Isolation and characterization of thymine-7-hydroxylase from *Rhodotorula glutinis* and iso-orotate decarboxylase from *Neurospora crassa*

by

Armend A. Axhemi

Submitted in Partial Fulfillment of the Requirements

For the Degree of

Masters of Science

In the

Chemistry

Program

Youngstown State University

May, 2005

Isolation and characterization of thymine-7-hydroxylase from *Rhodotorula glutinis* and iso-orotate decarboxylase from *Neurospora crassa*

Armend A. Axhemi

I hereby release this thesis to the public. I understand that this thesis will be made available from OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

Armend Axhemi 5/9/05

Armend A. Axhemi, Student Date

Approvals:

Jeffrey A. Smiley 5/9/05

Jeffrey A. Smiley, Ph. D., Thesis Advisor Date

Peter N 5/9/05

Peter Norris Ph.D., Committee Member Date

Thomas D. Kim 5/6/05

Thomas D. Kim, Ph. D., Committee Member Date

Peter J. Kasvinsky 5/9/05

Peter J. Kasvinsky, Dean of Graduate Studies Date

Thesis Abstract

N.crassa and *R. glutinis* are known to utilize a unique salvage pathway for the production of uracil. This is known as the pyrimidine salvage pathway. This pathway consists of four enzymatic steps to convert thymidine to uracil. The third enzyme of the pathway is thymine-7-hydroxylase (THase), which forms uracil-5-carboxylic acid from thymine by way of three oxidative reactions. The final enzyme is iso-orotate decarboxylase (IDCase), which forms uracil from iso-orotate through a decarboxylation reaction. The THase gene was transferred into an expression vector. Multiple culture conditions were used to optimize protein production. THase activity was measured using radioactive [^{14}C] thymine assays. The iso-orotate decarboxylase (IDCase) gene was also transferred into an expression system. Multiple culture conditions were used to optimize protein production. Various protein purification methods were used in an attempt to purify IDCase. The conserved cysteine residue Cys272 was changed by site-directed mutagenesis to Ala272; the mutant protein had only slightly diminished activity.

Acknowledgements

This thesis is dedicated to the loving memory of my mother and father, Zejgjane and Qazim Axhemi, whose love and support have made me the person that I am today. I thank them and my whole family back home and in the U.S. for helping me fulfill my career goals.

I would also like to acknowledge all of the Department of Chemistry at Youngstown State University for the knowledge that they have imparted on me and the guidance that they have provided through troubled times. Especially, I would like to thank Dr. Jeffrey A. Smiley for being such a great role model and research advisor in the positive attitude, ideas, knowledge and patience that he conveys on all of us. Finally, I thank the School of Graduate Studies at YSU for making all of this possible by providing me with a teaching assistantship.

Table of Contents

Chapter I: Introduction

Nucleotides are biological molecules that participate as intermediates in virtually all aspects of cellular metabolism. For example, in addition to regulating many anabolic and catabolic pathways, nucleotide triphosphates (NTPs) serve as the energy currency of cells where the energy stored in their phosphodiester bonds is used to propel thermodynamically unfavorable reactions. Yet, their central biological purpose is to serve as the building blocks of nucleic acids, the elements of heredity and the agents of genetic information transfer. Nucleic acids are linear polymers of nucleotides where the orderly sequence of four different bases can encode biological information.

Nucleotides are made up of three chemical components in equal amounts: a nitrogenous base linked to a ribose sugar via a glycosidic bond and a phosphate group linked to the pentose via the C-5' hydroxyl group. The two basic kinds of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). The pentose sugar in DNA is 2-deoxyribose; in RNA, it is ribose. DNA stores the genetic information in cells, while RNA serves in the transcription and translation of this information into proteins.

The identity of a particular nucleotide is determined by the nitrogenous base that it contains. These bases can be derivatives of either pyrimidine or purine rings. Pyrimidines are six-membered heterocyclic aromatic rings containing two nitrogen atoms. The common naturally occurring pyrimidines are cytosine, uracil, and thymine (5-methyluracil). Cytosine and thymine are the pyrimidines typically found in DNA, whereas cytosine and uracil are common in RNA (Figure 1-1).

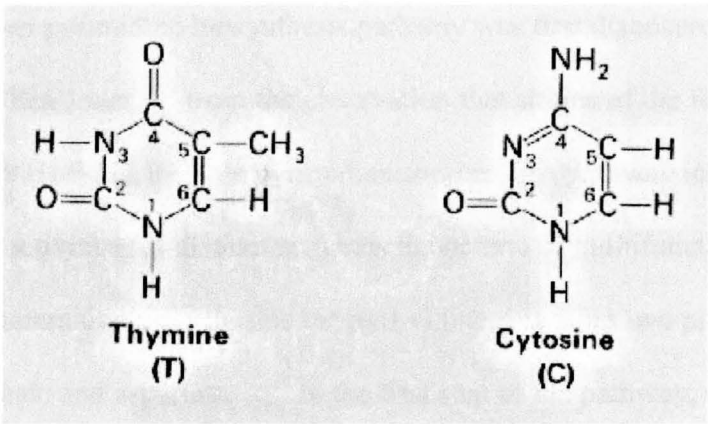


Figure 1-1. Thymine (5-methyluracil) and Cytosine

The purine ring structure is represented by the combination of a pyrimidine ring with a five-membered imidazole ring to yield a fused ring system. Adenine (6-amino purine) and guanine (2-amino-6-oxy purine), the two common purines, are found in both DNA and RNA (Figure 1-2).

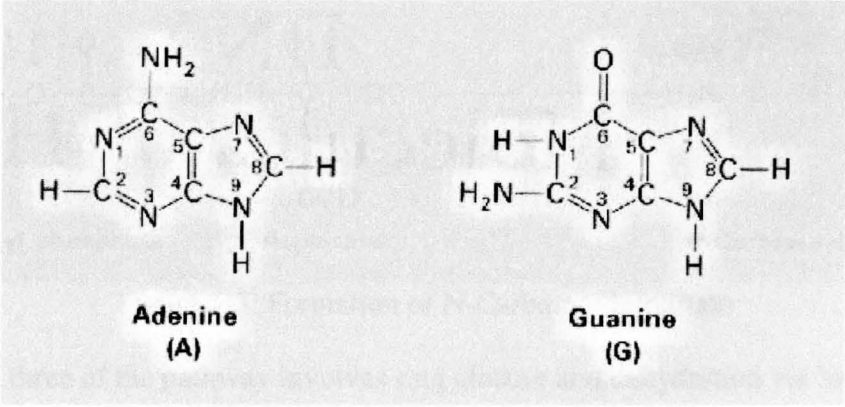


Figure 1-2. Adenine and Guanine

In both DNA and RNA, purine bases consistently pair up with pyrimidines to give the complementary base pairing that is so crucial to the central dogma of biology. In both molecules, cytosine consistently pairs up with guanine; meanwhile the adenine base discriminately pairs with thymine (5-methyluracil) in DNA and with uracil in RNA, respectively.

The *de novo* pyrimidine biosynthesis pathway was first discovered in *Neurospora crassa* by Mary Ellen Jones, [1] from the observation that strains of the fungi grew normally with orotic acid as the sole pyrimidine source. This pathway involves six different enzyme activities (as distinct enzymes in bacteria or multifunctional polypeptides in mammals) to synthesize the pyrimidine ring from two precursors, carbamoyl phosphate and aspartate. [2, 3] In the first step of the pathway, carbamoyl phosphate synthase II (CPSII) catalyzes the formation of carbamoyl phosphate from bicarbonate, glutamine, 2 ATP molecules and H_2O. In the second step, the condensation of carbamoyl phosphate and aspartate to yield *N*-carbamoylaspartate is catalyzed by aspartate transcarbamoylase, ATCase (Figure 1-3).

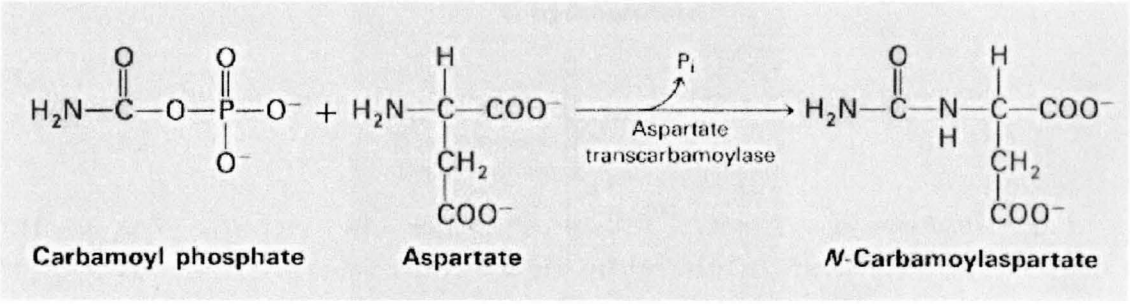


Figure 1-3. Formation of N-Carbamoylaspartate

Step three of the pathway involves ring closure and dehydration via linkage of the $-NH_2$ group introduced by carbamoyl phosphate with the former β-COO^- of aspartate; a reaction mediated by the enzyme dihydroorotase to yield dihydroorotate (DHO). In step four, DHO then undergoes an oxidation reaction by the transfer of two H atoms to either coenzyme Q (in mammals) or NAD^+-linked flavoproteins (in bacteria) with the mediation of DHO dehydrogenase to form orotate, which is a pyrimidine (Figure 1- 4).

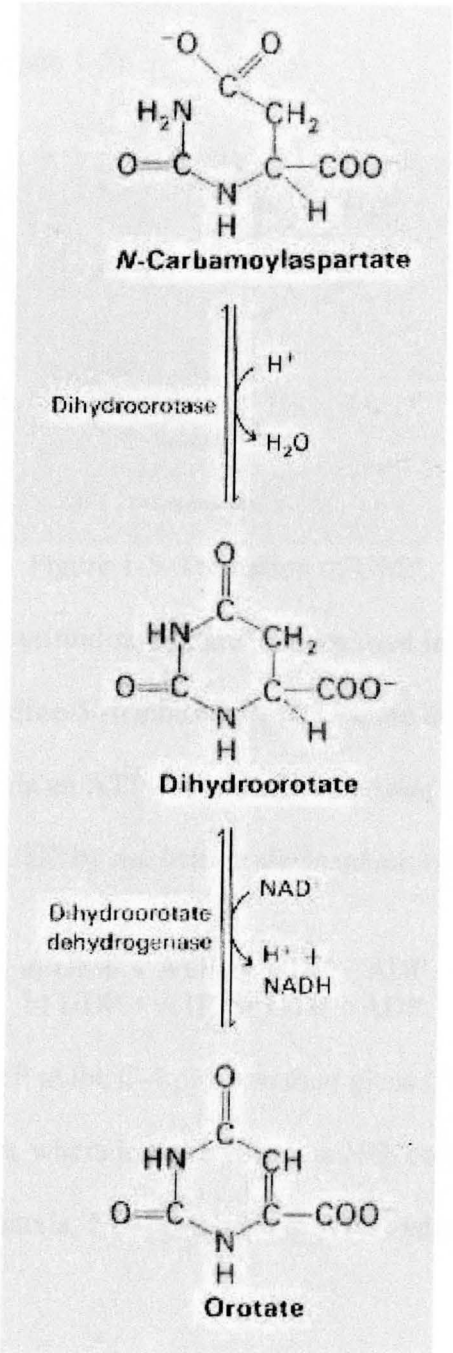


Figure 1-4. Formation of dihydroorotate and orotate.

Orotate then reacts with α-PRPP in the presence of orotate phosphoribosyltransferase, where ribose-5-phosphate is joined to its N-1 atom, giving the nucleotide orotidine-5'-monophosphate (OMP). The last step of the pathway is the

decarboxylation of OMP by OMP decarboxylase (ODCase) yielding uridine-5'-monophosphate (UMP) (Figure 1-5).

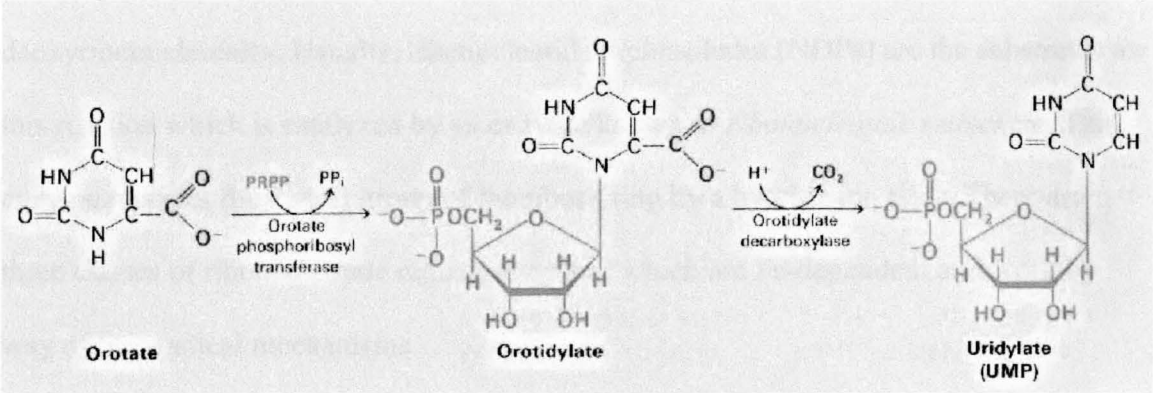


Figure 1-5. Formation of UMP.

The two prominent pyrimidine ribonucleotides used in RNA, uridine-5'-triphosphate (UTP) and cytidine-5'-triphosphate (CTP), are then derived from UMP. First, UDP is formed from UMP via an ATP-dependent *nucleoside monophosphate kinase (a)*. Then, UTP is formed from UDP by *nucleoside diphosphate kinase (b)*:

a) UMP + ATP ➔ UDP + ADP
b) UDP + ATP ➔ UTP + ADP

The amination of UTP at the C-4 position then gives CTP. This reaction is catalyzed by CTP synthetase, where in eukaryotes the NH_2 comes from the amide-N of glutamine; meanwhile in bacteria, NH_4^+ provides it. ATP hydrolysis provides the energy to drive this reaction.

Regulation of the *de novo* pyrimidine biosynthetic pathway is done by the allosteric regulation of the enzymes that catalyze the committed steps in this pathway. In *E. coli*, ATCase catalyzes the committed step and it is feedback-inhibited by the end product, CTP; meanwhile it is activated by ATP. In mammals, CPS-II catalyzes the

committed step. UDP and UTP are feedback inhibitors of CPS-II, while PRPP and ATP are allosteric activators. [2, 3]

In order to be incorporated into DNA, ribonucleotides must be reduced into deoxyribonucleotides. Usually, ribonucleotide diphosphates (NDPs) are the substrates for this reaction which is catalyzed by an enzyme known as *ribonucleotide reductase*. This enzyme replaces the 2'-OH group of the ribose ring by a hydride ion ($H:^-$). There are three classes of ribonucleotide reductases; all of which are Fe-dependent and work by way of free-radical mechanisms.

The synthesis of thymine (5-methyluracil) deoxyribonucleotide used in DNA proceeds from other pyrimidine deoxyribonucleotides, primarily dUDP and dCDP. Both dUDP and dCDP can lead to the formation of dUMP, the immediate precursor for dTMP synthesis:

From dUDP: dUDP → dUTP → dUMP → dTMP
From dCDP: dCDP → dCMP → dUMP → dTMP

Synthesis of dTMP from dUMP is then catalyzed by thymidylate synthase. This enzyme uses the folic acid derivative N^5, N^{10}-methylene-THF as a methyl donor and it methylates dUMP at the 5-position to create dTMP meanwhile releasing dihydrofolic acid (DHF) as a by-product, which is then recycled to form more of the N^5,N^{10}-methylene-THF (Figure 1-6).

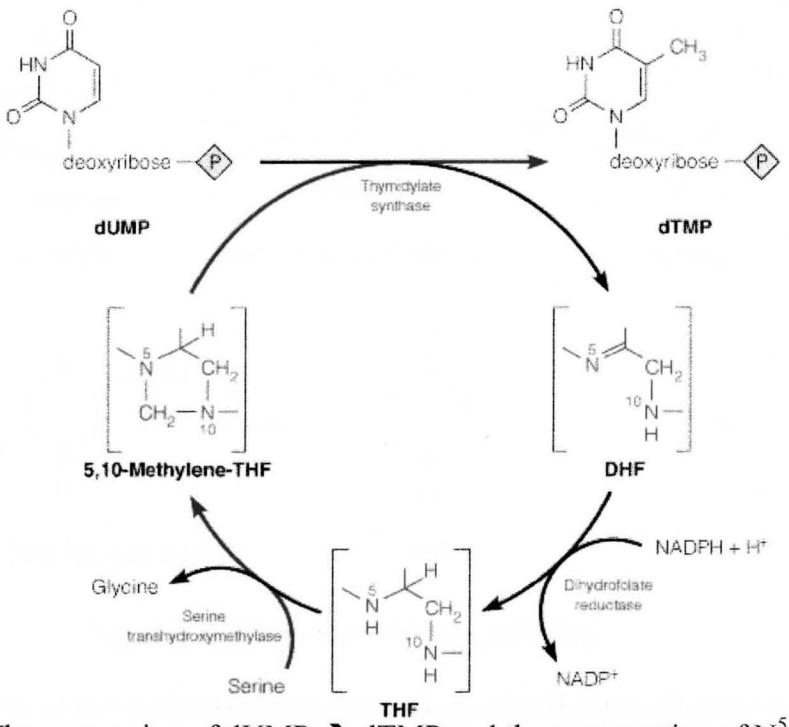


Figure 1-6. The conversion of dUMP ➔ dTMP and the regeneration of N^5,N^{10}-methylene-THF

The methylation of dUMP to dTMP is an irreversible, unidirectional reaction, and once the uracil ring has been methylated, there is no metabolic pathway in most organisms to demethylate the thymine back to uracil. This unidirectionality is usually considered a metabolic need for the regulation of thymidylate synthase activity and does not allow thymine nucleosides and nucleotides to be reused in the total pyrimidine pool.

Some organisms, however, have evolved a unique metabolic pathway for the conversion of thymidine to uracil, which can then be converted to UMP and re-enter the total pyrimidine nucleotide pool. This pathway was first identified in *Neurospora crassa* and other fungi (such as *Aspergillus nidulans* and *Rhodotorula glutinis*) by Palmatier *et al* upon the observation of the incorporation of ring-labeled thymidine predominantly into RNA.[4] The enzymes responsible for this metabolic conversion have thus been referred to collectively as the thymidine salvage pathway (Figure 1-7).

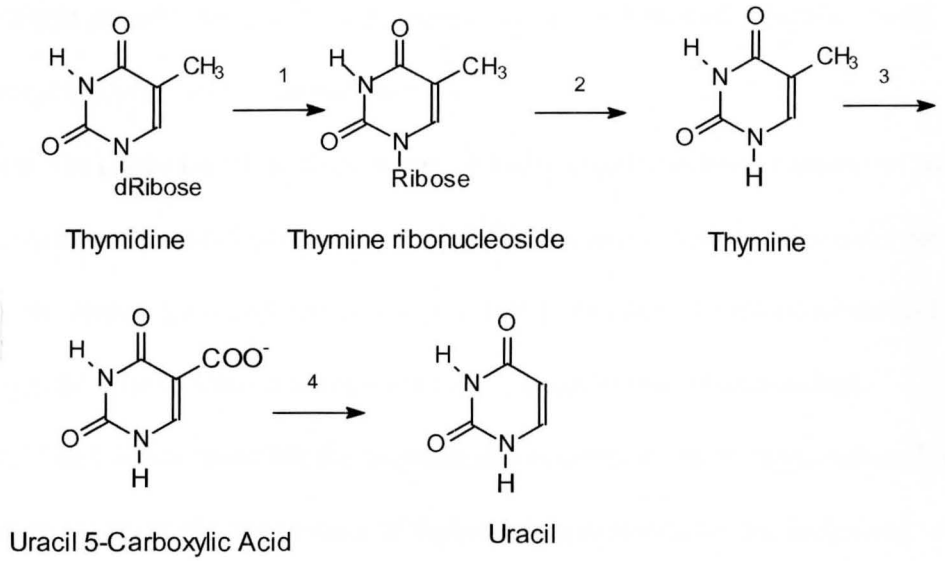


Figure 1-7. Thymidine salvage pathway.

The enzymes of this pathway disassemble thymidine into uracil (plus ribose and CO_2) through the following enzymatic steps: (1) oxidation of thymidine, a deoxyribonucleoside, to thymine ribonucleoside; (2) hydrolytic cleavage of the glycosidic bond to yield thymine and ribose; (3) a three-step oxidation of thymine to uracil-5-carboxylate (also referred to as isoorotate, IOA) and (5) decarboxylation of IOA to produce uracil.[5] Once uracil is formed, it can then be utilized by most organisms via its conversion to UMP by the enzyme uracil phosphoribosyltransferase.

The overall metabolic role of the thymidine salvage pathway is unknown and until recently none of the genes encoding the enzymes had been isolated, and only one of the enzymes, thymine hydroxylase, has been studied in any detail. Thymine-7-hydroxylase (THase) is the third enzyme of this pathway and it is responsible for converting thymine to uracil-5-carboxylic acid. The final enzyme is iso-orotate decarboxylase (IDCase), which forms uracil from iso-orotate through a decarboxylation reaction. In this

dissertation, I will present the results and evidence of the isolation and expression of these two fungal enzymes into bacterial systems.

Most of the previous research on thymine hydroxylase has been carried out using very small amounts of purified preparations from *Rhodotorula glutinis* since until now the gene had not been cloned and expressed in a bacterial system. Previous research has found that thymine hydroxylase is a non-heme Fe^{2+}, α-ketoglutarate-dependent dioxygenase, [6] and it is responsible for sequentially hydroxylating methyl, carbinol, and carbonyl carbon atoms in the conversion of thymine to iso-orotate by the following sequence (Figure 1-8):

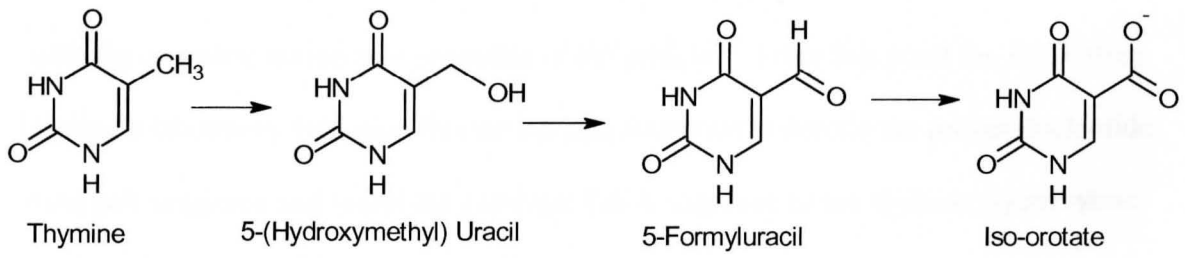


Figure 1-8. The reactions catalyzed by thymine hydroxylase

The overall reaction of thymine hydroxylase is:

$$\alpha\text{-Ketoglutarate} + O_2 + S \rightarrow \text{Succinate} + CO_2 + SO$$

(where S refers to the substrate)

In this case, one oxygen atom from O_2 is transferred to the pyrimidine's 5-substituent, with the remaining oxygen atom from O_2 transferred to α-ketoglutarate to yield succinate and CO_2. The above reaction occurs three times, as thymine is converted to 5-

(hydroxymethyl)uracil, 5-formyluracil and finally iso-orotate.[7, 8] The formation of uracil is then completed by the enzymatic decarboxylation of iso-orotate by IDCase.

Previous work has found that THase is not very substrate specific and that it can also catalyze epoxidations, oxidation of a thioether to a sulfoxide and a sulfone, and the oxidation of a methylamine to formaldehyde.[6] Furthermore, 5-ethynyluracil was found to be a mechanism-based inactivator of this enzyme. Mechanistically, thymine-7-hydroxylase is thought to work similarly as the heme Fe^{3+} dependent cytochrome P450s. This mechanism involves the formation of a high-valent iron-oxo species, which abstracts a hydrogen atom from the substrate.[6, 9]

However, of most importance to our goal is the research done by Thornburg, *et al.* where the author performed an N-terminal sequence analysis of the protein and came up with the complete amino acid sequence of the protein.[9] From this point on, Dr. Jeffrey Smiley's laboratory then used this amino acid sequence to decode the proper nucleotide base pair sequence and found the complete DNA sequence of the thymine hydroxylase gene by probing the cDNA library of *R. glutinis* with the constructed oligonucleotide. The isolated gene was sequenced and it gave a sequence similar to other Fe^{2+}, α-ketoglutarate-dependent dioxygenases. The isolated gene was then transferred into a pCal-n cloning vector and the pCal-THase plasmid was transferred into an *E. coli* BL21 expression system. The enzyme was over-expressed and it was used to run enzymatic assays.

Iso-orotate decarboxylase (IDCase), has been described only briefly in the scientific literature. Abbot *et al.* reported that the enzyme is not sensitive to EDTA, and that the addition of common cofactors (such as biotin) did not increase its measurable

activity.[4] A possible mechanism for the enzymatic decarboxylation of iso-orotate has been offered by Smiley *et al.*[5] The proposed mechanism involves an attack by a nucleophilic residue at the 6-position of the iso-orotate. This would form a β-keto acid intermediate which would make the carboxylate group a better leaving group, being released as CO_2, with subsequent (or concerted) reversal of the covalent bond-forming step. IDCase has also been observed to bind tightly to 5-nitrouracil and we postulate that it may be a mechanism-based inhibitor, based on the possible covalent attachment of the enzyme to its carbon-6. This mechanism is shown below (Figure 1-9):

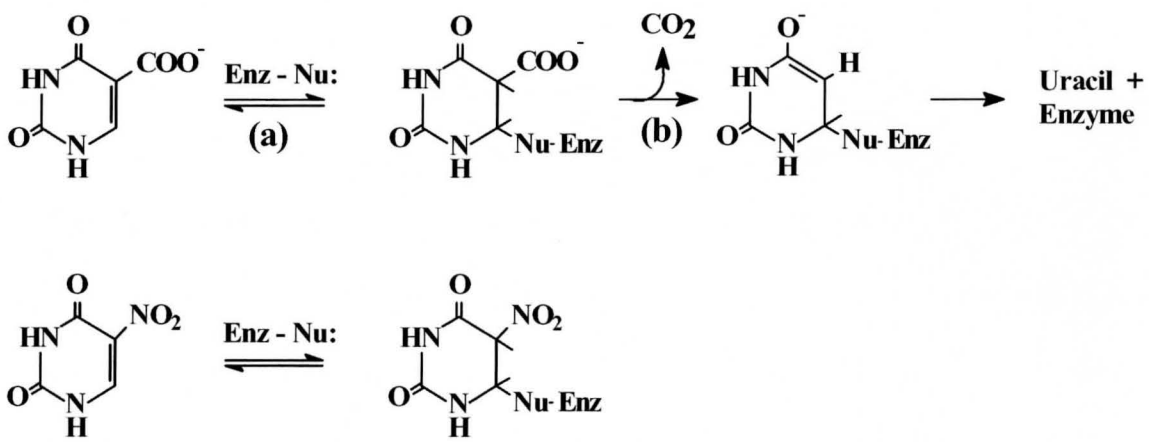


Figure 1-9. Proposed mechanism for covalent catalysis by IDCase and covalent attachment by 5-nitrouracil.

In addition to determining the function of the thymidine salvage pathway, our laboratory is attempting to gain insight into the catalytic mechanism of IDCase by protein purification, physical studies of the enzyme complexed with 5-nitrouracil, and isolation of the gene from one of a number of different organisms.

In this dissertation, the results of the isolation of the IDCase gene from *N. crassa* will be presented, along with the site-directed mutagenesis experiments performed in the

hopes of obtaining mechanistic information. Its expression into BL21 *E. coli* cells and the attempted purification by $AmSO_4$ fractionation and DEAE ion exchange chromatography will also be discussed.

Chapter II: THase Experiments

A. PCR: Transfer of THase gene from a *Rhodotorula glutinis* cDNA library to pCal-n cloning vector

Introduction

THase was cloned by first using the polymerase chain reaction (PCR) to amplify our target sequence. The PCR product (THase plasmid DNA) and the pCal-n expression vector were then cut by restriction endonucleases and ligated together by T4 DNA ligase to perform directional cloning. The THase-pCal plasmid was then used to transform BL21 *E. coli* cells. The transformed cells were then induced with IPTG to over-produce the enzyme, which was assayed by two different methods.

PCR is an *in vitro* technique used to dramatically amplify the amount of a specific DNA segment. In this method, a heat denatured (strand-separated) DNA sample is incubated with a heat-stable DNA Polymerase (from thermophilic bacteria), dNTPs (deoxynucleotide tri-phosphates), and two oligonucleotide primers designed to be complementary to the two 3'-ends of the specific DNA segment to be amplified.[3, 10] The primers are added in excess amounts of 1000 times or greater and they prime or "anneal" to the DNA segment at 55°C. This takes only about 20 seconds. The DNA polymerase then catalyzes the synthesis of the two complementary strands of the desired segment, doubling the concentration of DNA. This is done at 75°C, since it represents the temperature of the environment of thermophilic bacteria. The whole cycle requires only about two minutes to complete. After one cycle is complete, the DNA is then re-heated to 95°C to dissociate the DNA duplexes and then re-cooled so that the primers bind to both the newly synthesized and the old strands, and another round of DNA synthesis is completed. This technique dramatically amplifies the concentration of DNA, so that after

twenty cycles of PCR we increase the amount of DNA around one-million-fold, making it easier for us to work with the particular molecule and to move to the next step in our procedure.[2, 3, 10] Not only is the DNA amplified, but it also contains the ends that are synthesized into the primers. This becomes important in molecular cloning.

The purpose of molecular cloning is to insert our amplified DNA segment into an autonomously replicating DNA molecule, in our case the pCal-n plasmid, so that the DNA segment is replicated with the recombinant plasmid in a host organism, in our case *E. coli*, thus producing large quantities of thymine hydroxylase.

Plasmids are naturally occurring, circular, extra-chromosomal DNA molecules that contain all of the requirements for autonomous propagation in a host cell (a replicator, a selectable marker, and a cloning site). A replicator is an origin of replication, or *ori*, a site at which DNA replication is initiated. The selectable marker is typically a gene that confers antibiotic resistance, and is thus used to select for the cells that have the recombinant plasmid by growing them in certain antibiotics. Meanwhile, the cloning site is a sequence of nucleotides representing one or more restriction endonuclease cleavage sites. The pCal-n vector used in our experiments is shown below (Figure 2-1):

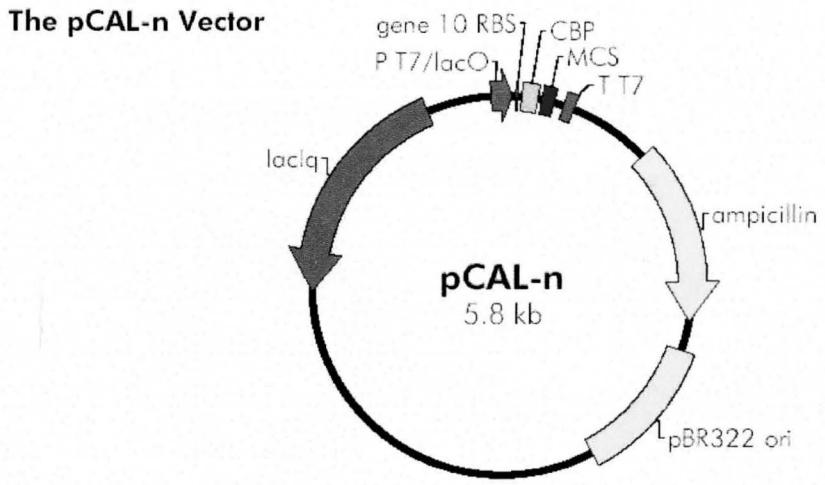


Figure 2-1. The pCal-n plasmid (5.8 kb) used in our cloning experiments.

Restriction endonucleases are enzymes that hydrolytically cleave polynucleotides at internal sites that contain a specific palindromic base sequence of four to eight base pairs. A palindromic sequence is one that contains exact twofold rotational symmetry (encodes the same forward and backward base sequence). In this project, restriction enzymes are used to cut the DNA and the plasmid at positions that are symmetrically staggered so as to create restriction fragments with complementary single stranded ends (*cohesive* or *sticky ends*). The complementary ends of the foreign DNA and the linearized plasmid will then cause ligation of the two, after which the base pairs are covalently joined by T4 DNA ligase. In our project, we will use two different restriction enzymes on the plasmid and our target DNA, so that the target ends have different overhangs ("sticky" ends) thus giving us unidirectional ligation with the help of DNA ligase. Unidirectional cloning is important because in order for replication to occur, in inserting our sequence into the plasmid the *promoter* must lie upstream of the gene that controls the expression of THase. Below is an illustration of the procedure used in gene cloning; note that in this case, only one restriction endonuclease was used (*Eco RI*):

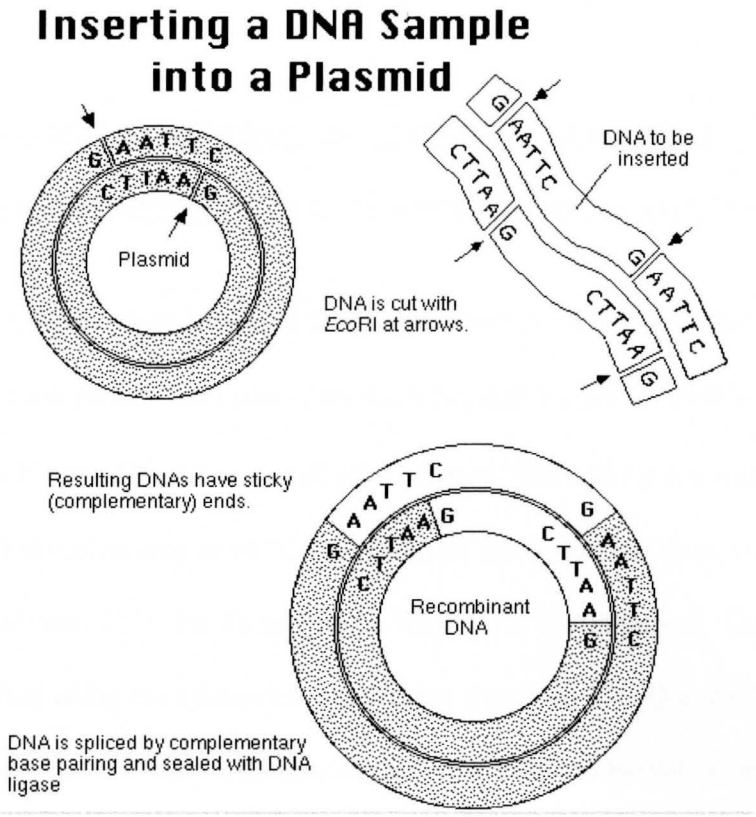


Figure 2-2 – Shows the preparation of a recombinant plasmid. (Source: http://web.mit.edu/esgbio/www/rdna/graphics/insert.gif)

Materials and Methods

Upon identification of a candidate THase gene, the coding sequence was amplified by PCR using primers designed for the ligation of the THase sequence into the NdeI-HindIII digested plasmid pCal-n (Stratagene). Primers used for this amplification were designed with 10 nucleotides of meaningless sequence at the 5' end (to facilitate restriction enzyme digestion of the amplified DNA), followed by six nucleotides of either the NdeI or HindIII restriction site, and sequence corresponding to the THase candidate

gene at the 3' end. These primers were designed with the following sequences (restriction sites underlined):

5' primer: 5'-TCCGTGAAGTCATATGGTCTCGTCTGGCATCGTC-3'

3' primer: 5'-TCCGTGAAATAAGCTTTCACTTCTTGGTGTAGGTCTC-3'

Amplification reactions included 5 ng of plasmid DNA carrying the candidate gene, 50 pmol of each primer, 200 μM of each dNTP, and 2 units of Vent® DNA polymerase (New England Biolabs) in 100 μL volumes. Thermal cycles were as follows: a single 5 min denaturation step at 94 °C; 35 cycles of the following three steps: 94 °C for 30 sec, 58 °C for 60 sec, 72 °C for 75 sec; and a final 72 °C 5 min period. The PCR product was purified using the QIAquick Nucleotide Removal Kit (Qiagen), digested with Nde I and HindIII restriction endonucleases (New England Biolabs), and ligated to Nde I- and HindIII-digested pCal-n with T4 DNA ligase (New England Biolabs). The ligation mix was used to transform XLI-Blue to ampicillin resistance; construction of the plasmid pCal- THase was confirmed by restriction digestion with NdeI and HindIII and agarose gel electrophoresis of the digested plasmid. The recombinant plasmid was then introduced into BL21 (DE3) Gold (Stratagene) cells (hereafter referred to as "BL21 cells").

E. coli Transformations:

pCal-THase was used to transform BL21 strains of *E. coli* by adding 5 μL of the prepared vector to 0.2 mL of $CaCl_2$ treated competent cells in chilled Falcon tubes. The tubes were allowed to sit in ice for 5 minutes and then they were heated at 42 °C for 2 minutes. After this short incubation, 1 mL of sterile LB media was added to the tubes and

the cells were allowed to grow for 30 min at 37°C. Following this incubation we then transferred 200 µL of these cells onto Ampicillin and Glucose containing LB agar plates, with overnight growth. Plasmid minipreps of the grown colonies showed positive transformations for a few selected colonies.

Results and Discussion

The XL1-Blue *E. coli* colonies that were grown in (+) ligation pCal-THase plates, were picked and grown overnight in 5 mL LB-medium starting cultures. These cultures were then used to make plasmid minipreps using the alkaline lysis method. The constructs were dissolved in TAE buffer and run in 0.8 % agarose gels. Figure 2-3 below shows one such gel, where one can observe a positive transformation in lane # 4, representing a band of DNA the same size as a pCal-THase plasmid.

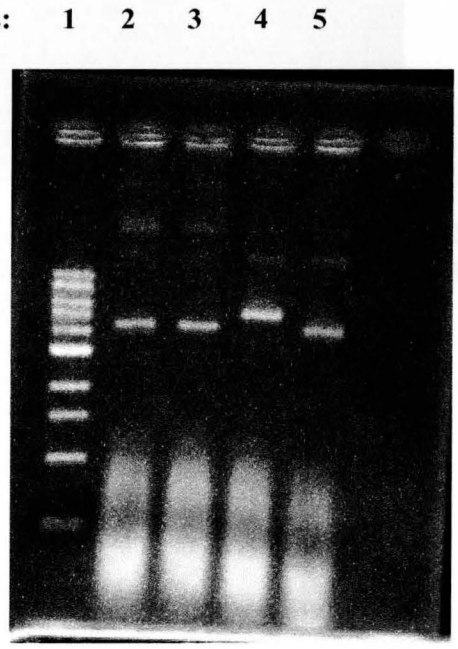


Figure 2-3. Agarose gel of plasmid minipreps of colonies that showed on (+) ligation plates. Lane 1 shows the 1 kbp DNA marker, lanes 2, 3, 5 show negative vector/insert ligations; meanwhile lane 4 shows the positive test.

The cultures that showed a positive ligation were then grown in 50 mL LB media and a large scale plasmid preparation was made using the Qiagen HiSpeed MidiKits. The construction of the pCal- THase plasmid was then confirmed by restriction digestion with NdeI and HindIII and agarose gel electrophoresis of the digested plasmid. Figure 2-4 shows these results.

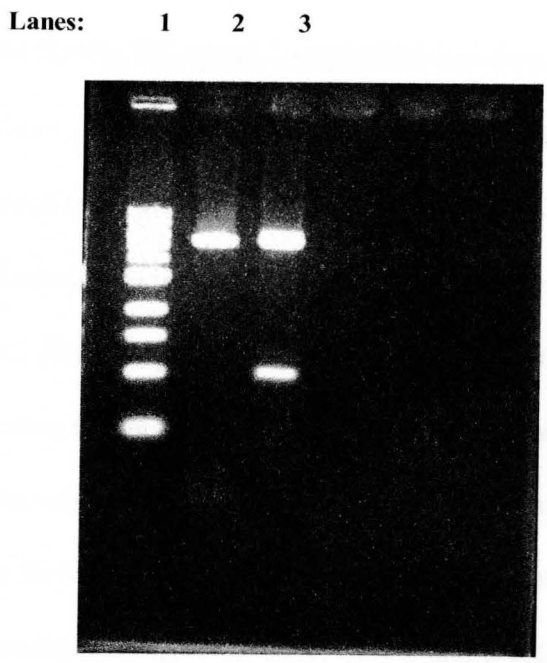


Figure 2-4. Agarose gel of NdeI-HindIII digested pCal-THase. Lane 1 shows the 1kbp DNA marker, lane 2 shows a NdeI-HindIII digested pCal-n vector, lane 3 is the digested plasmid (pCal and THase).

B. Production of protein in BL21 *E. coli* cells

Introduction

Materials and Methods

BL21 cells carrying the candidate plasmid pCal-THase were grown overnight at 37°C with shaking in LB media containing 50 μg mL^{-1} ampicillin. On the following day, a secondary culture of enriched media (35 g L^{-1} tryptone, 20 g L^{-1} yeast extract, 5 g L^{-1} NaCl, 10% glycerol, 100 μM $FeSO_4$, 50 μg mL^{-1} ampicillin) was inoculated with a 1:100 dilution of the overnight culture, and incubated at 37°C with shaking for 3 h. The secondary culture was removed from the 37°C incubator and cooled to room temperature (22-24°C), and protein production was induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG, Acros Organics) to a concentration of 0.2 mM. The culture was shaken at room temperature for 24 hours, at which time cells were harvested by centrifugation. The cell pellet was resuspended in 1/50 of the culture volume using a buffered solution containing 50 mM HEPES, pH 7.5, 1.0 mM dithiothreitol (DTT), 100 μM $FeSO_4$, 1.0 mM phenylmethylsulfonyl fluoride (PMSF), and 0.1% Triton X-100. The cell suspension was transferred into 2-mL screw cap tubes, to which 0.1 mm glass beads were also added. The cells were lysed using the Mini Bead Beater (BioSpec Products, Bartlesville, OK) using three 1- minute pulses of the instrument at high speed, interspersed with periods of cooling the lysate on ice. The cell lysate was transferred to a microcentrifuge tube and centrifuged for 5 min at 13,000 rpm in an Eppendorf microcentrifuge in 4°C refrigeration. The clarified lysate was removed from the pelleted cell debris and assayed for THase activity.

C. Assays for enzymatic activity

THase activity was measured using [2-^{14}C] thymine (Moravek Biochemicals) and determining the conversion of radiolabelled material into oxidized products. Labeled thymine was isotopically diluted to 240 cpm nmol^{-1}, in a final thymine concentration of 9 mM. Reaction volumes contained 100 µL of labeled thymine (220,000 cpm), 100 µL 4.5 mM α-ketoglutarate, 100 µL 23 mM sodium ascorbate, 100 µL 110 µM $FeSO_4$, 100 µL 500 mM HEPES pH 7.5, with protein solution and water added in varying amounts to total 1 mL. The reaction was run in open containers at 30°C; portions of the reaction were terminated at various time points by spotting 40 µL of the reaction mixture on the baseline of a hot silica TLC plate, keeping each sample spotted within a small area. Samples of unlabeled thymine and 5-hydroxymethyluracil (5-HMU, Aldrich) were co-spotted with each sample, to allow visualization of the substrate and product following separation. Reaction samples from various time points were separated using a solvent of 10% methanol/90% ethyl acetate, which gave R_f values of 0.6 and 0.4 for thymine and 5-HMU, respectively. Portions of the TLC sheet containing thymine or 5-HMU were cut out of the sheet and transferred to vials containing Scinti-Safe scintillation fluid (Fisher Scientific). THase activity was calculated from the decrease in cpm for thymine, or the increase in cpm for 5-HMU, over time. Total protein was determined using protein assay solution from Bio-Rad and bovine serum albumin (Pierce) as a standard.

Another THase assay used in our research was one in which we utilized HPLC in order to separate thymine from 5-Hydroxymethyluracil (HMU). Thymine was added to the growth media and the oxidized products were detectable in the culture supernatant over a period of IPTG induction of BL21 cells containing pCal-THase.

Results

E. coli BL21 cells carrying plasmid pCal-THase produced an abundance of protein with molecular weight approximately 37 kDa upon induction with IPTG. Shown below is an SDS-PAGE gel whole cells and cell lysates of over-expressed THase protein (Figure 2-5):

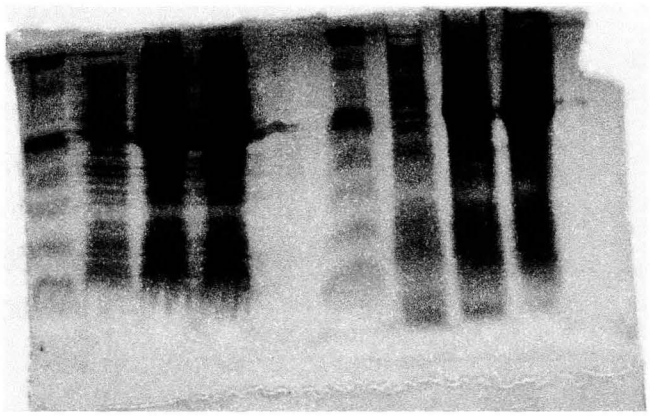


Figure 2-5. SDS-PAGE showing over-expressed THase bands at 37 kDa. Lanes # 1-4 show the MW marker, cell lysate at 6 hrs, 20 hrs and 24 hrs of IPTG induction; meanwhile lanes # 5-8 show the PRO$^-$ MW marker, whole cells at 0 hrs, 9 hrs and 24 hrs of induction, respectively.

Protein lysates from these cells converted [2-^{14}C] thymine to a new radioactive product, coincident on TLC with 5-HMU; additional cpms were present on the baseline of the TLC assay. From the assay data, THase activity of 1.3 nmol min^{-1} mg^{-1} was determined from the disappearance of thymine, or 0.64 nmol min^{-1} mg^{-1} was determined from the appearance of 5-HMU. The latter is likely an underestimation, since 5-HMU can be further oxidized to 5-formyluracil and again to iso-orotate. Regardless, the conversion of thymine to a product coincident with 5-HMU is not observable in *E. coli* cells without plasmid pCal-THase, and the identification of the THase gene seems to be established.

The HPLC assay used also showed THase activity present by way of measuring the increase in the concentration of HMU and the decrease in Thymine in the growth media of BL21 cells containing pCal-THase (Figure 2-6).

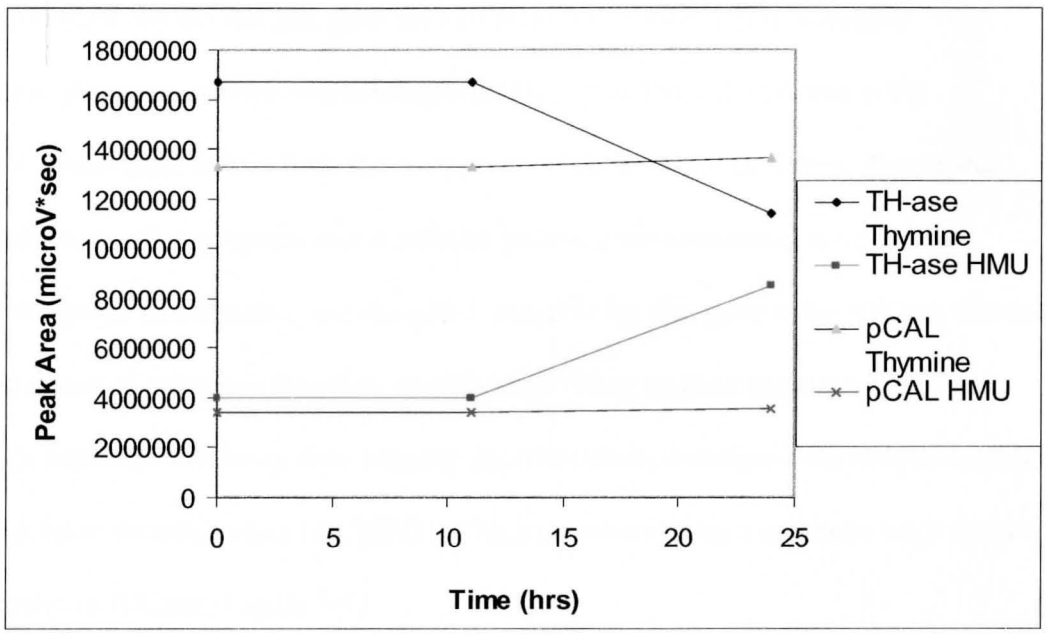


Figure 2-6. HPLC assay results showing a decrease in [thymine] and an increase in [HMU] in the growth medium of pCal-THase cells induced with IPTG.

From the figure above, one can see that we were able to measure THase activity in cells containing pCal-THase, however this activity was not present in cells containing only pCal. Unfortunately, the recombinant protein does not have especially high specific activity and optimization of activity awaits future efforts.

Chapter III: IDCase Experiments

A. Cloning of IDCase gene into pCal-n cloning vector for production of non-fusion-tagged protein

Introduction

Based on research done in Dr. Smiley's lab, the *N. crassa* gene NCU06417.1 appeared likely to be the IDCase gene for two reasons.[11] First, using sequence comparison programs at http://www.broad.mit.edu it was located adjacent to the presumed THase gene within both the *N. crassa* and *A. nidulans* genomes. Therefore, given that consecutive enzymes in a metabolic pathway are sometimes encoded by consecutive genes in a genome, we thought it possible for this gene to be IDCase. Second, these two genes of unknown function, neighboring THase in their respective genomes, had a high sequence similarity with another decarboxylase, 2-amino-3-carboxymuconate-6-semialdehyde decarboxylase (ACMSD), which decarboxylates a substrate very similar to iso-orotate in IDCase (Figure 3-1).

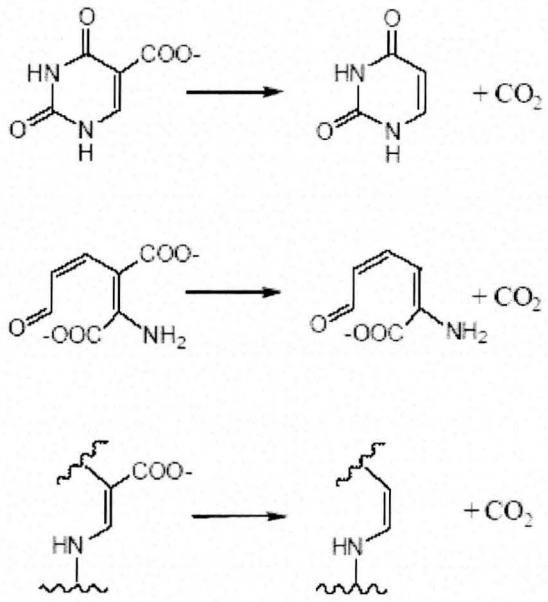


Figure 3-1. Comparison of the reactions catalyzed by IDCase (top) and ACMSD (middle), with structural similarity (bottom).

Given these reasons, the IDCase gene was initially cloned from *Neurospora crassa* by Dr. Smiley. IDCase assays (such as those described below) confirmed that the identified gene produced the desired enzyme activity. However, the original cloning produced an IDCase sequence fused to the Calmodulin Binding Polypeptide (CBP) or fusion protein. So, my task was to create a non-fusion pCal-IDCase plasmid. This was done by running a Polymerase Chain Reaction (PCR) using the NdeI and HindIII primers, then subsequently remove the sequence encoding for CBP. The sequences of the primers used are shown below:

5' primer (NdeI start):

5'-GGGAAAGGCATATGGGTTTCGGCGATGGCAAAG-3'

3' primer (HindIII stop):

5'-GGAAAGGAAAGCTTCTATCGGCACCCATCCCTCCTC-3'

The PCR reaction mix contained: 10 µL of HindIII primer (5 µM), 10 µL of NdeI primer (5 µM), 10 µL of pCal-fusion-IDCase (1:100 dilution of a previous Qiagen preparation), 10 µL of ThermoPol buffer (10 x concentration), 10 µL dNTPs (2 µM), 2 µL of VENT (*Taq* polymerase) and 48 µL sterile dH_2O to give a total volume of 100 µL. The thermal cycling program was set to run at: 94 °C for 30 sec, 58 °C for 60 sec, 72 °C for 90 sec; for a total of 35 cycles. The product of the PCR was purified using the QIAquick Nucleotide Removal Kit (Qiagen) and run on a 0.8 % agarose gel to verify a positive reaction. After verification, restriction digests were run of both pCal and amplified IDCase insert, with verification on a gel. The pCal-n digest contained: 9 µL of pCal-THase (1 µg), 3 µL NE Buffer 2, 1 µL NdeI, 1 µL HindIII, 16 µL dH_2O, meanwhile the PCR product digest contained: 25 µL PCR product, 3 µL NE buffer 2, 1 µL NdeI, 1

μL HindIII. The digestion reactions were run for 3-4 hours at 37°C and the products were spin-column purified using the QIagen method and run on a gel for verification (Figure 3-2).

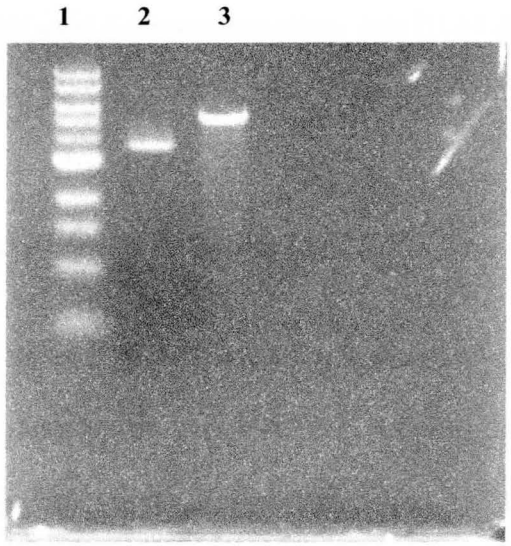


Figure 3-2. 0.8 % agarose gel of control digest of pCal-n vector. Lane 1 is DNA ladder, lane 2 uncut pCal and lane 3 is pCal cut with NdeI and HindIII.

The control ligation of twice-cut pCal contained: 6 μL of cut pCal (~300 ng), 1 μL buffer, 1 μL T4 DNA ligase, 2 μL sterile dH_2O; meanwhile that of twice cut PCR product contained: 3 μL cut PCR, 1 μL buffer, 1 μL T4 DNA ligase, 5 μL sterile dH_2O. The reactions were run for 20 min at room temperature. Once we obtained proof of positive control ligations we then ran the pCal (vector) and IDCase (insert) ligations. They contained: 6.5 μL of cut pCal (300 ng), 1.5 μL of cut PCR (100 ng), 1 μL ligase buffer, 1 μL T4 DNA ligase.

The ligation contents were used to transform XL1-Blue *E. coli* cells with our suspected pCal-NF IDCase plasmid using the $CaCl_2$ heat-shock method. The colonies that grew on the LB-Amp-Glucose plates were then used to inoculate starting cultures which were used to prepare plasmid minipreps using the alkaline lysis method. The

minipreps were run on 0.8% agarose gels and the ligations were verified. Then large-scale plasmid preparations were made using the QiagenHiSpeed MidiKit where we obtained a pCal-NF IDCase concentration of 45.5 µg / mL through its absorbance at 260 nm. This plasmid was then used for transformations into BL21 *E. coli* cells, which are better suited for protein expression.

B. Site-directed mutagenesis of IDCase Cys272Ala by overlap extension method

Introduction

The proposed mechanism of IDCase action involves an attack by a nucleophilic residue at the 6-position of the iso-orotate. It is likely that this residue can be one of the invariant cysteine residues found in the amino acid sequence of IDCase. Upon comparison with the suspected IDCase from *A. nidulans*, one cysteine residue, number 272 in the *N. crassa* sequence, was found within a highly conserved region in both sequences.

Translated *N. crassa* IDCase amino acid sequence (Cys272 underlined):

MGFGDGKVVDIHTHMYPPEYVKILESRTSIPLVRSFPGSPDPRLILLDTEVPILEEAEAAKA
RGETPANIPGRPLTKHYSSLNQKMHFMETHEIDISVVSLANPWLDFLSASEAGPIAESINA
DFSRM**C**EEK**C**GRLFFFAALPLTAPRDVILASIAHVSNLPY**C**RGVILGTSGLGKGLDDPDL
LPVFHALADAKLMIFLHPHYGLPNEVWGPRAKENYGHVLPLALGFPMETTIAVTRMYL
AGVFDQVPKLNMLLAHSGGTLPFLAGRIES**C**ILHDGHLHSAAGTKPKKTIWEVLSSQIYL
DAVVYSDVGLKAAVQASGPEGHERLMFGTDHPFFPPLGSDEEGEWESVTWNGAAVRK
AFGAEDGEDSEEGKKVRGVMGANAVRVLNLRRDGCR

Therefore, our objective was to conduct site-directed mutagenesis of these cysteine codons (TGT), individually, by converting them to chemically inactive alanine codons (GCC). The lack of enzymatic activity in one of the mutated proteins would then

contribute to the identification of the critical residue. For my specific project, we attempted to mutate the Cys residue at position 272 (underlined residue) of the amino acid sequence of IDCase. This was done by a method termed base substitution mutagenesis by PCR overlap extension (Higuchi et al., 1988).

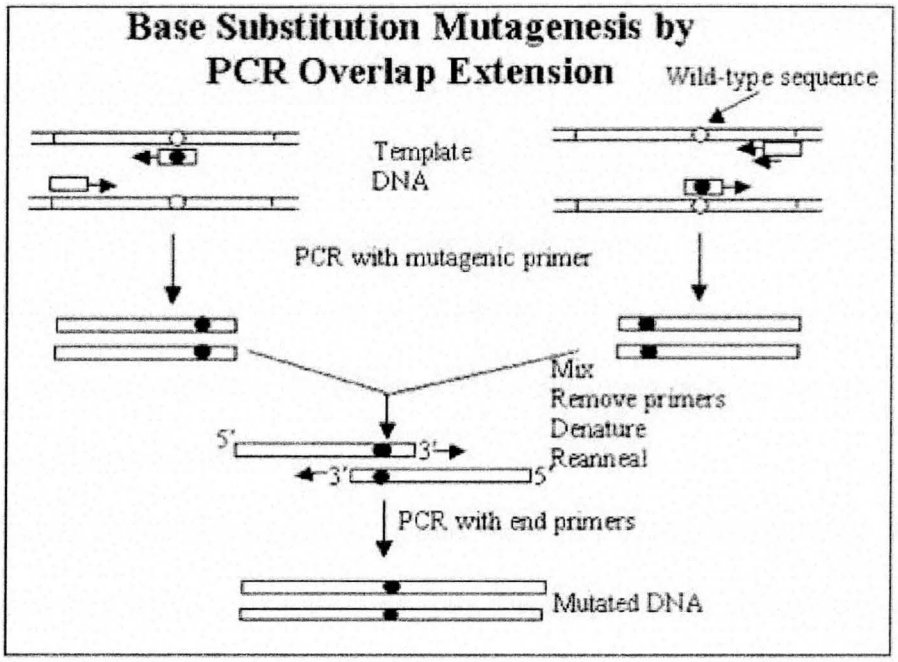


Figure 3-3. Depiction of base substitution mutagenesis by PCR overlap extension.

Materials and Methods

The overlap extension technique involved two separate PCRs, each using a non-mutagenic primer at the end of the sequence and a mismatched primer designed to introduce a point mutation at a specific point. In our case, PCR A contained the IDCase C272A forward primer and HindIII stop primer. Meanwhile, PCR B contained the IDCase C272A reverse primer and NdeI start primer:

IDCase C272A Forward: 5'-GGGAGGATAGAAAGCGCCATCTTGCATGATGGA-3'

IDCase C272A Reverse: 5'-TCCATCATGCAAGATGGCGCTTTCTATCCTCCC-3'

The mutagenic PCR reaction A contained: 10 µL IDCase C272A forward primer (5 µM), 10 µL IDCase HindIII stop primer (5 µM), 10 µL of pCal-IDCase-fusion (1:100 of qiagen), 10 µL of dNTPs (at 10 x concentration, 2 µM), 10 µL thermo polymerase buffer (10 x), 1 µL Vent polymerase and 49 µL sterile dH_2O to give a total volume of 100 µL. The mutagenic PCR reaction B contained: 10 µL IDCase C272A reverse primer (5 µM), 10 µL IDCase NdeI start primer (5 µM), 10 µL pCal-IDCase-fusion (1:100 of qiagen), 10 µL of dNTPs (2 µM), 10 µL ThermoPol buffer (10 x), 1 µL Vent polymerase and 49 µL sterile dH_2O to give a total volume of 100 µL. All these reactions were run at 3 x volumes. The contents were then loaded into 0.8% agarose gels and they are gel extract purified using the Qiagen method.

This resulted in two overlapping fragments (PCR A ~ 400 bp and PCR B ~ 800 bp) that contained the Cys272Ala base substitution. The fragments were then annealed together in a secondary PCR, the overlap extension, used to re-construct the complete mutated gene. PCR mix for overlap extension reaction contained: 15 µL of fragment A (smaller, 450 ng), 25 µL of fragment B (larger, 750 ng), 10 µL of 10x ThermoPol buffer, 10 µL of 10x dNTPs (2 µM), 1 µL Vent polymerase and 39 µL sterile dH_2O to give a total volume of 100 µL. This reaction was also run at 3 x volumes. The OE PCR thermal cycling program was set at: 7 min at 94°C, 1 min at 58°C, 75 seconds at 72°C; all for a total of 35 cycles. The OE PCR product was then Qiagen purified and run on a 0.8% agarose gel along with PCR fragments A and B (Figure 3-4).

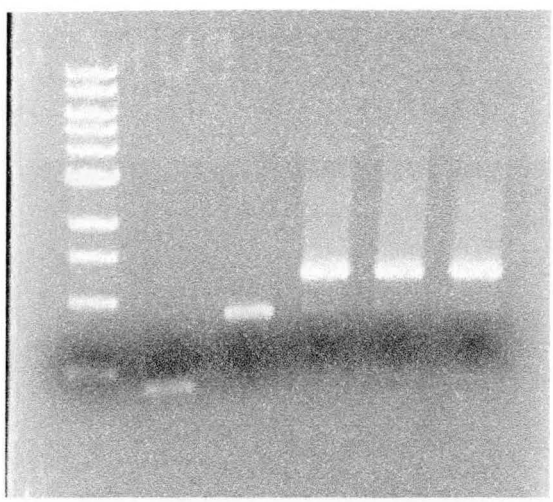


Figure 3-4. 0.8 % agarose gel of PCR A, PCR B and overlap extension product. Lane 1 -- DNA ladder; lane 2 -- PCR A (~400 bp); lane 3 -- PCR B (~ 800 bp); lanes 4-6 are the overlap extension product.

The OE PCR product represented a very faint band, so in order to increase its concentration another PCR of the suspected mutant IDCase was used. This PCR reaction contained: 10 µL NdeI start, 10 µL HindIII stop, 10 µL template (OE mutant), 10 µL ThermoPol buffer, 10 µL dNTP, 1 µL Vent polymerase and 49 µL sterile dH_2O to give a total volume of 100 µL. The reaction was run at 4x volumes and the PCR product was Qiagen purified.

Once we obtained a high enough concentration of our suspected mutant IDCase, we then performed restriction digests of both the mutant and a pCal-n vector and then ligated them to construct pCal-IDCase C272A. Prior to running the vector + insert ligation reactions, we ran test ligations using only insert and only vector. The only insert ligation contained: 6 µL mutant PCR (300 ng), 1 µL ligase buffer, 1 µL T4 DNA ligase, 2 µL dH_2O, meanwhile the only vector ligation contained: 8 µL vector (150 ng), 1 µL

buffer, 1 μL T4 DNA ligase. The reactions were run for 30 minutes at room temperature and then loaded on 0.8% agarose gels for verification (Figure 3-5).

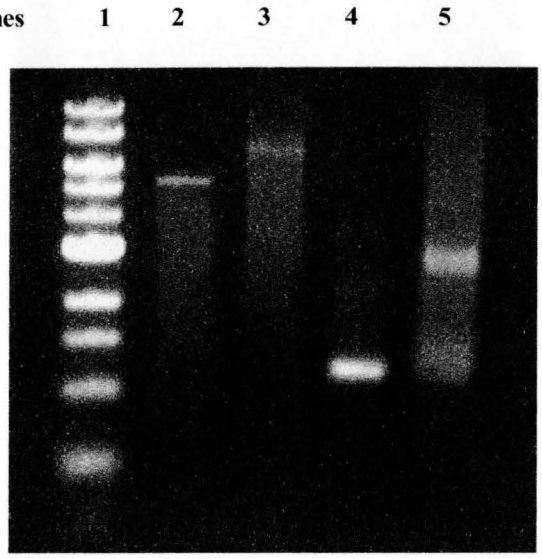


Figure 3-5. Control ligations of vector and insert in mutagenesis: lane 1 -- DNA ladder; lane 2 -- cut pCal (vector); lane 3 – ligated pCal; lane 4 – cut OE product (insert); lane 5 – ligated OE product.

Once we obtained proof of positive control ligations we then ran the pCal (vector) and mutant IDCase (insert) ligations. They contained: 20 μL of pCal (300 ng), 2 μL of insert (100 ng), 3 μL ligase buffer, 3 μL ligase, 2 μL sterile dH_2O. The ligation contents were used to transform XL1-Blue *E.coli* cells with our suspected mutant pCal-IDCase plasmid using the $CaCl_2$ heat-shock method. The colonies that grew on the LB-Amp-Glucose plates were then used to inoculate starting cultures which were used to prepare plasmid minipreps using the alkaline lysis method. The minipreps were run in 0.8% agarose gels and the ligation was verified (Figure 3-6).

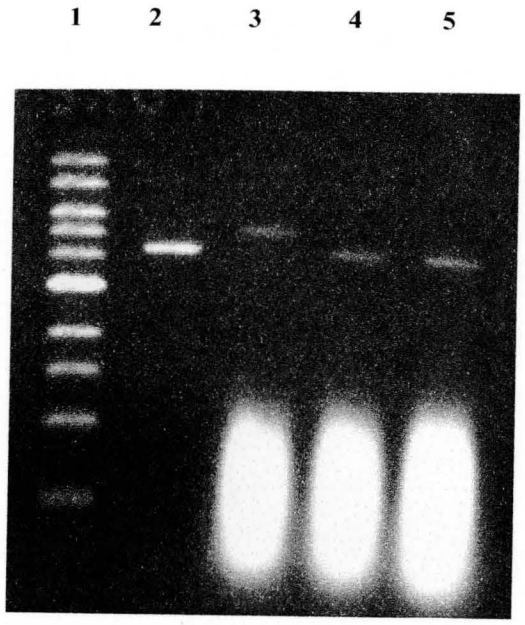


Figure 3-6. Successful positive ligation of pCal-mutant IDCase (lane 3).

Subsequently, large-scale plasmid preparations of the construct were made using the Qiagen method. This prep was used for transformations of the plasmid into BL21 *E. coli* cells, which are better suited for protein expression. Following these procedures, we then verified the overall construction of pCal-IDCase C272A by performing a restriction digest of both pCal-NF IDCase and pCal-mutantIDCase with NdeI/HindIII and running them side-by-side in a 0.8 % agarose gel (Figure 3-7).

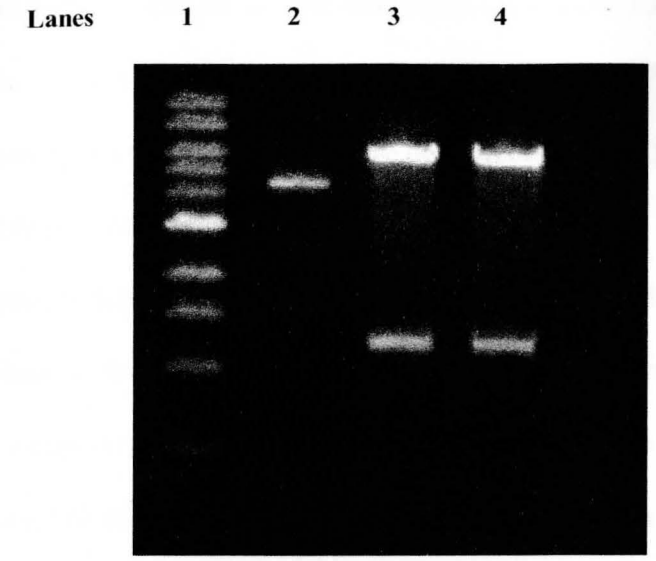


Figure 3-7. Restriction digest of mutant and wildtype pCal-IDCase with NdeI and HindIII. Lane 1 – DNA ladder; lane 2 – pCal, lane 3 – mutant pCal-IDCase; lane 4 – wildtype pCal-IDCase.

C. Verification of wildtype and mutant IDCase clones by DNA Sequencing

Introduction

DNA sequencing was invented in the late 1970's by Sanger and coworkers. It involves a modification of a normal cellular DNA replication process. As such, sequencing reactions require an oligonucleotide to use as a primer, a template to use for synthesizing a complementary strand, and dNTPs to use as substrates for synthesis. In our case, the templates were the mutant and wildtype pCal-IDCase plasmids. For each template, four different primers were used: pCal-n 5'Seq 11/34 (beginning of template), pCal-n 3'Seq 257/280 (end of template), IDCase forward AA218-226 primer and IDCase reverse AA 318-32 primer.

These reactions also required a way of identifying each base added to the growing DNA strand. This involved the use of "chain terminators", which are modified

nucleotides that can be incorporated into the new DNA strand, but do not permit continued synthesis of DNA. The chain-terminators are dideoxy-NTPs and they lack a 3′-hydroxyl group, thereby successfully terminating the polymerization process. With this method, it is possible to identify the nucleotide at each position since each of the four ddNTP's is fluorescently labeled with a specific dye. The bands are then detected (based on the fluorescent tags on the ddNTPs) by a laser as they are separated based on size through the use of a capillary electrophoresis as each one runs off the end of the gel. The result is a sequencing "electropherogram" (Figure 3-8), which is a tracing that shows the elution of the different fragments corresponding to the sequence.

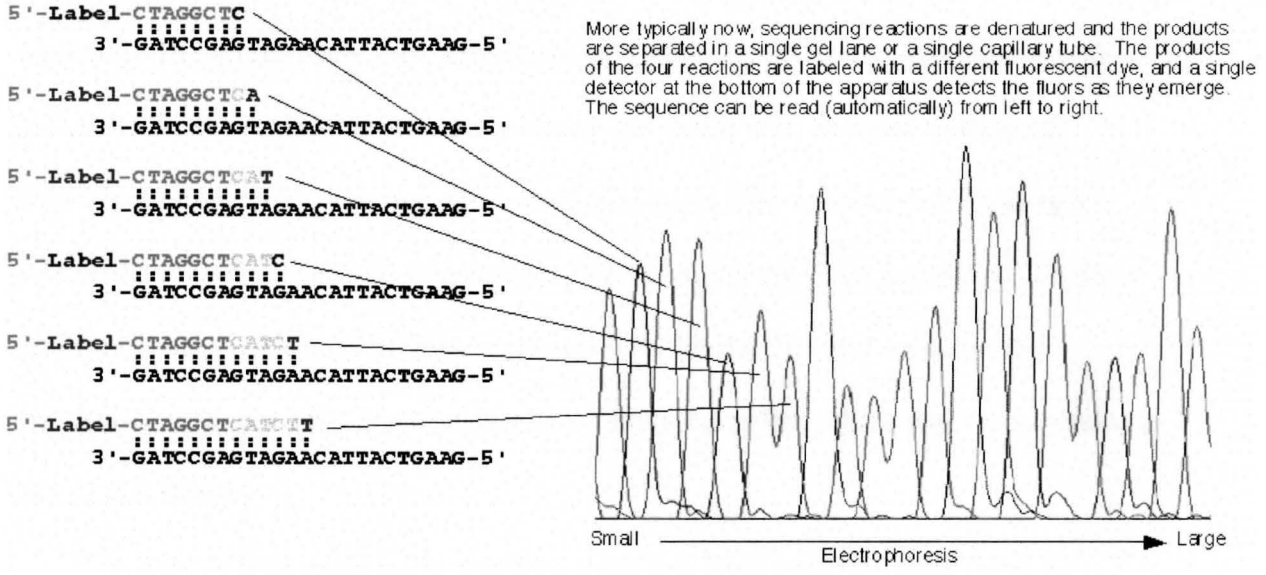


Figure 3-8. Illustration of an electropherogram (DNA sequencing result).

Materials and Methods

The four separate primers used in our sequencing reactions are:

pCal-n 5'Seq 11/34:
5'-CAC TAT AGG GGA ATT GTG AGC GGA-3'

pCal-n 3'Seq 257/280:

5'-GTG GCA GCA GCC AAC TCA GCT TCC 3'

IDCase Fwd AA218-226 primer, for sequencing 2nd half:
5'-TAC GGC CAC GTC TTG CCG CTT GCT CTC-3'

IDCase Rev AA 318-32 Primer, Rev. complimentary., for sequencing middle third:
5'-CAA TCT CTC GTG TCC CTC CGG TCC GCT-3'

Prior to sequencing our DNA, we first had to concentrate and purify our samples. This was done by performing a phenol-chloroform extraction of the plasmids. We took 500 µL of each plasmid and mixed with 500 µL of phenol (saturated with $CHCl_3$) by vortexing. The mixture was then centrifuged at 13,000 rpm. The aqueous supernatant contained the DNA. This layer was separated and mixed with an equal volume of chloroform, mixed by vortexing, and centrifuged again. The aqueous layer was again collected and to it we added 1/9 volume of 3M NaOAc at pH = 5.0 and mixed. To the resulting volume we added 2x volume of cold 95% EtOH. The mix was stored in the -20°C freezer for 2 hrs and centrifuged to obtain the pellet, which was re-suspended in 100 µL TE Buffer.

A total of nine sequencing reactions were run. Four reactions each for pCal IDCase WT and pCal IDCase C272A representing each of the four primers used. The last reaction was the pUC18 control template. Each reaction contained 12 µL of DTCS premix (10 x sequencing reaction buffer, dNTP mix, ddNTP dye terminator mix, polymerase enzyme), 2.0 µL of the appropriate primer (1.6 µM) and enough template and

sterile dH_2O to give a total volume of 20 µL. For pCal-Wildtype IDCase we used 2.6 µL (400ng) of template and 3.4 µL of sterile water, meanwhile for pCal-Mutant IDCase 3.2 µL (400 ng) of template and 2.8 µL of sterile dH_2O.

The thermal cycler was set to run at: 96 °C for 20 sec, 50 °C for 20 sec, 60 °C for 4 min, for 30 cycles followed by holding at 4 °C. After the thermal cycling, the sample products were then precipitated with ethanol and prepared for loading into the sequencer following the Beckman Coulter CEQ 2000 Dye Terminator Cycle Sequencing protocol.

Results

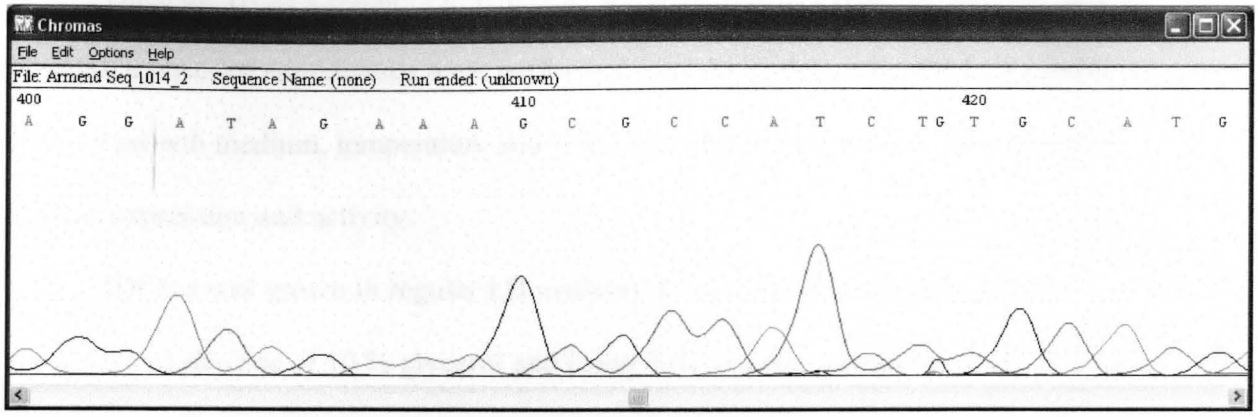
The gene sequence NCU06417.1 from *Neurospora crassa* was determined to encode for IDCase. The sequence contains 1185 nucleotides (395 codons) and is shown below. The bold, underlined codon represents the wildtype codon encoding for Cys272. In the mutagenesis project, this sequence has been changed to GCC, the codon encoding for alanine.

IDCase Wildtype sequence:

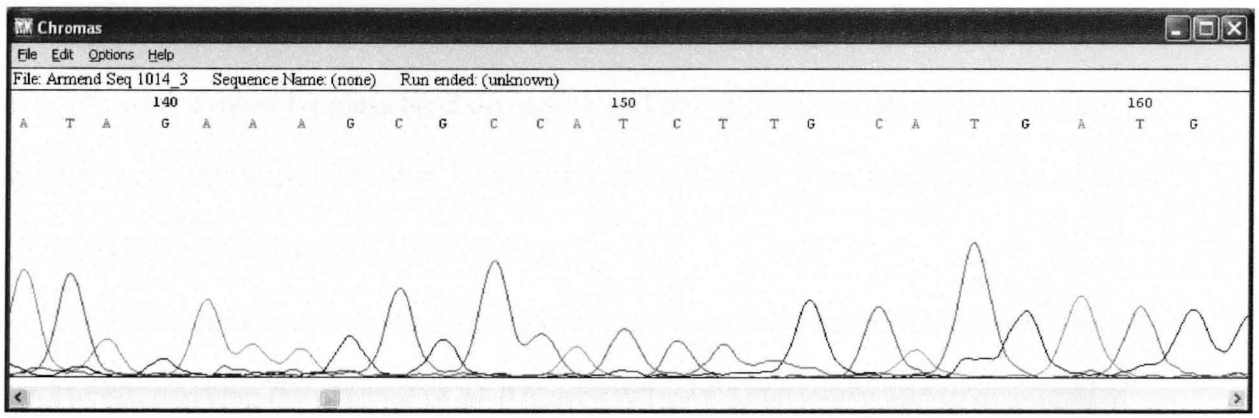
ATGGGTTTCGGCGATGGCAAAGTAGTGGACATCCACACCCACATGTACCCGCCCGAA
TACGTCAAGATCCTCGAATCCCGCACCTCCATCCCCCTCGTGCGCTCCTTCCCCGGCT
CCCCCGACCCGCGCCTCATCCTGCTCGACACCGAGGTTCCCATCCTCGAAGAAGCCG
AAGCAGCCAAAGCACGGGGCGAGACGCCCGCCAACATCCCCGGCCGGCCCTTGACC
AAGCACTACTCCTCCCTCAACCAAAAGATGCACTTTATGGAAACCCACGAAATCGAC
ATCTCCGTCGTCTCCCTCGCCAACCCCTGGCTTGACTTCTTGTCCGCCTCCGAAGCGG
GACCCATTGCCGAATCCATCAACGCCGACTTCAGCCGCATGTGTGAAGAGAAATGCG
GCCGGCTTTTCTTCTTTGCCGCCTTACCTCTGACCGCTCCCCGCGACGTCATCCTTGCC
TCCATCGCCCACGTTTCCAACCTGCCTTATTGCCGGGGCGTCATCCTCGGCACCTCCG
GCTTGGGAAAGGGGCTAGACGACCCCGATCTCTTGCCGGTTTTTCACGCCCTCGCCG
ACGCGAAGCTCATGATCTTCTTGCATCCGCACTATGGTCTGCCCAACGAAGTCTGGG
GGCCCCGCGCCAAGGAGAACTACGGCCACGTCTTGCCGCTTGCTCTCGGTTTTCCCAT
GGAAACCACCATTGCTGTCACGAGGATGTACCTGGCGGGGGTGTTTGATCAGGTGCC
GAAGCTAAACATGTTGCTGGCCCATTCGGGCGGGACGTTGCCGTTCCTGGCGGGGAG
GATAGAAAGCTGTATCTTGCATGATGGACACTTGCACTCTGCTGCGGGGACGAAGCC
GAAGAAGACGATTTGGGAGGTGTTGAGTAGTCAGATTTATCTTGACGCGGTGGTGTA
TAGCGATGTGGGGTTGAAGGCTGCCGTGCAAGCGAGCGGACCGGAGGGACACGAGA
GATTGATGTTTGGTACTGATCATCCGTTTTTCCCGCCTCTGGGGAGCGATGAGGAGG
GCGAGTGGGAGAGTGTCACGTGGAATGGGGCGGCGGTGAGGAAGGCGTTTGGAGCG
GAGGATGGGGAAGATAGTGAGGAGGGGAAGAAGGTAAGGGGAGTGATGGGCGCG
AATGCGGTTAGGGTGTTGAACTTGAGGAGGGATGGGTGCCGATAG

In the DNA sequencing reactions, we confirmed the identity of both the wildtype and mutant IDCase sequences, thereby proving that our cloning and mutagenesis experiments were successful. This is best shown in the two electropherogram results below:

Sequence data from Mutant IDCase using 3' primer (reverse complimentary), where codon GCC (412-414) now encodes for Alanine, from the previous Cysteine (TGT):



Sequence data from Mutant IDCase using new forward IDCase primer, where codon GCC (146-148) now encodes for Alanine, from the previous Cysteine (TGT):



Chapter IV: IDCase Experiments

A. Production of IDCase proteins in BL21 CodonPlus *E. coli* cells

Once the BL21 Codon(+) *E. coli* cells were transformed with the pCal-NF IDCase plasmid, a series of experiments were performed in order to determine the best conditions (type of growth medium, temperature and IPTG concentrations) suitable for maximum IDCase expression and activity.

IDCase was grown in regular LB medium, in enriched LB medium, in the presence and absence of 10 % glycerol and using different concentrations of IPTG over different periods of induction. BL21 cells containing the pCal-IDCase plasmid were also grown under different temperature conditions: at 37°C, room temperature, and 18°C. Different cell lysis buffer combinations were also used. After growth and IPTG induction, the cells were broken by glass bead disruption and the specific activity of IDCase in the lysates was determined using the $^{14}CO_2$ radioactivity assays. Also, the protein was run on SDS-PAGE gels in search of IDCase bands.

The conditions that produced the best specific activity of IDCase are as follows. BL21 cells carrying the plasmid pCal-IDCase (WT or C272A) were grown overnight at 37°C with shaking in LB media containing 50 μg mL^{-1} ampicillin. On the following day, a secondary culture of enriched media (35 g L^{-1} tryptone, 20 g L^{-1} yeast extract, 5 g L^{-1} NaCl, 50 μg mL^{-1} ampicillin) was inoculated with a 1:100 dilution of the overnight culture, and incubated at 37°C with shaking for 3 h. The secondary culture was removed from the 37°C incubator and cooled in ice buckets for 15 min, and protein production was induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG, Acros Organics) to a concentration of 0.8 mM. The culture was shaken at room temperature for ~18 hours,

at which time cells were harvested by centrifugation. The cell pellet was re-suspended in 1/50 of the culture volume using a buffered solution containing 50 mM HEPES, pH 7.5, 1.0 mM dithiothreitol (DTT), 2.0 mM phenylmethylsulfonyl fluoride (PMSF), 2.3 μM leupeptin, 3.0 μM pepstatin A and 0.1% Triton X-100.

Unlike with THase, the IDCase protein was not very soluble in solution and running it in an SDS-PAGE gel gave bands that corresponded to 42.9 kDa. However these bands did not increase in size as we increased the time of induction by IPTG, or as we induced with different IPTG concentrations (Figures 4-1 and 4-2).

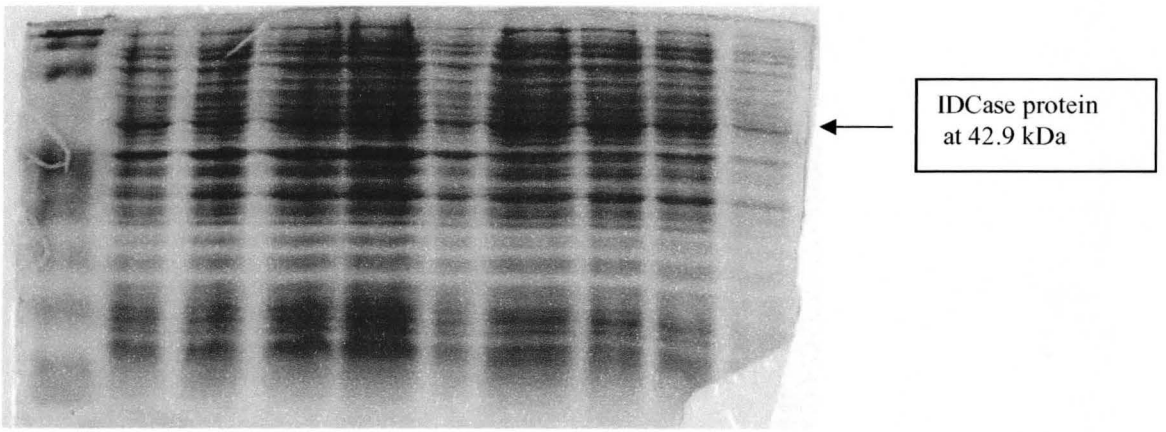


Figure 4-1. SDS-PAGE of IDCase protein over a time course of IPTG induction. Lanes 2-5 (cell lysate) and lanes 6-9 (whole cells) at 3, 6, 9, 18 hrs of induction.

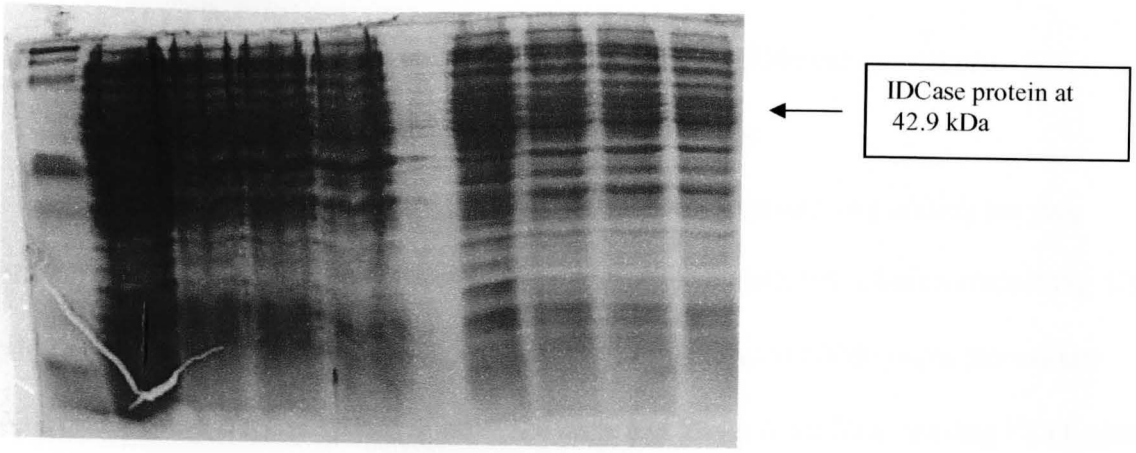


Figure 4-2. SDS-PAGE of IDCase protein with different amounts of IPTG. Lanes 2-5 (cell lysate) and lanes 6-9 (whole cells) at 0 mM, 0.2 mM, 0.5 mM and 0.8 mM IPTG.

B. $^{14}CO_2$ assays for enzymatic activity in a number of different conditions

Materials and Methods:

Iso-orotate decarboxylase (IDCase) activity was measured by adding varying amounts of IDCase (depending on its concentration) to a buffered solution containing 100 μM of [carboxy-^{14}C] labeled iso-orotate and averaging around 6000 counts per minute. As expected, the IDCase would remove the carboxyl group from IOA, leaving $^{14}CO_2$ and uracil. All reactions had a total volume of 500 μL and were run in 20 mL scintillation vials capped with a rubber septum containing a basket that contained wicks. The wicks were made of filter paper (Whatman) and were moistened with 2 M KOH to collect the CO_2 driven off by the reaction. In the cases that we tested IDCase as a cell lysate or part of an ammonium sulfate fraction the reaction contained: 390 μL dH_2O, 50 μL KH_2PO_4 buffer (pH = 7.0), 50 μL of 1mM [^{14}C] IOA and 10 μL of protein; meanwhile when we tested IDCase as part of ion-exchange fractions we used 250 μL of dH_2O and 150 μL of protein.

The reactions were initiated by adding the IDCase, were run for 5 minutes in a 30 °C shaking water bath (gentle shaking) and then quenched with 0.1 mL 2M HCl. The quenched reactions were allowed to sit for 2 hrs, allowing time for the wick to absorb all of the $^{14}CO_2$ released. The wicks were then removed and placed in a 70 °C oven for about twenty minutes or until dry. Once dry, each wick was individually inserted into a pre-counted vial containing scintillation fluid (ScintiSafe Econo 1) and the amount of ^{14}C present was measured with a scintillation counter.

Results and Discussion

The concentration of IPTG that gave most specific activity (nmol/min/mg) of IDCase was found to be 0.8 mM. The results of this finding are shown below (in Figure 4-3):

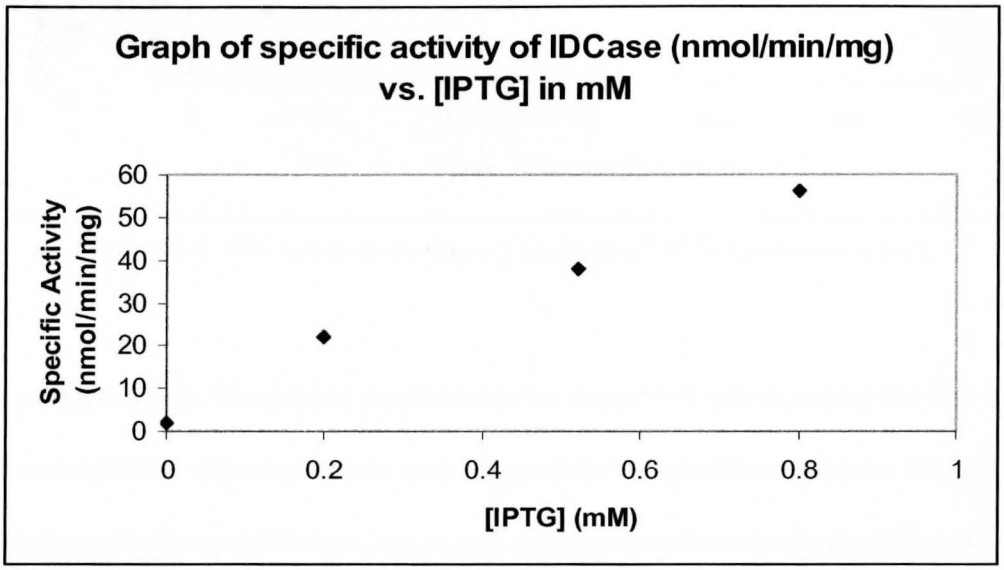


Figure 4-3. Graph of specific activity of IDCase (nmol/min/mg) vs. IPTG concentration in the growth cultures (mM).

Meanwhile, we also found out that the specific activity of IDCase increased with increasing time of induction by IPTG. The maximum activity was established at 19 hrs of induction, after which the activity started to drop, as shown below (Figure 4-4):

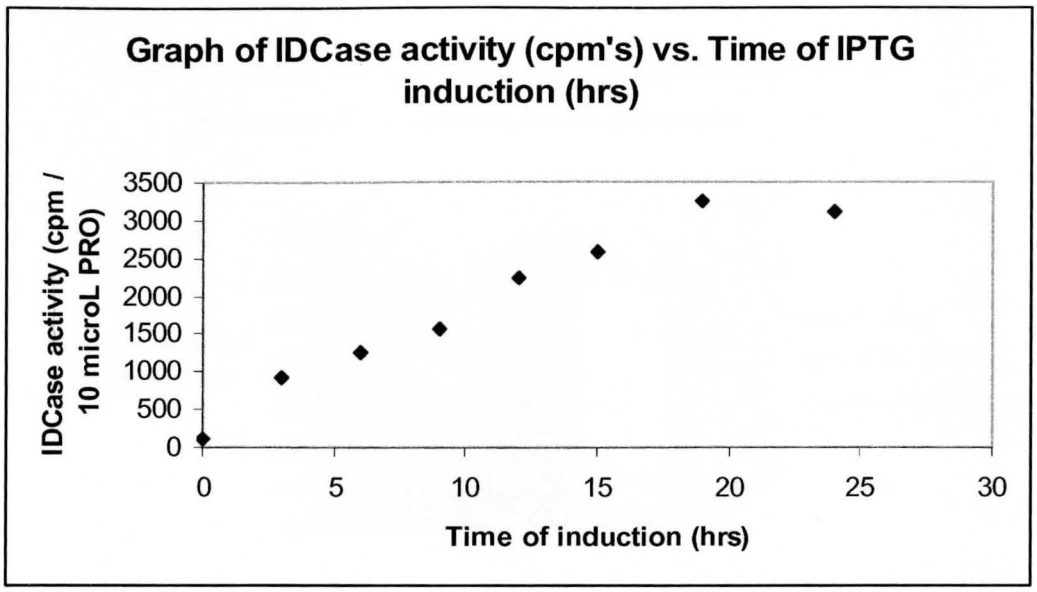


Figure 4-4. IDCase activity (cpm) vs. time of IPTG induction (hrs).

Most importantly, from these experiments we found out that changing the Cys residue into an Alanine, does not have a major impact in the specific activity of IDCase. This is best shown in the graph below, where we plot the specific activity of IDCase in lysates of cells carrying either pCal, pCal IDCase WT or pCal IDCase C272A (Figure 4-5).

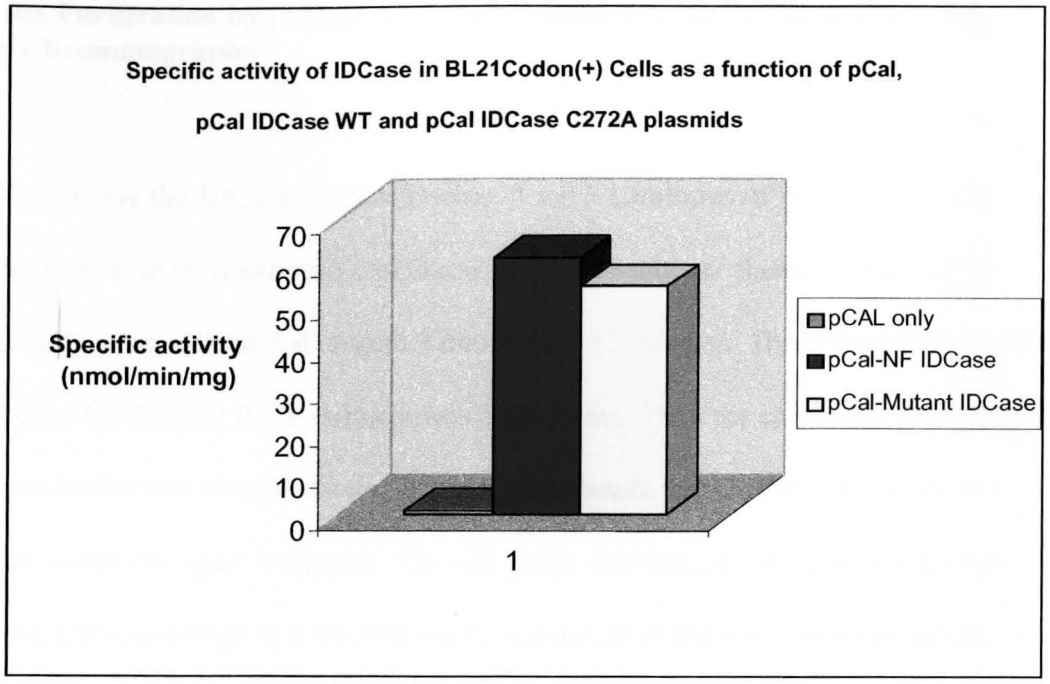


Figure 4-5. Specific activity of IDCase.

From the result above one can come to the conclusion that the mutated residue is not the nucleophilic residue that initiates the reaction by attacking at C-6 of iso-orotate. Instead there must be some other residue that acts as the nucleophile, if the proposed mechanism is operative.

C. IDCase Purification by $(NH_4)_2SO_4$ Fractionation, Ion-Exchange and Custom Affinity Chromatography

Materials and Methods:

To prepare the IDCase for purification, 4 x 0.5 L cultures of BL21 Codon (+) cells were grown under maximum conditions in 2 L Erlenmeyer flasks with plenty of shaking. Cells were collected at around 7,000 x g for 15 minutes. The average cell weight measurement for the four 0.5 L cultures was 5.81 grams. Then the cells were resuspended in cell lysis buffer and disrupted using 0.1 mm glass beads for 3 x 1 min intervals in a large bead beater (Biospec Products). The cell lysate was then centrifuged in a Sorvall RC5C Plus Ultracentrifuge at x 20, 000 x g to remove all of the unwanted cell debris. Some of the supernatant was kept on ice and saved for later quantification by $^{14}CO_2$ and Bradford assays.

To the rest of the soluble protein taken from the broken cells we added enough $(NH_4)_2SO_4$ (29.1 g per 100 mL) in order to give a 50 % saturated solution. The $AmSO_4$ was added slowly over a period of 30 min, while mixing by magnetic stirring in an ice bath. The resultant mixture was stirred for about 2 hours and then centrifuged at 20, 000 x g for 15 min. We saved 0.5 mL of the supernatant (50 % super) and threw away the pellet as it was found to contain no IDCase. To the rest of the 50 % supernatant we added enough $AmSO_4$ (26.8 g per 100 mL) to make a 90% saturated solution. Again, the $AmSO_4$ was added slowly with plenty of stirring in an ice bath. To achieve maximum equilibrium we sometimes left the mixture stirring overnight and then centrifuged at 20, 000 x g for 15 min. The supernatant was saved for Bradford and $^{14}CO_2$ assays, meanwhile the pellet was resuspended in a minimal volume (3-6 mL) of cell lysis buffer. The resuspended pellet was transferred to selective membrane dialysis bags and placed in

a beaker with 1 L of PBS buffer at pH = 7.5, slowly stirred at 4 °C and dialyzed overnight. After the dialysis, we measured the total protein concentration by Bradford Assays and measured IDCase specific activity by $^{14}CO_2$ assays.

The dialyzed protein was then concentrated by using an Amicon protein concentrator in preparation for the ion exchange column. This was done by using N_2 gas at 35 psi to push the protein solution past an ultrafiltration membrane of 44.5 mm diameter with an exclusion limit of 10, 000 Da (Millipore, Bedford, MA). The protein was concentrated at 4 °C to a volume of about 7-8 mL.

Prior to loading, the diethylaminoethyl (DEAE) sephacel anion exchange column was washed with three equivalents of column volume (3 x 40 mL) of 1 M NaCl to clean the column. Also, it was equilibrated by washing with 120 mL of 20 mM Tris Buffer, at a pH of 7.5. The buffer was applied using a flow adapter with the pump operating at a rate of 2 mL / min.

About 7 mL of concentrated protein were loaded on the DEAE column and samples were collected by the use of the fraction collector. The loaded protein had a pH of 7.5, at which point the IDCase would have an overall negative charge and it would stick to the positively charged DEAE resin. Once loaded, the column was then washed consecutively with increasing concentrations of buffered NaCl solutions in order to elute the IDCase off the column. The column was washed with three equivalents of column volume of: 20 mM Tris buffer, 0.2 M NaCl/Tris buffer, 0.4 M NaCl/Tris buffer, 0.6 M NaCl/Tris buffer, 0.8 M NaCl/Tris buffer and 1.0 M NaCl/Tris buffer. The elute was collected in 3-5 mL fractions in 13x100 mm test tubes and the fractions were checked for the presence of total protein by measuring the absorbance at 280 nm. The fractions that

contained significant amounts of protein were used to run IDCase $^{14}CO_2$ assays. Those fractions that showed considerable activity were pooled together and concentrated down.

Results and Discussion:

In the ion exchange fractions assayed we discovered that IDCase was eluting from the column towards the end of the 0.2 M NaCl wash and at the beginning of the 0.4 M NaCl wash, more specifically in fractions 42-62 (Figure 4-6). Furthermore, the majority of the unwanted protein (no IDCase) was eluting at the end of 0.4 M NaCl and throughout all of the 0.6 M NaCl wash. Given this, we thought that we were able to produce partially purified IDCase.

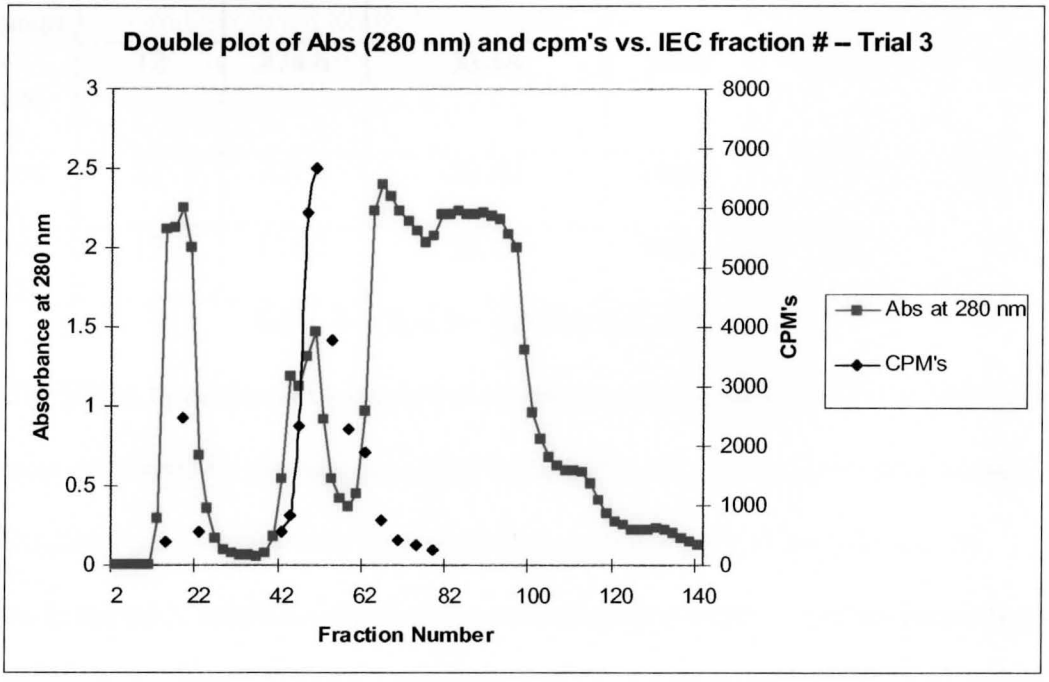


Figure 4-6. Double y-axis plot of Abs (280 nm) and IDCase activity (counts-per-minute) versus IEC fraction number

However, running SDS-PAGE gels of the pooled IEC fractions and the unpurified cell lysate proved otherwise as there was still a lot of unwanted protein in the IEC fractions. Moreover, making an IDCase purification table, such as the one shown below (Table 4-1) we discovered that the specific activity of the IEC pooled fractions was relatively the same as that of the cell lysate, indicating no purification.

Sample	Volume (mL)	Protein (mg)	Specific Activity (nmol/min/mg)	Total Activity	Purification Factor	Percent Yield (%)
Crude Lysate	38	683.7	**32.90**	**22500**		100
50 % $(NH_4)_2SO_4$ Supernatant	38	395.7	**45.69**	**18000**	1.4	80.4
90 % $(NH_4)_2SO_4$ Supernatant	34	63.72	**116.4**	**7400**	3.5	32.9
Non-Dialyzed Pellet	12	274.6	**33.65**	**9200**	1.0	41.0
Dialyzed Pellet	21	331.4	**30.64**	**10000**	0.93	45.1
Pooled IEC Fractions	12	178.7	**38.01**	**6800**	1.2	30.2

Table 4-1. IDCase purification table.

The IDCase purification scheme that we employed has encountered a number of problems as evidenced by the data presented in the table above. First, even after adding the $AmSO_4$ slowly and with plenty of stirring there was still a lot of specific activity remaining in the 90 % $AmSO_4$ supernatant, a troubling observation since this meant that we lost 30 % of our overall activity at this step. This could be due to the possibility that IDCase may be bound to another protein that does not easily precipitate out of solution, and that as we approach 90% saturation, some of the IDCase is still present in solution.

Furthermore, in going from the 50 % supernatant to the dialyzed pellet and then the IEC pooled fractions, the specific activity of IDCase hardly increased. This could be explained by a number of different factors. Given the gene sequence similarity between IDCase and ACMSD, and the possibility that ACMSD may be a metal binding enzyme, it could be that IDCase also requires a metal cofactor and that it is diffused out during the overnight dialysis thereby decreasing the efficiency of our enzyme. Another explanation, is that IDCase is degraded by proteases present in our solution, even though we do use many different protease inhibitors in our cell lysis buffer.

Whatever the case is, the IDCase purification scheme will be built upon and improved in the future projects in our lab until we conceive the most efficient procedure for the isolation of pure and highly active IDCase.

Chapter V: Conclusion

The THase gene from *R. glutinis* has been successfully cloned into a pCal expression system. Our laboratory has been the first to try a number of different growth conditions to produce the most abundant and active recombinant THase in a bacterial system. We were also able to design efficient radioactivity [2-^{14}C Thymine] and HPLC assays for THase.

The THase gene was used to isolate the IDCase gene from the *N. crassa* genome. The IDCase gene was successfully cloned into a pCal vector and multiple growth conditions were used to produce the highest specific activity. Unlike with THase, IDCase was not very soluble in solution and we were unsuccessful in making abundant amounts of protein observable through SDS-PAGE.

An IDCase purification scheme was designed where we used $AmSO_4$ fractionation and DEAE anion exchange chromatography as purification methods. This scheme will be built upon and improved until maximum purity and activity can be achieved.

The proposed nucleophilic cysteine 272 of IDCase was mutated into an Alanine residue. After the mutation, the activity of the mutant protein decreased only slightly bringing us to the conclusion that the original Cys 272 did not play a catalytic role. These results have led us to look at the amino acid sequence of IDCase more closely. Now we have found a very high sequence similarity between IDCase and certain amidohydrolase enzymes. In the amidohydrolases, the catalytic site is a group of histidine and carboxylate residues that activates a water molecule, making it a nucleophile. Therefore, we are now

in the process of proposing that a water molecule may be responsible for the nucleophilic attack at position C-6 of IOA, triggering a decarboxylation.

References

1. Jones, M.E. *Annu. Rev. Biochem.* **49**, 253-279.

2. Voet, D. & Voet, J. G. (2004) *Biochemistry.* Publisher: John Wiley & Sons, Inc.; 3rd edition.

3. Garret, R. H. & Grisham, C. M. (1999) *Biochemistry.* Publisher: Saunders College Publishing, 2nd edition.

4. Palmatier, R. D., McCroskey, R. P., and Abbott, M. T. *Journal of Biological Chemistry*, **245**, 6706-6710.

5. Smiley, J.A., Angelot, J.M., Cannon, R.C., Marshall, E.M., and Asch, D.K. *Analytical Biochemistry,* **266**, 85-92.

6. Thornburg, L. and Stubbe, J. (1993) *Biochemistry*, **32**, 14034-42.

7. Warn-Cramer, B., Macrander, L., and Abbott, M. (1983) *Journal of Biological Chemistry*, **258**, 10551-57.

8. Hsu, C., Saewert, M., Polsinelli, L., and Abbott, M. (1980) *Journal of Biological Chemistry*, **256**, 6098-6101.

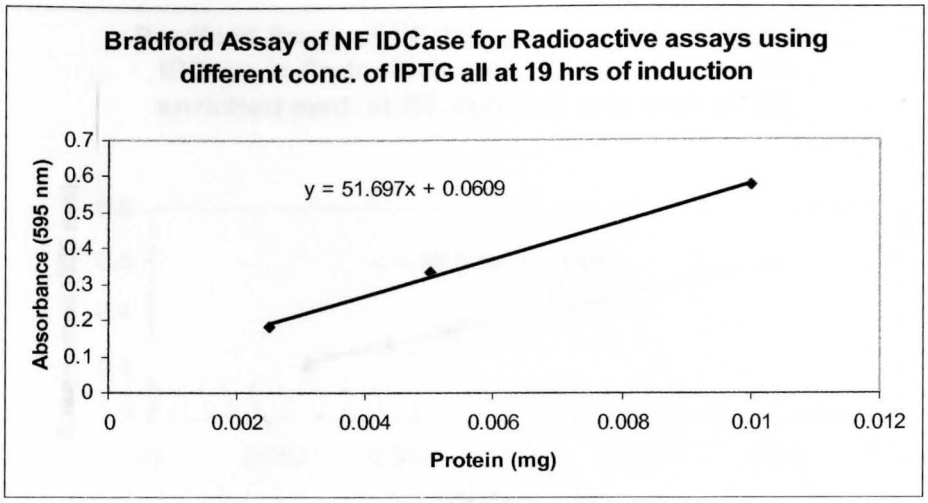
9. Thornburg, L., Lai, M., Wishnok, J., and Stubbe, J. (1993) *Biochemistry*, **32**, 14023-33.

10. Harvard Medical School, Boston, Massachusetts. (1997) *Current Protocols in Molecular Biology.* Publisher: John Wiley & Sons, Inc.

11. Smiley, J., Kundracik, M., Landfried, D., Barnes, V. and Axhemi, A. Genes of the thymidine salvage pathway: Thymine-7-hydroxylase from a *Rhodotorula glutinis* cDNA library and iso-orotate decarboxylase from *Neurospora crassa. Biochimica et Biophysica Acta*, Accepted for publication 4 February 2005.

Appendix A: Data used to obtain specific activity of IDCase (nmol/min/mg) vs. IPTG concentration (mM) (Figure 4-3 in thesis)

Bradford assay standard curve:



Protein concentrations of the individual samples:

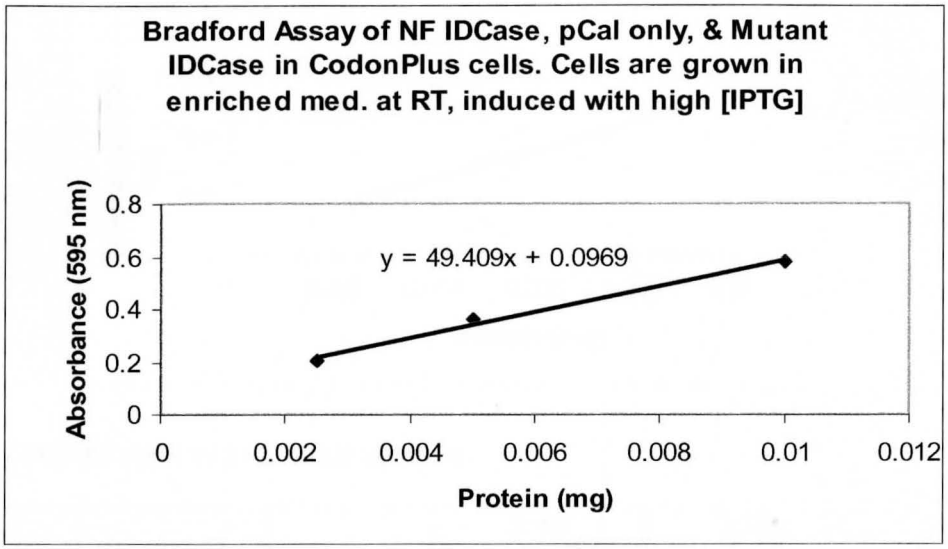
[IPTG] (mM)	Absorbance	Protein (mg)	1:10 conc. (mg /µL)	1:10 conc. (µg/µL)	Full [protein] (µg /µL)
0	0.49131	0.008325	0.001665	1.665	16.65
0.2	0.34561	0.005507	0.001101	1.101	11.01
0.52	0.24561	0.003572	0.0007145	0.7145	7.150
0.8	0.22718	0.003216	0.0006432	0.6432	6.430

Specific activities of the individual samples (10 µL of protein were used for each assay):

[IPTG] in (mM)	Average cpm's	nmol converted	nmol/min	[protein] (µg /µL)	protein in assay (mg)	Specific activity (nmol/min/mg)
0	197.3	1.644	0.3288	16.65	0.1665	1.97
0.2	1454.4	12.12	2.424	11.01	0.1101	22.01
0.52	1624	13.53	2.706	7.15	0.0715	37.85
0.8	2169.3	18.07	3.615	6.43	0.0643	56.23

Appendix B: Data used to obtain specific activities of IDCase in lysates of cells carrying either pCal, pCal IDCase WT or pCal IDCase C272A (Figure 4-5 in thesis).

Bradford assay standard curve:



Protein concentrations of the individual samples:

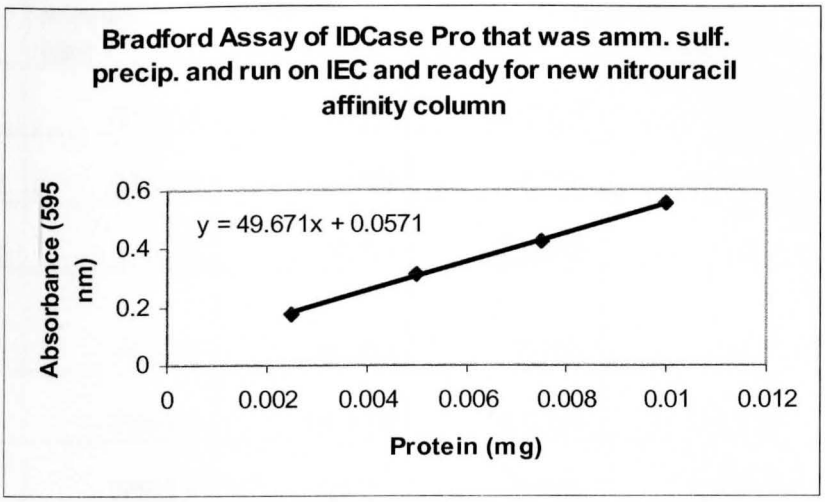
Sample	Absorbance	Protein (mg)	1:10 conc. (mg /µL)	1:10 conc. (µg/µL)	Full [protein] (µg /µL)
pCal at 21 hrs	0.30971	0.004307	0.0008614	0.8614	8.614
NF IDCase at 21 hrs	0.34359	0.004992	0.0009985	0.9985	9.985
Mutant IDCase at 21 hrs	0.22392	0.00257	0.0005141	0.5141	5.141

Specific activity of the individual samples (10 µL of protein were used for each assay):

Sample	Average cpm	nmol	Nmol / min	[protein] (µg /µL)	protein in assay (mg)	Spec. Activity (nmol/min/mg)
pCal at 21 hrs	45.4	0.3783	0.0756	8.614	0.086	0.8784
NF IDCase at 21 hrs	3620.7	30.172	6.034	9.985	0.099	60.43
Mutant IDCase at 21 hrs	1653.6	13.78	2.756	5.141	0.05141	53.61

Appendix C: Data used to obtain IDCase Purification Table (Table 4-1 in thesis)

Bradford Assay Standard Curve



Protein Concentration of individual samples:

Sample	Absorbance	Protein (mg)	1:10 Concentration (mg / µL)	1:10 Concentration (µg/µL)	Full Concentration (µg / µL)
0% supernatant	0.5039	0.00899	0.00179	1.799	17.99
50 % Supernatant	0.3157	0.00520	0.001041	1.041	10.41
90 % Supernatant	0.1036	0.000936	0.0001873	0.187	1.873
Non-Dialyzed Pellet	0.6254	0.01144	0.002288	2.288	22.88
Dialyzed Pellet	0.449	0.00788	0.001577	1.577	15.77
IEC Pooled Fractions.	0.4268	0.00744	0.001488	1.488	14.88

Specific Activity of the individual samples (10 µL of protein were used for each assay):

Sample	Average cpm	nmol converted	nmol/min	mg of PRO in assay	Specific activity (nmol/min/mg)
0% supernatant	3552.2	29.6	5.9203	0.17994	32.90
50 % Supernatant	2854.8	23.79	4.758	0.10413	45.69
90 % Supernatant	1309.3	10.91	2.1822	0.0187	116.44
Non-Dialyzed Pellet	4620.4	38.503	7.701	0.228	33.65
Dialyzed Pellet	2901.1	24.176	4.8352	0.1578	30.64
IEC Pooled Fractions.	3395.7	28.29	5.659	0.1488	38.01

IDCase Purification table:

Sample	Volume (mL)	Protein (mg)	Specific Activity (nmol/min/mg)	Total Activity	Purification Factor	% Yield	
Crude Lysate	38	683.772	32.9	22496.0		100	
50 % Super	38	395.694	45.69	18079.2	1.388	80.36	
90% Super	34	63.716	116.44	7419.09	3.539	32.97	
Non-Dialyz Pell	12	274.56	33.65	9238.94	1.022	41.06	
Dialyzed Pellet	21	331.38	30.641	10153.8	0.931	45.13	
Pooled IEC	12	178.68	38.01	6791.62	1.155	30.19	