (Obara et al., 2002). The mitochondria and chloroplast monodehydroascorbate reductase is dual-targeted by the use of multiple transcription initiation sites. Obara et al. (2002) showed that two MDAR mRNAs are produced from a single gene and give rise to two different proteins in which the longer transcript is transported to mitochondria and the shorter transcript to chloroplasts (Obara et al., 2002).

Figure 2.3 Mechanisms by which two proteins can be generated from a single gene. A gene with three exons (white, blue and pink boxes) is used to illustrate the production of two polypeptides, one of which contains a targeting sequence (hatched box) and one which does not. Introns are indicated by V-shaped lines. **A)** Alternative transcription starts, **B)** Alternative exons, **C)** Alternative translation starts (modified from Small et al., 1998).

2.2.2 Alternative Exons

Genes can be expressed in such a way that an internal exon encoding targeting information is or is not included in the mRNA depending on how the transcript is spliced. This mechanism results in the production of two transcripts, one of which starts in an intron of the longer transcript. In this way, it is possible for the second transcript to encode a protein with either an alternative N-terminal extension or one devoid of targeting information (Fig. 2.3B). There are only a small number of examples of this splicing which makes use of alternative 5' exons to encode N-terminal targeting sequences (Small et al., 1998). Which exon is used depends on where transcription is initiated (Fig. 2.3B). A good example is the human dUTPase gene, which can encode nuclear or mitochondrial isoforms depending on whether the exon encoding the mitochondrial targeting sequence is included in the mRNA (Lander et al., 1998).

2.2.3 Alternative Translation Starts

A third way in which dual-targeted protein products can be generated is by alternative translation. This occurs when a protein is encoded by a transcript that has two or more translation start site sequences in the same mRNA. This mechanism results in a single transcript being produced, but translation can occur from either the upstream AUG or the downstream AUG, resulting in two proteins with different (or absent) targeting signals (Fig. 2.3C). This type of dual-targeting mechanism is similar to alternative transcription start sites and it is hard to distinguish between mechanisms. Alternative translation starts seem to occur less frequently than alternative transcription starts in animals and fungi. Surprisingly, a high proportion of dual-targeted plant enzymes appear to be produced in this manner, including aminoacyl-tRNA synthetases (ALATS; Danpure et al., 1995).

In plants, ALATS are encoded in the nucleus and present activity in three cellular compartments; the cytosol, mitochondria, and chloroplasts (Danpure et al., 1995). Transcriptional analyses studies showed the presence of two potential translation initiation codons in some ALATS mRNAs. Translation from the upstream AUG would generate an N-terminal extension with features characteristic of mitochondrial targeting peptides while translation from the downstream AUG would not contain this targeting peptide. *In vitro* experiments confirmed that two polypeptides can be translated from a single ALATS transcript with most ribosomes initiating on the downstream AUG to give the shorter polypeptide, corresponding in size to the cytosolic enzyme (Mireau et al., 1996). It was concluded that the ALATS gene encodes both the cytosolic and the mitochondrial forms of ALATS, depending on which of the two AUG codons are used to initiate translation (Mireau et al., 1996). Silva-Filho et al. (2003) found that *Arabidopsis thaliana* THI1, a protein playing a role in the biosynthesis of thiamine (vitamin B1), is encoded by a single nuclear gene and directed simultaneously to mitochondria and chloroplasts from a single transcript. This gene contains two in-frame translational start codons which regulate the subcellular localization of THI1. When translation occurs from the first start site the product is localized to the chloroplast and conversely, when translation occurs from the second start site mitochondrial targeting is observed (Silva-Filho et al., 2003). Another protein utilizing a dual- targeting signal is protoporphyrinogen oxidase II, an enzyme necessary for the biosynthesis of chlorophyll in the chloroplasts and for heam in both chloroplasts and mitochondria. This protein also has two in-frame initiation codons. Two different proteins are made by the means of alternative translation in which the longer transcript is imported into chloroplasts and the shorter transcript into mitochondria (Watanabe et al., 2001).

2.2.4 Post-Translational Modifications

Post-translational modifications of proteins are common and often greatly affect tertiary structure and activity. It is possible to perceive that differential modification could lead to differential targeting. The post-translational modification of amino acids extends the range of functions of the protein by attaching other biochemical functional groups such as acetate, phosphate, various lipids and carbohydrates, or by making structural changes such as the formation of disulfide bridges. Protein structure can also effect the final destination of the protein as it is possible to envision the protein being folded in such a way as to bury the targeting signal (Karyniely and Pines**,** 2005). Modification of the protein in this manner may directly or indirectly affect the targeting properties through protein folding or binding.

It has been shown that Aky2, an adenylate kinase in yeast, has a dual location with the bulk of the enzyme residing in the cytoplasm and a minor fraction in the mitochondrial intermembrane space (Strobel et al., 2001). Depending on how the protein is folded either it is translated to completion and folds into an enzymatically active import-incompetent conformation that remains in the cytosol, or, during translation, but prior to the formation of a significant tertiary structure, it reaches a mitochondrial surface receptor and is internalized (Karneity and Pines, 2005).

2.3 Predicting Dual-Targeted Proteins

Different programs have been designed to predict the pre-sequence of mitochondrial or plastid targeting signals including Mitoprot (http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html; Claros and Vincens, 1996), TargetP (http://www.cbs.dtu.dk/services/TargetP/; Emanuelsson et al., 2000) and Predotar (http://urgi.versailles.inra.fr/predotar/predotar.html; Small et al., 2004). Predotar, TargetP and

MitoProt only predict N-terminal targeting sequences, and no method exists that will identify import signals present elsewhere in the protein, although such signals are known to exist (Emanuelsson et al., 2000). These programs predict protein location correctly approximately 80% of the time (Emanuelsson et al., 2000). The Predotar program is designed to predict dual-targeted proteins to the mitochondria and chloroplast but does not give a score for any other cellular organelles. Mitoprot does not give separate mitochondrial and plastid scores and cannot be used efficiently for predicting dual-targeted proteins. TargetP and Predotar give separate scores, and thus can be used for predicting dual-targeted proteins. However, it is not possible to assign unequivocally the sub-cellular localization of the mature proteins based on this kind of theoretical analysis. In practice the predictions are not reliable and experimental data is still required. For the dual-targeted proteins protoporphyrinogen oxidase II and THI1 (Watanabe et al., 2001; Silva-Filho et al., 2003), the longer protein is predicted to be targeted to the chloroplast as would be expected. However, the shorter version of these proteins is not predicted to be a mitochondrial targeted. Subsequent experimental data indicated this was not the case and they were indeed mitochondrial localized. This emphasizes the need for experimental data to unequivocally demonstrate protein targeting.

2.4 Isocitrate Dehydrogenase as a Model for Dual-Targeting

In plants, two different enzymes catalyzing the oxidative decarboxylation of isocitrate to 2-oxoglutarate can be distinguished by their cofactor specificity NAD^+-dependent isocitrate dehydrogenase (IDH; E.C.1.1.1.41) is located exclusively in the mitochondria where it plays a role in the TCA cycle. It catalyzes the third step of the cycle, the oxidative decarboxylation of isocitrate, producing 2-oxoglutarate (α-ketoglutarate) and CO_2 while converting NAD^+ to NADH (Horton et al., 2002). This is a two-step process, which involves oxidation of isocitrate

(a secondary alcohol) to oxalosuccinate (a ketone), followed by the decarboxylation of the carboxyl group to a ketone, forming α-ketoglutarate. The IDH step of the citric acid cycle, due to its large negative free energy change, is one of the irreversible reactions in the citric acid cycle and therefore must be carefully regulated to avoid unnecessary depletion of isocitrate and therefore an accumulation of α-ketoglutarate (Hodges et al., 2003). The reaction is stimulated by the simple mechanisms of substrate availability (isocitrate, NAD^+, Mg^{2+}/Ca^{2+}), product inhibition by NADH and α-ketoglutarate and competitive feedback inhibition by ATP (Horton et al., 2002). Other isoforms of the enzyme catalyze the same reaction, but use $NADP^+$ as a cofactor instead of NAD^+. $NADP^+$-dependent isocitrate dehydrogenase (ICDH; EC 1.1.1.42) belongs to a multi-isoenzymatic family, whose members have been located within the cytosol, plastids, mitochondria and peroxisomes (Gálvez et al., 1995). All are homodimeric enzymes, the subunit molecular mass being approximately 47 kDa. It has been proposed that the cytosolic ICDH isoform is responsible for producing the 2-oxoglutarate necessary for ammonia assimilation. The assimilation of ammonia is an example of a plant metabolic pathway that relies on enzymatic isoforms. Ammonia produced from nitrogen fixation, nitrate reduction, or photorespiration is incorporated into organic compounds by the glutamine synthetase/glutamate synthase (GS/GOGAT) cycle (Gálvez et al., 1995). This cycle requires a continuous supply of 2-oxoglutarate, which is synthesized by the action of cytosolic ICDH.

The cDNA encoding the mitochondrial isoform (GenBank accession number X96728.1) was isolated by Gálvez et al. (1998). The deduced amino acid sequence of this cDNA shared only 70-75% amino acid identity with that of the cytsolic gene. It was found that this cDNA contained two in-frame ATG start sites and it was hypothesized that this resulted in two targeting signals. Analysis of the deduced amino acid sequence appeared to demonstrate that the N-terminal targeting peptide was mitochondrial in nature. However, it is not possible to say

from the targeting signal sequence analysis to which sub-cellular compartment the protein is targeted. To address to this question, they transformed tobacco plants with two reporter constructs; one containing the entire coding sequence (from the first ATG) fused with a modified green fluorescent protein (GFP) and another construct containing only the coding sequence of the mature protein, also fused to a GFP. These experiments by Gálvez et al. (1998) found that the full-length construct with its targeting signal localized the protein to the mitochondria and surprisingly, weakly to the chloroplast, while the construct containing no targeting signal was retained in the cytoplasm.

The $NADP^+$-dependent ICDH being examined in this study is the mitochondrial isoform encoded by the cDNA isolated by Gálvez et al. (1998). One of the aims of this study was to clarify the results found in Gálvez et al. (1998) by recreating the two constructs that they utilized. In addition, another construct was created containing a theoretical second targeting signal beginning at the second start codon (predicted chloroplast targeting). Constructs were also created that possessed the same targeting signals described above but with and additional 50 amino acids of coding sequence for the mature protein. By creating these variations and thereby altering the targeting signals of this protein I was able to uncover a more in-depth and clearly developed picture of the mechanisms involved in the dual-targeting signals of the mitochondrial $NADP^+$-ICDH gene product.

3.0 MATERIALS AND METHODS

3.1 Plant Material and Growth Conditions

Tissue culture generated tobacco plants *(Nicotiana tabacum* L. cv Xanthi) for molecular analyses were obtained from Dr. Jonathon Page at the National Research Council of Canada - Plant Biotechnology Institute (Saskatoon, Saskatchewan). The propagation of these plants has been described previously (Horsch et al., 1985).

Tobacco plants (*Nicotiana tabacum* L. cv Petit Havana) utilized for transient expression experiments were germinated from seed in 100 mm diameter plastic pots containing a peat-soil mixture (Sun Gro, Sunshine mix; Vancouver, BC, Canada). A 20-20-20 complete fertilizer mix (Plant Prod; Brampton Canada) was added to the plants water source as required. Plants were grown at temperature regime of 23/18°C with a 16 /8 h light/dark cycle with a photon flux density of 170 µmol photons $m^{-2}s^{-1}$ provided by fluorescent tubes (Cool White, 215 W, F96T12/CW/VHO; Osram Sylvania Ltd., Mississauga, ON, Canada).

3.2 DNA and RNA Extraction

Genomic DNA was extracted from *N. tabacum* L. cv. Xanthi or Petit Havana leaves using the Qiagen DNasy easy plant mini kit (Qiagen; Mississauga, Ontario, Canada) following the manufactures recommendations. Samples were eluted in 50 µL of sterile water, DNA concentration was determined spectrophotometrically (Smart Spec Plus, Bio-Rad Laboraties) by measuring the absorbance of the DNA sample at 260 nm. The samples were stored in sterile microcenterfuge tubes at -20°C.

Total RNA was extracted from *Nicotiana tabacum* L. cv Xanthi or Petit Havana leaves using the Qiagen RNasy easy plant mini kit (Qiagen; Mississauga, Ontario, Canada) following

the manufacturer's recommendations. Samples were eluted in 50 µL of sterile water and their RNA concentrations were determined spectrophotometrically by measuring the absorbance of the RNA samples at 260 nm. The samples were stored in sterile microcentrifuge tubes at -20°C.

3.3 cDNA Synthesis

cDNA was generated from total RNA using the RevertAid™ Fermentus First Single Strand cDNA Synthesis kit (Fermentas; Burlington, Ontario, Canada). This was performed as described by the manufacturer as follows. Total RNA was diluted to 5 µg and placed in a sterile thin-walled PCR tube along with 1 µL of oligo $(dT)_{18}$ primer (0.5 µg/µL) and 8 µL of DEPC-treated water. This mixture was incubated at 70°C for 5 min and then placed on ice. The following components where then added: 4 µL 5X reaction buffer, 1 µL RiboLock™ Ribonuclease Inhibitor (20 U/µL) and 2 µL 10 mM dNTP (an equal mixture of 10 mM dATP, dGTP, dCTP and dTTP) mix and incubated for 5 min at 37°C. After this period, 1 µL of RevertAid™ H Minus M-MuLV Reverse Transcriptase (200 U/µL) was added bringing the total reaction volume to 20 µL. The mixture was placed at 42°C for 60 min and the reaction was then stopped by heating to 70°C for 10 min. All incubations occurred using a thermocycler (iCycler, Bio-Rad Laboraties). The single strand cDNA was then used as template DNA in the PCR reactions described in section 3.5.

3.4 Primer Design

The coding region of the tobacco mitochondrial $NADP^+$-ICDH cDNA (GenBank accession X96728.1) contains two in-frame ATG translation start sites. Oligonucleotide primers were designed to amplify various regions of the $NADP^+$-ICDH cDNA and these are indicated in Tables 3.1 and 3.2. This process also introduced EcoR1 and HindIII restriction cut

Table 3.1 Name and nucleotide sequence of primers utilized to amplify the entire coding region of the $NADP^+$-ICDH mature protein with altered targeting signals. The red underlined section indicates the start codon, blue italicized sections indicate added EcoRI (GAATTC) and HindIII (AAGCTT) restriction sites.

Primer Name	**Sequence (5'→3')**
DMc011f	CGC*GAATTC*ATGCTTACCACCCGACTCAGACTC
DMc012f	CGC*GAATTC*ATGGCTAGTGTTGCTTCTTTTATCTC
DMc013f	CGC*GAATTC*ATGAAAATCCGCGTTGAAAATCCT
DMc014r	CGC*AAGCTT*TACAACTGCGCAGGCACCGAGCTTCTC

Table 3.2 Name and nucleotide sequence of primers utilized to amplify regions encoding truncated proteins of $NADP^+$-ICDH with altered targeting signals. The red underlined section indicates the start codon, blue italicized sections indicate added EcoRI (GAATTC) and HindIII (AAGCTT) restriction sites.

Primer Name	**Sequence (5'→3')**
DMc011f	CGC*GAATTC*ATGCTTACCACCCGACTCAGACTC
DMc012f	CGC*GAATTC*ATGGCTAGTGTTGCTTCTTTTATCTC
DMc013f	CGC*GAATTC*ATGAAAATCCGCGTTGAAAATCCT
DMc030r	CGC*AAGCTT*CCACATAGACTTCAGCCCAAATT
DMc040r	CGC*AAGCTT*ATACTTCGTATCCAACTCTAG

sites at the 5' and 3' ends, respectively. All primers were synthesized commercially (Alpha DNA; Montreal, Quebec, Canada).

The first primer set was designed to amplify the coding region of the entire mature protein from the first ATG start site (both the putative mitochondrial and plastid targeting signals). This was designated mito-482 and used primer set DMc011/DMc014 (Table 3.1). The second amplified the coding region of the mature protein from the second ATG start site (only the putative plastid targeting signal). This was designated chloro-482 and used primer set DMc012/DMc014 (Table 3.1). The third amplified the coding region of the mature protein (no putative targeting signals). This was designated cyto-482 and used primer set DMc013/DMc014 (Table 3.1).

The next six primer sets was designed to utilized the same forward primers described above and were identical in terms of their putative targeting signals (Tables 3.1 and 3.2). However, two different reverse primers were utilized; DMc030 and DMc040 (Table 3.2). When translated, the resulting PCR products generated truncations at amino acid 160 if the DMc030 primer was used or at amino acid 138 if the DMc040 primer was used in the PCR amplification. The former PCR products were designated mito-160, chloro-148 and cyto-90 while the latter were designated mito-110, chloro-98 and cyto-40. These designations and those above were assigned based on the number of amino acids the translated PCR product contained.

3.5 Polymerase Chain Reaction

cDNA sequences were amplified by PCR using a thermostable DNA polymerase (*Pfu*-polymerase; Fermentas). The PCR mixture contained: 1-2 U *Pfu*-polymerase, 10 µL PCR buffer (1% [v/v] Triton X-100, 100 mM Tris-HCl [pH 8.8], 500 mM KCl, 15 mM $MgCl_2$), 2 µL 10 mM dNTPs (an equal mixture of 10 mM dATP, dGTP, dCTP and dTTP), 50 pM of each

forward and reverse primer, 1 µg cDNA or 1µg genomic DNA and RNase free water up to a total volume of 50 µL. The following cycling conditions were used on an iCycler (Bio-Rad Laboraties): 1) 2 min at 95°C for nuclease inactivation and complete DNA denaturation; 2) 30 cycles of 30 s at 94°C for DNA denaturation; 30 s at 55°C for annealing of oligonucleotide primers; 2 min at 72°C for new DNA synthesis (extension); and 3) 10 min at 72°C for completion of the last reaction.

3.6 Agarose Gel Electrophoresis

The amplified DNA fragments were analyzed by agarose gel electrophoresis. The fragments of DNA were separated by electrophoresis through a 1% (w/v) agarose gel in 1X TAE buffer. The working solution of 1X TAE buffer was made by diluting a 50X stock solution of TAE (242 g/L Tris, 57.1 mL/L acetic acid, 14.6 g/L disodium EDTA, pH 8.0) in deionized water. Ethidium bromide (0.5 µg/mL) was added to the gel which allowed visualization of the DNA under UV light. The DNA samples were mixed with 6X DNA loading buffer (Fermentas) prior to loading on the gel. Electrophoresis was carried out at 50 to 100V and stopped when the bromphenol blue migrated three quarters of the gel length. Gels were visualized on a short-wave UV-B trans-illuminator system (Gel Doc, Bio-Rad Laboraties) and the image stored electronically. The sizes of the DNA fragments were estimated by comparing them to standards of known size which were loaded in adjacent lanes.

3.7 Restriction Digestion

PCR products were digested with specific restriction endonucleases (up to 5 U of enzyme for 1 µg of DNA). The primers were designed to introduce a 5' EcoR1 and a 3' HindIII restriction site on the PCR products. A double restriction digest was performed by adding 10

µg of PCR product, 5 µL 2X Tango buffer (33 mM Tris-acetate [pH 7.9], 10 mM magnesium acetate, 66 mM potassium acetate, 0.1 mg/mL BSA), 2 µL EcoR1 (Fermentas), 2 µL HindIII (Fermentas) and nuclease free water up to 30 µL. This mixture was incubated for 1 hour at 37°C before inactivation at 65°C for 20 min. The digested fragments obtained were analyzed by agarose gel electrophoresis (section 3.6).

3.8 Extraction of DNA Fragments

DNA was resolved in a 1% (w/v) agarose gel and viewed with UV-light. An ethanol sterilized razor blade was used to cut out the appropriate region of the gel containing the DNA of interest. The DNA was then purified using the QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's instructions. Samples were eluted in 50 µL of sterile water, quantified spectrophotometrically (Smart Spec Plus, Bio-Rad Laboraties), and stored in sterile microcentrifuge tubes at -20°C.

3.9 Vector Preparation

The plasmid pBluescript KS+ (Stratagene; Kirkland, Washington, USA; Fig. A-1) was supplied by Dr. Peta Bonham-Smith (Department of Biology, University of Saskatchewan) as a glycerol stock in *Escherichia coli* DH5-α cells (Invitrogen; Burlington ON, Canada). This stock was grown in Luria-Bertani (LB) media (tryptone [10 g/L], yeast extract [5 g/L] and NaCl [5 g/L]). The media was sterilized by autoclaving (AMSCO eagle series 3021 gravity; Cheshire, CT, USA) for 20 min at 121°C. For antibiotic selection, an ampicillin stock (40 µg/mL) was filter sterilized using a 0.22 µm syringe filter (Millipore; Kankakee, Illinois, USA) and added to the cool media.

The *E. coli* stock was grown by inoculating 5 mL of LB media, containing 40 µg/mL ampicillin with 100 µL of bacterial culture and incubating at 37°C for 18 to 24 hours. Plasmid DNA was extracted (see section 3.10) and digested using EcoR1 (Fermentas) and HindIII (Fermentas) restriction enzymes (section 3.7). The linearized vector was analyzed by agarose gel electrophoresis (section 3.6) followed by purification by gel extraction (section 3.8).

3.10 Extraction of Plasmid DNA

Overnight *E. coli* cultures (5 mL) were centrifuged (Microfuge Lite, Beckman; Fullerton, California, USA) at 8,000 rpm for 2 min at room temperature to pellet the bacteria. Plasmid DNA was the extracted from the pelleted bacteria using the GeneJET Plasmid Miniprep Kit (Fermentas) according to the manufactures recommendations. Samples were eluted in 50 µL of sterile water, quantified spectrophotometrically (Smart Spec Plus, Bio-Rad Laboraties), and stored in sterile microcentrifuge tubes at -20°C.

3.11 Ligation Reaction

The amplified PCR products were directionally ligated into pBluescript KS+ to produce constructs containing the various fragments described previously (section 3.4). T4-DNA ligase (Invitrogen) was used to ligate DNA fragments. This reaction contained a linearized DNA vector (50-200 ng) and 2-5 times the molar concentration in excess of the DNA fragment (PCR product). This occurred in a 20 µL reaction containing with 2 µL of 10X ligation buffer (50 mM Tris-HCl, 10 mM $MgCl_2$ 5% [w/v] PEG-8000, 1 mM DTT, 1 mM ATP, pH 7.6), and 1 U T4-DNA ligase (Invitrogen) and water to volume. The reactions were performed at room temperature for 1 hour or at 4°C for 12 hours. The reactions were stopped by heat inactivating

the enzyme by incubation at 65°C for 10 min. The ligated DNA was used to transform *E. coli* competent cells.

3.12 Transformation of *Escherichia coli*

Prior to transformation, frozen *E. coli* DH5-α competent cells (Invitrogen) were allowed to thaw on ice for 2 min. Once thawed, the ligation reaction mixture (5 µL) was added to 50 µL of cells. The cells and plasmid DNA were incubated on ice for 30 min, then heat shocked for 2 min at 42°C and placed back on ice for a final 2 min to facilitate DNA uptake. After adding 1 mL of LB media, the cells were incubated at 37°C for 1 hour with shaking before being appropriately diluted (25-250 cfu/plate) and plated on LB plates containing ampicillin (40 µg/mL). After an overnight incubation period at 37°C, individual colonies on the plates were picked with a sterile toothpick, placed in 5 mL liquid LB-ampicillin media and grown overnight at 37°C. The bacterial culture was then subject to plasmid DNA extraction (see section 3.9), a double restriction digest (section 3.7) and gel electrophoresis (section 3.6) in order to confirm presence of an insert (PCR product). Glycerol stocks for long term storage were prepared by adding 0.9 mL of culture and 0.9 mL of 40% (v/v) glycerol to individual cryovials (Diamed; Missassuga, ON, Canada) and stored at -80°C.

3.13 Sequence Analyses

Confirmation of constructs was also performed by sequence analyses. Glycerol stocks or plasmid DNA were sent for sequencing at National Research Council of Canada - Plant Biotechnology Institute (Saskatoon, Saskatchewan). Nucleotide and amino acid alignment was performed using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html; Larkin et al., 2007). Deduced amino acid sequence was obtained using EMBOSS Transeq

(http://www.ebi.ac.uk/emboss/transeq/; Rice et al., 2000). Predictions of targeting were performed using Predator (http://www.inra.fr/Internet/Produits/Predotar/; Small et al., 2004).

3.14 Amplification of the YFP

The binary vector pVKHI8En6-ERD2-YFP (Batoko et al., 2000; Brandizzi et al. 2002*a*, 2002 *b*) (Fig. A-2), containing a region encoding a yellow fluorescent protein (YFP), was supplied as a purified plasmid by Dr. Federica Brandizzi (DOE Plant Research Laboratory, Michigan State University, East Lansing, MI, USA). PCR was performed using the primer set YFPf/YFPr (Table 3.3) and pVKHI8En6-ERD2-YFP as a DNA template. Cycling conditions were as described in section 3.5.The YFPf/YFPr primers were designed to amplify the YFP as well as to introduce 5' HindIII and 3' Spe1 restriction sites. Following PCR, the reaction mixture was subjected to a restriction digest using enzymes HindIII and Spe1. This was performed in a reaction containing 5 µL 2X Tango buffer (33 mM Tris-acetate [pH 7.9], 10 mM magnesium acetate, 66 mM potassium acetate, 0.1 mg/mL BSA), 2 µL Spe1 (Fermentas), 2 µL HindIII (Fermentas), 10 µg of PCR product, and nuclease free water up to 30 µL. This mixture was incubated for 1 hour at 37°C before being heat inactivated at 65°C for 20 min. The digested fragments obtained were analyzed by agarose gel electrophoresis (section 3.6) and the PCR product corresponding to the YFP (720 bp) was extracted from the gel using a commercially available kit (Qiagen) as described in section 3.8.

3.15 Creation of the pYellow Reporter Vector

A pRed (pRed-L23a-GST-GFP) binary vector (Fig. A-3; Degenhard and Bonham-Smith, 2008) was supplied by Dr. Peta Bonham-Smith (Department of Biology, University of Saskatchewan) as a glycerol stock in *E. coli* DH5-α cells (Invitrogen) and was used to generate

Table 3.3 Name and nucleotide sequence of primers utilized to amplify the YFP. The red underlined section indicates the start codon, blue italicized sections indicate added HindIII (AAGCTT) and Spe1 (ACTAGT) restriction sites.

Primer Name	Sequence (5'→3')
YFPf	CGC*AAGCTT*ATGGTGAGCAAGGGCGAGGAGCTG
YFPr	CGC*ACTAGT*TTACTTGTACAGCTCGTCCTAGCCGAG

the pYellow reporter vector (Fig. A-4). The bacterial culture was grown as described in section 3.9 and plasmid DNA extracted as previously described in section 3.8.

The pRed vector contains a red fluorescent protein (RFP) flanked by a 5' HindIII and a 3' Spe1 restriction site (Fig. A-3). The pRed vector (10 µg) was subjected to a HindIII/Spe1 digestion as described above in section 3.14 to release the RFP. The digestion was analyzed by agarose gel electrophoresis (section 3.6) and the linearized plasmid was gel purified using a commercially available kit (Qiagen) as described in section 3.8. The gel purified and digested YFP from section 3.14 was then directionally ligated (section 3.11) into the pRed vector and transformed into *E. coli* DH5α competent cells (Invitrogen) by heat shock (section 3.12). Positive colonies were selected and confirmed by a HindIII/Spe1 digestion (section 3.14) to confirm the presence of the 720 bp YFP insert. Glycerol stocks were created and stored at -80°C. This generated the binary reporter vector pYellow (Fig. A-4) which was subsequently gel purified (section 3.8) using a commercially available kit (Qiagen).

3.16 Creation of ICDH:YFP Fusion Constructs

The pYellow vector and the six PCR products (mito-160, chloro-148, cyto-90, mito-110, chloro-98 and cyto-40) described previously (section 3.4) were digested with EcoR1/HindIII (section 3.7), analyzed by agarose gel electrophoresis (section 3.6), and gel purified (section 3.8). The PCR products were then directionally ligated (section 3.11) into the linearized pYellow and used to transform *E. coli* DH5α competent cells (Invitrogen) by heat shock (section 3.12). Positive colonies were selected and subjected to an EcoRI/HindIII digestion (section 3.7) to confirm presence of inserts. Glycerol stocks were created and stored at -80°C. This gave rise to six ICDH:YFP fusion constructs which are presented in Table 3.4.

Table 3.4 ICDH:YFP fusion constructs utilized for confocal microscopy.

Name	Vector	Number of Amino Acids Residues	Predicted Localization
mito-160:YFP	pYellow	160	mitochondrial
chloro-148:YFP	pYellow	148	chloroplastic
cyto-90:YFP	pYellow	90	cytosolic
mito-110:YFP	pYellow	110	mitochondrial
chloro-98:YFP	pYellow	98	chloroplastic
cyto-40:YFP	pYellow	40	cytosolic

3.17 Transformation of *Agrobactium tumefaciens*

The fusion constructs presented in Table 3.4 were used to co-transform *Agrobacterium tumefaciens* (ElectroMAX LBA4404; Invitrogen) along with the helper vector pSoup (Fig. A-5; supplied by Dr. Peta Bonham-Smith, University of Saskatchewan) by electroporation. All ICDH:YFP fusion constructs in pYellow were co-electroporated with pSoup, which must be coresident in *A. tumefaciens* to provide the replication functions for pYellow (Hellens et al*.,* 2000). The plasmids were added to 20 µL of competent cells and eletroporation occurred using a Gene Pulser II apparatus with Pulse Control Plus and Capacitance Extender Plus modules (Bio-Rad Laboraties) according to the manufactures recommendations. The following conditions were used 2.0 kV, 200Ω, 25 µF with an electroporation cuvette (2 mm gap; VWR; Mississauga, ON, Canada). Immediately after electroporation, 1 mL of YM media (0.4 g/L yeast extract, 10.0 g/L mannitol, 0.1 g/L NaCl, 0.2 g/L $MgSO_4 \bullet 7H_2O$, 0.5 g/L $K_2HPO_4 \bullet 3H_2O$) was added to the cells followed by incubation at 28°C for 3 hours with shaking. The media was sterilized by autoclaving. For antibiotic selection, 100 µg/mL streptomycin and 50 µg/mL kanamycin were filter sterilized using a 0.22 µm syringe filter (Millipore) and added to the cool media. The cells were plated on the selective media and incubated for 2 days at 28°C. After the 2 day incubation period individual colonies were picked with a sterile toothpick, placed in 5 mL of liquid YM-kanamycin/streptomycin media and grown overnight at 28°C. Selective media and sequencing (data not shown) were used to confirm positive transformants. Glycerol stocks were created and stored at -80°C.

3.18 Transient Plant Transformation

Agrobacterium tumefaciens cultures from above (section 3.17) were grown from stocks in YM-kanamycin/streptomycin liquid media. The optical density was monitored at 600 nm

(OD_{600}) using a spectrophotometer (Smart Spec Plus, Bio-Rad Laboraties) until an OD_{600} of 0.20 was reached (Sparkes et al., 2006). At this time 1.5 mL of culture was centrifuged at 8,000 rpm for 5 min at room temperature to pellet the bacteria. The pelleted bacteria was suspended in 1 mL of infiltration buffer (2 mM $Na_3(PO_4)$, 50 mM MES, 100 μM acetosyringone and 50 mg/mL glucose). The suspensions were then injected into the epidermal region of mature tobacco leaves. Transient expression of the ICDH:YFP fusion proteins in tobacco epidermal cells required 2-3 days from infiltration to expression depending on the efficiency of the transformation. After the incubation period the leaves from the transfected tobacco were visualized using the confocal microscope (see section 3.19). *Arabidopsis* Sec12, a ER protein N-terminally fused to YFP was provided as an *A. tumefaciens* stock by Dr. Federica Brandizzi (DOE Plant Research Laboratory, Michigan State University) and also used to perform transient plant transformation as described above (daSilva et al., 2004).

3.19 Confocal Microscopy

Epidermal leaf cells of tobacco were transiently transfected with *A. tumefaciens* cultures harboring the pYellow reporter vector encoding either the mito-160:YFP, chloro-148:YFP, cyto-90:YFP, mito-110:YFP, chloro-98:YFP or cyto-40:YFP fusion proteins. After 48-72 hours the expression of the YFP was visualized using confocal microscopy (Department of Biology, University of Saskatchewan). Leaf tissues were examined by using an inverted Zeiss LSM 510 META laser scanning microscope (Jena, Germany) with a 63X water immersion objective. For imaging of the YFP constructs, the excitation line of an argon ion laser at 514 nm was used with a 530-600 nm band pass filter in the single-track facility of the microscope (Brandizzi et al., 2002*b*).

Mitochondria were visualized by incubating leaf disks in the presence of 50 mM Hepes (pH 7.0), 330 mM sorbitol, and 1 mM dihydrorhodamine-123 (Invitrogen) at room temperature for 1 hour (as described Gálvez et al., 1998). Excess dye was eliminated by washing the disks several times in dye-free buffer. Fluorescence from dihydrorhodamine-123 was observed by using the excitation line of an argon ion laser at 488 nm with a 505-530 nm band pass filter. Chloroplasts were identified by using chlorophyll autofluorescence which was observed using the excitation line of an argon ion laser at 488 nm with a long pass filter at 650 nm. YFP (yellow channel) and chlorophyll autofluorescence (red channel) were imaged simultaneously. Line switching was done using the multi-track facility of the microscope and settings were optimized for individual constructs. Acquired images were processed with Zeiss LSM Image Browser software. Post acquisition image processing was perfomred using Microsoft PowerPoint (Microsoft; Missassuaga, ON, Canada). Images reported in microscopy figures are representative of at least five independent experiments.

4.0 RESULTS

The following results reported in this chapter were obtained using the Xanthi cultivar of tobacco (*Nicotiana tabacum* L. cv. Xanthi). All results were repeated using *Nicotiana tabacum* L. cv. Petit Havana with identical findings. These results can be found in Appendix B.

4.1 Identification of Targeting Signals

The protein sequence of the mitochondrial isoform of ICDH demonstrated two N-terminal targeting signals when examined by bioinformatic analysis (Predotar). When the full protein amino acid sequence was examined with Predotar, the prediction software showed a 98% probability towards a mitochondrial targeting signal because it was enriched in basic residues at the N-terminal end. When the truncated amino acid sequence from the second start site was entered into Predotar software it was predicted to be targeted to the chloroplast (99%). However, it is not possible to assign unequivocally the sub-cellular localization of the mature protein based on this kind of theoretical analysis.

4.2 Generation of Full-Length Constructs

A PCR-based approach was utilized to directionally clone specific regions of the mitochondrial $NADP^+$-ICDH cDNA into pBluescript to examine the role of the putative targeting signals (Fig. 4.1A). In total, three individual fragments were generated (Fig. 4.1B). These fragments encompass the entire coding region of the protein with altered targeting signals. The first fragment was 1446 bp in size and spanned the coding region of the mature protein from the first translational start site which

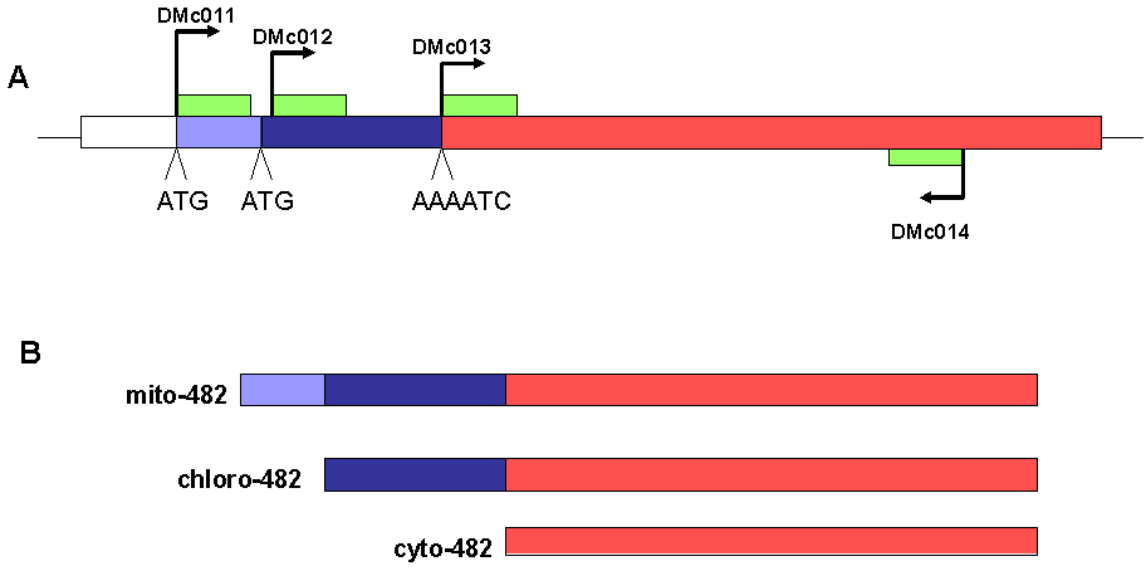

Figure 4.1 PCR strategy to generate full-length versions of $NADP^+$-ICDH with altered targeting signals. **A)** Graphical representation of $NADP^+$-ICDH nucleotide sequence. Primers are indicated by green sections, blue and light blue sections indicate predicted targeting regions, red section indicates coding region of mature protein, white section indicates untranslated region. **B)** Representation of possible PCR products using the primer sets described in Table 3.1.

included both the putative mitochondrial and chloroplast targeting signals. Upon translation this generated a protein of 482 amino acids in length that should be targeted to the mitochondrion. This product was designated mito-482 (Fig. 4.1B). The second fragment was 1410 bp in size and spanned the coding region of the mature protein from the second translation start site which included only the putative chloroplast targeting signal. This fragment was designated chloro-482 and should be presumably targeted to the chloroplast (Fig. 4.1B). The third fragment was 1233 bp in size, spanning the coding region of the mature protein with no putative targeting signals. An ATG translational start site was added during PCR primer design. This fragment was expected to be localized in the cytosol and was designated cyto-482 (Fig. 4.1B).

All three sets of primers amplified robustly resulting in distinct DNA products at approximately 1446, 1410 and 1233 bp when analyzed by agarose gel electrophoresis (Fig. 4.2A). Following digestion and purification, these DNA fragments (mito-482, chloro-482 and cyto-482) were directionally ligated into the Bluescript plasmid. Successful ligation and transformation was confirmed by restriction digestion with EcoRI/HindIII. This generated DNA products of the correct size as shown in Fig. 4.2B.

4.3 Sequence Analyses

The three pBluescipt plasmids harboring the PCR fragments (created in section 4.2) were sent for sequencing at National Research Council of Canada - Plant Biotechnology Institute (Saskatoon, Saskatchewan). The nucleotide sequences of the plasmid inserts were then aligned with *Nicotiana tabacum* L. cv. Xanthi mitochondrial $NADP^+$-ICDH to identify any nucleotide changes (Fig. 4.3). This sequence analysis demonstrated a 99.1% homology of the

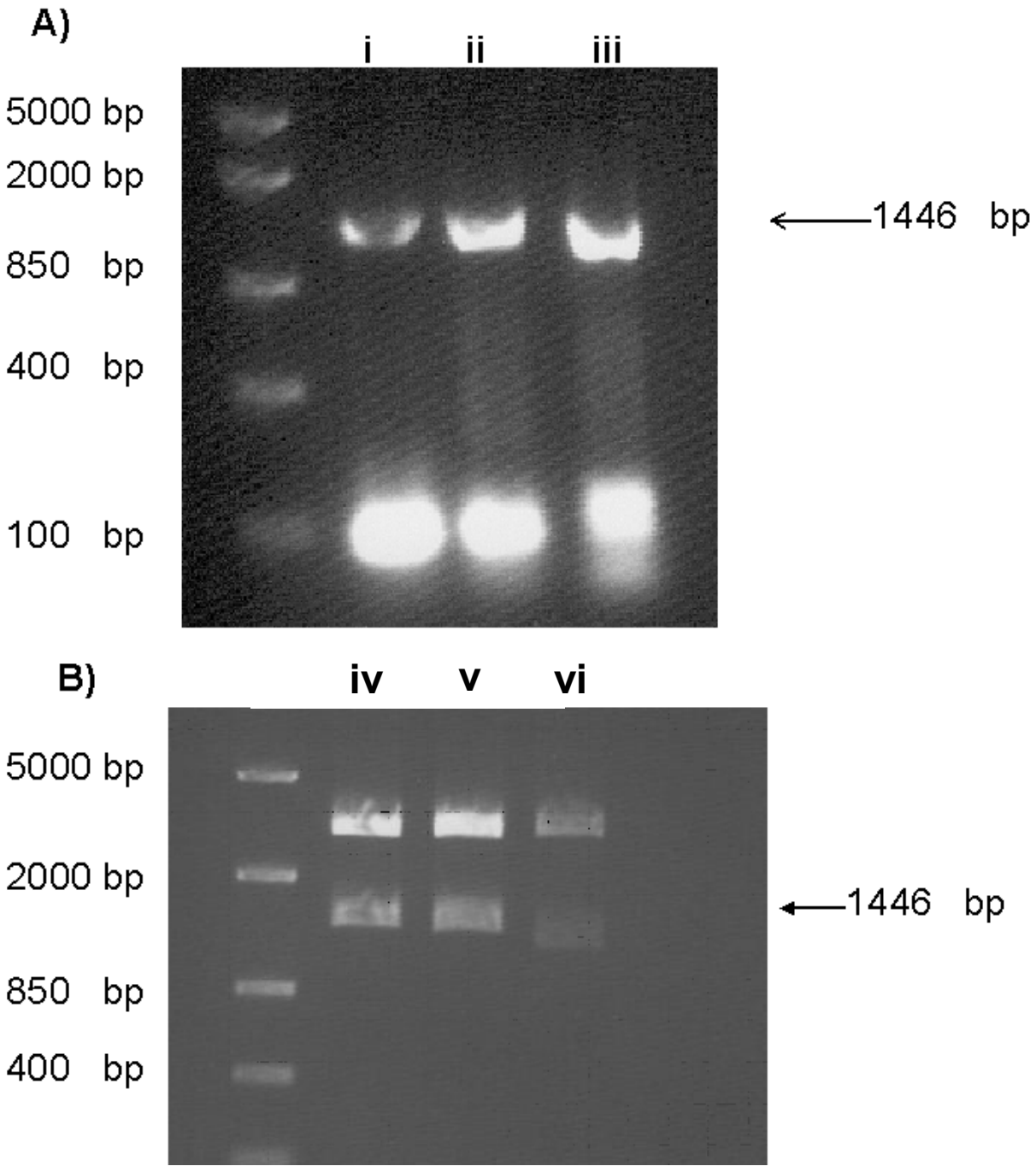

Figure 4.2 Agarose gel analyses of full-length PCR products of $NADP^+$-ICDH from *N. tabacum* L. cv Xanthi. **A)** PCR amplification of **i)** mito-482 **ii)** chloro-482 and **iii)** cyto-482 DNA fragments. **B)** Double restriction digested pBluescript containing the DNA fragments **iv)** mito-482, **v)** chloro-482 and **vi)** cyto-482 using EcoR1 and HindIII.

```
mito-482        ------------------------------------------------------------
chloro-482      ------------------------------------------------------------
cyto-482        ------------------------------------------------------------
NCBI            GCACGGCGACAAAGACAATAGAGAGTTGCTCAGAGCAGCAAAAACTAGCAACTGATTAAA 60


mito-482        ---------------ATGCTTACCACCCGACTCAGACTCCGGTGTTCCGCCATGGCTAGT 45
chloro-482      ------------------------------------------------ATGGCTAGT 9
cyto-482        ------------------------------------------------------------
NCBI            GCCTACAAAAGTGTTATGCTTACCACCCGACTCAGACTCCGGTGTTCCGCCATGGCTAGT 120


mito-482        GTTGCTTCTTTTATCTCATCTTCATCGGCTTCAACATCATCCGCAGTTACCAAAAACCTT 105
chloro-482      GTTGCTTCTTTTATCTCATCTTCATCGGCTTCAACATCATCCGCAGTTACCAAAAACCTT 69
cyto-482        ------------------------------------------------------------
NCBI            GTTGCTTCTTTTATCTCATCTTCATCGGCTTCAACATCATCCGCAGTTACCAAAAACCTT 180


mito-482        CCCTTTTCAATCATCTCCAATCGGCAACTGTTCAAGAACCGTGTTTATCTCCTCCACCGA 165
chloro-482      CCCTTTTCAATCATCTCCAATCGGCAACTGTTCAAGAACCGTGTTTATCTCCTCCACCGA 129
cyto-482        ------------------------------------------------------------
NCBI            CCCTTTTCCATCATTTCCAATCGGCAACTGTTCAAGAACCGTGTTTATCTCCTCCACCGA 240


mito-482        ATCCCCAATGCTTCAATTCGATGCTTCGCTTCCACTACAGCTTCGTCTAAAATCCGCGTC 225
chloro-482      ATCCCCAATGCTTCAATTCGATGCTTCGCTTCCACTACAGCTTCGTCTAAAATCCGCGTC 189
cyto-482        -------------------------------------------ATGTCTAAAATCCGCGTC 15
NCBI            ATCCCCAATGCTTCAATTCGATCGTTCGCTTCCACTACAGCTTCGTCTAAAATCCGCGTT 300
                                                           **************


mito-482        GAAAATCCTATTGTCGAAATGGACGGTGATGAAATGACGAGGGTTATATGGACAATGATC 285
chloro-482      GAAAATCCTATTGTCGAAATGGACGGTGATGAAATGACGAGGGTTATATGGACAATGATC 249
cyto-482        GAAAATCCTATTGTCGAAATGGACGGTGATGAAATGACGAGGGTTATATGGACAATGATC 75
NCBI            GAAAATCCTATTGTCGAAATGGACGGTGATGAAATGACGAGGGTTATATGGACAATGATC 360
                ************************************************************


mito-482        AAAGAGAAGCTAATATATCCTTATCTAGAGTTGGATACGAAGTATTACGATTTGGGGATA 345
chloro-482      AAAGAGAAGCTAATATATCCTTATCTAGAGTTGGATACGAAGTATTACGATTTGGGGATA 309
cyto-482        AAAGAGAAGCTAATATATCCTTATCTAGAGTTGGATACGAAGTATTACGATTTGGGGATA 135
NCBI            AAAGAGAAGCTAATATATCCTTATCTAGAGTTGGATACGAAGTATTACGATTTGGGGATA 420
                ************************************************************


mito-482        TTGAACCGTGATGCCACTGACGATCAAGTTACTGTTGAAAGTGCTGAGGCTACTCTTAAG 405
chloro-482      TTGAACCGTGATGCCACTGACGATCAAGTTACTGTTGAAAGTGCTGAGGCTACTCTTAAG 369
cyto-482        TTGAACCGTGATGCCACTGACGATCAAGTTACTGTTGAAAGTGCTGAGGCTACTCTTAAG 195
NCBI            TTGAACCGTGATGCCACTGACGATCAAGTTACTGTTGAAAGTGCTGAGGCTACTCTTAAG 480
                ************************************************************


mito-482        TATAATGTTGCTGTGAAATGCGCTACTATAACACCTGATGAGACCAGAGTTAAGGAATTT 465
chloro-482      TATAATGTTGCTGTGAAATGCGCTACTATAACACCTGATGAGACCAGAGTTAAGGAATTT 429
cyto-482        TATAATGTTGCTGTGAAATGCGCTACTATAACACCTGATGAGACCAGAGTTAAGGAATTT 255
NCBI            TATAATGTTGCTGTGAAATGCGCTACTATAACACCTGATGAGACCAGAGTTAAGGAATTT 540
                ************************************************************


mito-482        GGGCTGAAGTCTATGTGAAGAAGTCCCAATGGCACAATCAGAAACATTTTAAATGGTACT 525
chloro-482      GGGCTGAAGTCTATGTGAAGAAGTCCCAATGGCACAATCAGAAACATTTTAAATGGTACT 489
cyto-482        GGGCTGAAGTCTATGTGAAGAAGTCCCAATGGCACAATCAGAAACATTTTAAATGGTACT 315
NCBI            GGGCTGAAGTCTATGTGGAGAAGTCCCAATGCCACAATCAGAAACATTTTAAATGGTACT 600
                ***************** ************* ****************************


mito-482        GTTTTCCGGGAGCCTATACTATGCAAGAACGTCCCCAGAATTGTTCCTGGTTGGGAGAAA 585
chloro-482      GTTTTCCGGGAGCCTATACTATGCAAGAACGTCCCCAGAATTGTTCCTGGTTGGGAGAAA 549
cyto-482        GTTTTCCGGGAGCCTATACTATGCAAGAACGTCCCCAGAATTGTTCCTGGTTGGGAGAAA 375
NCBI            GTTTTCCGGGAGCCTATACTATGCAAGAACGTCCCCAGAATTGTTCCTGGTTGGAAGAAA 660
                ************************************************** *****


mito-482        CCCATTTGTATTGGTAGGCATGCTTTTGGTGACCAGTATCGTGCCACAGATGCAGTTATT 645
chloro-482      CCCATTTGTATTGGTAGGCATGCTTTTGGTGACCAGTATCGTGCCACAGATGCAGTTATT 609
cyto-482        CCCATTTGTATTGGTAGGCATGCTTTTGGTGACCAGTATCGTGCCACAGATGCAGTTATT 435
NCBI            CCCATTTGTATTGGTAGGCATGCCTTTGGTGACCAGTATCGTGCCACAGATGCAGTTATT 720
                ********************** ***********************************
```

```
mito-482        AATGGACCAGGAAAGCTCAAAATGGTTTTTGAGCCAGAAAATGGGGAAGCCCCTACGGAA 705
chloro-482      AATGGACCAGGAAAGCTCAAAATGGTTTTTGAGCCAGAAAATGGGGAAGCCCCTACGGAA 669
cyto-482        AATGGACCAGGAAAGCTCAAAATGGTTTTTGAGCCAGAAAATGGGGAAGCCCCTACGGAA 495
NCBI            AATGGACCAGGAAAGCTCAAAATGGTTTTTGAGCCAGAAAATGGGGAAGCCCCTACGGAA 780
                ************************************************************

mito-482        CTGGATGTTTATGATTTTAAAGGTCCAGGTGTTGCACTTGCCATGTACAATGTTGACCAG 765
chloro-482      CTGGATGTTTATGATTTTAAAGGTCCAGGTGTTGCACTTGCCATGTACAATGTTGACCAG 729
cyto-482        CTGGATGTTTATGATTTTAAAGGTCCAGGTGTTGCACTTGCCATGTACAATGTTGACCAG 555
NCBI            CTGGATGTTTATGATTTTAAAGGTCCAGGTGTTGCACTTGCCATGTACAATGTTGACCAG 840
                ************************************************************

mito-482        TCAATTCGAGCGTTTGCTGAATCATCAATGTCAATGGTATTTTCGAAGAAATGGCCTCTT 825
chloro-482      TCAATTCGAGCGTTTGCTGAATCATCAATGTCAATGGTATTTTCGAAGAAATGGCCTCTT 789
cyto-482        TCAATTCGAGCGTTTGCTGAATCATCAATGTCAATGGTATTTTCGAAGAAATGGCCTCTT 615
NCBI            TCAATTCGAGCGTTTGCTGAATCATCAATGTCAATGGCATTTTCGAAGAAATGGCCTCTT 900
                ********************************** **********************

mito-482        TATTTGAGTACAAAAAATACAATACTAAAGAAATACGATGGCAGGTTTAAGGACATTTTT 885
chloro-482      TATTTGAGTACAAAAAATACAATACTAAAGAAATACGATGGCAGGTTTAAGGACATTTTT 849
cyto-482        TATTTGAGTACAAAAAATACAATACTAAAGAAATACGATGGCAGGTTTAAGGACATTTTT 675
NCBI            TATTTGAGTACAAAAAATACAATACTAAAGAAATACGATGGCAGGTTTAAGGACATTTTT 960
                ************************************************************

mito-482        GAAGAGGTATATGAAGAGAAGTGGAAGCAACAGTTTGAGGAACACTCGATATGGTATGAG 945
chloro-482      GAAGAGGTATATGAAGAGAAGTGGAAGCAACAGTTTGAGGAACACTCGATATGGTATGAG 909
cyto-482        GAAGAGGTATATGAAGAGAAGTGGAAGCAACAGTTTGAGGAACACTCGATATGGTATGAG 735
NCBI            GAAGAGGTATATGAAGAGAAGTGGAAGCAACAGTTTGAGGAACACTCGATATGGTATGAG 1020
                ************************************************************

mito-482        CATAGATTGATAGATGACATGGTAGCTTATGCATTAAAAAGCGGGGGTGGATATGTTTGG 1005
chloro-482      CATAGATTGATAGATGACATGGTAGCTTATGCATTAAAAAGCGGGGGTGGATATGTTTGG 969
cyto-482        CATAGATTGATAGATGACATGGTAGCTTATGCATTAAAAAGCGGGGGTGGATATGTTTGG 795
NCBI            CATAGATTGATAGATGACATGGTAGCTTATGCATTAAAAAGCGGGGGTGGATATGTTTGG 1080
                ************************************************************

mito-482        GCATGCAAGAACTATGATGGAGATGTCCAGAGTGATCTGCTCGCTCAAGGATTTGGTTCT 1065
chloro-482      GCATGCAAGAACTATGATGGAGATGTCCAGAGTGATCTGCTCGCTCAAGGATTTGGTTCT 1029
cyto-482        GCATGCAAGAACTATGATGGAGATGTCCAGAGTGATCTGCTCGCTCAAGGATTTGGTTCT 855
NCBI            GCATGCAAGAACTATGATGGAGATGTCCAGAGTGATCTGCTCGCTCAAGGATTTGGTTCT 1140
                ************************************************************

mito-482        CTGGGCCTCATGACCTCTGTATTGTTATCTTCTGATGGCAAGACATTAGAAGCTGAAGCA 1125
chloro-482      CTGGGCCTCATGACCTCTGTATTGTTATCTTCTGATGGCAAGACATTAGAAGCTGAAGCA 1089
cyto-482        CTGGGCCTCATGACCTCTGTATTGTTATCTTCTGATGGCAAGACATTAGAAGCTGAAGCA 915
NCBI            CTGGGCCTCATGACCTCTGTATTGTTATCTTCTGATGGCAAGACATTAGAAGCTGAAGCA 1200
                ************************************************************

mito-482        GCTCATGGCACAGTAACCAGACATTTTCGGCTGCATCAAAAGGGTCAAGAAACTAGTACA 1185
chloro-482      GCTCATGGCACAGTAACCAGACATTTTCGGCTGCATCAAAAGGGTCAAGAAACTAGTACA 1149
cyto-482        GCTCATGGCACAGTAACCAGACATTTTCGGCTGCATCAAAAGGGTCAAGAAACTAGTACA 975
NCBI            GCTCATGGCACAGTAACCAGACATTTTCGGCTGCATCAAAAGGGTCAAGAAACTAGTACA 1260
                ************************************************************

mito-482        AATAGTGCTGCTTCTATTTTTGCATGGGCAAGGGGACTTGGACATAGGGCCCAGCTTGAT 1245
chloro-482      AATAGTGCTGCTTCTATTTTTGCATGGGCAAGGGGACTTGGACATAGGGCCCAGCTTGAT 1209
cyto-482        AATAGTGCTGCTTCTATTTTTGCATGGGCAAGGGGACTTGGACATAGGGCCCAGCTTGAT 1035
NCBI            AATAGTGTTGCTTCTATTTTTGCATGGGCAAGGGGACTTGGACATAGGGCCCAGCTTGAT 1320
                ******* ****************************************************

mito-482        GGGAACCAAAAGTTATCTGAATTTGTTCACGCCCTGGGAGCTGCTTGCGTTGGCACAATA 1305
chloro-482      GGGAACCAAAAGTTATCTGAATTTGTTCACGCCCTGGGAGCTGCTTGCGTTGGCACAATA 1269
cyto-482        GGGAACCAAAAGTTATCTGAATTTGTTCACGCCCTGGGAGCTGCTTGCGTTGGCACAATA 1095
NCBI            GGGAACCAAAAGTTATCTGAATTTGTTCACGCCCTGGAAGCTGCTTGCGTTGGCACAATA 1380
                ********************************** **********************

mito-482        GAGTCCGGGAAGATGACTAAGGATTTAGCTATATTGGTTCATGGACCCAAGGTATCAAGG 1365
chloro-482      GAGTCCGGGAAGATGACTAAGGATTTAGCTATATTGGTTCATGGACCCAAGGTATCAAGG 1329
cyto-482        GAGTCCGGGAAGATGACTAAGGATTTAGCTATATTGGTTCATGGACCCAAGGTATCAAGG 1155
NCBI            GAGTCCGGGAAGATGACTAAGGATTTAGCTATATTGGTTCATGGACCCAAGGTATCAAGG 1440
```

```
                ************************************************************

mito-482        GAACACTACTTGAATACTGAAGAATTTATTGATGCTGTAGCACAGAAACTTCAAGAGAAG 1425
chloro-482      GAACACTACTTGAATACTGAAGAATTTATTGATGCTGTAGCACAGAAACTTCAAGAGAAG 1389
cyto-482        GAACACTACTTGAATACTGAAGAATTTATTGATGCTGTAGCACAGAAACTTCAAGAGAAG 1215
NCBI            GAACACTACTTGAATACTGAAGAATTTATTGATCCTGTAGCACAGAAACTTCAAGAGAAG 1500
                ******************************* ************************

mito-482        CTCGGTGCCTGCGCAGTTGTA 1446
chloro-482      CTCGGTGCCTGCGCAGTTGTA 1410
cyto-482        CTCGGTGCCTGCGCAGTTGTA 1236
NCBI            CTCGGTGCCTGCGCAGTTGTA 1521
                *********************
```

Figure 4.3 ClustalW alignment of nucleotide sequence from full-length constructs of $NADP^+$-ICDH generated from *N. tabacum* L. cv Xanthi. Nucleotide sequence from *N. tabacum* L. cv Xanthi $NADP^+$-ICDH (GenBank accession X96728.1) was aligned with mito-482, chloro-482 and cyto-482. Differing nucleotide residues and the introduced ATG site in cyto-482 are highlighted in yellow, * indicates matching sequence.

three constructs to the expected nucleotide sequence. The alignment showed 13 changes at the nucleotide level which were present in all three constructs (Fig. 4.3).

In order to examine whether these 13 nucleotide changes had an effect on the protein, the sequences were translated and the deduced amino acid sequence aligned with the amino acid sequence present in the NCBI database (Fig. 4.4). It was found that eight of these 13 nucleotide changes resulted in changes at the amino acid level which are summarized in Table 4.1. This was consistent for all three constructs (Fig. 4.4) and resulted in a 99.1% identity and a 98.3% similarity with the $NADP^+$-ICDH protein. There was one amino acid change in the putative chloroplastic targeting signal (a serine to a cysteine). When examined using the Predotor software this change did not affect the putative mitochondrial or chloroplastic localization of the fragment. However, the most significant of these changes was at amino acid 161 where a tryptophan (nucleotide sequence TGG) had been replaced with a stop codon (nucleotide sequence TGA; Table 4.1).

4.4 Generation of Truncated Constructs

Based on the discovery of a stop codon at amino acid 161, another PCR strategy was designed which took advantage of the original 5' forward primers, but a new 3' primer was created to amplify the coding region of $NADP^+$-ICDH to just before the stop codon at amino acid 161 (Fig. 4.5A). This created a series of three fragments encoding truncated proteins with the same putative mitochondrial, chloroplastic and cytosolic targeting signals as above (Fig. 4.5B). An additional 3' primer was also utilized (Fig. 4.5A) to generate another series of three truncated proteins with the same putative mitochondrial, chloroplastic and cytosolic targeting signals as above (Fig. 4.5B). These proteins were truncated at amino acid 110 and were

```
mito-482      MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60
chloro-482    ------------MASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 48
cyto-482      ------------------------------------------------------------
NCBI          MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60

mito-482      IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
chloro-482    IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 108
cyto-482      ----------MKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 50
NCBI          IRSFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
                        **************************************************

mito-482      TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM#RSPNGTIRNILNGTVFREP 179
chloro-482    TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM#RSPNGTIRNILNGTVFREP 167
cyto-482      TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM#RSPNGTIRNILNGTVFREP 109
NCBI          TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSMWRSPNATIRNILNGTVFREP 180
              ************************************** ****.**************

mito-482      ILCKNVPRIVPGWEKPICIGRHAFGDQYRATDAVINGPGKLKMVFEPENGEAPTELDVYD 239
chloro-482    ILCKNVPRIVPGWEKPICIGRHAFGDQYRATDAVINGPGKLKMVFEPENGEAPTELDVYD 227
cyto-482      ILCKNVPRIVPGWEKPICIGRHAFGDQYRATDAVINGPGKLKMVFEPENGEAPTELDVYD 169
NCBI          ILCKNVPRIVPGWKKPICIGRHAFGDQYRATDAVINGPGKLKMVFEPENGEAPTELDVYD 240
              *************:************************************************

mito-482      FKGPGVALAMYNVDQSIRAFAESSMSMVFSKKWPLYLSTKNTILKKYDGRFKDIFEEVYE 299
chloro-482    FKGPGVALAMYNVDQSIRAFAESSMSMVFSKKWPLYLSTKNTILKKYDGRFKDIFEEVYE 287
cyto-482      FKGPGVALAMYNVDQSIRAFAESSMSMVFSKKWPLYLSTKNTILKKYDGRFKDIFEEVYE 229
NCBI          FKGPGVALAMYNVDQSIRAFAESSMSMAFSKKWPLYLSTKNTILKKYDGRFKDIFEEVYE 300
              ****************************.*********************************

mito-482      EKWKQQFEEHSIWYEHRLIDDMVAYALKSGGGYVWACKNYDGDVQSDLLAQGFGSLGLMT 359
chloro-482    EKWKQQFEEHSIWYEHRLIDDMVAYALKSGGGYVWACKNYDGDVQSDLLAQGFGSLGLMT 347
cyto-482      EKWKQQFEEHSIWYEHRLIDDMVAYALKSGGGYVWACKNYDGDVQSDLLAQGFGSLGLMT 289
NCBI          EKWKQQFEEHSIWYEHRLIDDMVAYALKSGGGYVWACKNYDGDVQSDLLAQGFGSLGLMT 360
              ************************************************************

mito-482      SVLLSSDGKTLEAEAAHGTVTRHFRLHQKGQETSTNSAASIFAWARGLGHRAQLDGNQKL 419
chloro-482    SVLLSSDGKTLEAEAAHGTVTRHFRLHQKGQETSTNSAASIFAWARGLGHRAQLDGNQKL 407
cyto-482      SVLLSSDGKTLEAEAAHGTVTRHFRLHQKGQETSTNSAASIFAWARGLGHRAQLDGNQKL 349
NCBI          SVLLSSDGKTLEAEAAHGTVTRHFRLHQKGQETSTNSVASIFAWARGLGHRAQLDGNQKL 420
              ************************************.***********************

mito-482      SEFVHALGAACVGTIESGKMTKDLAILVHGPKVSREHYLNTEEFIDAVAQKLQEKLGACA 479
chloro-482    SEFVHALGAACVGTIESGKMTKDLAILVHGPKVSREHYLNTEEFIDAVAQKLQEKLGACA 467
cyto-482      SEFVHALGAACVGTIESGKMTKDLAILVHGPKVSREHYLNTEEFIDAVAQKLQEKLGACA 409
NCBI          SEFVHALEAACVGTIESGKMTKDLAILVHGPKVSREHYLNTEEFIDPVAQKLQEKLGACA 480
              ******* ************************************.*************

mito-482      VV 481
chloro-482    VV 469
cyto-482      VV 411
NCBI          VV 482
              **
```

Figure 4.4 ClustalW alignment of deduced amino acid sequence from full-length constructs of NADP$^+$-ICDH generated from *N. tabacum* L. cv Xanthi. Deduced amino acid sequence from *N. tabacum* L. cv Xanthi NADP$^+$-ICDH (NCBI acession CAA65503) was aligned with mito-482, chloro-482 and cyto-482. Differing amino acid residues are highlighted in yellow a stop codon is indicated with a number sign (#),* indicates matching sequence.

Table 4.1 Amino acid changes observed in $NADP^+$-ICDH. Findings of this study were compared to the NCBI database sequence $NADP^+$-ICDH (accession CAA65503). The Xanthi cultivar of tobacco (*N. tabacum* L.) was used in both cases. * indicates conservative amino acid changes, ** indicates semi-conservative amino acid changes.

Amino Acid (database)	**Amino Acid (this study)**	**Amino Acid Position**
S (Serine)	C (Cysteine)*	63
W (Tryptophan)	Stop codon (TGA)	161
A (Alaine)	G (Glycine) *	166
K (Lysine)	E (Glutamic Acid) **	194
A (Alanine)	V (Valine)*	268
V (Valine)	A (Valine)*	398
E (Glutamic Acid)	G (Glycine)*	427
P (Proline)	A (Alaine)*	467

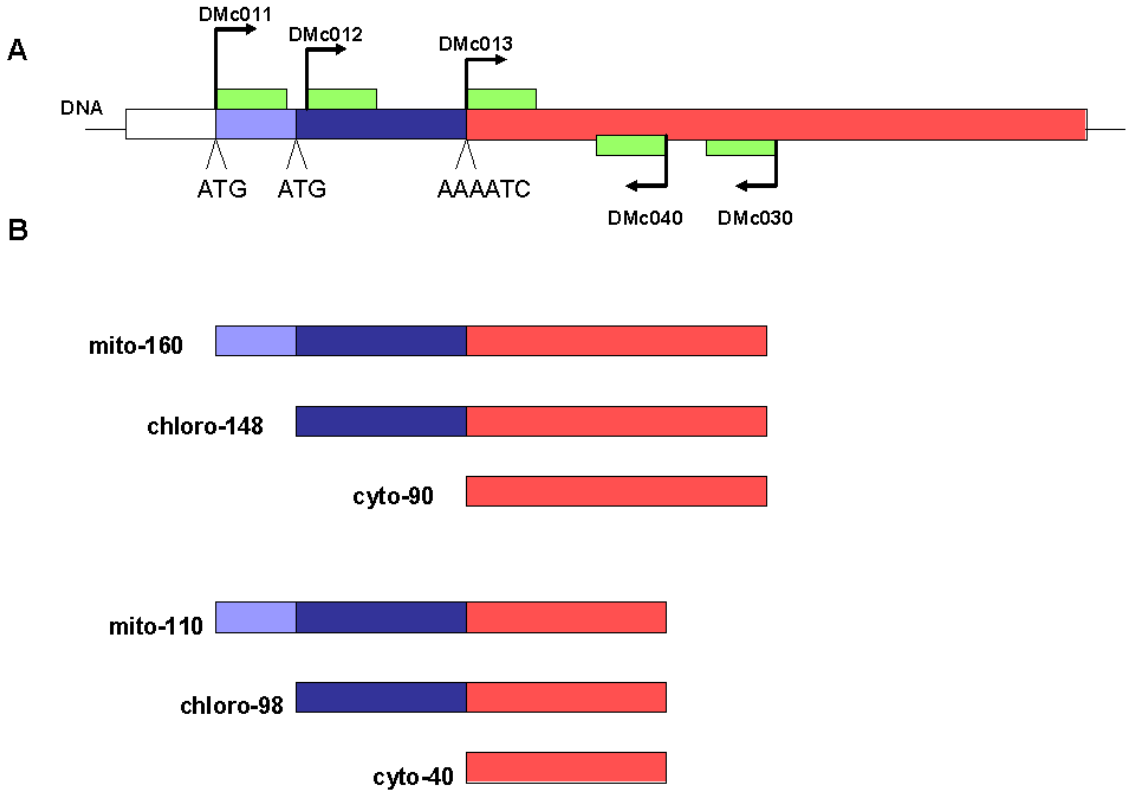

Figure 4.5 PCR strategy to generate truncated versions of NADP$^+$-ICDH with altered targeting signals. **A)** Graphical representation of NADP$^+$-ICDH nucleotide sequence. Primers are indicated by green sections, blue and light blue sections indicate predicted targeting regions, red section indicates coding region, white section indicates untranslated region. **B)** Representation of possible PCR products using the primer sets described in Table 3.2.

designed in this manner so as in part reproduce published results (Gálvez et al., 1998).

4.4.1 Constructs Truncated at Amino Acid 160

Constructs tructated at amino acid 160 were designed in order to circumvent the stop codon reported in section 4.3. The first fragment was 480 bp in size and spanned the coding region of the mature protein from the first translational start site which included both the putative mitochondrial and chloroplast targeting signals, but was truncated before the stop codon at amino acid 161. Upon translation this generated a protein predicted to be 160 amino acids in length that was hypothesized to be targeted to the mitochondrion. This fragment was designated mito-160 (Fig. 4.5B). The second fragment was 444 bp in size, spanning the coding region of the mature protein from the second ATG start site which included only the putative chloroplast targeting signal, but was truncated before a stop codon at amino acid 161. This fragment was designated chloro-148 and is predicted to be targeted to the chloroplast (Fig. 4.5B). The third fragment was 270 bp in size and spanned the coding region of the mature protein with no putative targeting signals, but was truncated before a stop codon at amino acid 161. An ATG translational start site was added during the PCR primer design. This fragment is expected to be localized in the cytosol and was designated cyto-90 (Fig. 4.5B).

All three sets of primers amplified robustly resulting in a distinct DNA products at approximately 483, 447, 270 bp when analyzed by agarose gel electrophoresis (Fig. 4.6A). Following digestion and purification, these fragments were used for further ligation into pBluescript resulting in the creation of mito-160, chloro-148 and cyto-90 constructs. Successful ligation and transformation was confirmed by restriction digestion with EcoR1/HindIII. This generated DNA products of the correct size as shown in Fig. 4.6B. This was also confirmed by

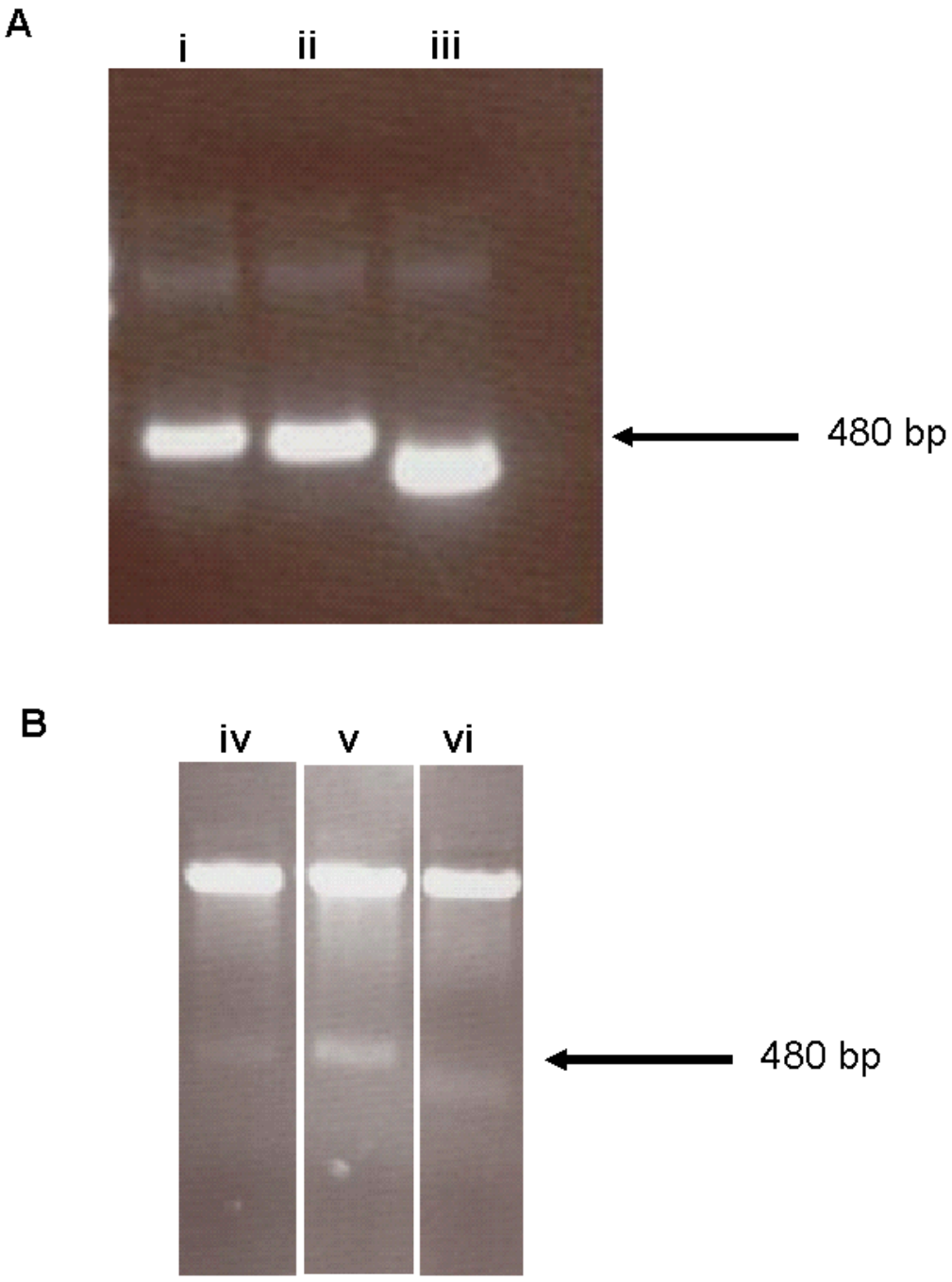

Figure 4.6 Agarose gel analyses of PCR products from *N. tabacum* L. cv Xanthi resulting in truncation at amino acid 160 of $NADP^+$-ICDH. **A)** PCR amplification of **i)** mito-160, **ii)** chloro-148 and **iii)** cyto-90 DNA fragments. **B)** Double restriction digested pBluescript containing the DNA fragments **iv)** mito-160, **v)** chloro-148 and **vi)** cyto-90 using EcoR1 and HindIII.

nucleotide sequence analysis (data not shown). Translation and alignment of the deduced amino acid sequence (Fig. 4.7A) demonstrated identical results as to those presented in Fig. 4.4 with the exception of truncation at amino acid 160 as expected.

4.4.2 Constructs Truncated at Amino Acid 110

Constructs tructated at amino acid 110 were also designed in order to replicate and build on results reported by Gálvez et al. (1998). The first fragment was 330 bp in size and spanned the coding region of the mature protein from the first ATG start site, which included both the putative mitochondrial and chloroplast targeting signals, but was truncated at amino acid 110. Upon translation, this generated a protein predicted to be 110 amino acids in length that should be targeted to the mitochondrion. This fragment was designated mito-110 (Fig 4.5B). The second fragment was 294 bp in size, spanning the coding region of the mature protein from the second ATG start site which included only the putative chloroplast targeting signal, but was truncated at amino acid 110. This fragment was designated chloro-98 and should be presumably targeted to the chloroplast (Fig 4.5B). The third fragment was 120 bp in size and spanned only the coding region of the mature protein with no putative targeting signals, but was truncated at amino acid 110. An ATG translational start site was added during PCR primer design. This fragment is expected to be localized in the cytosol and was designated cyto-40 (Fig 4.5B).

All three sets of primers amplified robustly resulting in a distinct DNA products at approximately 330, 294, 120 bp when analyzed by agarose gel electrophoresis (Fig. 4.8A). Following digestion and purification, these fragments were used for further ligation into pBluescript resulting in the creation of mito-110, chloro-98 and cyto-40 constructs. Successful

A

```
mito-160        MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60
chloro-148      ------------MASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 48
cyto-90         ------------------------------------------------------------
NCBI            MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60


mito-160        IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
chloro-148      IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 108
cyto-90         ----------MKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 50
NCBI            IRSFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
                          ************************************************

mito-160        TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM-------------------- 160
chloro-148      TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM-------------------- 148
cyto-90         TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM-------------------- 90
NCBI            TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSMWRSPNATIRNILNGTVFREP 180
                ************************************
```

B

```
mito-110        MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60
chloro-98       ------------MASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 48
cyto-40         ------------------------------------------------------------
NCBI            MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60

mito-110        IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKY---------- 110
chloro-98       IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKY---------- 98
cyto-40         ----------MKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKY---------- 40
NCBI            IRSFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
                          ************************************
```

Figure 4.7 ClustalW alignment of deduced amino acid sequence from truncated constructs of $NADP^+$-ICDH generated from *N. tabacum* L. cv Xanthi. Deduced amino acid sequence from *N. tabacum* L. cv Xanthi $NADP^+$-ICDH (NCBI acession CAA65503) was aligned **A**) mito-160, chloro-148 and cyto-90 and **B**) mito-110, chloro-98 and cyto-40. Differing amino acid residues are highlighted in yellow, ***** indicates matching sequence. The constructs in **B**) are designed, in part, based on Gálvez et al. (1998).

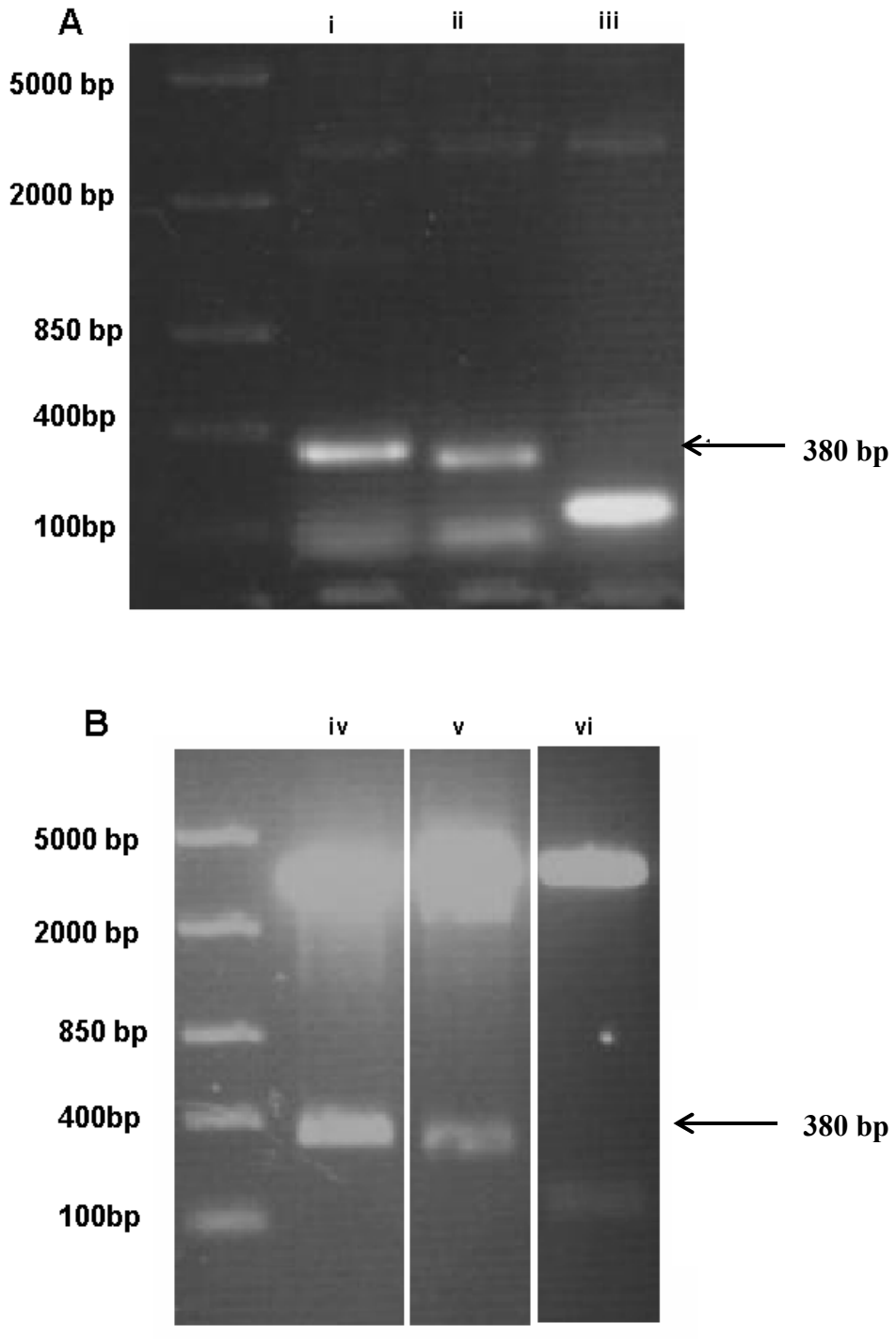

Figure 4.8 Agarose gel analyses of PCR products from *N. tabacum* L. cv Xanthi resulting in truncation at amino acid 138 of $NADP^+$-ICDH. **A)** PCR amplification of **i)** mito-110, **ii)** chloro-98 and **iii)** cyto-40 DNA fragments. **B)** Double restriction digested pBluescript containing the DNA fragments **iv)** mito-110, **v)** chloro-98 and **vi)** cyto-40 using EcoR1 and HindIII.

ligation and transformation was confirmed by restriction digestion with EcoRI/HindIII. This generated DNA products of the correct size as shown in Fig. 4.7B. This was also confirmed by nucleotide sequence analysis (data not shown). Translation and alignment of the deduced amino acid sequence (Fig. 4.7B) demonstrated identical results as to those presented in Fig. 4.4 with the exception of truncation at amino acid 110 as expected.

4.5 Creation of the pYellow Reporter Vector

Confocal microscopy studies required the creation of a reporter vector containing a YFP. This was required for visualization of the constructs in the mitochondria, chloroplasts and cytosol so as not to interfere with the fluorescence of the cellular markers for these compartments to be used in the imaging process.

4.5.1 PCR Amplification of the YFP

The binary vector pVKHI8En6-ERD2-YFP was used as a template along with the primer set YFPf/YFPr for PCR (Fig. A-2; Table 3.3). The PCR reaction was analyzed by agarose gel electrophoresis and demonstrated that the primer set amplified a DNA fragment of approximately 720 bp, corresponding to the YFP (Fig. 4.9; lane i). This PCR product was then digested with HindIII/Spe1, gel purified and used for the creation of pYellow.

4.5.2 pYellow Vector Construction

The pRed binary vector (pRed-L23a-GST-GFP; Degenhard and Bonham-Smith, 2008) (Fig. A-3) was used to generate pYellow. This vector (pRed) consists of the pGreen backbone which has been engineered with a RFP flanked by a 5' HindIII and a 3' Spe1 restriction site

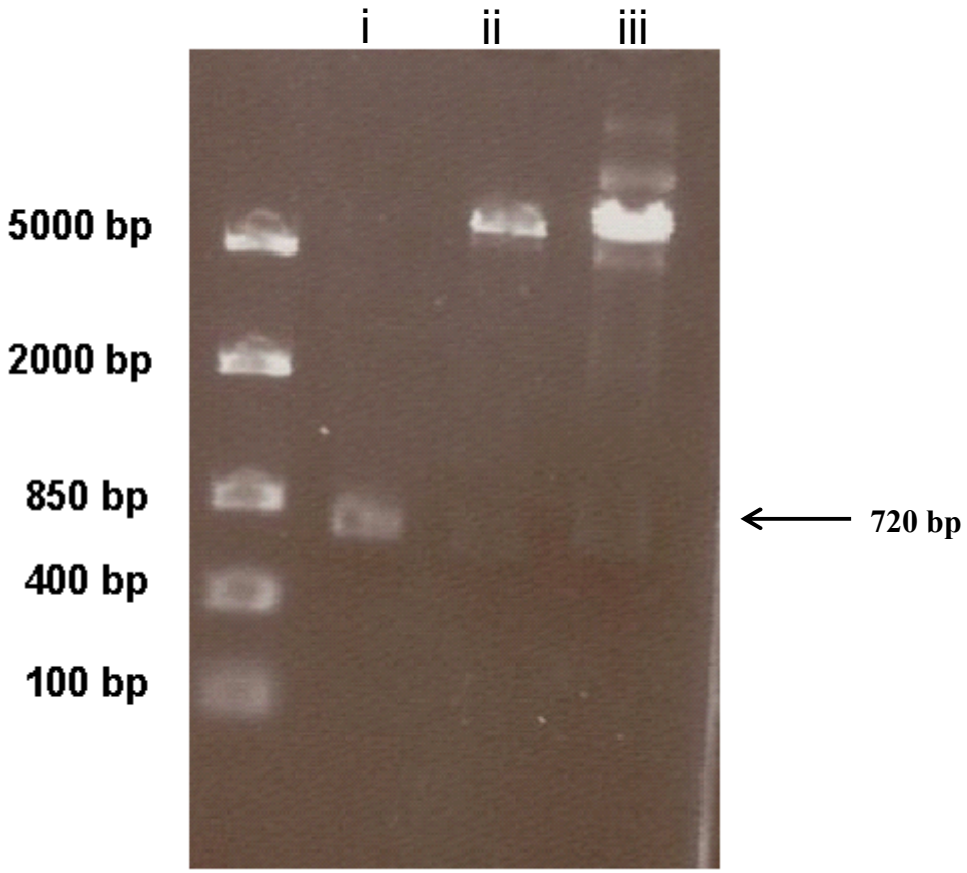

Figure 4.9 Agarose gel analyses of components used to construct the pYellow reporter vector. **i)** PCR amplification of YFP using primer set YFPf/YFPr. **ii)** Double restriction digest of pRed using the enzymes HindIII and Spe1. **iii)** Double restriction digestion of pYellow using the enzymes HindIII and Spe1.

(Fig. A-3). The pRed vector was subjected to a HindIII/Spe1 digestion to release the 715 bp RFP (Fig. 4.9; lane ii). The digestion was analyzed by agarose gel electrophoresis and the linearized plasmid gel purified.

The YFP PCR product previously generated was then directionally ligated into the digested pRed vector, thus generating pYellow (Fig. A-4). Following transformation, positive selection and plasmid DNA extraction, the presence of the YFP was confirmed by HindIII/Spe1 digestion and subsequent analysis by agarose gel electrophoresis (Fig. 4.9; lane iii). The digestion resulted in a band at 720 bp, the correct size for the YFP which was confirmed by sequencing (data not shown). This binary vector, pYellow, was then used for further ligations with the six PCR products generated in sections 4.4.1 and 4.4.2.

4.6 Generation of ICDH:YFP Fusion Constructs

The six PCR products generated earlier: mito-110, chloro-98, cyto-40, mito-160, chloro-146 and cyto-90, as well as pYellow, were subjected to a double digestion using the restriction enzymes EcoR1/HindIII. Following gel purification of both the linearized plasmid and PCR products, the PCR products were directionally ligated into pYellow, transformed into *E. coli* DH5α competent cells and positive transformants were selected. Successful ligation and transformation was confirmed by restriction digestion with EcoR1/HindIII (Fig 4.10). This was also confirmed by sequence analysis (data not shown). Following confirmation of the inserted fragment and the YFP fusion, the ICDH:YFP fusion constructs were named as described in Table 3.4 and used for *Agrobacterium* transformation and localization studies.

4.7 Co-Transformation of *Agrobacterium* and Transient Plant Transformation

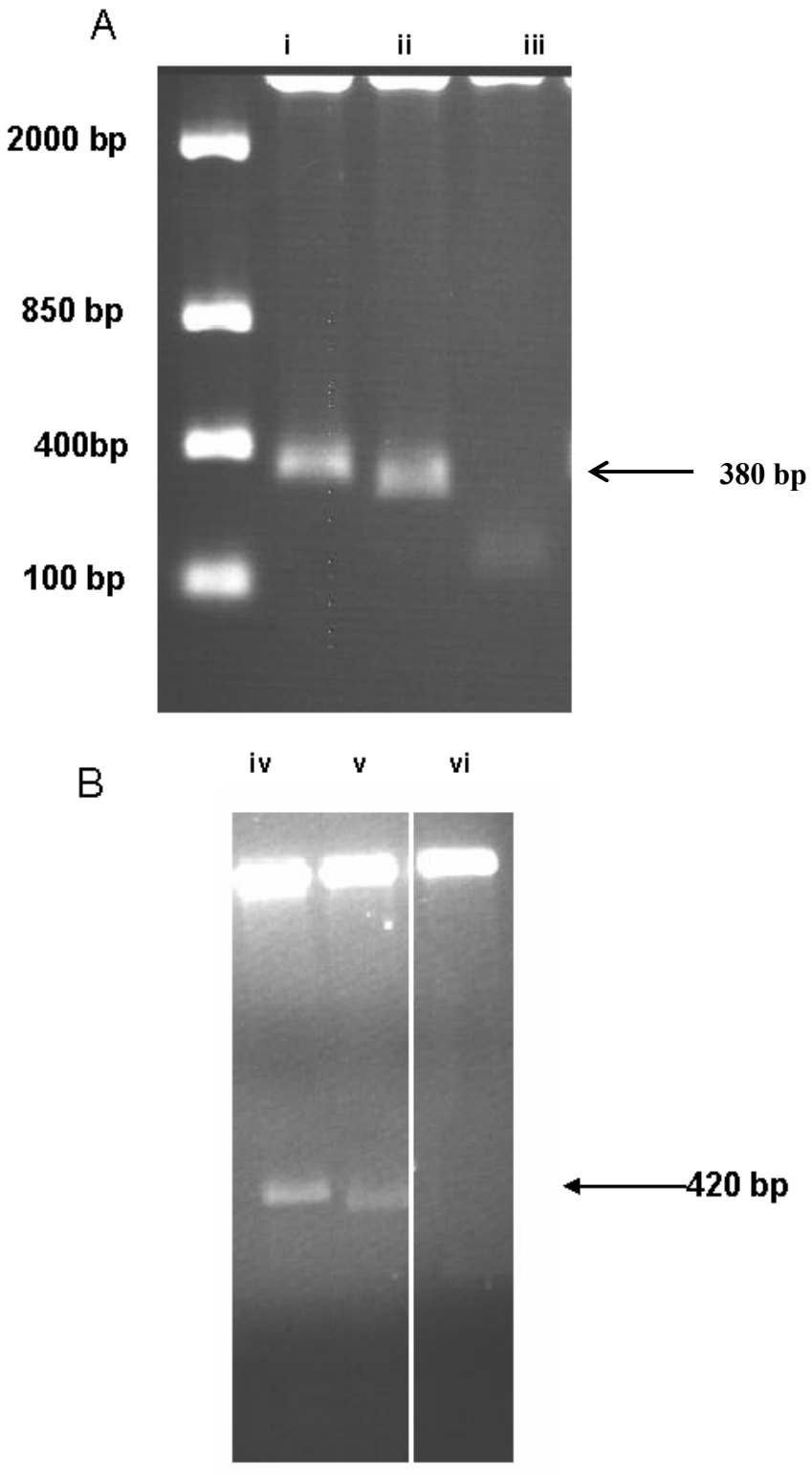

Figure 4.10 Agarose gel analyses of ICDH:YFP fusion constructs from *N. tabacum* L. cv Xanthi using EcoR1 and Hind III. **i)** mito-110, **ii)** chloro-98, **iii)** cyto-40, **iv)** mito-160, **v)** chloro-146 and **vi)** cyto-90.

Plasmid DNA isolated from overnight cultures of the 6 ICDH:YFP fusion constructs indicated in Table 3.4 and the helper plasmid pSoup (Fig. A-5) were separately co-transformed into *Agrobacterium* by electroporation. These cells were plated on selection media and individual colonies selected for growth in liquid media. Positive transformants were selected using antibiotic resistance and confirmed by sequencing (data not shown). Translation and alignment of the deduced amino acid sequence (Appendix C) demonstrated 100% identity to the $NADP^+$-ICDH sequence presented in Fig. 4.7 and that of the published EYFP sequence (GenBank accession AAF65454). In addition, the fusion between the two coding regions indicated that the various ICDH:YFP fusions were present in-frame within all the constructs. These line ups can be seen in Appendix C.

Once the *Agrobacterium* cultures reached a proper optical density they were used for transient plant transformation. All 6 ICDH:YFP fusion constructs were separately injected into different leaf lamina of tobacco plants, resulting in transient expression of the YFP fusion proteins. After 48-72 hours the transformed tobacco leaves were visualized using the confocal microscope.

4.8 Sub-Cellular Localization

Using confocal fluorescence microscopy, epidermal cells of transiently transformed tobacco plants were observed to examine the presence of an active YFP. This was performed 48-72 hours after transformation, depending on the efficiency of transformation, with all 6 fusion constructs to assess targeting to the mitochondria, chloroplasts or cytosolic locations.

4.8.1 Mitochondrial Localization

Upon comparison to mitochondria stained with dihydrorhodamine 123 (Fig. 4.11, panel i), the construct mito-110 appears to localize in bodies similar to that of mitochondria based on fluorescence (Fig. 4.11, panel ii). When chloroplast autoflourescence from the same sample (Fig. 4.11, panel v) is overlayed (Fig. 4.11, panel viii) it is clear that mito-110 is not localizing to the chloroplast. In contrast, mito-160 exhibits a fluorescence which appears to be originating from the chloroplasts (Fig. 4.11, panel iii). This is clearly the case when chloroplast autofluorescence (Fig. 4.11, panel vi) is overlayed (Fig. 4.11, panel ix).

4.8.2 Chloroplastic Localization

The YFP fluorescence of the two putative chloroplastic constructs chloro-98 and chloro-148 appears to be originating from the chloroplasts (Fig. 4.12, panels i and ii respectively). When chloroplast autofluorescence (Fig. 4.12 panels iii and iv) where overlayed the with the YFP fluorescence (Fig. 4.12, panels v and vi) it was obvious that these constructs are localizing to the chloroplasts.

4.8.3 Cytosolic Localization

Upon comparison with the ER protein marker Sec12 (Fig. 4.13, panel i) was used as a positive control to optimize the YFP settings on the confocal microscope.I It appears that the YFP fluorescence from the constructs cyto-40 and cyto-90 is observed in the cytosol (Fig. 4.13, panels ii and iii). When chloroplast autofluorescence (Fig. 4.13, panels iv, v, vi) is overlayed with the YFP fluorescence (Fig. 4.13, panels vii, viii, ix) the cytosolic localization was confirmed by the exclusion of chloroplastic and ER localization.

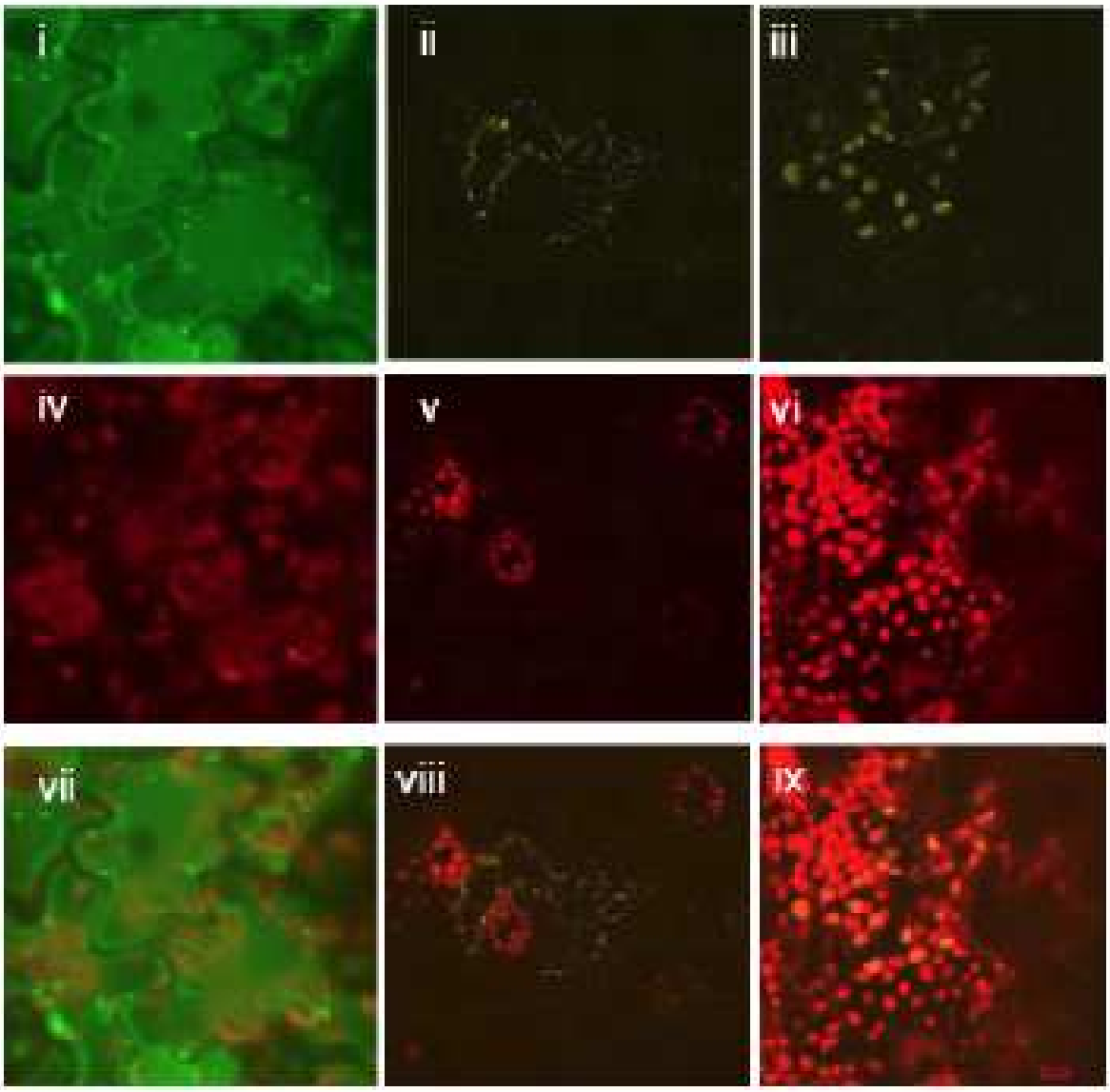

Figure 4.11 Localization of putative mitochondrial ICDH:YFP fusion proteins. Confocal images of tobacco (*N. tabacum* L. cv Petit Havana) leaf epidermal cells after *A. tumefaciens* infiltration. Infiltration occurred when cultures had reached an OD_{600} = 0.20. **(i)** mitochondria stained with dihydrorhodiamine 123 **(ii)** mito-110 **(iii)** mito-160. Panels **(iv), (v) and (vi)** represent chloroplast autofluorescence. Panels **(vii)**, **(viii)** and **(ix)** are merged images of **(i)** and **(iv)**, **(ii)** and **(v)**, and **(iii)** and **(vi)** respectively.

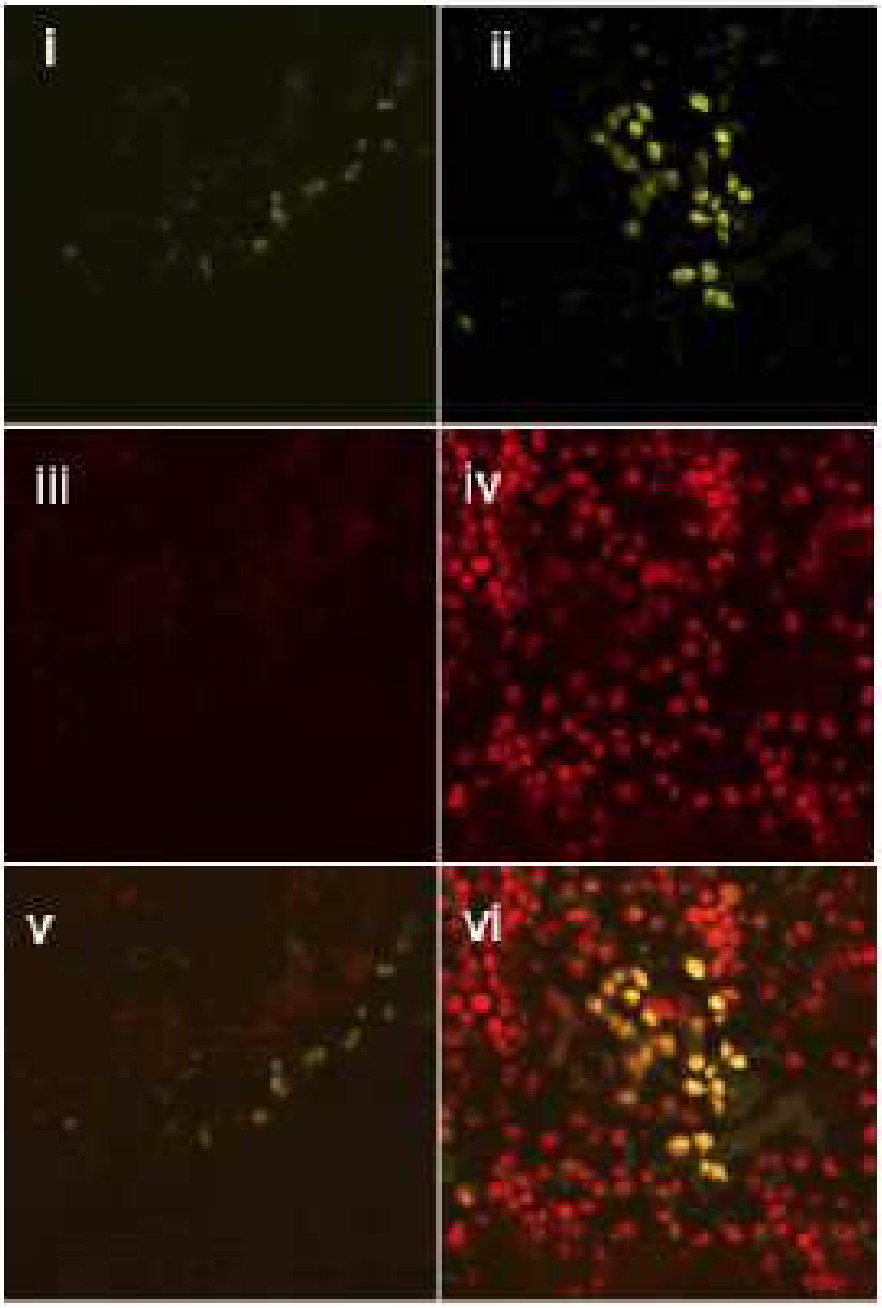

Figure 4.12 Localization of putative chloroplastic ICDH:YFP fusion proteins. Confocal images of tobacco (*N. tabacum* L. cv Petit Havana) leaf epidermal cells after *A. tumefaciens* infiltration. Infiltration occurred when cultures had reached an OD_{600} = 0.20. **(i)** chloro-98 **(ii)** chloro-148. Panels **(iii)** and **(iv)** represent chloroplast autofluorescence. Panels **(v)** and **(vi)** are merged images of **(i)** and **(iii)** and **(ii)** and **(iv)** respectively.

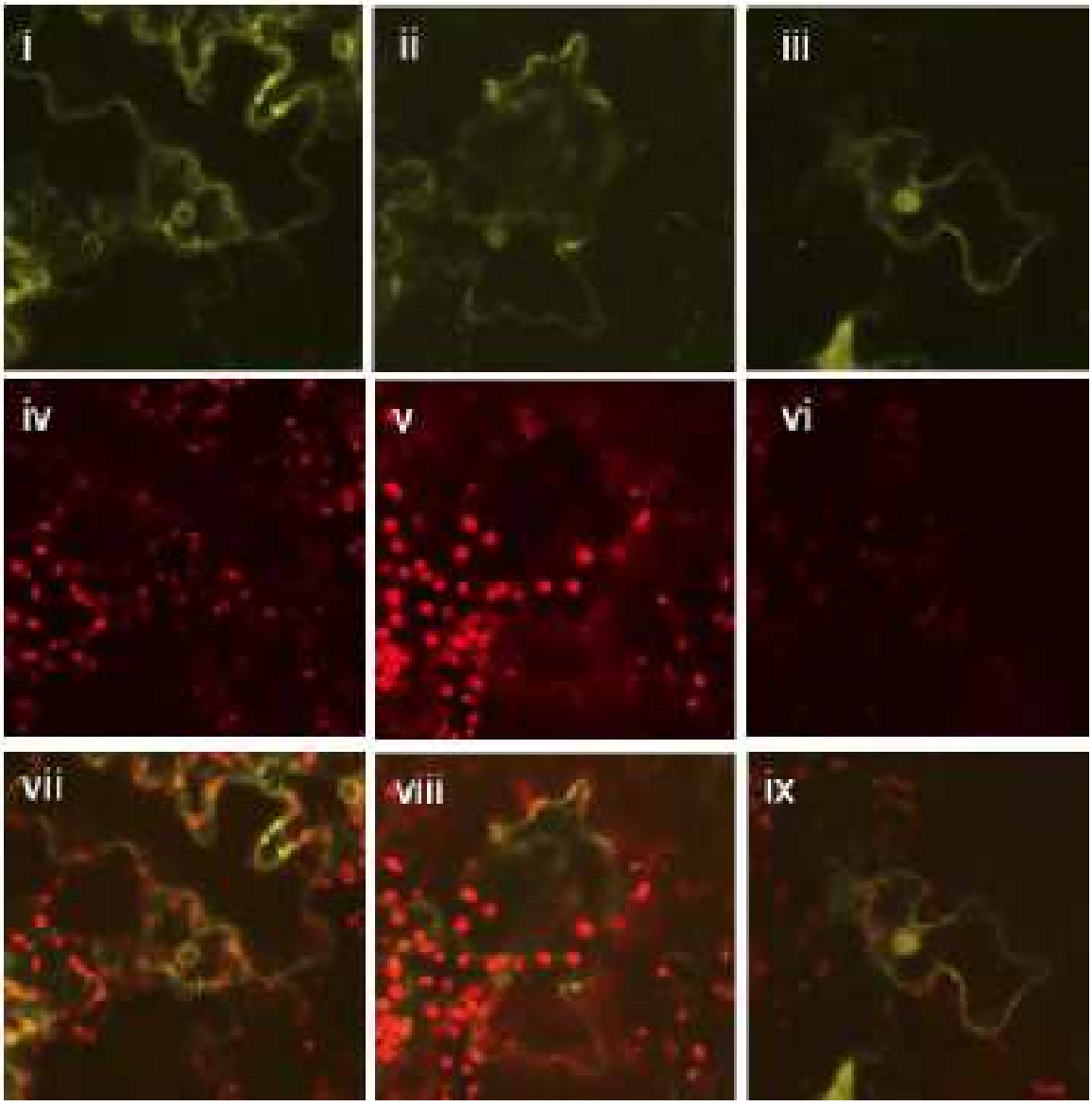

Figure 4.13 Localization of putative cytosolic ICDH:YFP fusion proteins. Confocal images of tobacco (*N. tabacum* L. cv Petit Havana) leaf epidermal cells after *A. tumefaciens* infiltration. Infiltration occurred when cultures had reached an OD_{600} = 0.20. **(i)** Sec12 **(ii)** cyto-40 **(iii)** cyto-90. Panels **(iv)**, **(v)** and **(vi)** represent chloroplast autofluorescence. Panels **(vii)**, **(viii)** and **(ix)** are merged images of **(i)** and **(iv)**, **(ii)** and **(v)**, and **(iii)** and **(vi)** respectively.

5.0 DISCUSSION

$NADP^+$-dependent isocitrate dehydrogenase ($NADP^+$-ICDH; EC 1.1.1.42) exists in multiple isoforms in several cellular compartments including the cytosol, chloroplasts and mitochondria. This enzyme functions in the oxidative decarboxylation of isocitrate, producing alpha-ketoglutarate (α-ketoglutarate) and CO_2 while converting $NADP^+$ to NADPH (Horton et al., 2003).

Gálvez et al. (1996) first isolated a cDNA encoding the cytosolic isoform of $NADP^+$-ICDH (ICDH1) from a tobacco cell suspension cDNA library using a soybean cDNA probe (Gálvez et al., 1996). Subsequently, this cDNA library was screened with a homologous probe and a new cDNA (GenBank accession number X96728.1) was isolated (Gálvez et al., 1998). The deduced amino acid sequence of this cDNA shared only 70-75% amino acid identity with other known plant $NADP^+$-ICDH sequences. This cDNA was found to contain two in-frame ATG translational start sites and it was hypothesized that this may result in a mitochondrial signal based on analyses of the deduced amino acid sequence. To address to this question, they transformed tobacco plants with two constructs; one containing the entire coding sequence (from the first ATG) and another construct containing only the coding sequence of the mature protein. Both of these constructs were fused with a modified green fluorescent protein (GFP). These experiments by Gálvez et al. (1998) found that the full-length targeting signal localized the protein to the mitochondria and weakly to the chloroplast, while the construct containing no targeting signal was retained in the cytoplasm. They concluded that the cDNA that was isolated encoded the mitochondrial isoform of $NADP^+$-ICDH. This was subsequently confirmed by others working with purified protein from isolated mitochondria (Gray et al., 2004; Gray and McIntosh, unpublished).

In this thesis, I examined the localization of the mitochondrial $NADP^+$-ICDH enzyme and tested the hypothesis that this isoform is dual-targeted to both the mitochondrion and the chloroplast. This was accomplished by generating full-length and truncated constructs of $NADP^+$-ICDH with altered targeting signals. A yellow fluorescent protein (YFP) reporter vector was created and these constructs were fused to the YFP and used for transient transformation of tobacco leaves. The sub-cellular localization of the subsequent ICDH:YFP fusion proteins was then determined using confocal microscopy. Using this approach I was able to develop an in-depth and clear picture of the targeting process of the mitochondrial isoform of $NADP^+$-ICDH.

5.1 Amino Acid Changes

To examine the targeting of $NADP^+$-ICDH, primers were developed to amplify specific sections of the $NADP^+$-ICDH cDNA that would putatively result in mitochondrial, chloroplastic and cytosoic localization. These PCR products were successfully cloned and sequenced. Sequencing revealed 13 nucleotide changes when compared to the sequence present in the NCBI data base. These changes were present in all three constructs irrespective of putative localization. The 13 nucleotide changes resulted in eight amino acid changes. When the eight amino acid changes were examined, it was found that the deduced protein shared a 98.3% similarity and a 99.1% identity to the NCBI database sequence. Included were seven conservative amino acids changes (represented with a * in Table 4.1) and one semi-conservative amino acid change (represented with ** in Table 4.1).

The most significant of these amino acid changes was a stop codon introduced at amino acid 161 (counting from the first ATG site) as a result of a single nucleotide change. Interestingly, both tobacco cultivars used in this study (Xanthi and Petit Havana) displayed this

stop codon (for Petit Havana results see Appendix B). The NCBI database sequence indicates that amino acid 161 should be a tryptophan (W), encoded by the nucleotide sequence TGG. My sequence analysis revealed the nucleotide sequence in this position to be TGA, resulting in a stop codon at the amino acid 161 position. There is only a one nucleotide difference between these two codons. Our clones were sequenced multiple times in two cultivars (Xanthi and Petit Havana) all revealing the same result. Amplification and sequencing was performed using genomic DNA as a template revealing the same nucleotide and amino acid changes (data not shown). The NCBI database entry for the $NADP^+$-ICDH cDNA (accession X96782.1) was deposited by Gálvez et al. (1998) some 10 years ago from *Nicotiana tabacum* L. cv. Xanthi. It is possible that in these earlier days of sequencing, a mistake in sequencing was made. In this study, the analysis was based onmultiple sequencing reactions, always yielding the same results. Therefore it is assumed that the results are correct with respect to the presence of a stop codon in the two cultivars tested.

It is clear that the $NADP^+$-ICDH mature protein contains a stop codon at the amino acid 161 position. This has been found in all the constructs created in this work. The question then arises as to why this in-frame stop codon occurs? The fact that this stop codon is present in-frame in the coding region of the mature protein is an enigma. Upon several sequenecing reactions, in multiple constructs, it became clear that this was not an artifact, but rather a definite stop codon. The mass of the purified $NADP^+$-ICDH protein from tobacco has been found to be 43 kDa (Gray et al., 2004) and reflects the predicted mass of the protein. This is intriguing as a recognized termination codon would not allow a mature protein of this mass to be observed.

While it is clear there is a stop codon at the amino acid 161 position, it is possible that we are seeing ‘leaky’ termination of the protein, meaning only partial termination may occur.

Namy et al. (2001) reported that in *Saccharmyces cerevisiae* upstream and downstream sequences can influence the translation termination efficiency. Bonetti et al. (1995) identified the CAA STOP CAA read through motif in the STE6 gene as a leaky mutation. These workers showed that the nucleotide sequence CAA up and downstream of the stop codon act together to promote a high read-through level (Bonetti et al., 1995). The base pairing of the surrounding nucleotides resulted in destabilized secondary structures in the ribosome and these changes which would affect binding of release factors. The $NADP^+$-ICDH that was the subject of this study contains the sequence ATG STOP AGA. Although my analysis did not indicate any of these read-through motifs, the fact that they exist provides precedence for this type of read-through and may provide one explanation as to how $NADP^+$-ICDH of the correct mass has been purified and exists in spite of a stop codon at the amino acid 161 position. It is possible that the CAA read through motif is one of many combinations making termination read-through possible.

RNA editing is a post-transcriptional process by which the primary structure of an RNA molecule is altered by nucleotide insertions, deletions or alterations. The resulting RNA contains genetic information that was not present in the gene from which it was originally transcribed. There have been examples in the literature were a terminal stop codon was created in-frame by RNA editing (Wintz, 1991). Wintz (1991) cloned and sequenced the cDNA from the *atp9-1* gene of *Petunia hybrida* and found that ten C residues at different positions in the gene, are changed into U in the mRNA. One of these changes resulted in a stop codon. The protein deduced from the cDNA sequence is three amino acids shorter than the one predicted from the gene. This was the first report of a stop codon introduced into a plant mitochondrial mRNA by RNA editing (Wintz, 1991). It is possible that this same mechanism is responsible

for the creation of the stop codon observed in this work although this is difficult to reconcile without genomic sequence data.

The constructs generated in this study only contained the mature protein with altered N-terminal targeting signals; it is possible that the 5' and 3' untranslated regions (UTRs) are needed for proper translation. There are examples in the literature where the UTRs of genes are required to create proteins of the correct mass. The addition of selenocysteine is one such example. Selenocysteine is encoded by a UGA codon, which is normally a stop codon (Schomburg et al., 2004). It should be noted that this is the stop codon we observed in this study. The UGA codon is altered to encode selenocysteine by the presence of a SECIS element (SElenoCysteine Insertion Sequence) in the mRNA (Hesketh, 2004; Copeland et al., 2001). The SECIS element is defined by characteristic nucleotide sequences and secondary structure base-pairing patterns. In Archaea and in eukaryotes, the SECIS element is in the 3' UTR of the mRNA, and can direct multiple UGA codons to encode selenocysteine residues (Schomburg et al., 2004). It is possible that a similar process occurs with $NADP^+$-ICDH. As we do not have the 5' or 3' UTRs in our constructs, we may be missing the regulatory regions necessary to form a complex that would cause the stop codon to encode an amino acid.

5.2 Sub-Cellular Localization

In this study, I have taken advantage of a modified YFP (EYFP) and generated a reporter vector (pYellow) to assess the sub-cellular localization of the protein encoded by the cDNA of $NADP^+$-ICDH (GenBank accession X96728.1). Two sets of ICDH:YFP fusion constructs were developed; one of which was truncated at amino acid 160 and the other at amino acid 110. Each of these sets of constructs contained three separate constructs with different targeting signals (either mitochondrial, chloroplastic or no targeting signal at all).

These constructs were created from both Xanthi and Petit Havana cultivars (for Petit Havana results see Appendix B).

In the current study, the experiments of Gálvez et al. (1998) were repeated with a truncated construct of the mature protein containing the predicted mitochondrial targeting signal (mito-110). It was found that this construct appeared to be targeted to the mitochondria when compared to mitochondria stained with dihydrorhondamine 123 (Fig 4.11, panels i and ii). This was expected and corresponds with the results of Gálvez et al. (1998).

In contrast, when a slightly longer construct of the mature protein containing the predicted mitochondrial targeting signal (mito-160) was observed under the confocal microscope an interesting observation was noted. The mito-160 construct appeared to be chloroplastic in its sub-cellular localization. This was determined by individually observing the YFP fluorescence (Fig 4.11, panel iii), the chloroplastic autofluorescence (Fig 4.11, panel vi) and by overlaying the two (Fig 4.11, panel ix). By doing this overlay, we see a perfect correspondence between the YFP signal and the chloroplastic autofluorescence. This targeting occurred in both tobacco cultivars, Xanthi and Petit Havana. This construct contains the exact same targeting signal as the mito-110 construct, but a very different localization of the two YFP fusion proteins occured. The only difference between the two proteins is an extra 50 amino acids in the mito-160 construct. It is clear that the addition of these 50 amino acids results in altered targeting. It is interesting that when the same 50 amino acids where added to the truncated construct, chloro-148, we did not see any change in targeting when compared to the other cholorplastic construct (chloro-98). Since we observed two targeting destinations from one apparent targeting signal in the mitochondrial constructs, it is not possible to assign a sub-cellular destination for the full-length targeting signal.

When the putative chloroplastic constructs were examined (chloro-148 and chloro-98) it was clear that localization was to the chloroplast in both cases. This was hypothesized, as the constructs created from the second ATG translational start site were predicted to be chloroplastic in nature by the Predator prediction software. This was determined by individually observing the YFP fluorescence (Fig 4.12 panel i and ii), the chloroplastic autofluorescence (Fig 4.12, panel iii and iv) and by overlaying the two images (Fig 4.12, panels v and vi). The overlay illustrates that there is a perfect alignment between the YFP signal and the chloroplastic autofluorescence. This was the case in both Xanthi and Petit Havana cultivars. Hodges and co-workers also report on this topic in a 2003 review article (Hodges et al., 2003). Using $NADP^+$-ICDH fused to a GFP reporter in tobacco and *Arabidopsis*, it was suggested that constructs containing the putative chloroplastic targeting sequence (from the second translational start site) do not localize to the chloroplasts (Hodges et al., 2003). However, this was reported using unpublished data and little information is provided with regard to the constructs utilized or their design. Our results would indicate the opposite, as we see clear chloroplastic targeting. This is the first study that demonstrates $NADP^+$-ICDH contains a chloroplastic targeting signal, which leads to the assumption that the gene encoding the mitochondrial isoform of $NADP^+$-ICDH also results in an isoform of $NADP^+$-ICDH that is dual-targeted to the chloroplast. This work was only performed using the targeting signals and a truncated portion of the mature protein. It is assumed that the full-length protein is targeted in the same manner, but it can not be unequivocally stated that this is the case.

Constructs containing no targeting signal (cyto-90 and cyto-40) appeared to localize in the cytosol of the tobacco cells. This was predicted as these constructs contain no sub-cellular targeting signal. This is illustrated in Fig. 4.13 (panels ii and iii), where the YFP signals are clearly fluorescing in the cytosol. This localization is exactly what was predicted, as these

constructs contained no targeting signals. As a result, the proteins being translated in the cytosol were not targeted to any organelle and thus stayed in the cytosol. This was the case in both cultivars, Xanthi and Petit Havana.

Differential targeting could be explained if the additional 50 amino acids contained a chloroplastic targeting signal. However, when the 50 amino acids were entered into targeting signal prediction software alone or with the rest of the $NADP^+$-ICDH truncated sequence, no targeting was predicted. This idea that the additional 50 amino acids could be a chloroplastic targeting signal was also refuted with the targeting of the cyto-90 construct. If this were the case and the additional 50 amino acids contained a chloroplastic targeting signal we would have seen chloroplastic targeting in this construct, which was not the case as cytosolic localization was observed.

It is also possible that these 50 amino acids cause a conformational change in the protein. The translocase complexes located in the outer and inner mitochondrial membranes (TOM and TIM) and chloroplast envelope membrane (TOC and TIC) sorting mechanisms are not completely understood. It is clear however, that organelle-specific receptors recognize protein signal peptides and guide the protein into the import pore of the cellular organelle. After import, the targeting peptide is cleaved off by the mitochondrial or the stromal processing peptidase (Bhushan et al., 2006; Jarvis, 2004). It is possible that by the addition of the 50 amino acids, the translated protein is folded in such a way that the mitochondrial targeting signal is unable to be recognized by the translocase of the outer mitochondrial membrane (TOM). If the TOM does not recognize the targeting signal, the protein is not granted entry into the cellular organelle (Jarvis, 2004). This conformational change may also present the chloroplastic targeting peptide in a favorable manor, resulting in chloroplastic localization.

6.0 CONCLUSIONS AND FUTURE STUDIES

6.1 Conclusions

The research conducted for this thesis focused on the localization of mitochondrial $NADP^+$-ICDH from tobacco (*Nicotiana tabacum* L. cv. Petit Havana) leaves and tested the hypothesis that this isoform is dual-targeted to both the mitochondrion and chloroplast. This involved creating constructs of the protein with altered putative targeting signals and fusing them with a yellow fluorescent protein (YFP) for sub-cellular localization studies using the confocal microscope.

During this process it was discovered that a stop codon existed at amino acid position 161 of the mature protein which is not consistent with the deduced amino acid sequence derived from the cDNA enoding this isoform of $NADP^+$-ICDH. This was observed in several different constructs from 2 individual cultivars of tobacco and confirmed by multiple sequencing reactions. To circumvent this stop codon, truncated constructs of the mature $NADP^+$-ICDH protein were created, also with altered targeting signals and fused to a YFP. The first set was truncated at amino acid 161, just before the identified stop codon. The other set was truncated at amino acid 110 which corresponds to the constructs used by Gálvez et al. (1998) who confirmed that the protein was targeted to the mitochondria.

When these ICDH:YFP fusion proteins were examined using the confocal microscope, some very interesting results were generated. The assumption that constructs cyto-90 and cyto-40, were theoretically targeted to the cytosol was found to be accurate. The hypothesis that constructs chloro-148 and chloro-98 were targeted to the chloroplast was also found to be true. It was hypothesized that the constructs mito-160 and mito-110 would be targeted to the mitochondria. This was found to be true only for the mito-110 construct. In contrast, the mito-

160 construct was found to be targeted to the chloroplast. This however has been attributed to the addional 50 amino acids present on this construct. This research has provided evidence that $NADP^+$-ICDH is targeted to both the mitochondria and the chloroplast, containing two possible targeting signals; the first signal being mitochondrial in nature and encoded from the first ATG translational start site and the second signal being chloroplastic and encoded from the second ATG translational start site. We also show indication that regions of the mature $NADP^+$-ICDH protein are able to influence targeting as is demonstrated by the differential targeting witnessed by the addition of 50 amino acids.

6.2 Future Studies

In light of the stop codon which was discovered in this thesis, it would be interesting to perform a set of experiments in order to examine the localization of the full-length protein. Such experiments would either require the mutagenization of the stop codon at amino acid 161 or perhaps the addition of the 5' and 3' UTRs to the constructs. The generation of a full-length protein, which occurs *in vivo*, would assist in determining if read-through was occurring and if so, the level of read through efficiency.

It would be interesting to attach the targeting signals examined in this study to other proteins and determine their sub-cellular location. These types of studies have been performed previously and would involve amplifying both the mitochondrial and the chloroplastic targeting signals and attaching them individually to a protein of choice. The protein would preferentially be a cytosolic protein with no targeting signal or a protein with the targeting signal removed. It would also be interesting to place the targeting signals examined in this study up-stream or down-stream of other targeting signals and determine the strength of the $NADP^+$-ICDH targeting signal. Based on my results, I would hypothesize that both the putative mitochondrial

and chloroplastic targeting signals of the $NADP^+$-ICDH would be able to localize passenger proteins to their respective organelles but this would have to be proven experimentally.

The purification and sequencing of $NADP^+$-ICDH protine would be beneficial as it would enable us to determine whether the stop codon observed in this study actually encodes a selenocysteine, or results in a read-through event. It would be important to examine the possibility that the nucleotides surrounding the discovered stop codon are able to influence translational read-through. This examination would determine whether the up- and down-stream nucleotides cause a read-through motif like the one reported by Bonetti et al. (1995). One could place the CAA stop CAA motif (Bonetti et al., 1995) around the stop codon found in this study and examine the read-through efficiency. It would also be interesting to place the ATG stop AGA motif, found in this study, around the stop codons found in other genes to determine whether read-through occurs. The protein sequence may also give insight into the other amino acid changes witnessed and help to determine the apparent discrepancy of the sequencing results found here as compared to that reported in the NCBI database.

One key experiment that should be performed is a determination of the number of $NADP^+$-ICDH genes in the genomic DNA. To do this a southern blot would be performed. This would determine how many copies of the $NADP^+$-ICDH are present in the genome. If indeed multiple copies of the gene are found it may provide an explanation as to why we are seeing a stop codon present in the DNA. It could be possible that this gene is a pseudo of $NADP^+$-ICDH and that the primers designed have an greater affinity towards binding this cryptic gene instead of the active $NADP^+$-ICDH.

It would also be interesting to do the same set of localization experiments performed in this thesis but transfect a different tissue. Different tissues have a different abundance of cellular organells; the amount of organelle may influence the localization results. Therefore, it

is possible that the created constructs may undergo different targeting in a tissue other then the leaf.

7.0 REFERENCES

Abe Y, Shodai T, Muto T, Mihara K, Torii H, Nishikawa S, Endo T, Kohda D (2000) Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. Cell 100: 551-560

Bauer MF, Hofmann S, Neupert W, Brunner M (2000) Protein translocation into mitochondria: role of TIM complexes. Trends Cell Biol **10:** 25-31

Batoko H, Zheng HQ, Hawes C, Moore I (2000) A Rab1 GTPase is required for transport between the endoplasmic reticulum and Golgi apparatus and for normal Golgi movement in plants. Plant Cell **12**: 2201-2217

Becker T, Jelic M, Vojta A, Radunz A, Soll J, Schleiff E (2004) Preprotein recognition by the Toc complex. EMBO J **23**: 520-530

Brett D, Pospisil H, Valcárcel J, Reich J, Bork P (2001) Alternative splicing and genome complexity. Nature Genet **30**: 29-30

Bhushan S, Kuhn C, Berglund A, Roth C, Glaser E (2006) The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and miss-sorting. FEBS Lett **580**: 3966-3972

Bonetti B, Fu LW, Moon J, Bedwell DM (1995) The efficiency of translation termination is determined by a synergistic interplay between upstream and downstream sequences in *Saccharomyces cerevisiae*. J Mol Biol **251**: 334-345

Brandizzi F, Frangne N, Marc-Martin S, Hawes C, Neuhaus JM, Paris N (2002*a*) The destination for single-pass membrane proteins is influenced markedly by the length of the hydrophobic domain. Plant Cell **14:** 1077-1092

Brandizzi F, Snapp EL, Roberts AG, Lippincott-Schwartz J, Hawes C (2002*b)* Membrane protein transport between the endoplasmic reticulum and the Golgi in tobacco leaves 10 is energy dependent but cytoskeleton independent: evidence from selective photobleaching. Plant Cell **14**: 1293-1309

Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem **241**: 779-786

Creissen G, Reynolds H, Xue Y, Mullineaux P (1995) Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplasts and mitochondria in transgenic tobacco. Plant J **8**: 167-175

Copeland PR, Stepanik VA, Driscoll DM (2001) Insight into mammalian selenocysteine insertion: domain structure and ribosome binding properties of sec insertion sequence binding protein 2. Mol Cell Biol **21**: 1491-1498

Cunillera N, Boronat A, Ferrer A (1997) The Arabidopsis thaliana FPS1 gene generates a novel mRNA that encodes a mitochondrial farnesyl-diphospate synthase isoform. J Biol Chem **272**: 15381-15388

Danpure CJ (1995) How can the products of a single gene be localized to more than one intracellular compartment? Trends Cell Biol **5**: 230-238

daSilva LL, Snapp EL, Denecke J, Lippincott-Schwartz J, Hawes C, Brandizzi F (2004) Endoplasmic reticulum export sites and Golgi bodies behave as single mobile secretory units in plant cells. Plant Cell **16:** 1753-1771

Degenhard RF, Bonham-Smith P (2008) Arabidopsis ribosomal proteins RPL23aA and RPL23aB are differentially targeted to the nucleolus and are disparately required for normal development. Plant Physiol **147**: 128-142

Ellis R (1981) Chloroplast proteins: synthesis, transport and assembly. Annu Rev Plant Physiol **32**: 11-137

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol **300**: 1005-1016

Gálvez S, Gadal P (1995) On the function of the NADP-dependent isocitrate dehydrogenase isoenzymes in living organisms. Plant Sci **105**: 1-14

Gálvez S, Hodges M, Decottignes P, Bismuth E, Lancien M, Sangwan RS, Dubois F, LeMarechal P, Crétin C, Gadal P (1996) Identification of a tobacco cDNA encoding a cytosolic NADP-isocitrate dehydrogenase, Plant Mol Biol **30:** 307-320

Gálvez S, Roche O, Bismuth E, Brown S, Gadal P, Hodges M (1998) Mitochondrial localization of a NADR-dependent isocitrate dehydrogenase isoenzyme by using the green fluorescent protein as a marker. Proc Natl Acad Sci USA **95**: 7813-7818

Gray GR, Villarimo A, Whitehead C, McIntosh L (2004) Trangenic tobacco (*Nicotiana tabacum* L.) plants with increased expression levels of mitochondrial $NADP^+$-dependent isocitrate dehydrogenase: Evidence implicating the enzyme in the redox activation of the alternative oxidase. Plant Cell Physiol **45**: 1413-1425

Hammen PK, Gorenstein DG, Weiner (1996) Amphiphilicity determines binding properties of three mitochondrial presequences to lipid surfaces. Biochem **35**: 3772-3781

Hesketh J (2004) 3'-Untranslated regions are important in mRNA localization and translation: lessons from selenium and metallothionein. Biochem Soc Trans **32**: 990-993

Hines V, Schatz G (1993) Precursor binding to yeast mitochondria. A general role for the outer membrane protein Mas70p. J Biol Chem **268**: 449-454

Hurt EC, Pesold-Hurt B, Schatz G (1984) The cleavable prepiece of an imported mitochondrial protein is sufficient to direct cytosolic dihydrofolate reductase into the mitochondrial matrix. FEBS Lett **178**: 306-310

Hodges M, Flesch V, Gálvez S, Bismuth E (2003) Higher plant $NADP^+$-dependent isocitrate dehydrogenases, ammonium assimilation and NADPH production. Plant Physiol Biochem **41**: 577-585

Hörmann F, Soll J, Bölter B (2007) The chloroplast protein import machinery: a review. Methods Mol Biol. **390**:179-93

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT (1985) A Simple and general method for transferring cloned genes into plants. Science **227**: 1229-1231

Horton H, Morgan L, Ochs R, Rawn JD, Scrimgeour G (2002) Principles of Biochemistry. Ed 3 Pearson Publishing, Toronto, pp 382-390

Jarvis P (2008)Targeting of nucleus-encoded proteins to chloroplasts in plants. New Phytol. **179**:257-85

Jarvis P, Soll J (2001) Toc, Tic, and chloroplast protein import. Biochim Biophys Acta **1541**: 64-79

Karniely S, Pines O (2005) Single translation- dual destination: mechanisms of dual protein targeting in eukaryotes. EMBO Rep **6**: 420-425

Koehler CM (2000) Protein translocation pathways of the mitochondrion. FEBS Lett **476**: 27-31

Kouranov A, Chen X, Fuks B, Schnell DJ (1998) Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. J Biol Chem **143**: 991-1002

Komiya T, Sakaguchi M, Mihara M (1996) Cytoplasmic chaperones determine the targeting pathway of precursor proteins to mitochondria. EMBO J **15**: 399-407

Ladner RD, Caradonna SJ (1998) The human dUTPase gene encodes both nuclear and mitochondrial isoforms. J Biol Chem **272**: 19072-19080

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) ClustalW and ClustalX version 2. Bioinformatics **23**: 2947-2948

Lang BF, Gray MW, Burger G (1999) Mitochondrial genome evolution and the origin of eukaryotes. Annu Rev Genet **33**: 351-397

Lee CM, Sedman J, Neupert W, Stuart RA (1999) The DNA helicase, Hmi1p, is transported into mitochondria by a C-terminal cleavable targeting signal. J Biol Chem **274**: 20937-20942

Lister R, Hulett JM, Lithgow T, Whelan J (2005) Protein import into mitochondria: origins and functions today. Mol. Membr. Biol. **22**: 87-100

Mackenzie S (2005) Plant organellar protein targeting: a traffic plant still under construction. Trends Cell Biol **15**: 548-553

Margulis L (1981). Symbiosis in Cell Evolution. W. H. Freeman, San Francisco

Mihara K, Omura T (1996) Cytosolic factors in mitochondrial protein import. Experientia **52:** 1063-1068

Millar A, Whelan J, Small I (2006) Recent surprises in protein targeting to mitochondria and plastids. Curr Opin Plant Biol **9**: 610-615

Mireau H, Lancelin D, Small I (1996) The same arabidopsis gene encodes both cytosolic and mitochondrial alanyl-tRNA synthetases. Plant Cell **8**: 1027-1039

Namy O, Hatin I, Rousset JP (2001) Impact of the six nucleotides downstream of the stop codon on translation termination. EMBO Rep **2**: 787-793

Neupert W (1997) Protein import into mitochondria. Ann Rev Biochem **66**: 863-917

Nunnari J, Fox TD, Walter P (1993) A mitochondrial protease with two catalytic subunits of nonoverlapping specificities. Science **262**: 1997-2004

Pfanner N, Geissler A (2001) Versatility of the mitochondrial protein import machinery. Nature Rev Mol Cell Biol **2**: 339-349

Peeters N, Small I (2001) Dual targeting to mitochondria and chloroplasts. Biochim Biophys Acta **1541**: 54-63

Obara K, Sumi K, Fukuda H (2002) The use of multiple transcription start sites causes the dual targeting of Arabidopsis putative monodehydroascorbate reductase to both mitochondria and chloroplasts. Plant Cell Physiol **43**: 697-705

Ramage L, Junne T, Hahne K, Lithgow T, Schatz G (1993) Functional cooperation of mitochondrial protein import receptors in yeast. *EMBO J.* **12**: 4115-4123

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. **16**: 276-277

Schnell DJ (2000) Functions and origins of the chloroplast protein-import machinery. Essays Biochem. **36**:47-59

Schreier P, Seftor E, Schell J, Bohnert H (1985) The use of nuclear-encoded sequences to direct the light-regulated synthesis and transport of a foreign protein into plant chloroplasts. EMBO J **4**: 25-32

Schomburg L, Schweizer U, Köhrle J (2004) Selenium and selenoproteins in mammals: extraordinary essential enigmatic. Cell Mol Life Sci **61**: 1988-1995

Silva-Fiho M (2003) One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. Curr Opin in Plant Biol **6**: 589-595

Small I, Peeters N, Legeai F, Lurin C (2004) Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics **4**: 1581-1590

Small I, Wintz H, Akashi K, Mireau H (1998) Two birds with one stone: genes that encode products targeted to two or more compartments. Plant Mol Biol **38**: 265-277

Smith M (2006) Protein import into chloroplasts: an ever-evolving story. Can. J. Bot. **84**: 531–542

Söllner T, Griffiths G, Pfaller R, Pfanner N, Neupert W (1989) MOM19, an import receptor for mitochondrial precursor proteins. Cell **59**: 1061-1070

Soll J, Schleiff E (2004) Protein import into chloroplasts. Nature **5**: 198-209

Sparkes IA, Runions J, Kearns A, Hawes C (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nature Protoc **1**: 2019-2025

Strobel G, Zollner A, Angemayr M, Bandlow W (2001) Competition of spontaneous protein folding and mitochondrial import causes dual subcellular location of major adenylate kinase. Mol Biol Cell **13**: 1439-1448

Watanabe N, Che FS, Iwano M, Takayama S, Yoshida S, Isogai A (2001) Dual targeting of spinach protorporphyrinogen oxidase II to mitochondria and plastids is due to alternative translation initiation in single transcripts. J Biol Chem **276**: 20474-20481

Wintz H, Hanson MR (1991) A termination codon is created by RNA editing in the petunia atp9 transcript. Curr Genet **19**: 61-64

Zhang XP, Glaser E (2002) Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone**.** Trends Plant Sci **7**: 14-21

APPENDIX A

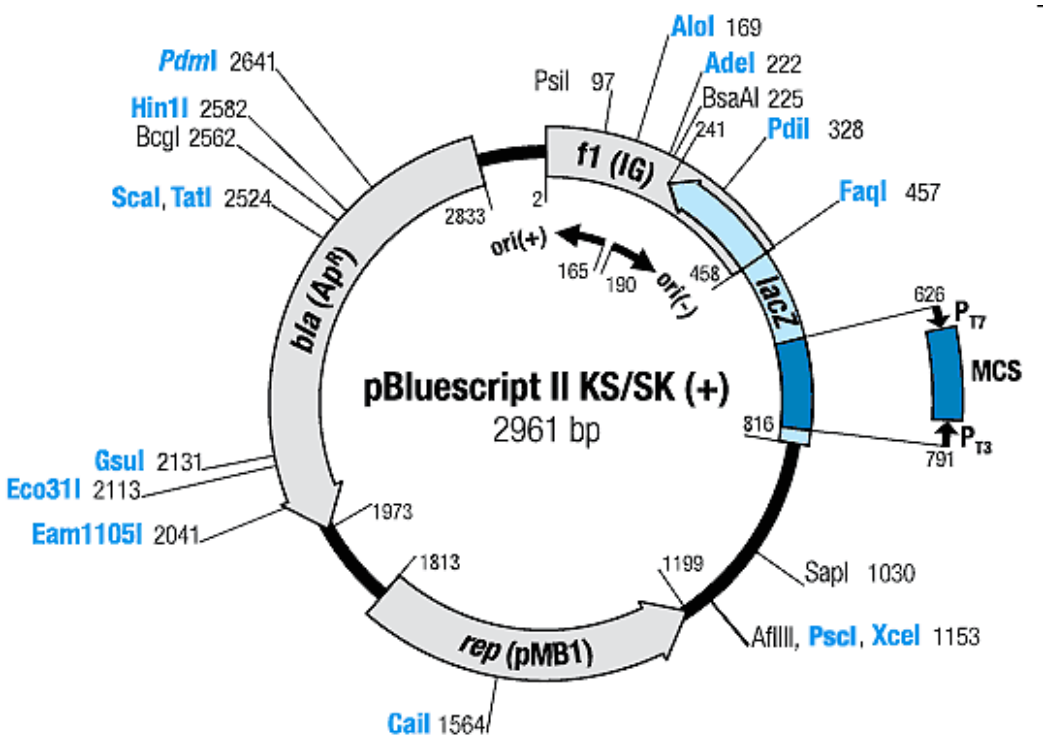

Figure A-1 Graphical representation of pBluescript KS (+).

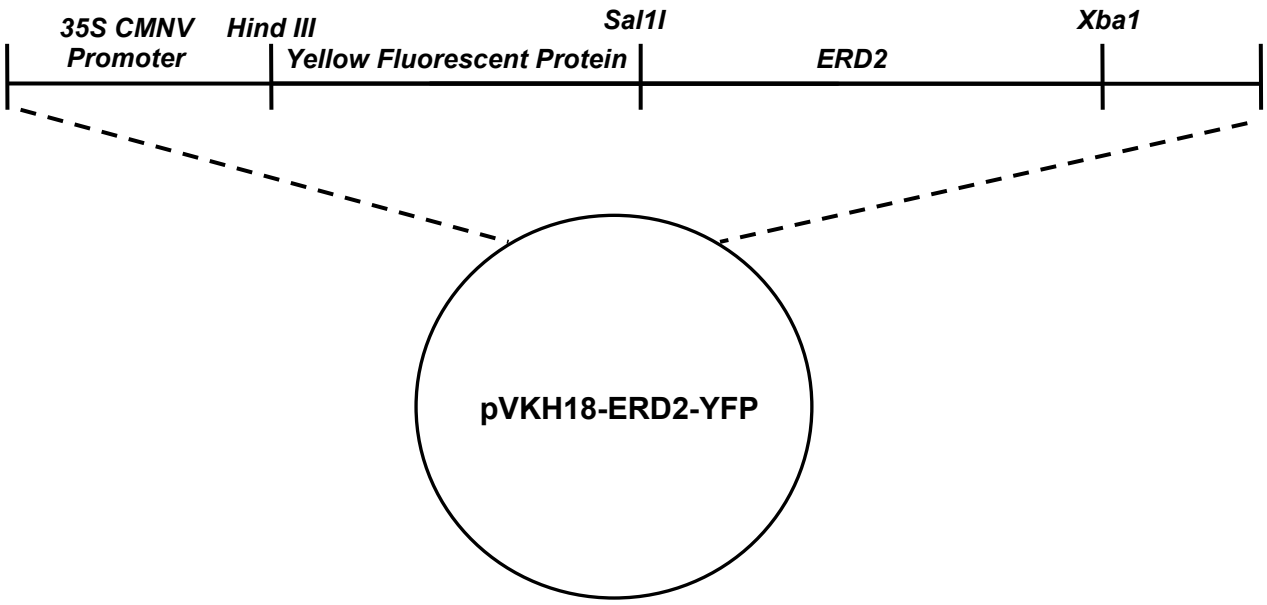

Figure A-2 Graphical representation of pVKH18-ERD2-YFP. Dashed lines indicate inserted casette into the multiple cloning site (MCS).

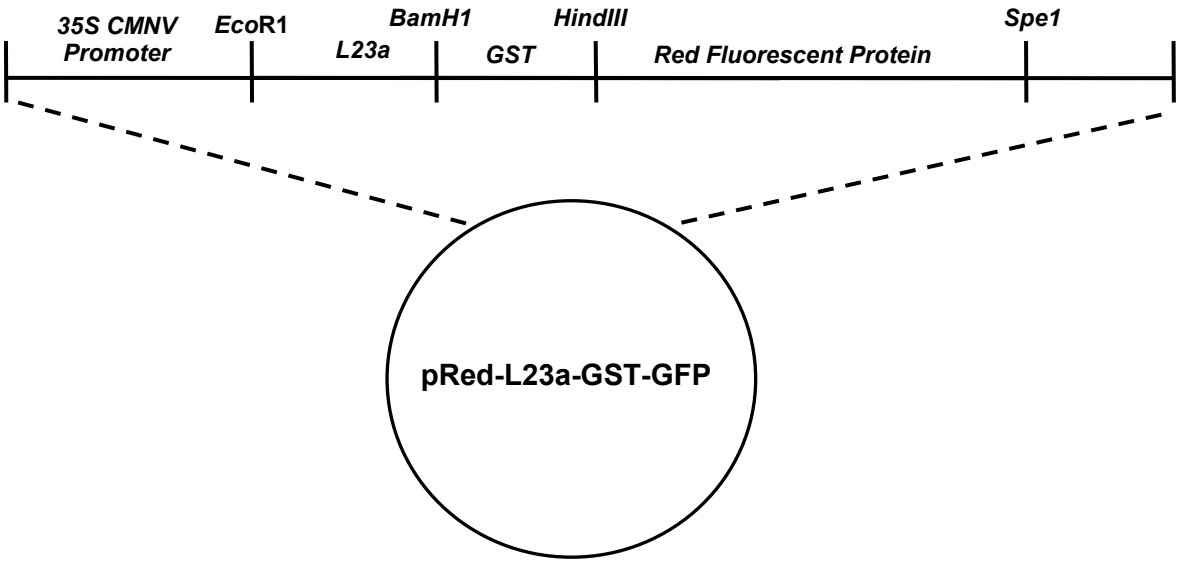

Figure A-3 Graphical representation of pRed-L23a-GST-GFP (Degenhard and Bonham-Smith, 2008). Dashed lines indicate inserted casette into the MCS; solid circle indicates pGreen backbone.

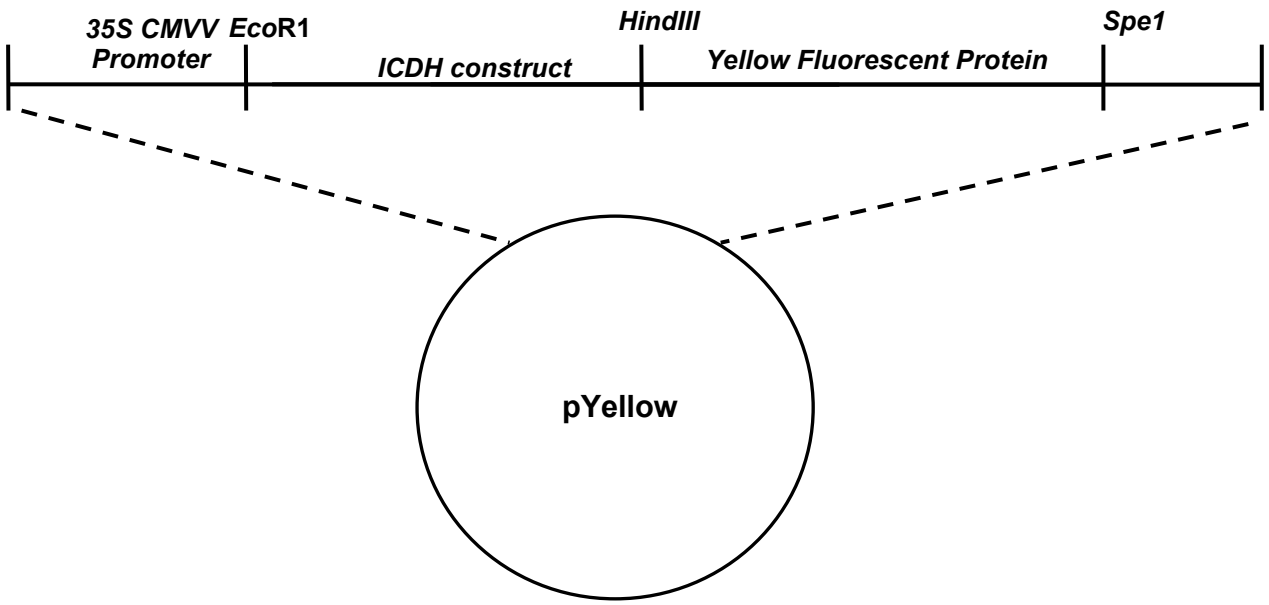

Figure A-4 Graphical representation of the pYellow reporter vector. Dashed lines indicate inserted casette into the MCS.

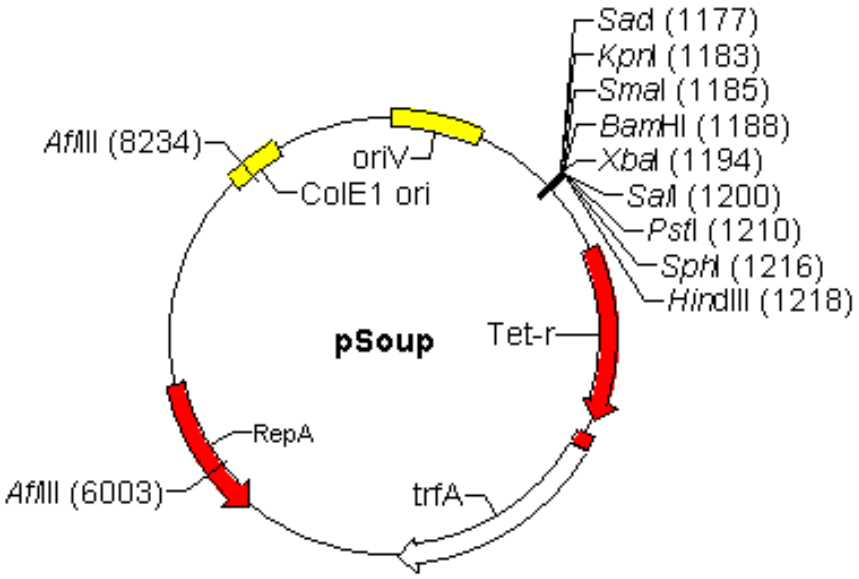

Figure A-5 Graphical representation of pSoup.

APPENDIX B

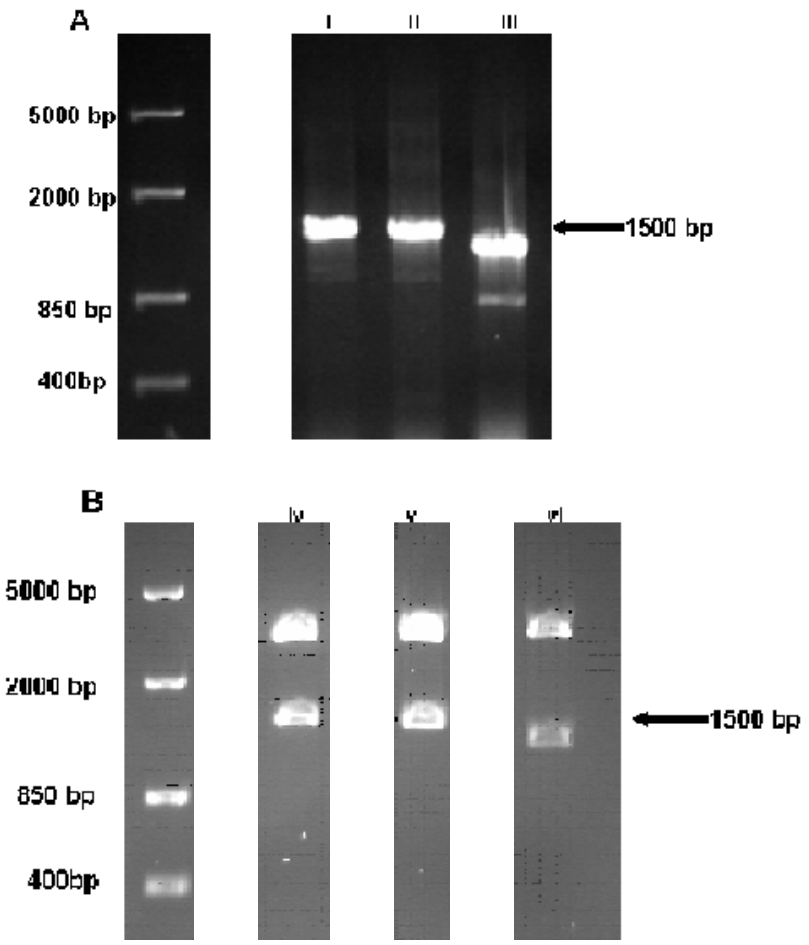

Figure B-1 Agarose gel analyses of full-length PCR products of $NADP^+$-ICDH from *N. tabacum* L. cv Petit Havana **A)** PCR amplification of **i)** mito-482 **ii)** chloro-482 and **iii)** cyto-482 DNA fragments. **B)** Double restriction digested pBluescript containing the DNA fragments **iv)** mito-482, **v)** chloro-482 and **vi)** cyto-482 using EcoR1 and HindIII.

```
mito-482        ------------------------------------------------------------
chloro-482      ------------------------------------------------------------
cyto-482        ------------------------------------------------------------
NCBI            GCACGGCGACAAAGACAATAGAGAGTTGCTCAGAGCAGCAAAAACTAGCAACTGATTAAA 60

mito-482        ---------------ATGCTTACCACCCGACTCAGACTCCGGTGTTCCGCCATGGCTAGT 45
chloro-482      ------------------------------------------------ATGGCTAGT 9
cyto-482        ------------------------------------------------------------
NCBI            GCCTACAAAAGTGTTATGCTTACCACCCGACTCAGACTCCGGTGTTCCGCCATGGCTAGT 120

mito-482        GTTGCTTCTTTTATCTCATCTTCATCGGCTTCAACATCATCCGCAGTTACCAAAAACCTT 105
chloro-482      GTTGCTTCTTTTATCTCATCTTCATCGGCTTCAACATCATCCGCAGTTACCAAAAACCTT 69
cyto-482        ------------------------------------------------------------
NCBI            GTTGCTTCTTTTATCTCATCTTCATCGGCTTCAACATCATCCGCAGTTACCAAAAACCTT 180

mito-482        CCCTTTTCAATCATCTCCAATCGGCAACTGTTCAAGAACCGTGTTTATCTCCTCCACCGA 165
chloro-482      CCCTTTTCAATCATCTCCAATCGGCAACTGTTCAAGAACCGTGTTTATCTCCTCCACCGA 129
cyto-482        ------------------------------------------------------------
NCBI            CCCTTTTCCATCATTTCCAATCGGCAACTGTTCAAGAACCGTGTTTATCTCCTCCACCGA 240

mito-482        ATCCCCAATGCTTCAATTCGATGCTTCGCTTCCACTACAGCTTCGTCTAAAATCCGCGTC 225
chloro-482      ATCCCCAATGCTTCAATTCGATGCTTCGCTTCCACTACAGCTTCGTCTAAAATCCGCGTC 189
cyto-482        ------------------------------------------ATGTCTAAAATCCGCGTC 15
NCBI            ATCCCCAATGCTTCAATTCGATCGTTCGCTTCCACTACAGCTTCGTCTAAAATCCGCGTT 300
                                                            **************

mito-482        GAAAATCCTATTGTCGAAATGGACGGTGATGAAATGACGAGGGTTATATGGACAATGATC 285
chloro-482      GAAAATCCTATTGTCGAAATGGACGGTGATGAAATGACGAGGGTTATATGGACAATGATC 249
cyto-482        GAAAATCCTATTGTCGAAATGGACGGTGATGAAATGACGAGGGTTATATGGACAATGATC 75
NCBI            GAAAATCCTATTGTCGAAATGGACGGTGATGAAATGACGAGGGTTATATGGACAATGATC 360
                ************************************************************

mito-482        AAAGAGAAGCTAATATATCCTTATCTAGAGTTGGATACGAAGTATTACGATTTGGGGATA 345
chloro-482      AAAGAGAAGCTAATATATCCTTATCTAGAGTTGGATACGAAGTATTACGATTTGGGGATA 309
cyto-482        AAAGAGAAGCTAATATATCCTTATCTAGAGTTGGATACGAAGTATTACGATTTGGGGATA 135
NCBI            AAAGAGAAGCTAATATATCCTTATCTAGAGTTGGATACGAAGTATTACGATTTGGGGATA 420
                ************************************************************

mito-482        TTGAACCGTGATGCCACTGACGATCAAGTTACTGTTGAAAGTGCTGAGGCTACTCTTAAG 405
chloro-482      TTGAACCGTGATGCCACTGACGATCAAGTTACTGTTGAAAGTGCTGAGGCTACTCTTAAG 369
cyto-482        TTGAACCGTGATGCCACTGACGATCAAGTTACTGTTGAAAGTGCTGAGGCTACTCTTAAG 195
NCBI            TTGAACCGTGATGCCACTGACGATCAAGTTACTGTTGAAAGTGCTGAGGCTACTCTTAAG 480
                ************************************************************

mito-482        TATAATGTTGCTGTGAAATGCGCTACTATAACACCTGATGAGACCAGAGTTAAGGAATTT 465
chloro-482      TATAATGTTGCTGTGAAATGCGCTACTATAACACCTGATGAGACCAGAGTTAAGGAATTT 429
cyto-482        TATAATGTTGCTGTGAAATGCGCTACTATAACACCTGATGAGACCAGAGTTAAGGAATTT 255
NCBI            TATAATGTTGCTGTGAAATGCGCTACTATAACACCTGATGAGACCAGAGTTAAGGAATTT 540
                ************************************************************

mito-482        GGGCTGAAGTCTATGTGAAGAAGTCCCAATGGCACAATCAGAAACATTTTAAATGGTACT 525
chloro-482      GGGCTGAAGTCTATGTGAAGAAGTCCCAATGGCACAATCAGAAACATTTTAAATGGTACT 489
cyto-482        GGGCTGAAGTCTATGTGAAGAAGTCCCAATGGCACAATCAGAAACATTTTAAATGGTACT 315
NCBI            GGGCTGAAGTCTATGTGGAGAAGTCCCAATGCCACAATCAGAAACATTTTAAATGGTACT 600
                ***************** ************* ****************************

mito-482        GTTTTCCGGGAGCCTATACTATGCAAGAACGTCCCCAGAATTGTTCCTGGTTGGGAGAAA 585
chloro-482      GTTTTCCGGGAGCCTATACTATGCAAGAACGTCCCCAGAATTGTTCCTGGTTGGGAGAAA 549
cyto-482        GTTTTCCGGGAGCCTATACTATGCAAGAACGTCCCCAGAATTGTTCCTGGTTGGGAGAAA 375
NCBI            GTTTTCCGGGAGCCTATACTATGCAAGAACGTCCCCAGAATTGTTCCTGGTTGGAAGAAA 660
                ************************************************** *****

mito-482        CCCATTTGTATTGGTAGGCATGCTTTTGGTGACCAGTATCGTGCCACAGATGCAGTTATT 645
chloro-482      CCCATTTGTATTGGTAGGCATGCTTTTGGTGACCAGTATCGTGCCACAGATGCAGTTATT 609
cyto-482        CCCATTTGTATTGGTAGGCATGCTTTTGGTGACCAGTATCGTGCCACAGATGCAGTTATT 435
NCBI            CCCATTTGTATTGGTAGGCATGCCTTTGGTGACCAGTATCGTGCCACAGATGCAGTTATT 720
                ********************** ***********************************
```

```
mito-482        AATGGACCAGGAAAGCTCAAAATGGTTTTTGAGCCAGAAAATGGGGAAGCCCCTACGGAA 705
chloro-482      AATGGACCAGGAAAGCTCAAAATGGTTTTTGAGCCAGAAAATGGGGAAGCCCCTACGGAA 669
cyto-482        AATGGACCAGGAAAGCTCAAAATGGTTTTTGAGCCAGAAAATGGGGAAGCCCCTACGGAA 495
NCBI            AATGGACCAGGAAAGCTCAAAATGGTTTTTGAGCCAGAAAATGGGGAAGCCCCTACGGAA 780
                ************************************************************

mito-482        CTGGATGTTTATGATTTTAAAGGTCCAGGTGTTGCACTTGCCATGTACAATGTTGACCAG 765
chloro-482      CTGGATGTTTATGATTTTAAAGGTCCAGGTGTTGCACTTGCCATGTACAATGTTGACCAG 729
cyto-482        CTGGATGTTTATGATTTTAAAGGTCCAGGTGTTGCACTTGCCATGTACAATGTTGACCAG 555
NCBI            CTGGATGTTTATGATTTTAAAGGTCCAGGTGTTGCACTTGCCATGTACAATGTTGACCAG 840
                ************************************************************

mito-482        TCAATTCGAGCGTTTGCTGAATCATCAATGTCAATGGTATTTTCGAAGAAATGGCCTCTT 825
chloro-482      TCAATTCGAGCGTTTGCTGAATCATCAATGTCAATGGTATTTTCGAAGAAATGGCCTCTT 789
cyto-482        TCAATTCGAGCGTTTGCTGAATCATCAATGTCAATGGTATTTTCGAAGAAATGGCCTCTT 615
NCBI            TCAATTCGAGCGTTTGCTGAATCATCAATGTCAATGGCATTTTCGAAGAAATGGCCTCTT 900
                ********************************** **********************

mito-482        TATTTGAGTACAAAAAATACAATACTAAAGAAATACGATGGCAGGTTTAAGGACATTTTT 885
chloro-482      TATTTGAGTACAAAAAATACAATACTAAAGAAATACGATGGCAGGTTTAAGGACATTTTT 849
cyto-482        TATTTGAGTACAAAAAATACAATACTAAAGAAATACGATGGCAGGTTTAAGGACATTTTT 675
NCBI            TATTTGAGTACAAAAAATACAATACTAAAGAAATACGATGGCAGGTTTAAGGACATTTTT 960
                ************************************************************

mito-482        GAAGAGGTATATGAAGAGAAGTGGAAGCAACAGTTTGAGGAACACTCGATATGGTATGAG 945
chloro-482      GAAGAGGTATATGAAGAGAAGTGGAAGCAACAGTTTGAGGAACACTCGATATGGTATGAG 909
cyto-482        GAAGAGGTATATGAAGAGAAGTGGAAGCAACAGTTTGAGGAACACTCGATATGGTATGAG 735
NCBI            GAAGAGGTATATGAAGAGAAGTGGAAGCAACAGTTTGAGGAACACTCGATATGGTATGAG 1020
                ************************************************************

mito-482        CATAGATTGATAGATGACATGGTAGCTTATGCATTAAAAAGCGGGGGTGGATATGTTTGG 1005
chloro-482      CATAGATTGATAGATGACATGGTAGCTTATGCATTAAAAAGCGGGGGTGGATATGTTTGG 969
cyto-482        CATAGATTGATAGATGACATGGTAGCTTATGCATTAAAAAGCGGGGGTGGATATGTTTGG 795
NCBI            CATAGATTGATAGATGACATGGTAGCTTATGCATTAAAAAGCGGGGGTGGATATGTTTGG 1080
                ************************************************************

mito-482        GCATGCAAGAACTATGATGGAGATGTCCAGAGTGATCTGCTCGCTCAAGGATTTGGTTCT 1065
chloro-482      GCATGCAAGAACTATGATGGAGATGTCCAGAGTGATCTGCTCGCTCAAGGATTTGGTTCT 1029
cyto-482        GCATGCAAGAACTATGATGGAGATGTCCAGAGTGATCTGCTCGCTCAAGGATTTGGTTCT 855
NCBI            GCATGCAAGAACTATGATGGAGATGTCCAGAGTGATCTGCTCGCTCAAGGATTTGGTTCT 1140
                ************************************************************

mito-482        CTGGGCCTCATGACCTCTGTATTGTTATCTTCTGATGGCAAGACATTAGAAGCTGAAGCA 1125
chloro-482      CTGGGCCTCATGACCTCTGTATTGTTATCTTCTGATGGCAAGACATTAGAAGCTGAAGCA 1089
cyto-482        CTGGGCCTCATGACCTCTGTATTGTTATCTTCTGATGGCAAGACATTAGAAGCTGAAGCA 915
NCBI            CTGGGCCTCATGACCTCTGTATTGTTATCTTCTGATGGCAAGACATTAGAAGCTGAAGCA 1200
                ************************************************************

mito-482        GCTCATGGCACAGTAACCAGACATTTTCGGCTGCATCAAAAGGGTCAAGAAACTAGTACA 1185
chloro-482      GCTCATGGCACAGTAACCAGACATTTTCGGCTGCATCAAAAGGGTCAAGAAACTAGTACA 1149
cyto-482        GCTCATGGCACAGTAACCAGACATTTTCGGCTGCATCAAAAGGGTCAAGAAACTAGTACA 975
NCBI            GCTCATGGCACAGTAACCAGACATTTTCGGCTGCATCAAAAGGGTCAAGAAACTAGTACA 1260
                ************************************************************

mito-482        AATAGTGCTGCTTCTATTTTTGCATGGGCAAGGGGACTTGGACATAGGGCCCAGCTTGAT 1245
chloro-482      AATAGTGCTGCTTCTATTTTTGCATGGGCAAGGGGACTTGGACATAGGGCCCAGCTTGAT 1209
cyto-482        AATAGTGCTGCTTCTATTTTTGCATGGGCAAGGGGACTTGGACATAGGGCCCAGCTTGAT 1035
NCBI            AATAGTGTTGCTTCTATTTTTGCATGGGCAAGGGGACTTGGACATAGGGCCCAGCTTGAT 1320
                ******* ****************************************************

mito-482        GGGAACCAAAAGTTATCTGAATTTGTTCACGCCCTGGGAGCTGCTTGCGTTGGCACAATA 1305
chloro-482      GGGAACCAAAAGTTATCTGAATTTGTTCACGCCCTGGGAGCTGCTTGCGTTGGCACAATA 1269
cyto-482        GGGAACCAAAAGTTATCTGAATTTGTTCACGCCCTGGGAGCTGCTTGCGTTGGCACAATA 1095
NCBI            GGGAACCAAAAGTTATCTGAATTTGTTCACGCCCTGGAAGCTGCTTGCGTTGGCACAATA 1380
                ********************************** **********************

mito-482        GAGTCCGGGAAGATGACTAAGGATTTAGCTATATTGGTTCATGGACCCAAGGTATCAAGG 1365
chloro-482      GAGTCCGGGAAGATGACTAAGGATTTAGCTATATTGGTTCATGGACCCAAGGTATCAAGG 1329
cyto-482        GAGTCCGGGAAGATGACTAAGGATTTAGCTATATTGGTTCATGGACCCAAGGTATCAAGG 1155
NCBI            GAGTCCGGGAAGATGACTAAGGATTTAGCTATATTGGTTCATGGACCCAAGGTATCAAGG 1440
```

```
                ************************************************************

mito-482        GAACACTACTTGAATACTGAAGAATTTATTGATGCTGTAGCACAGAAACTTCAAGAGAAG 1425
chloro-482      GAACACTACTTGAATACTGAAGAATTTATTGATGCTGTAGCACAGAAACTTCAAGAGAAG 1389
cyto-482        GAACACTACTTGAATACTGAAGAATTTATTGATGCTGTAGCACAGAAACTTCAAGAGAAG 1215
NCBI            GAACACTACTTGAATACTGAAGAATTTATTGATCCTGTAGCACAGAAACTTCAAGAGAAG 1500
                ******************************* ************************

mito-482        CTCGGTGCCTGCGCAGTTGTA 1446
chloro-482      CTCGGTGCCTGCGCAGTTGTA 1410
cyto-482        CTCGGTGCCTGCGCAGTTGTA 1236
NCBI            CTCGGTGCCTGCGCAGTTGTA 1521
                *********************
```

Figure B-2 ClustalW alignment of nucleotide sequence from full-length constructs of $NADP^+$-ICDH generated from *N. tabacum* L. cv Petit Havana. Nucleotide sequence from *N. tabacum* L. cv Xanthi $NADP^+$-ICDH (GenBank accession X96728.1) was aligned with mito-482, chloro-482 and cyto-482. Differing nucleotide residues and the introduced ATG site in cyto-482 are highlighted in yellow, * indicates matching sequence.

```
mito-482      MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60
chloro-482    ------------MASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 48
cyto-482      ------------------------------------------------------------
NCBI          MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60

mito-482      IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
chloro-482    IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 108
cyto-482      ----------MKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 50
NCBI          IRSFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
                        ************************************************

mito-482      TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM#RSPNGTIRNILNGTVFREP 179
chloro-482    TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM#RSPNGTIRNILNGTVFREP 167
cyto-482      TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM#RSPNGTIRNILNGTVFREP 109
NCBI          TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSMWRSPNATIRNILNGTVFREP 180
              ************************************** ****.**************

mito-482      ILCKNVPRIVPGWEKPICIGRHAFGDQYRATDAVINGPGKLKMVFEPENGEAPTELDVYD 239
chloro-482    ILCKNVPRIVPGWEKPICIGRHAFGDQYRATDAVINGPGKLKMVFEPENGEAPTELDVYD 227
cyto-482      ILCKNVPRIVPGWEKPICIGRHAFGDQYRATDAVINGPGKLKMVFEPENGEAPTELDVYD 169
NCBI          ILCKNVPRIVPGWKKPICIGRHAFGDQYRATDAVINGPGKLKMVFEPENGEAPTELDVYD 240
              *************:************************************************

mito-482      FKGPGVALAMYNVDQSIRAFAESSMSMVFSKKWPLYLSTKNTILKKYDGRFKDIFEEVYE 299
chloro-482    FKGPGVALAMYNVDQSIRAFAESSMSMVFSKKWPLYLSTKNTILKKYDGRFKDIFEEVYE 287
cyto-482      FKGPGVALAMYNVDQSIRAFAESSMSMVFSKKWPLYLSTKNTILKKYDGRFKDIFEEVYE 229
NCBI          FKGPGVALAMYNVDQSIRAFAESSMSMAFSKKWPLYLSTKNTILKKYDGRFKDIFEEVYE 300
              ***************************.********************************

mito-482      EKWKQQFEEHSIWYEHRLIDDMVAYALKSGGGYVWACKNYDGDVQSDLLAQGFGSLGLMT 359
chloro-482    EKWKQQFEEHSIWYEHRLIDDMVAYALKSGGGYVWACKNYDGDVQSDLLAQGFGSLGLMT 347
cyto-482      EKWKQQFEEHSIWYEHRLIDDMVAYALKSGGGYVWACKNYDGDVQSDLLAQGFGSLGLMT 289
NCBI          EKWKQQFEEHSIWYEHRLIDDMVAYALKSGGGYVWACKNYDGDVQSDLLAQGFGSLGLMT 360
              ************************************************************

mito-482      SVLLSSDGKTLEAEAAHGTVTRHFRLHQKGQETSTNSAASIFAWARGLGHRAQLDGNQKL 419
chloro-482    SVLLSSDGKTLEAEAAHGTVTRHFRLHQKGQETSTNSAASIFAWARGLGHRAQLDGNQKL 407
cyto-482      SVLLSSDGKTLEAEAAHGTVTRHFRLHQKGQETSTNSAASIFAWARGLGHRAQLDGNQKL 349
NCBI          SVLLSSDGKTLEAEAAHGTVTRHFRLHQKGQETSTNSVASIFAWARGLGHRAQLDGNQKL 420
              ************************************.***********************

mito-482      SEFVHALGAACVGTIESGKMTKDLAILVHGPKVSREHYLNTEEFIDAVAQKLQEKLGACA 479
chloro-482    SEFVHALGAACVGTIESGKMTKDLAILVHGPKVSREHYLNTEEFIDAVAQKLQEKLGACA 467
cyto-482      SEFVHALGAACVGTIESGKMTKDLAILVHGPKVSREHYLNTEEFIDAVAQKLQEKLGACA 409
NCBI          SEFVHALEAACVGTIESGKMTKDLAILVHGPKVSREHYLNTEEFIDPVAQKLQEKLGACA 480
              ******* ************************************.*************

mito-482      VV 481
chloro-482    VV 469
cyto-482      VV 411
NCBI          VV 482
              **
```

Figure B-3 ClustalW alignment of deduced amino acid sequence from full-length constructs of $NADP^+$-ICDH generated from *N. tabacum* L. cv Petit Havana. Deduced amino acid sequence from *N. tabacum* L. cv Xanthi $NADP^+$-ICDH (NCBI acession CAA65503) was aligned with

mito-482, chloro-482 and cyto-482. Differing amino acid residues are highlighted in yellow. A stop codon is indicated with a number sign (#), * indicates matching sequence.

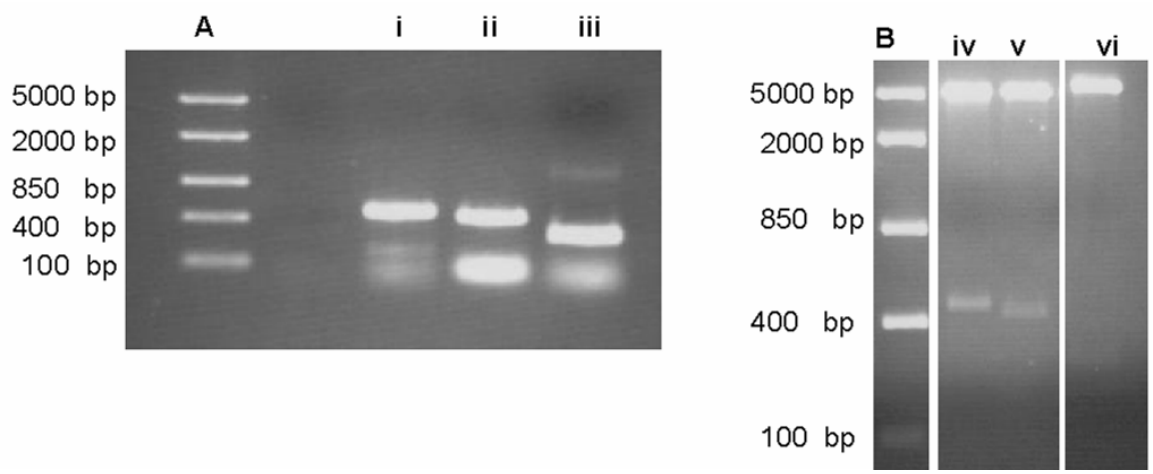

Figure B-4 Agarose gel analyses of PCR products from *N. tabacum* L. cv Petit Havana resulting in truncation at amino acid 160 of $NADP^+$-ICDH. **A)** PCR amplification of **i)** mito-160, **ii)** chloro-148 and **iii)** cyto-90 DNA fragments. **B)** Double restriction digested pBluescript containing the DNA fragments **iv)** mito-160, **v)** chloro-148 and **vi)** cyto-90 using EcoR1 and HindIII.

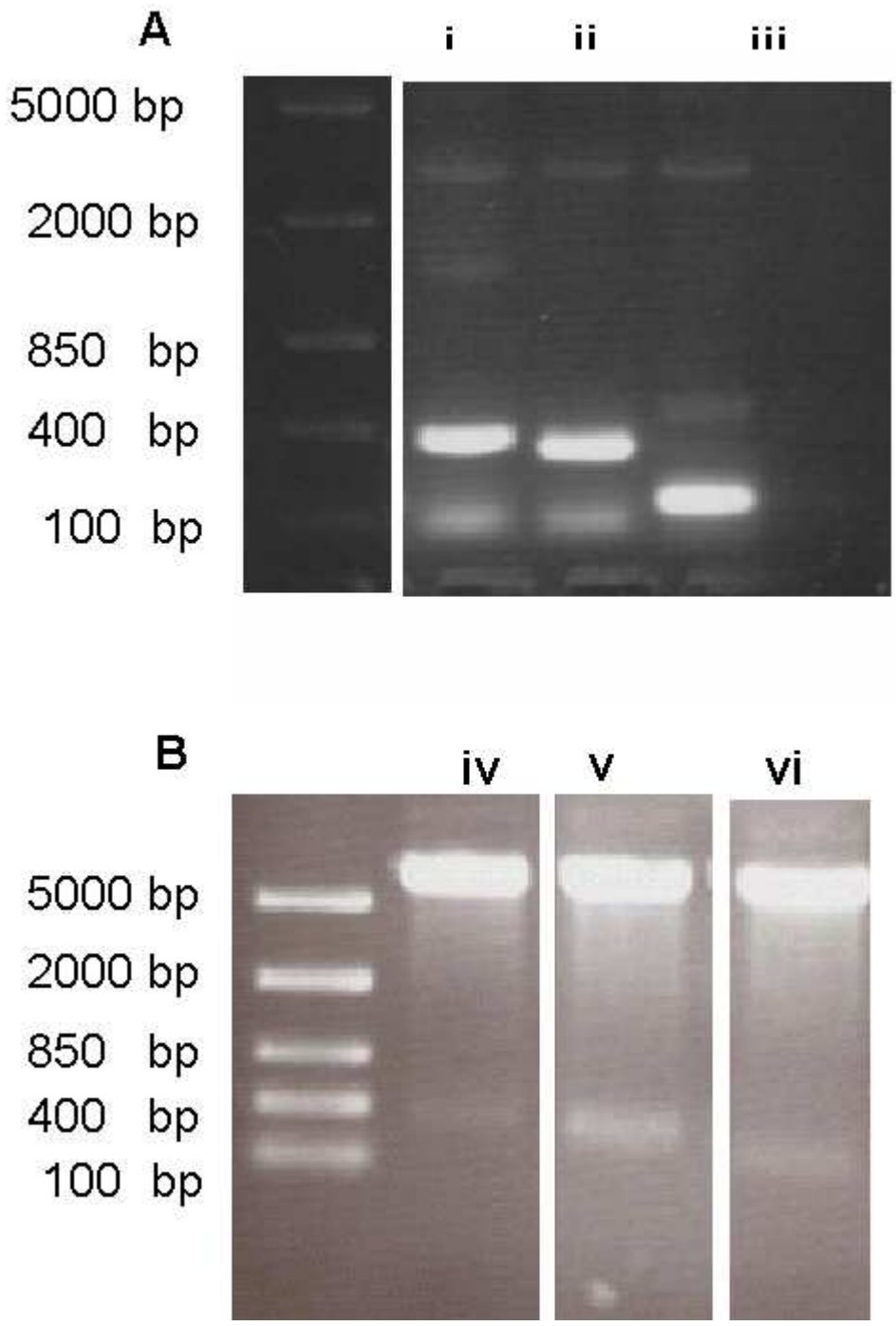

Figure B-5 Agarose gel analyses of PCR products from *N. tabacum* L. cv Petit Havana resulting in truncation at amino acid 138 of $NADP^+$-ICDH. **A)** PCR amplification of **i)** mito-110, **ii)** chloro-98 and **iii)** cyto-40 DNA fragments. **B)** Double restriction digested pBluescript containing the DNA fragments **iv)** mito-110, **v)** chloro-98 and **vi)** cyto-40 using EcoR1 and HindIII.

A

```
mito-160        MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60
chloro-148      ------------MASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 48
cyto-90         ------------------------------------------------------------
NCBI            MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60


mito-160        IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
chloro-148      IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 108
cyto-90         ----------MKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 50
NCBI            IRSFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
                          ************************************************

mito-160        TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM-------------------- 160
chloro-148      TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM-------------------- 148
cyto-90         TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM-------------------- 90
NCBI            TDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSMWRSPNATIRNILNGTVFREP 180
                ************************************
```

B

```
mito-110        MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60
chloro-98       ------------MASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 48
cyto-40         ------------------------------------------------------------
NCBI            MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60

mito-110        IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKY---------- 110
chloro-98       IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKY---------- 98
cyto-40         ----------MKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKY---------- 40
NCBI            IRSFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRDA 120
                          ************************************
```

Figure B-6 ClustalW alignment of deduced amino acid sequence from truncated constructs of $NADP^+$-ICDH generated from *N. tabacum* L. cv Petit Havana. Deduced amino acid sequence from *N. tabacum* L. cv Xanthi $NADP^+$-ICDH (NCBI acession CAA65503) was aligned **A**) mito-160, chloro-148 and cyto-90 and **B**) mito-110, chloro-98 and cyto-40. Differing amino acid residues are highlighted in yellow, * indicates matching sequence. The constructs in **B**) are designed, in part, based on Gálvez et al. (1998).

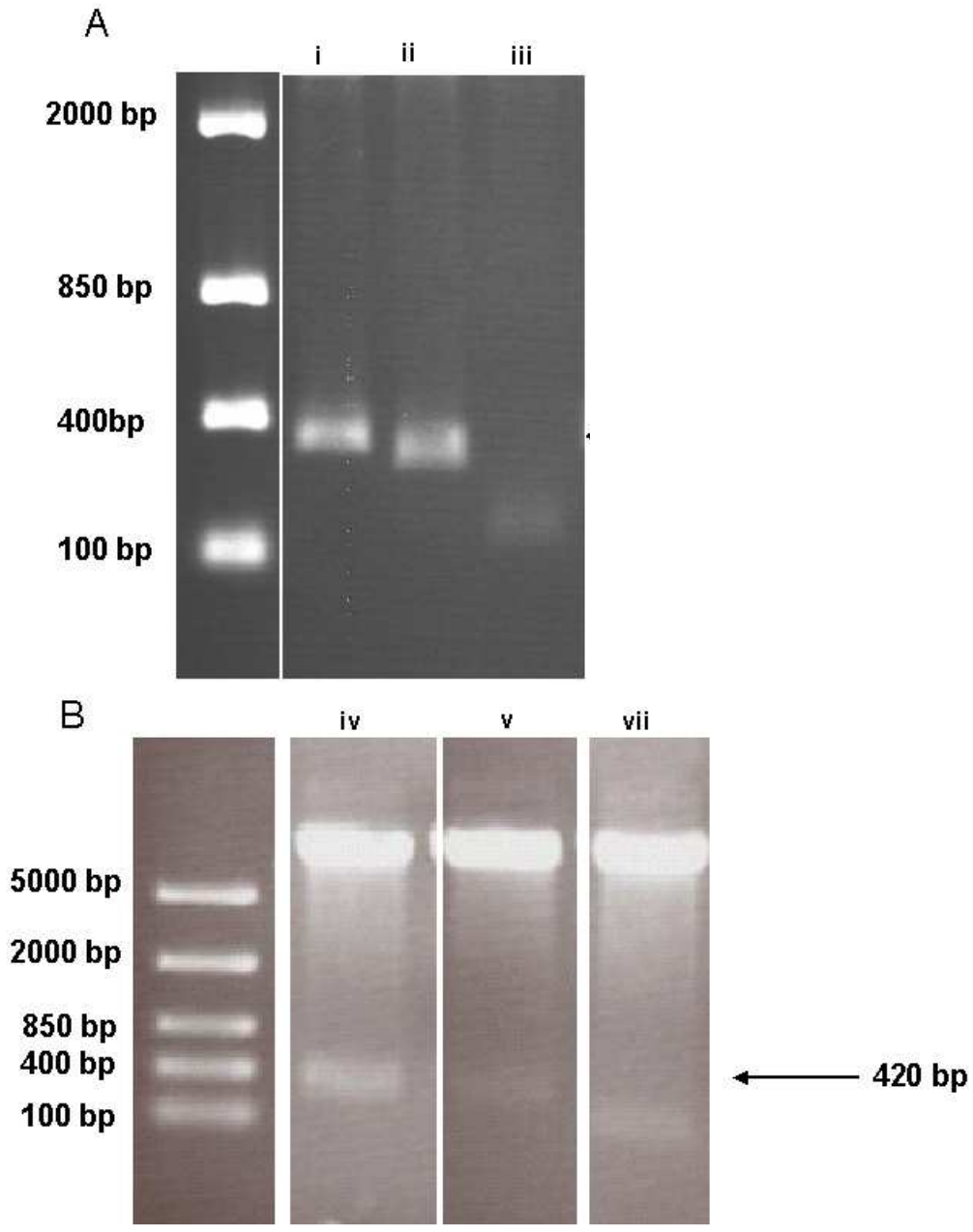

Figure B-7 Agarose gel analyses of ICDH:YFP fusion constructs from *N. tabacum* L. cv Petit Havana using EcoR1 and Hind III. **i)** mito-110, **ii)** chloro-98, **iii)** cyto-40, **iv)** mito-160, **v)** chloro-146 and **vi)** cyto-90.

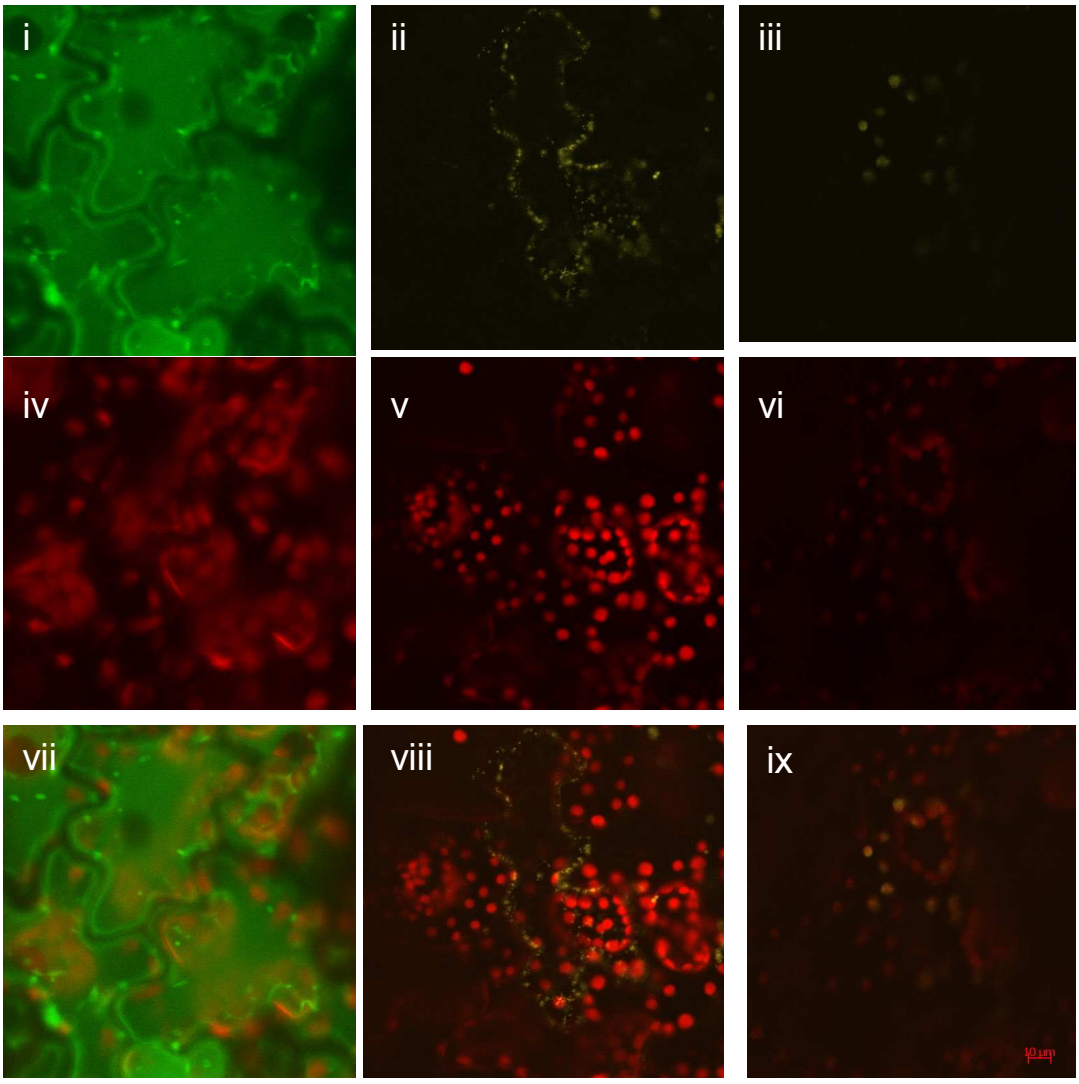

Figure B-8 Localization of putative mitochondrial ICDH:YFP fusion proteins generated from (*N. tabacum* L. cv Petit Havana). Confocal images of tobacco (*N. tabacum* L. cv Petit Havana) leaf epidermal cells after *A. tumefaciens* infiltration. Infiltration occurred when cultures had reached an OD_{600} = 0.20. **(i)** mitochondria stained with dihydrorhodiamine 123 **(ii)** mito-110 **(iii)** mito-160. Panels **(iv)** and **(v)** represent chloroplast autofluorescence. Panels **(vii)**, **(viii)** and **(ix)** are merged images of **(i)** and **(iv)**, **(ii)** and **(v)**, and **(iii)** and **(vi)** respectively.

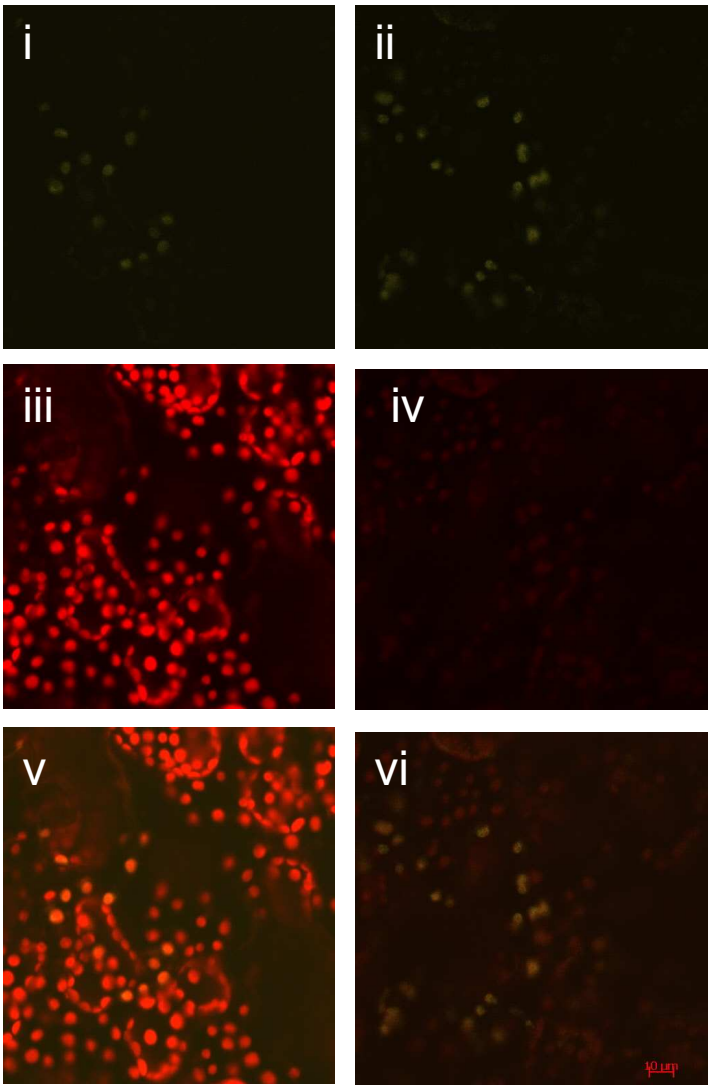

Figure B-9 Localization of putative chloroplastic ICDH:YFP fusion proteins generated from (*N. tabacum* L. cv Petit Havana). Confocal images of tobacco (*N. tabacum* L. cv Petit Havana) leaf epidermal cells after *A. tumefaciens* infiltration. Infiltration occurred when cultures had reached an OD_{600} = 0.20. **(i)** chloro-98 **(ii)** chloro-148. Panels **(iii)** and **(iv)** represent chloroplast autofluorescence. Panels **(v)** and **(vi)** are merged images of **(i)** and **(iii)** and **(ii)** and **(iv)** respectively.

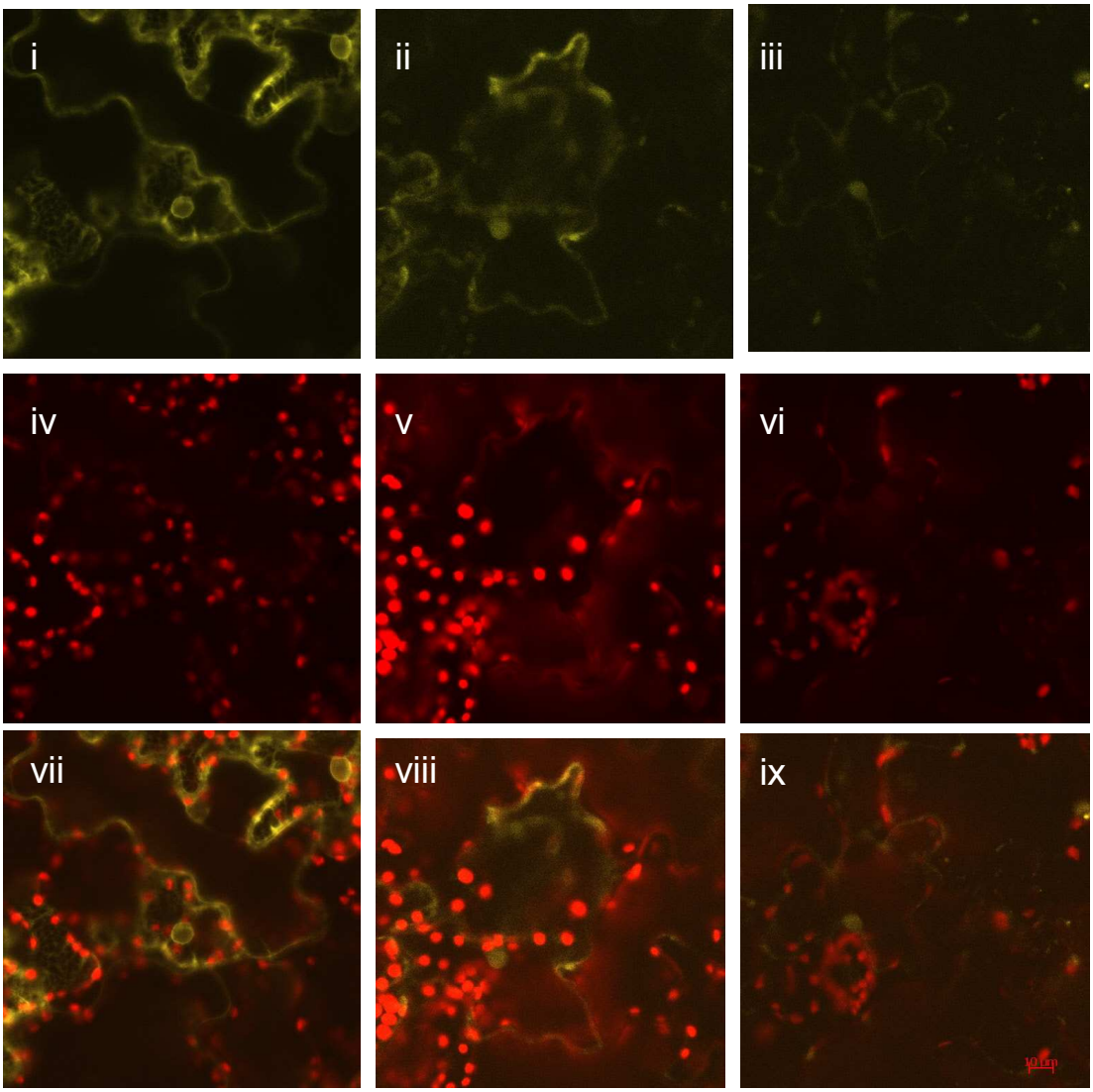

Figure B-10 Localization of putative cytosolic ICDH:YFP fusion proteins generated from (*N. tabacum* L. cv Petit Havana). Confocal images of tobacco (*N. tabacum* L. cv Petit Havana) leaf epidermal cells after *A. tumefaciens* infiltration. Infiltration occurred when cultures had reached an OD_{600} = 0.20. **(i)** Sec12 **(ii)** cyto-40 **(iii)** cyto-90. Panels **(iv)**, **(v)** and **(vi)** represent chloroplast autofluorescence. Panels **(vii)**, **(viii)** and **(ix)** are merged images of **(i)** and **(iv)**, **(ii)** and **(v)**, and **(iii)** and **(vi)** respectively

APPENDIX C

```
NCBI            MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60
mito-160        MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 60
chloro-148      ------------MASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNAS 48
cyto-90         ------------------------------------------------------------
EYFP            ------------------------------------------------------------

NCBI            IRSFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRD 120
Mito-160        IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRD 120
chloro-148      IRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRD 108
cyto-90         ----------MKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYYDLGILNRD 49
EYFP            ------------------------------------------------------------

NCBI            ATDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSM                    160
mito-160        ATDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSMKLMVSKGEELFTGVVPILV 180
chloro-148      ATDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSMKLMVSKGEELFTGVVPILV 168
cyto-90         ATDDQVTVESAEATLKYNVAVKCATITPDETRVKEFGLKSMKLMVSKGEELFTGVVPILV 109
EYFP            -------------------------------------------MVSKGEELFTGVVPILV 19
                                                           *****************

mito-160        ELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPD 240
chloro-148      ELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPD 228
cyto-90         ELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPD 169
EYFP            ELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLQCFARYPD 79
                ************************************************************

mito-160        HMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDPLVNRIELKGIDFKEDGNI 300
chloro-148      HMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDPLVNRIELKGIDFKEDGNI 288
cyto-90         HMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDPLVNRIELKGIDFKEDGNI 229
EYFP            HMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDPLVNRIELKGIDFKEDGNI 139
                ************************************************************

mito-160        LGHKLEYNYNSHNVYIMANKQKNGIKVNFKIRHNIEDGSVQLANHYQQNTPIGDGPVLLP 360
chloro-148      LGHKLEYNYNSHNVYIMANKQKNGIKVNFKIRHNIEDGSVQLANHYQQNTPIGDGPVLLP 348
cyto-90         LGHKLEYNYNSHNVYIMANKQKNGIKVNFKIRHNIEDGSVQLANHYQQNTPIGDGPVLLP 289
EYFP            LGHKLEYNYNSHNVYIMANKQKNGIKVNFKIRHNIEDGSVQLANHYQQNTPIGDGPVLLP 199
                ************************************************************

mito-160        DNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 402
chloro-148      DNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 390
cyto-90         DNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 331
EYFP            DNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 241
                **********************************
```

Figure C-1 ClustalW alignment showing ICDH:YFP fusion in constructs truncated at amino acid 160 generated from *N. tabacum* L. cv Petit Havana. Deduced amino acid sequence from *N. tabacum* L. cv Xanthi $NADP^+$-ICDH (NCBI acession CAA65503) and EYFP (NCBI acession AAF65454) were aligned with mito-160, chloro-148 and cyto-90. Differing amino acid residues are highlighted in yellow, * indicates matching sequence.

```
NCBI             MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNA 60
mito-110         MLTTRLRLRCSAMASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNA 60
chloro-98        ------------MASVASFISSSSASTSSAVTKNLPFSIISNRQLFKNRVYLLHRIPNA 48
cyto-40          ------------------------------------------------------------
EYFP             ------------------------------------------------------------

NCBI             SIRSFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKY          110
mito-110         SIRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYKLMVSKGEE 120
chloro-98        SIRCFASTTASSKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYKLMVSKGEE 108
cyto-40          -----------MKIRVENPIVEMDGDEMTRVIWTMIKEKLIYPYLELDTKYKLMVSKGEE 49
EYFP             ------------------------------------------------MVSKGEE 9
                                                                  *******

mito-110         LFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGY 180
chloro-98        LFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGY 168
cyto-40          LFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGY 109
EYFP             LFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGY 69
                 ************************************************************

mito-110         GLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDPLVNRIELK 240
chloro-98        GLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDPLVNRIELK 228
cyto-40          GLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDPLVNRIELK 169
EYFP             GLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDPLVNRIELK 129
                 ************************************************************

mito-110         GIDFKEDGNILGHKLEYNYNSHNVYIMANKQKNGIKVNFKIRHNIEDGSVQLANHYQQNT 300
chloro-98        GIDFKEDGNILGHKLEYNYNSHNVYIMANKQKNGIKVNFKIRHNIEDGSVQLANHYQQNT 288
cyto-40          GIDFKEDGNILGHKLEYNYNSHNVYIMANKQKNGIKVNFKIRHNIEDGSVQLANHYQQNT 229
EYFP             GIDFKEDGNILGHKLEYNYNSHNVYIMANKQKNGIKVNFKIRHNIEDGSVQLANHYQQNT 189
                 ************************************************************

mito-110         PIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 352
chloro-98        PIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 340
cyto-40          PIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 281
EYFP             PIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 241
                 ***************************************************
```

Figure C-2 ClustalW alignment showing ICDH:YFP fusion in constructs truncated at amino acid 138 generated from *N. tabacum* L. cv Petit Havana. Deduced amino acid sequence from *N. tabacum* L. cv Xanthi $NADP^+$-ICDH (NCBI acession CAA65503) and EYFP (NCBI acession AAF65454) were aligned with mito-110, chloro-98 and cyto-40. Differing amino acid residues are highlighted in yellow, * indicates matching sequence.