ABSTRACT

Title of Document: IODOTYROSINE DEIODINASE FROM SELECTED PHYLA ENGINEERED FOR BACTERIAL EXPRESSION

Jennifer Marilyn Buss
Doctor of Philosophy, 2012

Directed By: Professor Steven E. Rokita
Department of Chemistry and Biochemistry

Iodide is a well known halogen necessary for development. The majority of iodide processing in biological systems occurs in the thyroid gland. Iodide salvage is essential to thyroid hormone metabolism and metabolic regulation. The DEHAL1 gene product iodotyrosine deiodinase (IYD) is responsible for deiodination of the mono- and diiodotyrosine byproducts of thyroid hormone synthesis (triiodothyronine and thyroxine, T3 and T4, respectively). IYD is a membrane-bound flavoprotein comprised of three domains with the catalytic domain belonging to the NADPH oxidase/flavin reductase structural superfamily. This enzyme required engineering for expression of soluble protein in *E. coli* and was characterized using CD spectra, kinetic rate constants, binding constants of substrates, and crystal structure. Analysis

of the crystal structure of IYD indicates a dimer with an active site comprising of both monomers and orienting the C-I bond of iodotyrosine substrate stacking above the N5 of the flavin mononucleotide (FMN) required for activity. The crystal structure also identifies an active site lid that distinguishes IYD from other proteins in the NADPH oxidase/flavin reductase superfamily. Three amino acids (E153, Y157, and K178) on the active site lid form hydrogen bonding and electrostatic contacts with the zwitterionic portion of the substrate. Mutation to any of these three amino acids significantly decreases substrate-binding affinity and enzymatic activity. Homologous sequences of IYD were identified in other organisms and four sequences as representatives from their phyla were expressed in *E. coli*. Zebrafish, lancelets, honeybees, and sea anemones each have a protein that acts as a deiodinase.

IODOTYROSINE DEIODINASE FROM SELECTED PHYLA ENGINEERED FOR BACTERIAL EXPRESSION

By

Jennifer Marilyn Buss

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor Steven E. Rokita, Chair
Professor Douglas A. Julin
Professor David Fushman
Assistant Professor Nicole LaRonde-LeBlanc
Professor Steven Mount

Dedication

To my family, for their endless support and encouragement.

Acknowledgements

I'd like to begin by thanking my advisor, Dr. Steve Rokita. Thank you for your continued support. Thank you for challenging me to be a better scientist, writer, and thinker. For teaching me to tell a story and a constant reminder to always check the pH.

Thank you to my committee Professors Julin, Fushman, LaRonde-LeBlanc, and Mount. I appreciate your input and your expertise. To past members of my committee and other faculty who helped make me a better scientist: Dr. Gerratana, Dr. Cropp, and Dr. Beckett.

A special thank you to Dr. Jim Watson. We began as coworkers doing late night experiments and we've ended with you as my boss and late night grading. Thank you for keeping it real and putting it all in to perspective.

Thank you to all the collaborators that have made my research possible. Dr. Nicole LaRonde-LeBlanc, Dr. Seth Thomas, Dr. Lee Chuenchor for solving crystal structures of our protein that has made my work possible. Dr. Hamza and Dr. Xiao the use of their tissue culture hoods.

To Rokita lab members past and present, thank you. To the endless friends I have made. The support that you have provided has been unbelievable. Kathy, words cannot express our time together in school and in the real world. Petrina, Crazy nights adding just a drop more fun. Mike – Go Blue Hens! Patrick, less is more.

To the Departmental staff, who have been the moving parts behind closed doors. Thank you, Tia, for keeping everything running smoothly.

Last but not least, I thank my parents, who have called me Almost Doctor for entirely too many years. The day has come. Thank you for all your encouragement, through the ups and downs, for the best mattress, and the free advice. To my brother and sister, for keeping me laughing, either at them or at myself.

Table of Contents

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Abbreviations

BLAST	Basic Local Alignment Search Tool
DIT	diiodotyrosine
EC	Enzyme Commission
FMN	flavin mononucleotide
FRP	flavin reductase P
ID	iodothyronine deiodinase
IPTG	isopropyl β-D-1-thiogalactopyranoside
IYD	iodotyrosine deiodinase
IYD(Δtm)	truncated iodotyrosine deiodinase
IYD(Δtm)DM	truncated double cysteine mutant iodotyrosine deiodinase
His_6	polyhistadine affinity tag
k_{cat}	catalytic rate
K_D	dissociation constant
K_M	Michaelis-Menten constant
MIT	monoiodotyrosine
NADPH	nicotinamide adenine dinucleotide phosphate
NCBI	National Center for Biotechnology Information
NIS	Na^+/I^- symporter
PVDF	polyvinylidine fluoride
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TRX	thioredoxin

Chapter 1: Introduction

1.1 Iodide Salvage

Iodide is a critical micronutrient for mammalian health, and deficiencies in either dietary iodide or iodide metabolism may lead to hypothyroidism, goiter, and developmental defects. The World Health Organization has identified iodide deficiency as a leading cause of impaired cognitive development in children resulting in a drop of approximately 15 IQ points (*1*). Over a billion people worldwide are estimated to lack sufficient access to iodide despite the modest requirement for this micronutrient (recommended daily allowance of 150 μg) (*1*). Iodide is quite scarce in land-locked environments. However, even seawater has a very low concentration of iodide (under 1 μM), which is well below that of the other halogens (*2*). In seawater, brown algae, a major food in marine diets, accumulates iodide (*3*).

Iodide is utilized by the thyroid to synthesize the thyroid hormone 3,3',5-triiodothyronine (T3) and its precursor 3,3',5,5'-tetraiodothyronine (T4) (Figure 1.1). Almost a century ago, the thyroid was determined to be essential for development and metamorphosis. Endogenous thyroid hormone metabolism has been extensively studied beginning as far back as the 1920s. Removal of thyroids in amphibian cultures does not result in metamorphosis. However, injecting thyroid particulates or iodotyrosines (mono- or diiodotyrosine, MIT or DIT, respectively) did allow for metamorphosis in amphibians (*4*). Here, MIT and DIT serve as the iodide source for the organisms. Interestingly, in the same study, larvae that were fed bromotyrosines showed no change in anatomic transformations. This indicates that bromide cannot be substituted for iodide to complete thyroid hormone metabolism (*4*). More recently,

studies in lancelets, the most basal chordates, identified thyroid hormone receptors that bind thyroid hormone derivatives (*5*). Chordates utilize thyroid hormones in conjunction with their respective receptors to regulate metamorphosis (*6, 7*). The process of metamorphosis and its key players are conserved through chordate lineage.

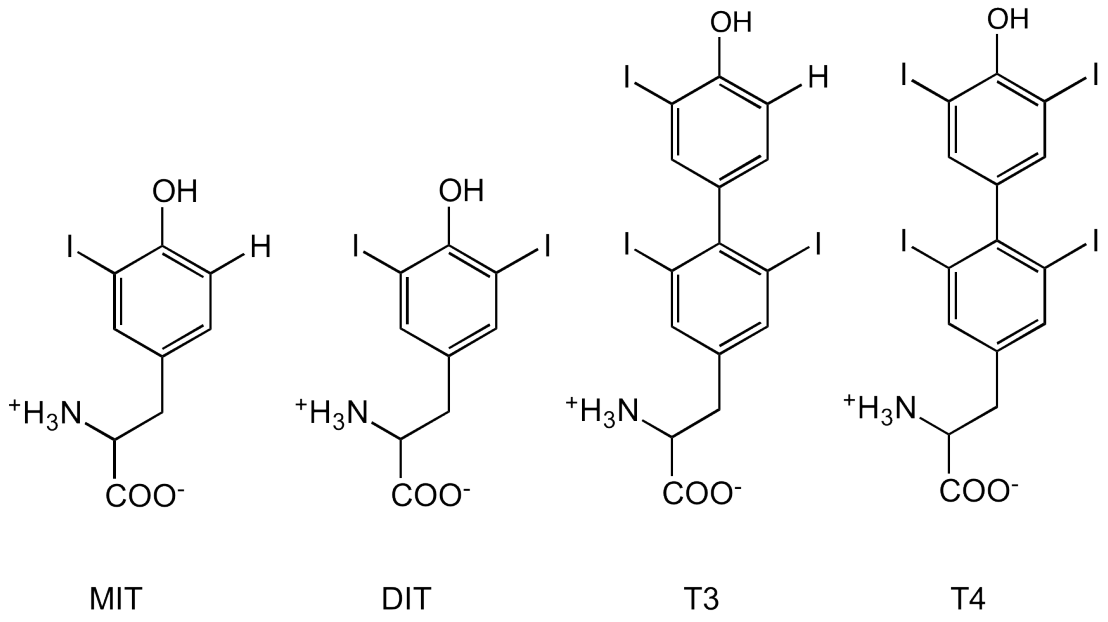


Figure 1.1 Monoiodotyrosine and diiodotyrosine are the by-products formed during T3 and T4 biosynthesis.

The Chordate thyroid gland emerged through evolution. Vertebrates have a specific follicular thyroid gland to gather iodide and produce thyroid hormones. Cephalochordata and Urochordata have a region of their body, identified as an endostyle, that functions as a thyroid (*8*). More primitive organisms, such as sand dollars (*Peronella japonica*), do not have a defined region but do have indications of a use for thyroid hormones, specifically in metamorphosis (*9*). Thyroid function can be traced back as far as Cyanobacteria in its ability to use iodide as an antioxidant (*10*). As organisms developed and transferred from the relatively iodide-rich water environment to iodide-deficient land, they had to compensate for the lack of available

iodide. This adaptation is evidenced by an iodide reserve in the thyroid gland. Iodide in the vertebrate thyroid is used to produce iodinated thyroid hormones.

The production of thyroid hormones is very well characterized in mammalian systems. First, the sodium iodide symporter (NIS) transports iodide into the thyroid cells (Figure 1.2). NIS, located in the basal membrane of thyroid follicular cells, concentrates iodide (40-fold) from the plasma (*11*). Iodide is then transported into the colloid where iodination of tyrosine residues on thyroglobulin occurs via thyroid peroxidase. Thyroglobulin is the major constituent of the thyroid colloid and its mono- and diiodinated tyrosine residues serve as intermediates in the formation of T3 and T4. Approximately a quarter of the ~220 tyrosine residues found in thyroglobulin are available for iodination. Some of the proximal iodinated tyrosines undergo a peroxidase-mediated conjugation to form T4. Ultimately, each thyroglobulin polypeptide produces one T4, while seven MITs and six DITs remain and just one out of every three thyroglobulin polypeptides produces one T3 rather than T4 (*12*). Thyroid stimulating hormone initiates endocytosis of thyroglobulin into the follicular thyroid cells. Subsequent proteolysis of thyroglobulin releases T3 and T4, as well as the ubiquitous by-products MIT and DIT.

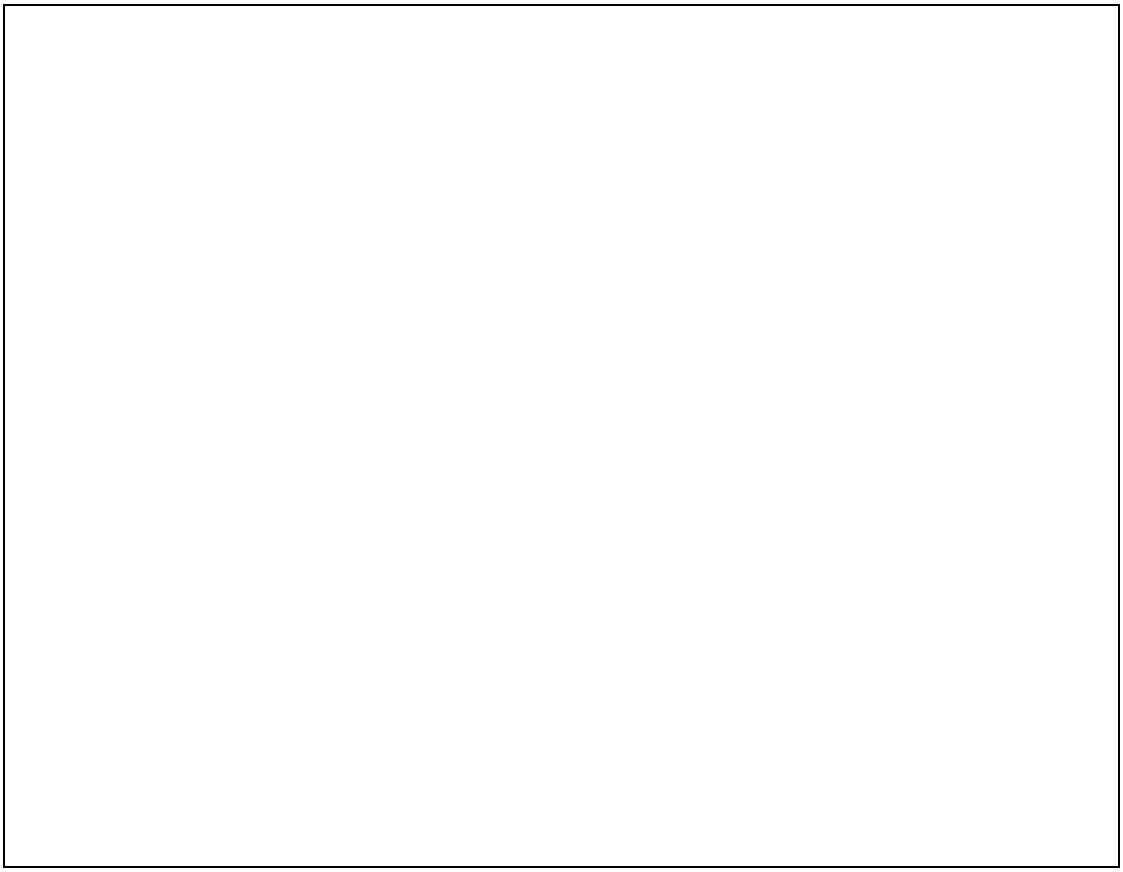


Figure 1.2 Iodide uptake, transport, metabolism and salvage in follicular thyroid cells. Adapted from (*13*).

Iodide in the thyroid is stored as iodinated tyrosine residues of thyroglobulin. Upon proteolysis, these amino acids are released from the protein as waste. Reincorporation of iodide into thyroglobulin requires free iodide rather than iodinated tyrosines. MIT and DIT released during proteolysis of thyroglobulin are deiodinated by iodotyrosine deiodinase (IYD) (Figure 1.3). This recycling is critical for thyroid hormone production; otherwise, MIT and DIT would be excreted as waste, exponentially increasing our iodide requirement (*14*).

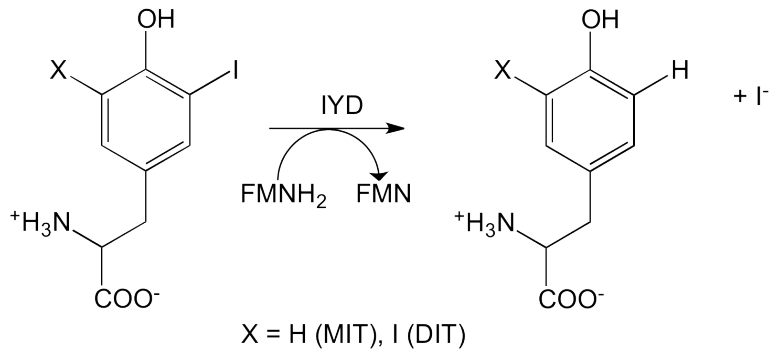


Figure 1.3 IYD catalyzes the deiodination of MIT and DIT.

1.2 Cysteines and deiodination

Iodothyronine deiodinases (ID) catalyze the reductive deiodination of thyroid hormones T3 and T4 and their derivatives (*15*). T3 is 10-fold more potent for regulating metabolism than T4 (*15*), and therefore the enzyme that catalyzes this reaction (ID) is biologically relevant. ID contains an essential active site selenocysteine, which is involved in catalysis. Mutation of this selenocysteine to cysteine reduces the catalytic activity of ID by 100-fold (*15*). The net reaction promoted by ID relies on the reducing power of thiols to regenerate its active site selenocysteine from a proposed selenyl iodide reaction (*16*).

IYD was originally thought to deiodinate its substrate by a similar mechanism to ID (*17*). Mammalian IYD has two conserved cysteine residues that were thought to be responsible for catalysis (*18*). A three-dimensional model of the catalytic domain predicted that the two cysteine nucleophiles were in close proximity to each other and FMN creating an active site (*19*). These two cysteines could react in a concerted fashion similar to the selenocysteines in ID. This theory was proven wrong when mutation of IYD's two cysteines resulted in catalytically active protein (*18*). Single mutations and the double cysteine mutation each released iodide as a result of

substrate turnover. This proved that cysteines are not responsible for removal of iodine from MIT or DIT. IYD must dehalogenate its substrate through a different reductive process.

1.3 IYD is a dehalogenase

There are three known reductive dehalogenation reactions in aerobic organisms, all of which are structurally independent. The first example, described above, is the deiodination of thyroid hormone by ID. This enzyme belongs to a thioredoxin superfamily that, unlike thioredoxin, has active site selenocysteines responsible for catalysis (*20*). The second enzyme, tetrachlorohydroquinone dehalogenase, dechlorinates tetrachlorohydroquinone by nucleophilic attack of a cysteine through the oxidation of glutathione (*21, 22*). Tetrachlorohydroquinone dehalogenase is a member of the glutathione S-transferase structural superfamily. IYD has yet a third mechanism of action to promote dehalogenation; we already know that cysteines are not plausibly responsible. IYD is the only dehalogenase to utilize FMN (*23, 24*) and is a member of the NAD(P)H oxidase/flavin reductase structural superfamily (*19*). Turnover requires reduction of FMN by NADPH *in vivo* (*25*) and can be reduced *in vitro* by dithionite.

IYD is the primary mechanism of iodide recycling in mammals (*13*). Recent advances in IYD expression and purification allowed for in depth investigation of enzymatic properties. Specifically, IYD was found to not only act as a deiodinase, but also as a much more versatile dehalogenase (*26*). Chlorinated and brominated tyrosine substrates were also found to have the same effect on the FMN spectrum insinuating that their dehalogenation was also promoted by oxidation of reduced

FMN. However, the addition of fluorinated tyrosine to reduced IYD did not promote oxidation of FMN, and thus IYD does not have the catalytic power to reduce a C-F bond.

Both chlorotyrosine and bromotyrosine are generated by peroxidases *in vivo* and used as markers for respiratory problems (*27, 28*). IYD is not only found in the thyroid, but traces are found in other organs such as the kidney and liver (*29*). So it is possible that IYD can dehalogenate the halogenated tyrosines as they circulate thorugh the blood in either the liver or the kidney. Exposure to some brominated compounds has shown reduction of thyroid hormone concentrations (*30*). It is curious that IYD has the ability to complete these reactions while brominated tyrosines, unlike iodinated tyrosines, did not have a morphological effect on organisms as previously reported (*4*).

1.4 IYD has unique function within its structural superfamily

In its native state, IYD is a membrane bound protein that can be described in three structural domains (Figure 1.4). Its N-terminus binds the protein to the membrane, the C-terminus is the catalytic domain, and these two functional domains flank an intermediate domain with no known structure or function aside from connecting the two. The catalytic domain has been assigned to the NADPH oxidase/flavin reductase structural superfamily (*19*). Upon removal of the N-terminal membrane domain (residues 2-33), IYD retains activity albeit without its ability to be reduced by NADPH (*18*). Expression of the truncated protein in Sf9 cells produced sufficient amounts of protein for crystallography.

Figure 1.4 IYD is comprised of three domains: a transmembrane domain (1-24), an intermediate domain (25-82), and a catalytic NADPH oxidase/flavin reductase domain.

IYD(Δtm), a transmembrane-deleted, soluble form, was crystallized with and without substrate present. It has a α-β fold that forms a domain-swapped dimer, which is characteristic of proteins in the superfamily (*31*) (Figure 1.5). The dimer interface is comprised of one helix from each monomer that forms an 'X' or a criss-cross at the boarder. The N- and C-terminal extensions from each polypeptide wrap around the opposite monomer, consistent within the NADPH oxidase/flavin reductase superfamily. Two equivalent active sites are comprised of amino acids from each polypeptide chain indicating that dimer formation is necessary for co-factor binding and catalytic activity. The co-crystals of IYD(Δtm)•MIT contained two regions of electron density that were unavailable from IYD(Δtm). The new electron density implicates an active site lid that forms a helix-turn-helix covering the active site and sequesters the substrate and flavin from exposure to solvent.

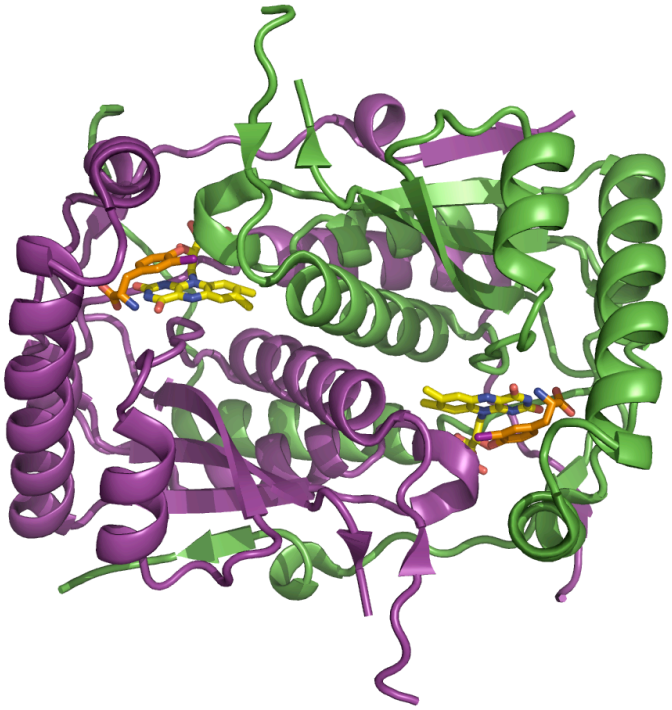


Figure 1.5 An overview of the IYD homodimer crystallized in the presence of MIT substrate. Flavin is in yellow and MIT substrate is orange.

Until recently, all proteins in the superfamily belonged to two sub-classes generally represented by NADH oxidase from *Thermus thermophilus* (*32*) and FRP from *Vibrio harveyi* (*33*). The closest structural neighbor is BluB (PDB ID: 2ISK), a bacterial enzyme responsible for the degradation of its FMN cofactor (*34*). BluB and IYD differentiate themselves from the superfamily based on their differences in both structure and catalytic function. Specifically, IYD and BluB retain sequence that forms the active site lid in a region that does not share homology with the proteins in the other two classes of the superfamily (Figure 1.6). For NOX, the lid is formed from a central region of its polypeptide. For FRP, the C-terminus is extended creating the active site lid.

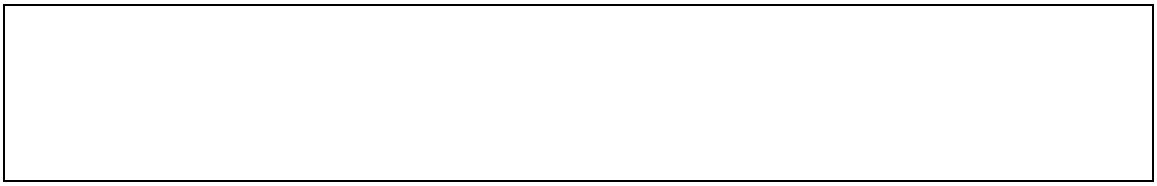


Figure 1.6 Secondary structure alignment of representatives of the NADPH oxidase/flavin reductase superfamily. IYD (PDB ID: 3FBD) and BluB (PDB ID: 2ISL) now define a third subclass of the NADH oxidase/flavin reductase superfamily. NADH oxidase (NOX, PDB ID: 1NOX) and FRP (PDB ID: 2BKJ) illustrate the α-β fold for the original two subclasses of this superfamily. The boxed regions indicate the sequences that form the active site lids. The dotted lines indicate spacing inserted for alignment. Structural assignments were derived from crystallographic data.

Despite the vast structural similarities (root mean square deviation of only 3.1 Å), BluB and IYD only share 19% sequence identity and 43% sequence similarity. When we compare IYD to NOX or FRP having 42% sequence similarity of the crystallized domain (amino acids 66-273 of the mouse IYD sequence). This leads to questions of the relationship between sequence, structure, and function. The relationship between these has been elusive over the decades since protein structures were initially solved. In this superfamily, there is low sequence identity and high structure similarity. In general, protein sequence similarity/identity is equated to protein structure similarity. IYDs connection in this superfamily of bacterial enzymes leads to questions involving the evolution of the protein through animal lineage.

1.5 Similarity of thyroidal enzymes in model organisms

For years, less complex organisms have been extensively studied to model human behavior and anatomy. Mice are frequently used as a model organism to study

human diseases because of their availability, short life cycle, and commonalities in mammalian development. Many experiments that would be unethical to perform in humans are performed in mice. Other organisms that are easier to maintain are used to study development including *Caenorhabditis elegans* (roundworms), *Nematostella vectensis* (sea anemone), and *Strongylocentrotus purpuratus* (sea urchin). A variety of organisms from different phyla have been used to study evolution including *Branchiostoma floridae* (lancelet), *Drosophila melanogaster* (fruit fly), and *Danio rerio* (zebrafish).

Many of these organisms have evidence of thyroid hormone and homologous proteins to those found in the human thyroid. Specifically, the sodium iodide symporter, thyroid peroxidase, and thyroglobulin are all present in chordates. Thyroid hormone related proteins, including the transport protein transthryetin and activating enzyme iodothyronine deiodinase, are also thought to be present based on predicted proteins extrapolated from the genome. Some of the lower, less complex organisms that do not have a thyroid or thyroid-like gland, specifically species in the subphylums Echinodermata and Cnidaria, have both protein sequence that is homologous to IYD (57% identity) and an indication of thyroid hormones.

Thyroidectomies were performed on amphibians to study the effect of lacking a thyroid (*4*). Thyroidless frogs did not morph without the presence of iodide in their diet (*4*). More recently, IYD was detected in the olfactory epithelium, cerebellum, pituitary glands, and epithelial cells of tadpoles (*7*). In zebrafish, mRNA of IYD was detected as early as 38 hours postfertilization in developing thyroid tissue and localization of IYD to the differentiated thyroid and liver cells after 100 hours

postfertilization (*35*). After metamorphosis, reduced expression of functional differentiation markers (IYD, thyroid peroxidase, and thyroglobulin) was observed (*35*) meaning expression of these proteins is only seen in larvae, not in the adults.

1.5 Specific Aims

Characterization of IYD has been long overdue since it was originally identified over 60 years ago. Only over the last six years has significant progress has been made understanding IYD catalysis. My research has focused on making the protein available through expression in bacterial culture and tracing the evolution of IYD from a bacterial oxidase/reductase to a dehalogenase.

1) For years, recombinant IYD was expressed as insoluble protein in *E. coli*. Isolation of IYD in a large scale was essential to overcome the cost/time commitment to mammalian cell expression. The use of fusion proteins was investigated in hopes to aid folding in *E. coli* to continue mechanistic analysis.

2) The engineered protein expressed in *E. coli* had to be compared to the wild-type expressed in insect or yeast cells in order to use the engineered protein as a model for the wild-type. The biochemical and biophysical properties were examined of truncated IYD.

3) Three amino acids were suggested to be at least partially responsible for substrate binding. The requirement for these amino acids was investigated through site-directed mutagenesis to determine which of these amino acids impacted substrate binding and catalytic turnover.

4) Structurally, IYD belongs to a bacterial superfamily and has its own unique function within this group. The origin of IYD was studied through phylogenetic and bioinformatics analysis. Heterologous expression of homologous IYD proteins allowed for evaluation of deiodinase activity.

Chapter 2: Engineering IYD for isolation from *E. coli*

2.1 Introduction

The initial isolation of IYD required detergent extraction of thyroid microsomes (*36*). Purification was accomplished in very low yield and followed by precipitation and multiple chromatographic procedures (*24*). Limited proteolysis by trypsin was subsequently discovered to remove the N-terminal membrane anchor of IYD and release a soluble and active domain from thyroid microsomes (*19*). However, neither the purity nor stability of this fragment was sufficient for mechanistic and structural studies. Only after identification of the IYD gene could it be cloned and expressed in cell culture (*19, 37*). HEK 293 cells supported mutagenesis studies on IYD requiring only catalytic amounts of protein and demonstrated that deletion of residues 2-33 generated a soluble and stable form of the enzyme IYD(Δtm) (*18*). Further truncation of IYD to its core domain homologous to the NADPH oxidase/flavin reductase superfamily resulted in insoluble and inactive protein when expressed in HEK 293 (*38*). Subsequent expression of the truncated gene in *E. coli* only resulted in insoluble protein (*38*).

Heterologous expression of proteins in bacterial hosts has been made available through the use of recombinant DNA technology. A high level of expression of recombinant proteins frequently results in misfolded and insoluble protein. A variety of fusion proteins exist to aid solubility and purification (*39*). Common fusions glutathione-S-transferase, NusA, and thioredoxin were expressed N-terminal of IYD but failed to produce soluble protein (*38*).

To overcome the folding problems evident with *E. coli* and the limited material available from mammalian cell culture, methylotropic yeast *Pichia pastoris* was the next host of choice for deiodinase expression. Expression of a soluble derivative of the *Mus musculus* gene IYD(Δtm) was not detected in the cell lysate after denaturing PAGE, Coomassie staining and Western blotting due to the discrepancy in codon usage between *P. pastoris* and HEK 293 (*40*). A synthetic gene was constructed to overcome the use of rare codons in *P. pastoris*. Recombination and expression of optimized IYD in *P. pastoris* produced detectable deiodinase activity although the yield of enzyme was small (*41*).

IYD(Δtm) was expressed very efficiently in Sf9 insect cells utilizing the baculovirus expression vector system of the Bac-to-Bac® kit as described previously (*31*). This host provided ample quantities of protein for crystallographic studies (*31*). The information yielded from the structure instigated a series of investigations requiring site-directed mutagenesis. Although the cell line was sufficient for wild-type protein, expressing many variations of mutations is inconvenient. Each genetic mutation requires repackaging of baculovirus DNA prior to infection of Sf9 cells. Despite the favorable results from expression in Sf9, preparation of mutants in this system is not ideal because of their demand for time and materials.

The first mutations to the IYD gene made to test the proposed cysteine mechanism were expressed in HEK 293 (*18*). Two conserved cysteine residues (C217 and C239) found in the catalytic domain were thought to be responsible for catalysis (*19, 38*). Mutation of these cysteines to alanines resulted in active enzymes indicating that neither cysteine nor the combination were responsible for deiodination (*17, 18*).

Improper oxidation or reduction of the thiol side chain of cysteines often complicates expression of recombinant proteins in bacteria (*42, 43*). Once catalysis was discovered to be independent of cysteines, efforts were reinitiated to express such mutants in *E. coli*.

The struggle to use *E. coli* for production has continued despite some recent advances in eukaryotic cell lines. This chapter will describe the many fusion protein expression variations for *Mus musculus* IYD in *E. coli*. Insertion of the IYD gene into plasmids containing thioredoxin (pET32a), glutathione-S-transferase (pGEX4T-1), and SUMO (pET28-SUMO) fusion proteins was tested for soluble expression of the truncated IYD gene. Ultimately, mutation of the cysteines and fusion of thioredoxin to IYD has allowed for soluble expression in *E. coli*.

2.2 Experimental Procedures

2.2.1 Materials

Oligodeoxynucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA). All enzymes were purchased from New England Biolabs (Ipswich, MA). OneShot Top10 *Escherichia coli* were purchased from Invitrogen (Carlsbad, CA) and Rosetta2 (DE3) *Escherichia coli* were purchased from Novagen (San Diego, CA). The pET plasmids and antibodies for Western blotting were purchased from Novagen. All other reagents were obtained at the highest grade available and used without further purification.

2.2.2 General Methods

DNA isolation was performed using either a Qiaprep Mini Kit (Qiagen, Valencia, CA) or a GeneJet Plasmid Miniprep Kit (Fermentas, Glen Burnie, MD). PCR reactions were performed using an Eppendorf Mastercycler. Agarose gel electrophoresis (horizontal) was performed according to Ausubel (*44*) using 125 V and a standard of Mass Mix DNA Ladder (Fermentas). Ligations and dephosphorylations were performed under standard conditions using T4 ligase and Antarctic phosphatase, respectively (New England Biolabs, Ipswich, MA). *E. coli* transformations were performed according to Ausubel (*44*) with an Eppendorf Electroporator 2510 (1700 V, 1 mm gap cuvette), and samples were selected against LB plates containing the appropriate antibiotic. DNA sequencing was performed by Geneway Research (Hayward, CA) and the University of Maryland Biotechnology Institute (College Park, MD).

Discontinuous SDS-PAGE gels (12 % acrylamide resolving and 5 % stacking) and Laemmli running buffer were prepared according to standard procedures (*45*) and run according to Ausubel (*44*) using 200 V and a Mini Protean 3 gel system (Bio-Rad, Hercules, CA). All protein gels were stained with Coomassie Brilliant Blue. Electrophoretic transfer of proteins from SDS-PAGE gels to PVDF membranes (Invitrogen) for western blotting was performed with a Bio-Rad Mini Trans-Blot Cell according to the manufacturer's directions. Western Blot was performed according to Ausubel (*44*) using a His•Tag monoclonal antibody, a Goat Anti-Mouse IgG alkaline phosphatase conjugate and fluorescence of ECF (GE Healthcare Bio-Sciences Corp., Piscataway, NJ) fluorescence detected by a Storm 860 (GE Healthcare Bio-Sciences

Corp.). UV measurements were obtained with a Hewlett-Packard 8453 spectrophotometer (Palo Alto, CA).

2.2.3 Subcloning of IYD in *E. coli* expression vectors

The mouse IYD gene (GenBank accession number: BC023358) was cloned into respective expression vectors and transformed into Rosetta2 *E. coli*. Many of these vectors included fusion proteins (glutathione transferase, NusA, thioredoxin, SUMO) to aid in folding of mouse IYD.

A truncated (2-33) and his-tagged (His_6) variant of the mouse IYD gene (IYD(Δtm)) and the gene with two cysteine to alanine mutations (C217A and C239A (IYD(Δtm)DM)) were amplified from the pcDNA3.1(+) plasmid constructed previously (*18*) using primers 5'-AAGCTTAAGCTTGGATCCGCCACCATG GCTCAAGTTCAGCCC-3' and 5'-CTCGAGCTCGAGCTAATGGTGATGGTG ATGGTGTACTGTCACCATGATC-3'. The resulting PCR product was digested with BamHI and XhoI. The genes were individually ligated to various plasmids with T4 ligase and used to transform OneShot Top10 cells. DNA was extracted from colonies resistant to the appropriate antibiotics and sequenced to confirm the presence of the proper insert. Plasmids were then used to transform electrocompetent Rosetta2 *E. coli* for expression.

Cloning of IYD into pET21a

The vector pET21a and PCR products IYD(Δtm) or IYD(Δtm)DM were digested with BamHI and XhoI. The resulting two fragments (linear pET21a and one gene) were ligated with T4 ligase overnight at 16 °C and transformed into Top10 *E.*

coli. The gene insertion product was confirmed by restriction mapping with PstI and PflmI and DNA sequencing.

Cloning of IYD into pET32a

The vector pET32a and PCR products IYD(Δtm) or IYD(Δtm)DM were digested with BamHI and XhoI. Linear vector and genes were ligated and transformed. Insertion of the gene in pET32a was confirmed by digestion with PstI and double digestion with BamHI and StuI as well as DNA sequencing.

Cloning of IYD into pGEX4T-1

Digestion of plasmids pET32a-IYD(Δtm), pET32a-IYD(Δtm)DM, and pGEX4T-1 by BamHI and XhoI was performed. The genes IYD(Δtm) and IYD(Δtm)DM were ligated to linear pGEX4T-1 and transformed in to OneShot Top10 *E. coli*. Proper insertion of the gene was confirmed by digestion with PstI and double digestion with BamHI and StyI.

Cloning of IYD into pET28-SUMO

The pET28-SUMO vector and pET32a-IYD(Δtm)DM were digested with BamHI and XhoI. The IYD(Δtm) gene was ligated to linear pET28-SUMO and transformed into OneShot Top10 *E. coli*. Proper insertion was confirmed by sequencing with the T7 promoter primer.

Mutagenesis of polyhistidine tags

After construction of the pET32a-IYD(Δtm)DM plasmid expression of the thioredoxin fusion protein TRX-IYD(Δtm)DM would contain two polyhistidine affinity tags. The pET32a-IYD(Δtm)DM vector codes for two His_6: one is inherent in

the commercial pET32 plasmid at the C-terminus of the fusion protein thioredoxin and the second His_6 is at the C-terminus of the mouse IYD gene. After proteolysis with enterokinase, both the thioredoxin fusion and IYD(Δtm)DM would contain a His_6. Separation of the mixture of these two in solution is no longer trivial. Both of the 18 base pair hexahistidine sequences were mutated in order to ease purification of thioredoxin from IYD. Two subsequent site-directed mutagenesis reactions were run requiring mutation of four DNA bases causing a change in three amino acids to remove the His_6 which was C-terminal of TRX. The first PCR reaction of site-directed mutagenesis used forward primer 5'-GCCGGTTCTGGCCATATGCACAA TATTCATCTTCATTCTTCTGGTCTG-3' and its reverse complement containing a SspI site to change the second, third, and fifth His to Asn, Ile, and Leu, respectively. The second PCR reaction required forward primer 5'- GCCGGTTC TGGCCATATGTACAATATGCAGCTTAATTCTTCTGGTCTG-3' containing a BsrGI site and its reverse complement creating the final amino acid sequence Tyr-Met-Asn-Gln-Leu-Asn. Following the PCR reaction, the solution was treated with DpnI for four hours at 37°C to remove the template plasmid. The remaining plasmids were transformed into One Shot Top10 *E. coli*. Restriction digestion and DNA sequencing confirmed each round PCR mutagenesis followed by expression in Rosetta 2 *E. coli*.

Removal of C-terminal polyhistidine tag

The C-terminal polyhistidine tag was removed by site directed mutagenesis. The forward primer 5'- CTGGACCAGATCATGGTGACCGTA**TAG**CATCACC ATCACCATTAG-3' and reverse complement changed two DNA bases to include a

BstEII restriction site and a stop codon (in bold) immediately following the gene. Following the PCR reaction, the solution was treated with DpnI for four hours at 37 °C to remove the template plasmid. The remaining plasmids were transformed into OneShot Top10 *E. coli*. Restriction digestion with BstEII and DNA sequencing confirmed mutagenesis prior to expression in Rosetta2 *E. coli*.

2.2.4 Expression in *E. coli*

For expression in *E. coli*, Rosetta2 cells containing a vector with IYD gene were grown in LB media with the appropriate antibiotic at 37 °C to an OD_{600} of 0.7. The cells were cooled to 18 °C (30 min), induced by addition of 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and incubated with shaking for 4 hrs (18 °C). Cells were harvested by centrifugation at 3,000 x *g* for 5 min. The cell pellet was resuspended in 500 mM sodium chloride, 50 mM sodium phosphate (pH 8), 10 mM imidazole and lysed by three passages through a French press at approximately 10,000 psi. Lysates were centrifuged at 20,000 x *g* for one hour and the supernatant filtered (0.22 μm). The insoluble protein pellet was resuspended to a volume equal to that of the supernatant.

Attempts to aid protein folding in vivo

Rosetta2 cells containing pET21a-IYD(Δtm) were grown in LB at 37 °C with shaking. Final concentrations of 5 mM betaine, 0.5 M NaCl, or 10 mM benzyl alcohol were added to cells at OD_{600} = 0.5 (*46, 47*). The flasks were allowed to shake at 18 °C for twenty minutes before the addition of 0.1 – 1 mM IPTG for four hour induction.

Another method to aid solubility is to heat shock the cells prior to induction (*48*). This was tested by growing Rosetta2 cells containing pET21a-IYD(Δtm) to an OD_{600} = 0.5 then moved to a 42 °C water bath for 30 minutes. The cells were then induced with 0.5 mM IPTG and incubated at 18 °C for four hours (shaking). Cells were harvested by centrifugation at 5000 x *g* for 5 minutes.

Refolding studies were conducted on IYD(Δtm)DM expressed from pET21a in Rosetta2 *E. coli*. The pellet containing improperly folded IYD was resuspended in 30 mL denaturing lysis buffer (6 M guanidine HCl, 20 mM sodium phosphate (pH 7.8), 500 mM NaCl). This was centrifuged to remove any remaining insoluble proteins. The solublized protein was dialyzed overnight at 4 °C in three separate buffers (Buffer A: 50 mM Tris (pH 8.2), 21 mM NaCl, 0.88 mM KCl, 0.015 mM FMN; Buffer B: 50 mM sodium phosphate (pH 7.4), 100 mM KCl, 0.015 mM FMN; Buffer C: 50 mM Tris (pH 7.4), 100 mM KCl, 2 mM PEG-4000, 0.015 mM FMN) at two different concentrations of protein. The dialyzed protein was centrifuged for thirty minutes at 18,000 × *g* at 4 °C. The soluble and insoluble proteins were analyzed by gel electrophoresis.

2.2.5 Purification of IYD(Δtm)DM

TRX-IYD(Δtm)DM was purified on a Hi-Trap HP column (1 mL) chelated with Ni^{2+} using an AKTA FPLC (GE Healthcare Bio-Sciences Corp.). Soluble lysates were applied to the affinity column, washed with 5 column volumes of wash buffer (500 mM sodium chloride, 50 mM sodium phosphate pH 8.0, 20 mM imidazole) and eluted with a linear gradient of 20-300 mM imidazole in wash buffer.

The pooled fractions were dialyzed against 0.2 M Tris-HCl pH 7.4. Enterokinase was then added to the dialyzed protein, and the resulting solution was incubated overnight at 18 °C. IYD(Δtm)DM was separated from TRX by anion exchange (Mono Q, GE Healthcare Bio-Sciences Corp) using a wash of 5 column volumes of 0.2 M Tris-HCl pH 7.4 and elution with a linear gradient of 0-1 M NaCl in 0.2 M Tris-HCl (pH 7.4). Fractions containing the deiodinase were identified by SDS-PAGE and then pooled and dialyzed against 10 mM potassium phosphate pH 7.4.

The concentration of enzyme-bound FMN was determined by A_{450} (ε_{450} 12,500 M^{-1} cm^{-1}) (*49*) and the concentration of deiodinase was likewise determined by A_{280} using an extinction coefficient of 38,000 M^{-1} cm^{-1} calculated by ProtParam (*50*). The contribution at 280 nm from FMN was subtracted from the total absorbance at 280 nm to calculate the protein concentration.

2.3 Results and Disucssion

Efforts to induce soluble protein expression in *E. coli* resulted in many plasmids with different fusion proteins. Only two of the plasmids resulted in soluble protein that was then used to express and purify IYD. A list of each plasmid constructed is provided in Appendix A.

2.3.1 IYD expressed as insoluble protein in *E. coli*

Expression of non-fusion IYD(Δtm) and its double mutant from pET21a at 18 °C yielded insoluble protein (Figure 2.1). The addition of salt, betaine, and benzyl alcohol to the growth media did not have considerable affect on the solubility of

IYD(Δtm) (figure not shown). However, heat shock treatment (42 °C) slightly increased the amount of soluble IYD(Δtm) seen in Figure 2.1. This serves as a good indication that heating cells to 42 °C prior to induction with IPTG could work in conjunction with another method that also increased the amount of soluble protein produce *in vivo*.

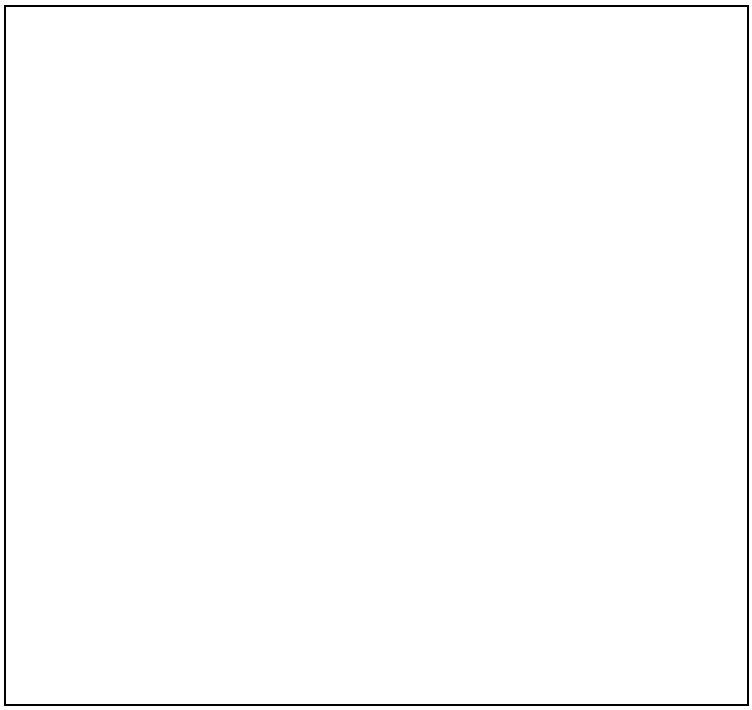


Figure 2.1 Denaturing PAGE analysis of IYD(Δtm) (~30 kD) expression from pET21 with and without heat shock prior to induction by IPTG. Cells were incubated at either 18 °C or 42 °C for 30 minutes prior to induction. Lanes are labeled for molecular weight standards (M), untransfected cells (Un), whole cell lysate (W), soluble (S) and pellet (P) after lysis and centrifugation.

Refolding of IYD(Δtm)DM from pET21a was accomplished by dialyzing two protein concentrations (4 μM and 20 μM) into either potassium phosphate or Tris buffers, at various pH, with different salts, and excess FMN. None of the six samples show significant amount of soluble protein (Figure 2.2). Obtaining soluble protein by refolding with these buffers was not possible.

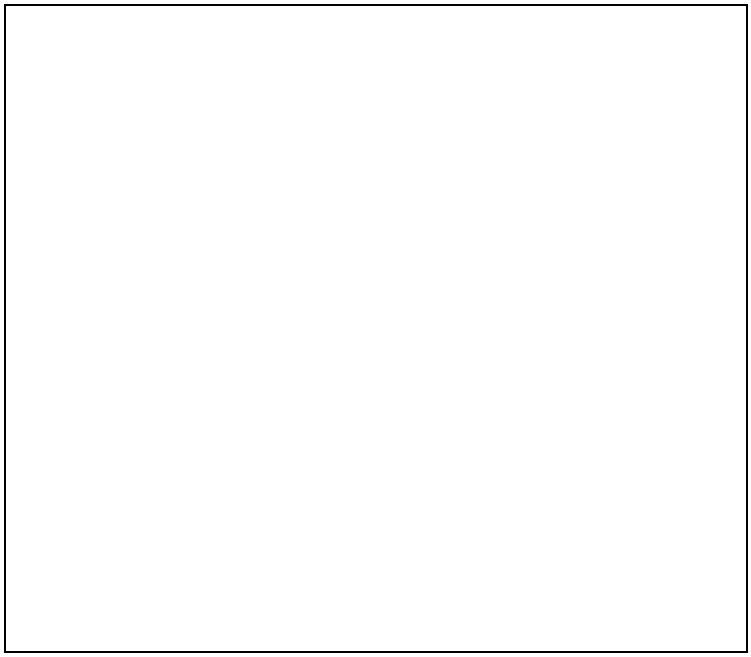


Figure 2.2 Denaturing PAGE analysis of refolding of IYD(Δtm) (high: 20 μM and low 4 μM) expressed from pET21 in three buffer conditions (Buffer A: 50 mM Tris (pH 8.2), 21 mM NaCl, 0.88 mM KCl, 0.015 mM FMN; Buffer B: 50 mM sodium phosphate (pH 7.4), 100 mM KCl, 0.015 mM FMN; Buffer C: 50 mM Tris (pH 7.4), 100 mM KCl, 2 mM PEG-4000, 0.015 mM FMN) Lanes include marker (M), soluble (S) and pellet (P).

IYD(Δtm) and IYD(Δtm)DM genes were cloned into the pGEX4T-1 vector and expressed in Rosetta2 cells. As seen in Figure 2.3, both constructs (IYD(Δtm) and IYD(Δtm)DM) were expressed as insoluble protein. This fusion protein was expressed concurrently as vectors with other fusion proteins were expressed. The remaining constructs were evaluated before further conditions were attempted to aid *in vivo* folding.

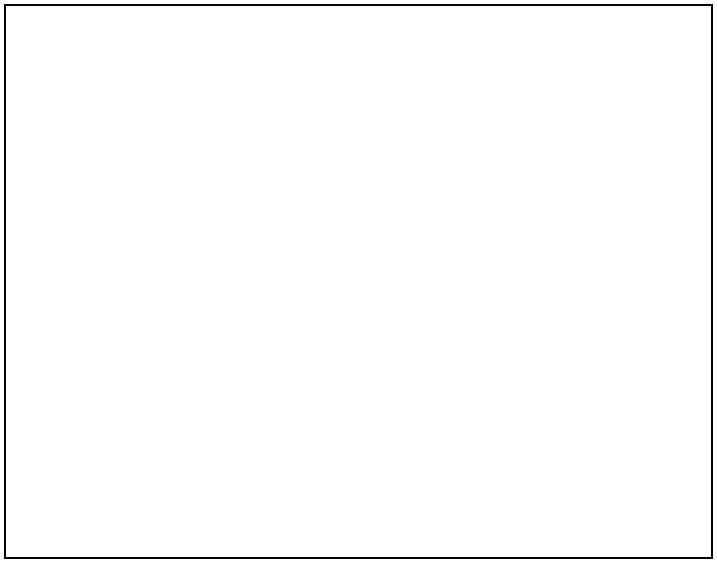


Figure 2.3 Denaturing PAGE analysis of pGEX4T-1-IYD(Δtm) expressed in Rosetta2 *E. coli.* Expressed protein is indicated by the arrow at ~50 kD. Lanes include marker (M), soluble supernatant (S), and pellet (P) after lysis and centrifugation.

2.3.2 IYD(Δtm)DM expressed as soluble protein in *E. coli*

After various attempts of expressing IYD(Δtm) in *E. coli* using fusion proteins provided here and from previous studies (*38*), the only soluble constructs were using the pET32a vector with thioredoxin (TRX) fusion and pET28-SUMO with SUMO fusion (Figure 2.4). Each plasmid expresses a fusion protein with a polyhistidine sequence, a proteolytic cleavage site, and IYD(Δtm)DM followed by a C-terminal polyhistidine sequence. Soluble protein (TRX-IYD(Δtm)DM and SUMO-IYD(Δtm)DM) can be found after expression at 18 °C in the lanes labeled in lane S in Figure 2.5. Following proteolysis to remove the fusion protein, His_6 is present on both the fusion protein and IYD(Δtm)DM.

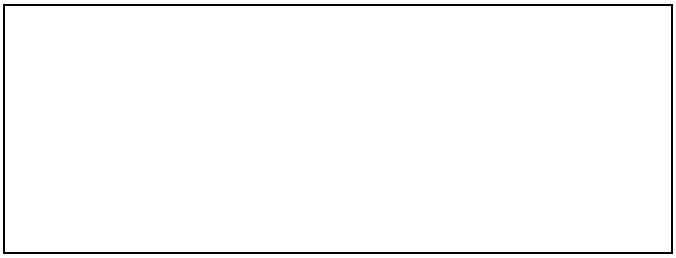


Figure 2.4 Schematic of fusion protein and IYD(Δtm)DM. (A) Thioredoxin fusion protein expressed from pET32a and (B) SUMO fusion protein expressed from pET-28-SUMO. Abbreviations indicate TRX: thioredoxin gene, SUMO: psmt3 gene, His_6: hexa-histidine sequence, Ent: enterokinase protease site, and Ulp: SUMO protease site.

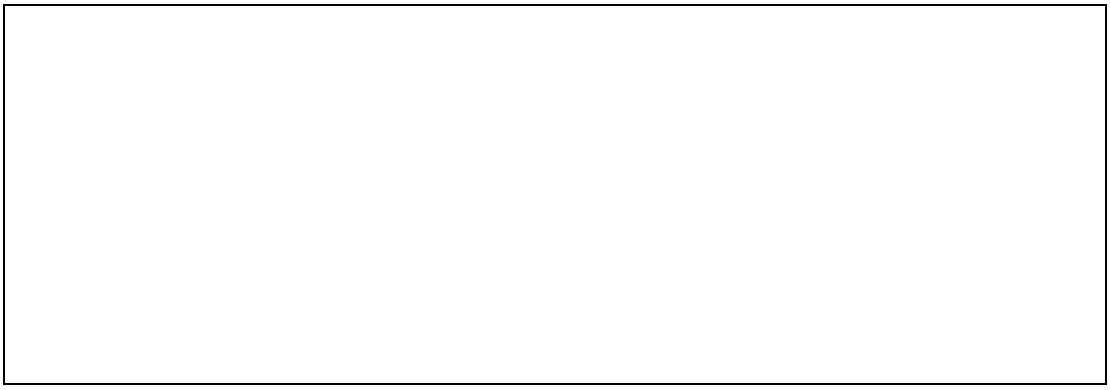


Figure 2.5 Denaturing PAGE analysis of fusion-IYD(Δtm)DM expressed as soluble protein (A) from pET32a and (B) from pET28-SUMO in Rosetta2 *E. coli.* Lanes include marker (M), whole cell lysate (W), soluble (S) and pellet (P) after lysis and centrifugation.

Since His_6 is present on both of the cleaved proteins, a Hi-Trap chelating column is not an advantage to separate the mixture of IYD and the fusion protein in solution. The removal of one of the two His_6 sequences would ease the separation of IYD from the fusion protein. The fusion protein has two His_6 tags and it is possible that one of the two or both had affinity for the column so the initial Hi-Trap purification protocol would be preserved. Unfortunately, removal of either

polyhistidine tag in expression of pET32a-IYD(Δtm)DM resulted in insoluble protein. The insoluble IYD(Δtm)DM protein without the C-terminal His_6 can be seen in Figure 2.6. The protein expressed without the intermediate His_6 linking thioredoxin with IYD(Δtm)DM can be found in the pellet in Figure 2.7. Neither construct produced soluble protein when one of the two His_6 was removed. Affinity of IYD(Δtm)DM to a Ni^{+2} column will not be the ideal way to complete its purification because the fusion protein will also have affinity to the column.

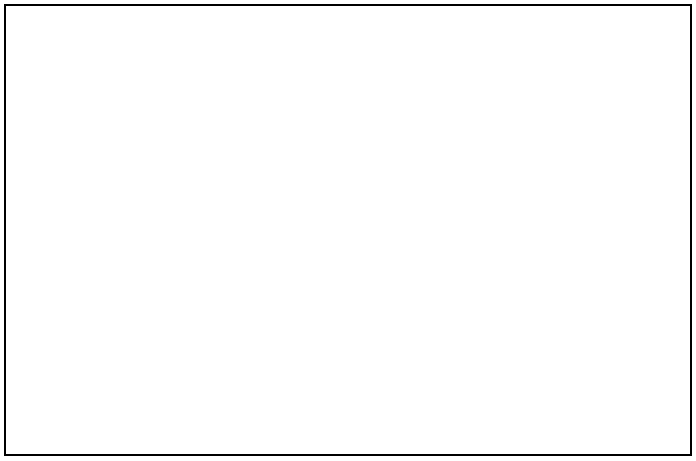


Figure 2.6 Denaturing PAGE analysis of TRX-IYD(Δtm)DM expressed in pET32a without a C-terminal His_6. Protein is indicated by an arrow at ~50 kD. Lanes include whole cell lysate (W), soluble (S) and insoluble (P) proteins after lysis and centrifugation, and marker (M).

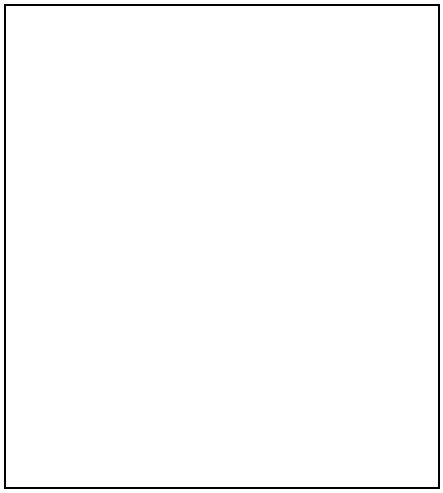


Figure 2.7 Denaturing PAGE analysis of TRX-IYD(Δtm)DM expressed in pET32a without the intermediate His_6. Fusion protein is indicated by the arrow at ~50 kD. Lanes represent marker (M), whole cell lysate (W), and soluble (S) and insoluble (P) after lysis and centrifugation.

Since each of the fusion proteins (TRX and SUMO) and IYD(Δtm)DM have a His_6 sequence, a HiTrap Chelating column can no longer support purification. Other protein purification methods are available including anion or cation exchange based on overall charge (pI) and gel filtration based on size. In order to best assess which fusion protein to continue to use for production of IYD(Δtm)DM, we can consider the chemical differences between the fusion proteins in comparison to our protein of interest. The two fusion proteins are both dimers of about 30 kD and IYD(Δtm)DM is a dimer that is approximately 60 kD. Gel filtration can separate either fusion protein from IYD(Δtm)DM. Separating proteins of these sizes will require a long column and a slow flow rate. This process can be long requiring a lot of time on a machine that is in high demand. The major difference in the two fusion proteins is their pI (Table 2-1). The charge difference between thioredoxin and IYD(Δtm)DM allows for separation on the MonoQ anion exchange column whereas the similarities in pI of SUMO and IYD(Δtm)DM would cause the two proteins to elute together. Therefore,

separate IYD(Δtm)DM from the mixture with its fusion protein, the thioredoxin fusion system (pET32a) was selected for production of IYD(Δtm)DM.

Table 2-1. pI of proteins calculated by ProtParam(*51*).

Protein	pI
IYD(Δtm)DM His_6	6.1
Thioredoxin His_6	5.4
His_6 SUMO	6.2

Isolation of the desired IYD(Δtm)DM required two chromatographic steps. First, its fusion protein was isolated by affinity of the His_6 sequence. TRX-IYD(Δtm)DMHis_6 eluted off a HiTrap Chelating HP column with imidazole over a range of 10 fractions with two distinct peaks which contained the fusion protein as evidence by a band at 50 kD on SDS-PAGE. The latter of these peaks was yellow indicating bound FMN while the other was clear. Since both TRX and IYD(Δtm)DM are dimers, it is possible that the first protein that eluted (peak that was not yellow) was the fraction that had TRX as a dimer. This would mean that IYD(Δtm)DM was in monomeric form and did not have a bound FMN. This theory is also supported by the fact that the intermediate His_6 between TRX and IYD(Δtm)DM would perhaps have less affinity to the column than a C-terminal His_6. The pooled fractions were simultaneously dialyzed and digested with enterokinase. IYD(Δtm)DM His_6 was next separated from TRX His_6 using a MonoQ anion exchange column to yield over 10 mg enzyme / L of culture with an acceptable purity of greater than 80%. This preparation

included a small fraction (27 %) of deiodinase that had been further truncated by enterokinase (lane MonoQ in Figure 2.8). This would arise from cleavage at the *N*-terminus since the *C*-terminal His_6 was still detected by Western blotting (Figure 2.8B). Collectively, the pooled fractions of deiodinase contained 1.9 FMN per native enzyme dimer.

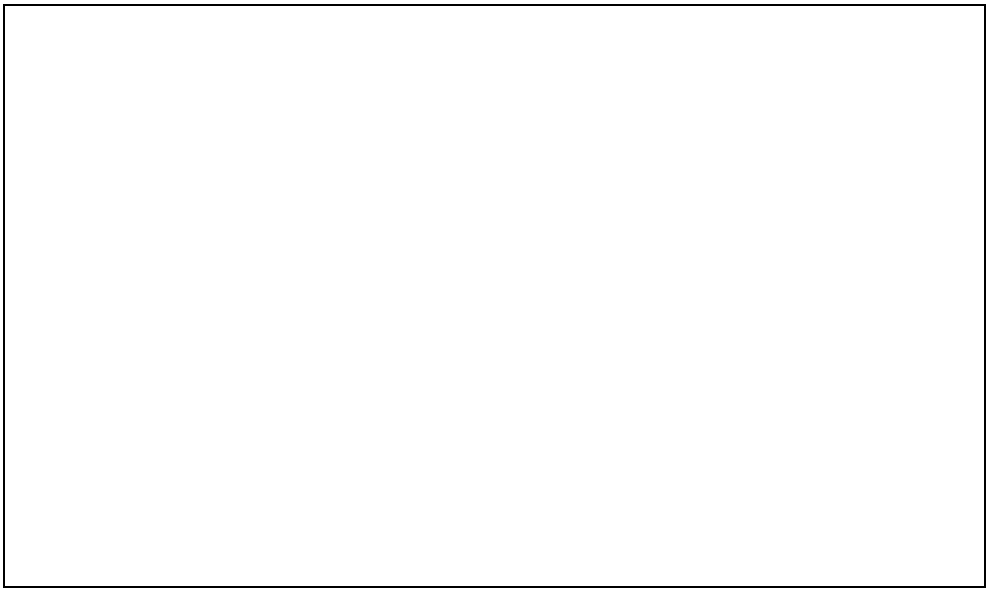


Figure 2.8 IYD(Δtm)DM purified from pET32a in Rosetta2 *E. coli.* (A) Denaturing PAGE analysis of TRX-IYD(Δtm)DM expressed in pET32a. Lanes identify the marker (M), whole cell lysate (W), soluble and insoluble proteins in the pellet after lysis and centrifugation (S and P, respectively). Ni^{+2} and MonoQ indicate proteins collected after HiTrap and MonoQ columns. (B) Western with anti-His antibodies after purification by MonoQ.

2.4 Conclusions

Two cysteines are present in the NADH oxidase/flavin reductase domain of all mammalian iodotyrosine deiodinases characterized to date. The native oxidation state of these residues was not initially known and disulfide bonding was considered a

possible cause for misfolding and inclusion body formation when various constructs of this enzyme were first expressed in *E. coli* (*38*). Contrary to our original expectation, a C217A; C239A double mutant (DM) of the deiodinase retained significant catalytic activity (*18*). Since these two cysteines that perhaps were causing misfolding are not necessary, this gene could then perhaps provide soluble protein when expressed in *E. coli*. Thus, expression of the enzyme was again attempted in *E. coli* using various fusion proteins with IYD(Δtm)DM. Ultimately, soluble protein was expressed from a gene encoding an *N*-terminal thioredoxin fusion and polyhistidine tags at the *C*-terminus and in the linker region between the TRX and IYD(Δtm)DM gene (Figure 2.7A). Use of other fusion proteins in place of TRX, removal of either His_6 sequence, or lack of cysteine to alanine mutation resulted in expression of only insoluble protein.

Chapter 3: Characterization of IYD

3.1 Introduction

Initial reports of iodotyrosine deiodinase describe extraction and purification of the enzyme from bovine thyroids (*36*). From the small amounts of homogeneous protein (0.75 mg) recovered from thyroid tissue (1 kg), the authors began biochemical characterization (*23, 24*). This allowed for initial examination including molecular weight, flavin content, pH dependence of catalysis, specific activity, and reduction potentials of oxidized and reduced enzyme. Complete characterization of IYD requires large amounts of pure protein. Once the DNA sequence of IYD was discovered, the gene was manipulated for expression in HEK 293, *Pichia pastoris*, Sf9, and *E. coli* (*40*). DNA technologies allow for DNA manipulation to create recombinant proteins.

Presented here is the first time that the IYD gene was engineered for expression in *E. coli* after years of different attempts (*38, 40*). Heterologous expression in Sf9 and *P. pastoris* supported expression of significant amounts (mg) of soluble IYD (*40*). Although these cell lines offered enough protein for initial characterization, they are inconvenient for studies using site directed mutagenesis. These expression systems are time consuming and not amenable to mutagenesis to further study the protein. Soluble expression in *E. coli* was the key step to create mutants. In order to do this, however, we first must validate that the mutant protein expressed in *E. coli* in Chapter 2 is enzymatically comparable to the wild-type

counterpart expressed in human, yeast, and insect cell lines. Specific activity, substrate binding, and the structures are all compared herein.

The wild-type enzyme without its membrane domain (IYD(Δtm)) expressed in Sf9 cells was crystallized. The first holo-enzyme structure was solved by sulfur phasing and additional co-crystal structures were solved by molecular replacement (*31*). The structure confirmed the placement of IYD in the NADPH oxidase/flavin reductase structural superfamily. The enzyme forms a homodimer with two individual active sites which each bind a molecule FMN and substrate.

Pure enzyme allows for calculations of accurate enzyme kinetics and substrate binding calculations. Michaelis-Menten kinetic constants were obtained for IYD(Δtm) through the use of a discontinuous assay measuring the release of radioactive iodide from DIT. Equilibrium binding of substrate to IYD(Δtm) to calculate a binding constant was performed via the inherent fluorescence of FMN bound to the protein. The values for IYD(Δtm) are the basis for understanding the enzyme and are used as a comparison for the validity of specific mutations.

It is especially important that we observe wild-type-like enzymatic features for this double cysteine mutant IYD(Δtm), noted as IYD(Δtm)DM. Expression in *E. coli* is only a useful tool if the protein can be used as a model of wild-type enzyme to make mutations to study mechanistic properties. Therefore, IYD(Δtm)DM must display similar enzymatic parameters and similar structure to IYD(Δtm) in order to make additional mutations to IYD(Δtm)DM to investigate the biochemical properties of the protein.

3.2 Experimental

3.2.1 Materials

Protein was expressed and purified as previously described (Chapter 2). UV measurements were obtained with a Hewlett-Packard 8453 spectrophotometer (Palo Alto, CA). Circular dichroism spectra were collected with a Jasco J-810 spectropolarimeter (Easton, MD). A Hitachi F-4500 fluorescence spectrophotometer was used to collect fluorescence spectra (Schaumburg, IL). ^{125}I used for radiolabeling was purchased from Perkin Elmer (Waltham, MA). Radiation was counted on a Packard 1600 TR liquid scintillation counter.

3.2.2 Circular Dichroism

CD spectra were taken of IYD(Δtm) and IYD(Δtm)DM. Ten measurements were made of 300 μl protein in 10 mM potassium phosphate (pH 7.4) with and without 300 μM MIT in a 0.1 cm cuvette. Spectra were collected at two protein concentrations (4 μM and 60 μM) for the far-UV (190 nm – 300 nm) and visible wavelengths (300 nm – 500 nm), respectively. Spectra were averaged and buffer contributions were subtracted from data.

3.2.3 Fluorescence quenching

Substrate dissociation constants were calculated based on the change in fluorescence of flavin bound to protein upon titration of substrate. Fluorescence measurements were recorded at λ_{em} 525 nm (λ_{ex} 450 nm) two minutes after each addition of MIT (0-300 μM). Data was normalized by dividing the observed

fluorescence from the initial fluorescence intensity and plotted against substrate concentration. Dissociation constants (K_D) were calculated according to Equation 3-1 in Origin 7.0 (*52*).

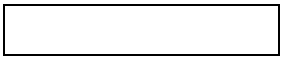
Equation 3-1

3.2.4 Deiodinase activity

Deiodination of ^{125}I radiolabeled diiodotyrosine by IYD(Δtm)DM was measured by a discontinuous assay (*19, 53*). Briefly, solutions (radiolabeled diiodotyrosine (20,000 cpm), 100 mM potassium phosphate (pH 7.4), 200 mM KCl, 50 mM β-mercaptoethanol, 0.50 mM methimidazole, and 30 μM FMN) containing enzyme. Reactions were initiated by dithionite and incubated for 30 minutes at ambient temperature with unlabeled DIT (0 – 300 μM). The reactions were loaded on a cation exchange column and eluted with two column volumes of 10 % (v/v) acetic acid. A sample of the reaction solution prior to separation (dpm1), as well as the two collected fractions collected after ion exchange (dpm2 and dpm3, each one column volume) were counted. The rate (V) was calculated by subtracting the ratio of eluted ^{125}I (F) from the background radiation (F_0) as seen in equations 3-2 and 3-3, where t is the time (hours) after initiation and n is the moles of DIT (nmol). The value is multiplied by a coefficient of 2 because two deiodination sites exist on the DIT substrate used in the assay but statistically only one of the two is labeled with ^{125}I.

Kinetic constants were calculated by fitting the data of three independent assays to the Michaelis-Menten equation in Origin 7.0.

Equation 3-2



Equation 3-3

3.2.5 Crystallization of IYD(Δtm)DM

IYD(Δtm)DM crystallized using the hanging drop diffusion method at one part enzyme (10 mg/ml; 10 mM potassium phosphate, pH 7.4) to one part reservoir solution containing 0.2 M ammonium acetate, 0.1 M BisTris (pH 6.5), and 45% v/v 2-methyl-2,4-pentanediol. IYD(Δtm)DM•MIT co-crystals formed at 20 °C in 15% w/v PEG 10,000, 20% glycerol, 0.1 M citric acid (pH 5.5). Crystal diffraction data was collected using a Brucker Microstar H2 generator with a Proteum Pt135 CCD detector at 100 K at a wavelength of 1.54178 Å and 0.9795 Å for the protein complex with MIT. Data were integrated and scaled using Proteum. Molecular replacement was performed by PHASER in the CCP4 program suite. The structures were refined by iterations in COOT. Refinement statistics are in Table 3-1.

Table 3-1. Crystallographic parameters of IYD(Δtm)DM and its co-crystal with MIT.

	IYD(Δtm)DM	IYD(Δtm)DM • MIT
PDB Code:	3TO0	3TNZ
Data Collection		
Space group	*P*31	*P*3
Cell dimesions		
a, b, c (Å)	87.270, 87.270, 62.725	108.98, 108.98, 49.39
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 120.00
Molecules/ Asymmetric Unit	2	2
Wavelength (Å)	1.54187	0.9795
Resolution (Å)	19.88-2.66	50-2.25
R_{sym} (last shell)	0.105	0.15
$I/\sigma I$	73	12.4
completeness (%)	99.71	99.7
Redundancy	3.7	8
Refinement		
Resolution (Å)	28.566-2.655	26.5-2.248
R_{work}/R_{free}(%)	18.23/23.69	17.2/20.0
Number of protein residues per monomer	221/222	221/221
Number non-protein atoms		
Ligand	80	128
Solvent	50	298
Mean B-factors $(Å)^2$	6.14	10.05
RMS deviations		
Bond lengths (Å)	0.015	0.014
Bong angles (°)	1.639	1.384
Ramachandran plot		
Most favorable	94.96%	97.00%
Additionally allowed	4.76%	2.50%
Disallowed	0.28%	0.50%

3.3 Results and Discussion

Earlier studies characterized IYD that was extracted from microsomes, solublized by trypsin digest, and extensively purified. K_M values, calculated by double reciprocal plots, were determined to be 2.5 μM for DIT (*24*). Enzyme turnover was approximately 10 min^{-1} for DIT (*24*). It was in this study that the addition of thiols (50 mM β-mercaptoethanol) and high ionic strength (200 mM KCl) were noted to increase the observed enzyme activity (*24*). Subsequent studies noted the reduction of the enzyme by stoichiometric amounts of dithionite and enzyme could not be activated by photoreduction with EDTA (*23*). Oxidation-reduction potentials were calculated at -412 mV for the oxidized and -190 mV for the reduced enzyme (*23*). These values are impressively similar to our findings despite the authors' now archaic protein purification methods. This can serve as a good model that comparable results are obtained from homogenated protein solubilized by trypsin digest and heterologously expressed protein.

Here, we have biochemically characterized heterologously expressed and purified IYD. The protein was expressed with an N-terminal thioredoxin fusion protein in *E. coli*. Proteolytic digestion and further purification yielded 11 mg/L at an acceptable 80% purity with a 1:1 FMN:IYD ratio (Chapter 2). This protein was used for the experiments described below.

3.3.1 CD of flavoprotein iodotyrosine deiodinase

Circular dichrosim is a technique used to study the secondary structure of proteins. Structural changes induced by denaturation or ligand binding can be

monitored by changes in the spectra. CD spectra in the far-UV region of IYD(Δtm) and IYD(Δtm)DM are shown in Figure 3.1 and Figure 3.2. The α-helical and β-sheet content of each protein as predicted by K2D2 (*54*) (57 % and 8 %,) is within error of the calculated content from the crystal structure (3GB5; 42 % and 9 %). Mutation of the cysteine residues did not alter the amount of α-helical or β-sheet (compare panel A and B in Figure 3.1), as predicted by K2D2. This was expected based on detectable activity from the previous expression of these mutants in HEK 293 cells (*18*).

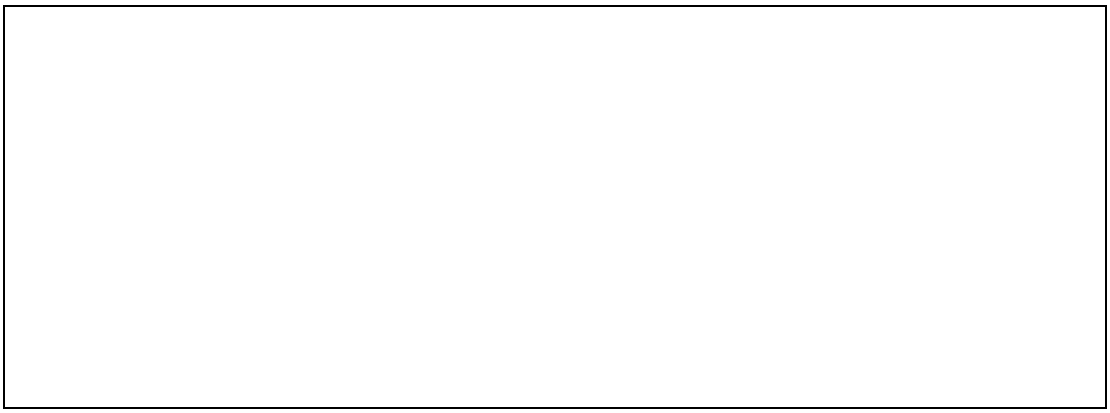


Figure 3.1 Far-UV CD of IYD(Δtm). Protein (4.5 μM) with (red) and without (black) 300 μM MIT. Protein was expressed in A) *E. coli* (C217A, C239A) and B) *P. pastoris*.

CD spectra collected in the visible region presents useful information about the flavin environment. Each flavoprotein has distinct signature CD spectra in the visible range (300-500 nm) based on the amino acids and secondary structure surrounding the flavin and oxidation state of the flavin (*55*). Protein for the visible spectra was 15-fold more concentrated than the far-UV data in order to detect significant change in ellipticity. IYD(Δtm) has positive peaks at 340 and 450 nm. The same peaks are observed in both IYD(Δtm) and IYD(Δtm)DM. From this, we can

conclude that the flavin environment is the same in IYD(Δtm) and IYD(Δtm)DM; same orientation, oxidation state, exposure to solvent (Figure 3.2, black lines in panel A and B). Every flavoprotein has a distinct CD spectra depending on the amino acids surrounding the flavin (*56*). Seeing these two proteins have the exact same CD spectra indicate that they have the same amino acids surrounding and stabilizing the bound flavin.

More interestingly, upon the addition of saturating substrate, a significant change in ellipticity is observed. A new negative peak is observed at 350 nm in the presence of substrate. This is thus a change in the flavin environment, perhaps due to the stacking of the MIT aromatic ring over the isoallaxozine ring of FMN.

The similarity in visible region CD spectra when comparing IYD(Δtm) and IYD(Δtm)DM designate that the flavin is equivocally bound to the wild-type and mutant proteins. Upon substrate binding to either wild-type or mutant protein, we observe the same shift in the ellipticity in the visible CD spectra. The difference in the CD spectra between samples with and without substrate present indicate that the chemical environment surrounding bound FMN changes upon the addition of substrate. We see the same shift in wild-type and mutant proteins with substrate present indicating that both proteins provide for the same interaction between FMN and MIT.

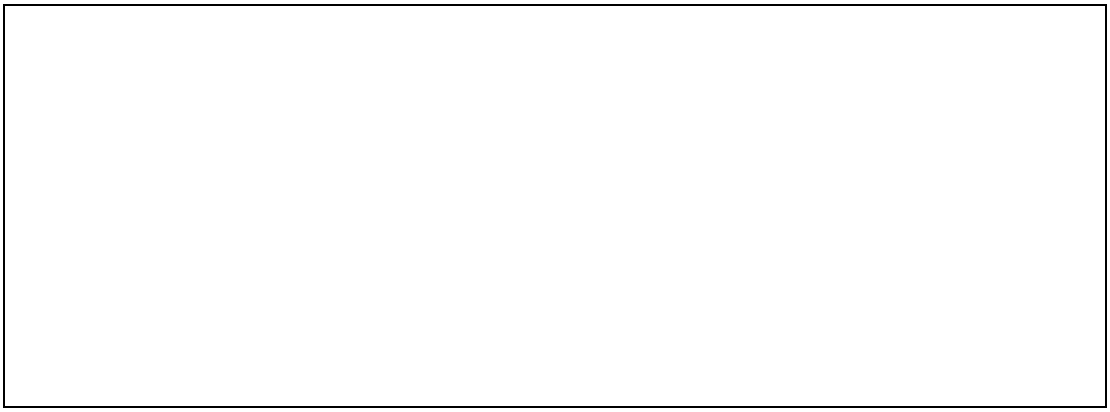


Figure 3.2 CD of visible region of IYD. Protein (60 μM) with (red) and without (black) 300 μM MIT. Protein was expressed in A) E. coli (C217A, C239A) and B) P. pastoris.

3.3.2 Kinetics of IYD(Δtm)DM

Kinetic parameters of heterologously expressed and purified IYD have been previously published. Removal of the transmembrane domain resulted in no change in the k_{cat}/K_M values for reactions initiated with dithionite when expressed in HEK 293 cells (*18*). However, this truncation did remove the ability of NADPH to reduce FMN in vitro (*18*). Expression of the truncated enzyme in Sf9 and *P. pastoris* provided similar kinetic values to each other as well as to the comparable sequence expressed in HEK 293. Previously, the double cysteine mutant was expressed as the full length gene in HEK 293 cells and displayed two fold lower k_{cat}/K_M (*18*).

Deiodination of IYD(Δtm)DM was found to have similar kinetic parameters as the enzymes expressed in HEK293 cells. A comparison can be made of IYD(Δtm) and IYD DM expressed in HEK 293 and Sf9 to IYD(Δtm)DM expressed in *E. coli* as seen in Table 3-2. Regardless of the version of the protein, IYD has very similar kinetic constants. IYD DM expressed in HEK 293 and IYD(Δtm)DM expressed in *E.*

coli have virtually the same K_M and k_{cat}. This is further indication that IYD(Δtm)DM will be a useful model to use for further mechanistic studies.

Table 3-2. Catalytic properties of iodotyrosine deiodinase (Mus musculus) derivatives.

		DIT[a]			MIT
Source	Enzyme	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (min^{-1} μM^{-1})	K_D (μM)
HEK293	IYD[b]	7.1 ± 0.9	8 ± 3	0.89 ± 0.35	-
HEK293	IYD (Δtm)[b]	5.8 ± 0.6	6 ± 2	0.95 ± 0.33	-
HEK293	IYD DM[b]	16 ± 2	42 ± 7	0.38 ± 0.08	-
Pichia	IYD (Δtm)	6.9 ± 1.3	19 ± 3	0.36 ± 0.09	
Sf9	IYD (Δtm)	4.5 ± 0.7	9 ± 1	0.49 ± 0.09	0.09 ± 0.04[c]
E. coli	IYD (Δtm)DM	9.3 ± 1.6	40 ± 5	0.23 ± 0.05	2 ± 0.2

[a]Kinetic measurements based on deiodination were fit to Michaelis-Menten kinetics. [b]Data from HEK 293 expression were determined previously (*18*). [c]From McTamney and Rokita (*26*).

3.3.3 Equilibrium binding to IYD(Δtm)DM

Dissociation constants of MIT to IYD(Δtm)DM were calculated by Origin 7.0 using titration curve data (Table -2). The two cysteine mutations reduce the affinity of the enzyme for substrate compared to the wild-type enzyme, IYD(Δtm) (Table -2). The assay for the mutated protein was performed at 3x higher concentration than that of wild-type. When the assay was performed at a lower protein concentration (1.5 μM), only 30% decrease in fluorescence was observed. Upon increasing the protein concentration (4.5 μM), eighty percent of the fluorescence was quenched with saturating amounts of substrate. This phenomenon could be due to the structural stability of the mutated enzyme. It is possible that the protein would not be a stable dimer at lower concentrations, that the K_D of dimerization is higher than that of wild-

type. This is not confirmed by experimental data but could be reviewed by sedimentation equilibrium experiments. The protein concentration does not affect the calculation of the substrate binding constant. Regardless of the protein concentration necessary to observe the change in fluorescence, there is a twenty-fold decrease in MIT affinity when comparing IYD(Δtm)DM to IYD(Δtm). This is a change worth noting, but not significant enough to retract the claim that the protein is a good model.

Figure 3.3 Binding titration of MIT to IYD(Δtm)DM . Ligand binding was monitored by the change in fluorescence of the flavin bound to IYD using λ_{ex} of 450 nm and λ_{em} of 527 nm. Solutions of IYD (4.5 μM) in 10 mM potassium phosphate (pH 7.4) at 25 °C were titrated with MIT (0-250 μM). Three independent titrations were recorded and the fluorescence intensities were corrected for the slight dilution caused by the addition of ligand, normalized, and plotted against MIT concentration.

3.3.4 Crystal structure of IYD(Δtm)DM

The crystal structures of IYD(Δtm)DM and its complex with MIT were characterized to ensure that the cysteine to alanine substitutions did not significantly

perturb the enzyme structure. C217 is located at the interface of the homodimer, and C239 is closer to the protein surface but still protected from solvent (Figure 3.4). Removal of the two cysteines did not affect protein crystallization. In fact, IYD(Δtm) and IYD(Δtm)DM both crystallized with 0.2 M ammonium acetate, 0.1 M BisTris (pH 6.5 and 5.5, respectively) and 45 % v/v 2-methyl-2,4-pentanediol. The structures of IYD(Δtm)DM and IYD(Δtm)DM•MIT were solved by molecular replacement and remained very consistent with the parent structures In the absence of MIT (Figure 3.4), the active site appeared very accessible to solvent due to a lack of detectable structure in two surrounding regions of the polypeptide.

Figure 3.4. IYD(Δtm)DM crystal structure. IYD(Δtm)DM (PDB ID: 3TO0) crystallized in the absence of substrate. Monomers are distinguished by color. Disordered regions (156-177 and 195-208) are not shown due to lack of electron density. Cysteines from 3GFD (in gray) are shown as yellow spheres.

In the presence of MIT (Figure 3.5), an active site lid comprised of a helix and loop (residues 156-177) was detected from the otherwise unstructured regions. The

crystal with MIT had electron density for these amino acids unlike the crystal of the holo-enzyme. This indicates that either the active site lid is folded and flexible or that it is unfolded and unstructured. Either way, electron density was not detected until substrate was present. This lid effectively sequesters the substrate-flavin complex from solvent.

The prospective use of *E. coli* expressed IYD(Δtm)DM•MIT as a model of native enzyme was validated by the very low RMSD of 0.262 Å when overlaid with the original IYD(Δtm)•MIT (3GFD) structure (Figure 3.4). The small deviations in folding are not localized but rather distributed throughout the three dimensional structure. If one region was perturbed more than the rest of the structure, IYD(Δtm)DM would not be a good model. Deviation of less than one angstrom across the whole structure is insignificant which is an important trait of IYD(Δtm)DM as a model for IYD(Δtm).

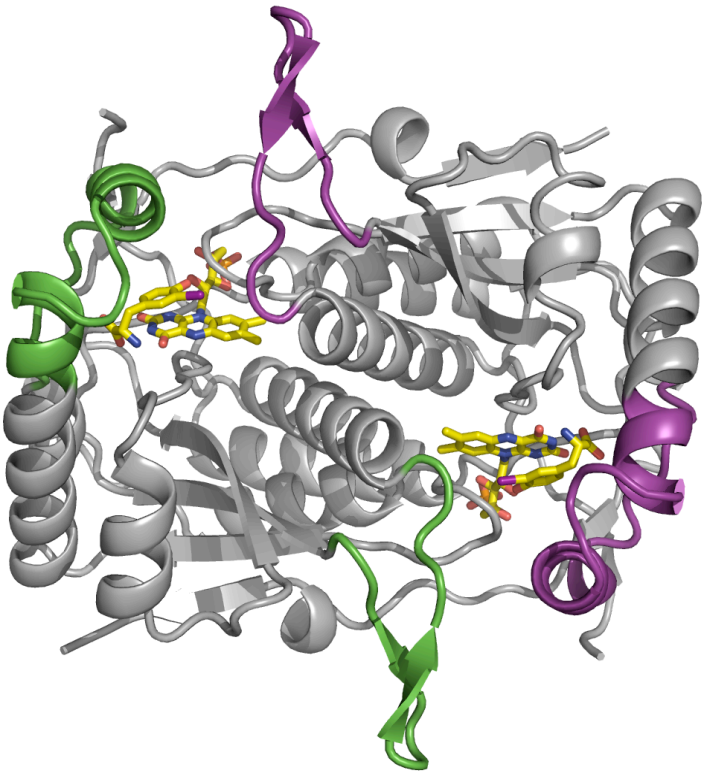


Figure 3.5 Crystal structure of IYD(Δtm)DM bound to MIT. IYD(Δtm)DM (PDB ID: 3TNZ) Structure associated with substrate binding is in purple and green with respect to the monomeric unit. MIT and FMN are in yellow.

Most importantly, the active site characteristics of IYD(Δtm) and its C217A;C239A mutant are essentially identical. Their co-crystals both revealed the same overall contacts between the protein, FMN and MIT (*31*). The aromatic portion of MIT stacks over the isoalloxazine ring of FMN, and its phenolic –OH forms hydrogen bonds to the 2'-hydroxyl group of the FMN ribose and the Ala126 backbone nitrogen. The carbon of MIT bonded to the iodide is only slightly more distant from the C4a of FMN than that detected in the parent enzyme expressed in Sf9 (3.73 Å vs. 3.65 Å, respectively) (*31*). The zwitterionic arm of MIT is bound through a network of polar interactions including hydrogen bonding to the N-3 and O^4 of the

flavin ring and the side chains of three amino acid residues (Glu153, Tyr157, Arg178) as part of the active site lid (Figure 3.6).

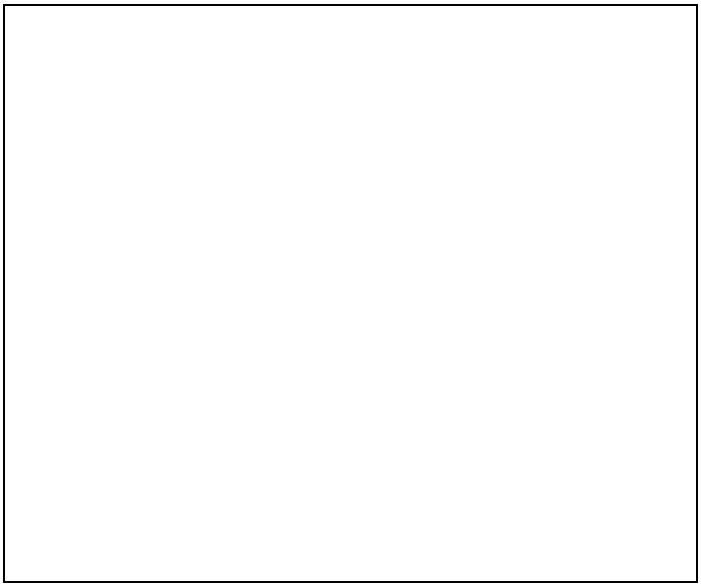


Figure 3.6 IYD(Δtm)DM active site overlayed on IYD(Δtm). MIT interactions with IYD and FMN are shown through polar contacts within 4Å indicated by dashed lines. Each monomer of 3TNZ is colored in green or purple and 3GFD is gray.

3.4 Conclusion

As the similarity in the crystal structure displays, the cysteine-free protein expressed in a soluble form in *E. coli* is an alternative to tedious expression systems necessary to express wild-type enzyme. Despite the slight deviation in kinetics and affinity for substrate, the double mutant (IYD(Δtm)DM) expressed from *E. coli* is an acceptable form of the protein. The protein expressed in *E. coli* can act as an alternative to the Sf9 and yeast expression systems to easily create mutations to study the mechanism of action. It is concluded that the cysteine mutations would allow for subsequent investigations of the catalytic mechanism.

Chapter 4: Substrate coordination to the active site lid

4.1 Introduction

Engineering IYD for expression in *E. coli* not only simplifies its preparation but also expedites its mutagenesis. These advances were applied first to measure the individual contributions of three residues that were expected to be critical for substrate coordination. IYDs unique function within its structural superfamily makes the study of its activity and substrate recognition much more interesting.

Active site lid structures in flavoproteins are common and do not share overall structural similarities or function. Flavoproteins with active site lids act as oxidases (*57*), dehydrogenase (*32*), bioluminescence (*58*), and oxidoreductases (*59*). Specifically in the NADPH oxidase/flavin reductase superfamily, the proteins have an active site lid (*33, 60*). The flavin reductase P (FRP) from *V. harveyi* has a nine residue sequence (201-209) that did not have electron density upon solving the crystal structure (*58*). This region, and specifically two residues R203 and K208, bind NADPH, although no crystal structures of this enzyme have been solved with NADPH as the ligand (*33*). Other proteins in the superfamily of NOX proteins have crystal structures that show electron density in the structure surrounding the active site while others do not.

The dynamic processes of IYD catalysis involve the binding of substrate to the enzyme to form the enzyme-substrate complex, the catalytic deiodination (turnover), and the release of products. The active site lid is involved with each of these steps. Proper orientation of the substrate in the active site pocket is coordinated by the hydrogen bonding contacts between the zwitterionic arm of the substrate and

protein residues on the active site lid (*31*) (Chapter 3). The proximity of the two ligands (MIT and FMN) is critical to forming the catalytically active complex prior to deiodination. A network of hydrogen bonds between side chains on the lid and residues on the stable dimer aids in both formation of the active complex and sequestering the MIT-FMN interaction from solvent. Following catalysis, the products must be released. The active site lid has to be flexible to open away, perhaps like a hinge, to liberate free iodide and tyrosine.

IYD's active site lid is not positioned over the active site without substrate present as indicated by the lack of electron density in IYD(Δtm), PDB ID 3TO0. The active site lid forms a helix-turn-helix (residues 152-178) to sequester the substrate binding pocket from solvent. Not only does the lid have helical structure but the amino acids R100, N156, M161, R165, W165, D168, L169, K171, L237 from one chain and V205, H206, Y207, and Y208 from the other chain form a network of hydrogen bonds to deprive the closed active site of any external (solution) exchanges. None of these amino acids, except R100 and L237, have detectable structure in the holoenzyme crystal structure as indicated by the lack of electron density in 3TO0. Each is a part of what is creates the active site lid. The lack of electron density and formation of a sequestered active site is most prevalently seen in the difference between the right and left surface representations of the crystal structure in Figure 4.1. Specifically, functional groups of only two residues (Y207 to D168) form an H-bond over the active site. These two residues are found on opposite monomers, reinforcing the observation that both monomers are critical for active site formation.

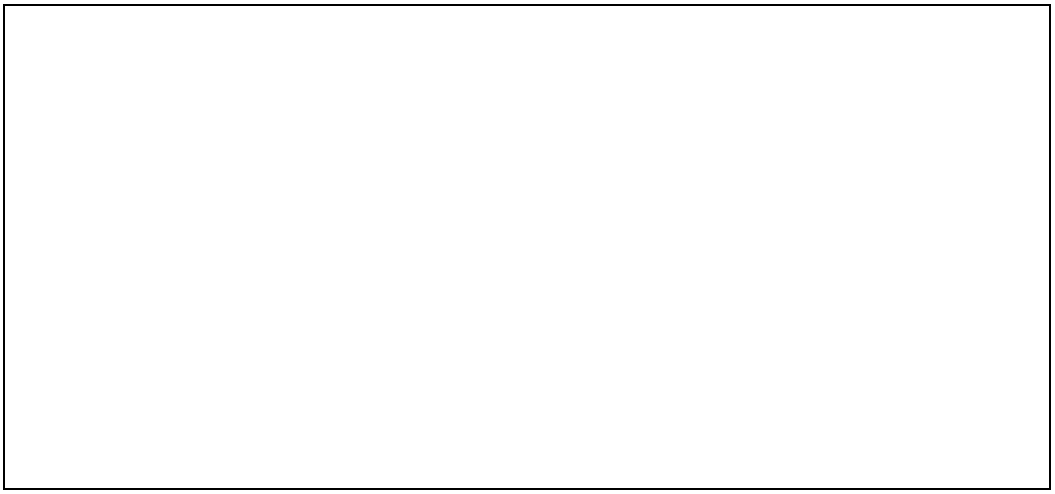


Figure 4.1 The surface properties of IYD(Δtm)DM (left) and its complex with MIT (right) as calculated using vacuum electrostatics in PyMOL (*61*).

The tight pocket precludes any other substrates from fitting into the active site. The addition of the lid to condense over the substrate would direct the placement of the substrate to the FMN and block anything larger than diiodotyrosine in the active site. Members of the NADPH oxidase/flavin reductase structural superfamily bind NADPH (*59*) but it would not fit in the known space of IYDs active site. IYD needs NADPH to reduce FMN *in vivo*, as indicated by the lack of reduction in whole cell lysates in CHO cells compared to reduction in HEK 293 cells (*18*). Prior to the crystal structure, NADPH binding was not predicted for IYD (*25*) and a lack of fluorescence quenching by NADPH confirmed this (*41*). Similarly, a peptide sequence containing an iodinated tyrosine residue was unable to be deiodinated (unpublished work from Rokita lab). Other iodinated tyrosine substrates T3 and T4 also would not fit in IYDs active site (*31*) and are likewise not processed by IYD (*62*). The distinctive function of IYD and its substrate specificity encourage the study active site lid within the superfamily.

The clamping of the active site lid prevents larger substrates from entering the pocket. On the lid, only three amino acids hydrogen bond to the substrate. These residues are E153, Y157, and K178 (Figure 4.2). These three are responsible for orienting the substrate over the flavin to facilitate deiodination. We can discern the role of each of the three amino acids by mutating each to disrupt hydrogen bonding.

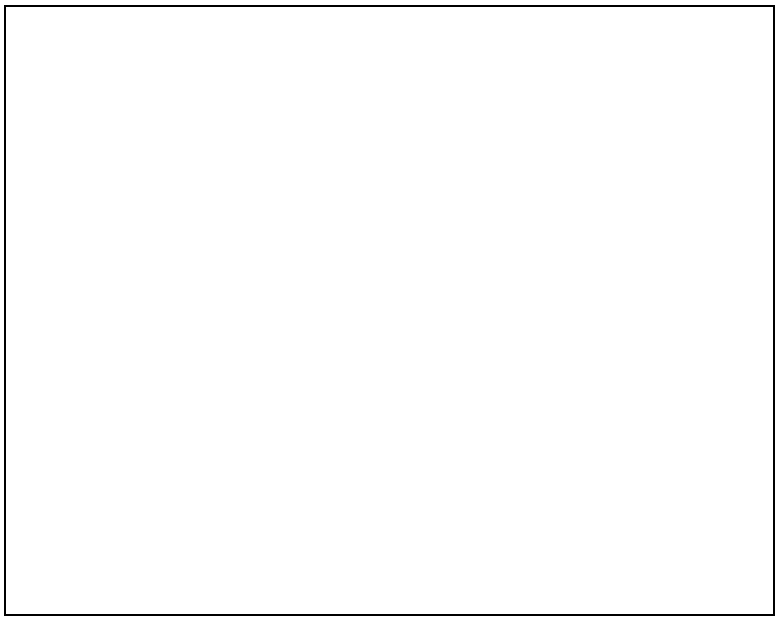


Figure 4.2 The active site of IYD(Δtm)DM. Amino acids that make hydrogen bonding contacts to the MIT substrate are shown in sticks and labeled. The amino acid substrate MIT and cofactor FMN are in yellow. PDB ID: 3TNZ.

Conservative mutations were made to the three residues that make hydrogen-bonding contacts to the zwitterionic portion of the substrate. Glutamate interacts with the amino terminus of the amino acid substrate. To disrupt this interaction, the carboxyl group of glutamate should be sacrificed. Glutamate at 153 was mutated to glutamine, which retains the side chain length and replaces the functional carboxyl group with an amine. The second amino acid, tyrosine at 157, contacts the carboxyl end of the substrate. Tyrosine was mutated to phenylalanine, which removes the hydroxyl functional group on the aromatic ring. Lysine at 178 was mutated to

glutamine to disrupt hydrogen bonding interactions between the amino group of lysine and the carboxyl group of the substrate. Conservative mutations were made to retain the number of carbons but would disrupt the hydrogen bonding ability of the side chain to the zwitterionic substrate. The protein characterized in Chapter 3, IYD(Δtm)DM, serves as the parent enzyme for mutagenesis as it is expressed easily in *E. coli*, which is an ideal system for introducing mutations. It is predicted that each of these mutations will negatively affect substrate binding and catalysis.

4.2 Materials and Methods

4.2.1 General Methods

General methods follow the same protocols described in Chapter 3. Proteins were expressed in Rosetta 2 *E. coli* and purified via two chromatographic steps in the same manner as described in Chapter 2. IYD(Δtm) was expressed in *Pichia* and purified as previously described (*40, 41*). UV measurements were obtained with a Hewlett-Packard 8453 spectrophotometer (Palo Alto, CA). CD measurements were made on a Jasco J-810 spectropolarimeter (Easton, MD). Ten spectra were collected at 20 °C in a 0.1 cm pathlength quartz cuvette and averaged. Ligand binding to IYD was monitored by the change in bound flavin fluorescence using λ_{ex} of 450 nm and λ_{em} of 527 nm as reported previously (*26*). Catalytic deiodination of [^{125}I]-diiodotyrosine (DIT) was determined by discontinuous measurement of [^{125}I]-iodide release described previously using DIT concentrations ranging from 1.0 to 50 μM and dithionite as the reductant (1 % w/v) (*19, 53*). Assays were performed in triplicate,

averaged, and data were fit to Michaelis-Menten kinetics using Origin 7.0 (Northampton, MA).

4.2.2 Site-directed Mutagenesis

Mutations to the gene were generated by site-directed mutagenesis of the pET32A plasmid encoding TRX-IYD(Δtm)DM. E153Q was introduced by the forward primer 5'-GAAGAGGAG**C**AAGAAATTAATTACATGAAAAGGATG GGAAAGCGATGGG-3' and its reverse complement. Y157F was created by the forward primer 5'-GAGGAGGAAGAAATAAATT**T**CATGAAAAGGATGGGAA AGCGATGGG-3' and its reverse complement. K178Q was formed by using the forward primer 5'-AGAACCAACTGGATT**C**AGGAGTACTTGGACACCGCC CCAGTTCTGATCCT -3' and its reverse complement. The mutated codon is underlined and the amino acid difference is in bold. The desired mutation was confirmed by sequencing and plasmids were used to transform electrocompetent Rosetta 2(DE3) *E. coli* for expression.

4.3 Results and Discussion

Numerous active site residues are likely responsible for catalysis (indirectly) and binding of substrate and FMN. Previous studies identified key amino acids involved in FMN binding by sequencing the IYD gene from humans that had thyroid disease (*14*). Later crystallographic studies identified E153, Y157 and K178 as key to substrate recognition (*31*).

The most conservative substitutions E153Q, Y157F and K178Q were generated individually by standard oligonucleotide-directed mutagenesis. The

mutants alter the charged atoms from the amino acid that creates the hydrogen bond between the protein and the zwitterionic tail of the substrate. The K178Q mutant of IYD(Δtm)DM expressed in *E. coli* but was insoluble and inactive. This was not pursued further. However, the two remaining mutants alternatively containing E153Q and Y157F were expressed in a soluble form. The ability of these mutants to bind and deiodinate substrate was examined through circular dichroism, equilibrium binding, and release of radioactive iodide from DIT.

4.3.1 Confirmation of soluble and folded protein

When mutations that are intended to disrupt ligand binding are made to proteins, it is imperative that the expressed soluble protein is folded the same with respect to the parent protein. Here, mutations were made to the active site lid of parent IYD(Δtm)DM, presented earlier. Of the three mutations created, only two (E153Q and Y157F) successfully folded in *E. coli*. The third mutation, K178Q, was expressed as insoluble protein in *E. coli*, despite similar efforts to aid solubility described in Chapter 2. The two purified enzymes were confirmed to have a 1:1 FMN to protein molar ratio. This is important because the mutant proteins must bind flavin in a constructive manner to perform catalysis. Proper protein folding and expected flavin environment are validated by circular dichroism (CD).

Even though many of the proteins were soluble, their ability to fold in an active conformation must be confirmed. CD spectra of each mutant were collected and overlaid for comparison of α-helical and β-sheet content (Figure 4.3). They each display appropriate negative peaks at 208 nm and 222 nm for standard secondary

structure. Likewise, the CD spectra in the visible region (300 – 500 nm) show positive peaks at 340 nm and 450 nm, as well as a negative peak at 365 nm. Both of the mutants' CD spectra display corresponding peaks indicating that the flavin environment is identical in each of the proteins. Any loss in catalytic activity cannot be attributed to misfolded protein or a difference in flavin environment.

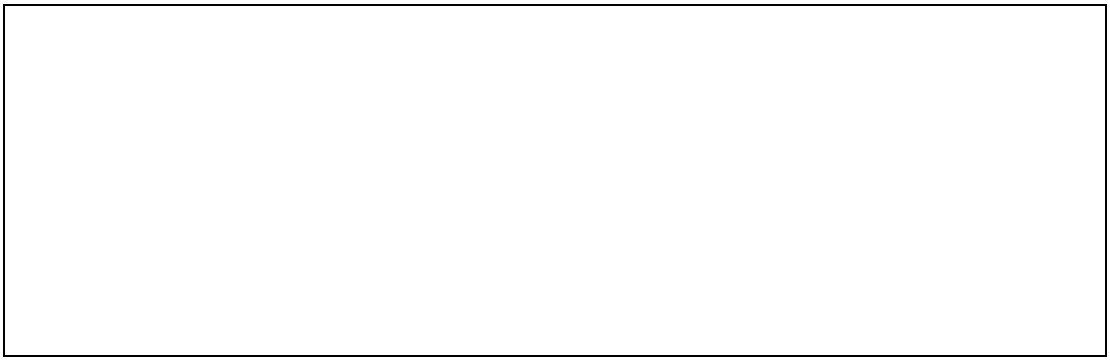


Figure 4.3 CD spectra of mutations of IYD(Δtm). Black lines indicate IYD(Δtm), green lines indicate Y157F, and red lines indicate E153Q. Proteins are in10 mM potassium phosphate (pH 7.4) at 25 °C in a 0.1 cm cuvette. a) Far-UV spectra taken of 4.5 μM protein. b) Visible spectra of 60 μM protein.

4.3.2 Deiodinase activity of mutants

Mutation to these two sites reduced the enzymes' ability to deiodinate DIT. The loss of the phenolic group (Y157F) increased the k_{cat} and K_M values for deiodination by more than 7-fold and decreased the k_{cat}/K_M value more modestly by less than 40% (Table , Figure 4.4). In contrast, loss of the carboxylic acid provided by E153 reduced the deiodinase activity to an undetectable level. Loss of FMN from the active site was not responsible for these changes since its occupancy remained constant at one per active site in these two mutants enzymes, identical to the parent IYD(Δtm)DM. Additionally, CD spectra of the deiodinase mutants indicated that their extent of α-helix and β-sheet structures or the flavin environment remained

constant (Figure 4.3a). This means that the inability for the mutated protein to catalyze deiodination is directly related to the amino acid substitution.

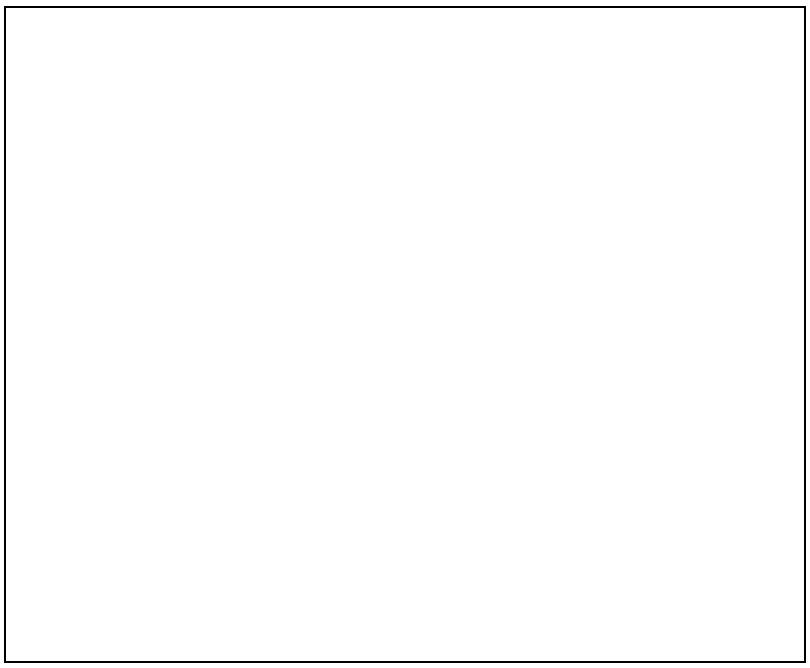


Figure 4.4 Specific activity of IYD(Δtm)DM and its mutants. Three data sets were measured and their averages and standard deviations are shown. Kinetic constants were obtained by fitting the data to Michaelis-Menten kinetics using Origin 7.0.

Table 4-1. Catalytic properties of iodotyrosine deiodinase (*Mus musculus*) derivatives[a].

		DIT[a]			MIT
Source	Enzyme	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (min^{-1} μM^{-1})	K_D (μM)
Sf9	IYD (Δtm)	4.5 ± 0.7	9 ± 1	0.49 ± 0.09[c]	0.09 ± 0.04[b]
E. coli	IYD (Δtm)DM	9.3 ± 1.6	40 ± 5	0.23 ± 0.05[c]	2 ± 0.2
E. coli	Y157F	65 ± 16	440 ± 170	0.15 ± 0.07[c]	40 ± 10
E. coli	E153Q	-	-	-	>1000

[a]Kinetic measurements based on deiodination were fit to Michaelis-Menten kinetics. [b]From McTamney and Rokita (*26*). [c]Propagation of error was calculated from the equation in Appendix B.

Active site binding of MIT was measured independently from enzyme turnover by the fluorescent properties of the bound flavin (Figure 4.4). Titrations

were limited by the solubility of MIT in 10 mM potassium phosphate (pH 7.4). The E153Q mutant expressed no measurable binding affinity for MIT, which could explain its lack of catalytic activity (Table -1, Figure 4.4). Even the Y157F mutant weakened the binding of MIT by 20-fold relative to its IYD(Δtm)DM parent (Figure 4.5). Almost the same extent of binding affinity was also lost by the substitution of the two cysteines for alanines despite their distance from the active site (Table -1). The twenty-fold decrease in this case indicates that Y157 contacts are important but their absence does not disrupt the ability for substrate to bind.

Kinetic and equilibrium binding studies to date indicate that E153 is required but not solely responsible for substrate binding. In contrast, Y157 contributes to substrate binding and catalysis but is not essential. The loss of a hydrogen bond provided by Y157 weakens MIT binding yet increases k_{cat}. Perhaps, the loss of binding affinity is the result of faster release of substrate and similarly faster release of the product tyrosine during turnover. The lid lacks a hydrogen bond that helps stabilize the productive formation for catalysis. Without this hydrogen bond, the lid is perhaps less established and would take less energy to release after catalysis.

The lack of hydrogen bonding between the zwitterionic arm and the mutated residues from the active site lid might prevent proper orientation of the aromatic substrate to stack directly with FMN. In this case the active site lid is directing the substrate into the catalytic conformation. The substrate fits comfortably *perfectly* into this cavity with the lid creating an environment unexposed to solvent. The degree by which the active site lid can do this is governed by the few contacts made by E153 and Y157 to the zwitterionic arm of the substrate. It is evident by the decreased

substrate affinity and activity that the hydrogen bonding of the charged arm to the N-3 and O^4 of FMN alone are insufficient for proper orientation of the substrate as neither are directly involved with catalysis.

These initial substitutions did not affect the CD spectrum in the visible region based on the flavin absorbance (Figure 4.6). Both IYD(Δtm) expressed in *Pichia* and IYD(Δtm)DM expressed in *E. coli* generated equivalent spectra in the absence of MIT and responded similarly after addition of MIT. Subsequent mutation of Y157F had minimal effect on CD spectra, but not surprisingly, mutation of E153Q dramatically diminished response to MIT (Figure 4.5).

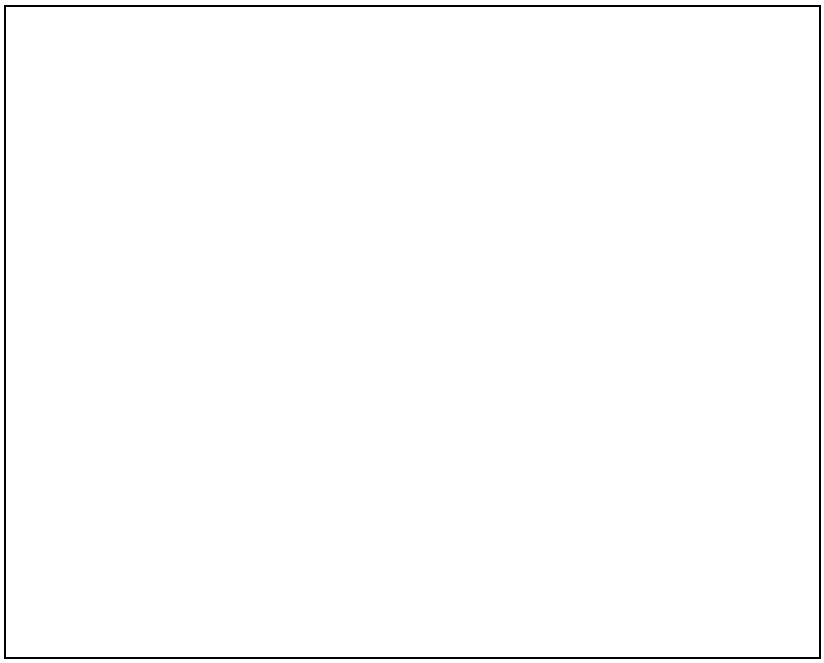


Figure 4.5 Equilibrium binding curves of IYD(Δtm)DM and its mutants. Ligand binding was monitored by the change in fluorescence of the flavin bound to IYD using λ_{ex} of 450 nm and λ_{em} of 527 nm. Solutions of IYD (4.5 μM) in 10 mM potassium phosphate (pH 7.4) were titrated with MIT to the extent possible. Three independent titrations were recorded, the fluorescence intensities were corrected for the slight dilution caused by the addition of ligand, normalized by dividing the fluorescence by the initial fluorescence and plotted against ligand concentration. Dissociation constants were calculated as prescribed in the literature from a nonlinear best fit to Eq. using Origin 7.0.

The substrate stacks over the flavin, creating a change in the active site environment. The CD spectra in visible wavelengths will change based on a shift in the interactions with the flavin (*56*). The shift could be caused by a conformational change in the protein structure or due to ligand binding. Differences in the visible UV CD are indicative of differences in the mode of binding of ligand to protein or a conformational shift essential for biological function (*63*). A change in the CD spectra of IYD(Δtm) after saturation with MIT is observed as a shift to a negative peak at 350 nm in the visible wavelengths (flavin absorption). The shift was observed for Y157F but not for E153Q. In the active site of Y157F, MIT must form a similar interaction with FMN as the parent IYD(Δtm)DM based on the sensitivity of CD in these wavelengths. As for E153Q, the persistent positive peak at 340 nm and negative peak at 370 nm in the presence of substrate indicates that no interaction occurs between the FMN and MIT. After the lack of activity and fluorescence quenching, it is no surprise that the visible CD spectra of E153Q with MIT does not change (Figure 3.6). This is consistent with the idea that CD can detect the connection between ligand binding and protein conformation.

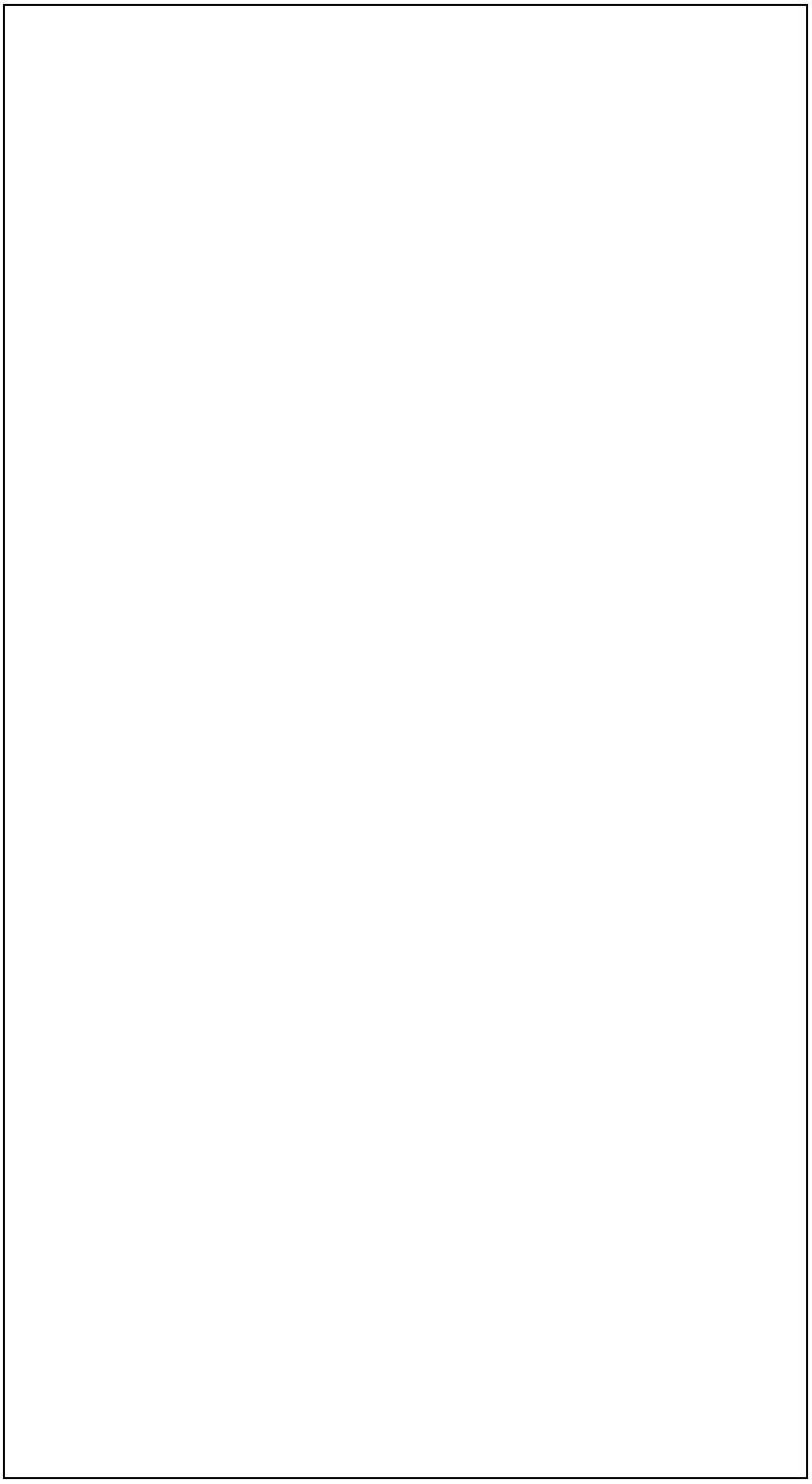


Figure 4.6 Visible UV CD spectra taken of IYD (60 μM) in 10 mM potassium phosphate (pH 7.4) in the presence and absence of 300 μM MIT at 25 °C in a 0.1 cm cuvette. a) parent protein IYD(Δtm)DM, b) IYD(Δtm)DM Y157F, c) IYD(Δtm)DM E153Q.

Y157 and K178 form hydrogen bonds to the carboxylic acid of the substrate. Removing only one of these two contacts, as with Y157F, would still allow for an interaction with the carboxylic acid of the substrate. We can assume that K178 is sufficient for stabilizing this interaction. This is in comparison to E153 that forms hydrogen bonds with the amine on the substrate. E153 is the only amino acid that hydrogen bonds to this group. Therefore, in order to observe substrate binding, thus catalysis, the active site lid must contact both zwitterionic groups.

These residues must be involved in orienting the substrate, stabilizing the active site lid and indirectly activating the FMN for catalysis. Dissecting their role has been facilitated by expression of the deiodinase in *E. coli* for routine site-directed mutagenesis. Perhaps a fine balance is maintained between dynamic and fixed states of the lid region. This active site lid is unique to IYD within its structural superfamily. These residues are defining factors to discriminate IYD from its oxidoreducase superfamily.

Chapter 5: Expression of homologous IYD

5.1 Introduction

IYD is the only known mammalian member of the NADH oxidase/flavin reductase structural superfamily (*19, 31*). Other enzymes in the superfamily are bacterial flavoproteins that function as NADPH oxidases or flavin reductases (*32, 33*). Until the crystal structure of IYD was solved, the enzymes in the structural superfamily were thought to belong to one of the two subclasses based on the primary structure of the active site lid domain. The active site lid domain of the bacterial enzymes is either between β2 and αE or after β5 while IYDs active site lid is αC to αDE (Figure 5.1). The location of the active site lid can distinguish between the three classes of enzymes in the NADPH oxidase/flavin reductase superfamily. IYD forms a third subclass of enzymes in this structural superfamily because the placement of its active site lid domain differs from the other two classes (*31*).

IYD has a unique function of dehalogenation within the NADPH oxidase/flavin reductase superfamily. Not only does IYD deiodinate its native iodinated substrate but it dehalogenates bromo- and chlorotyrosines (*26*). The enzyme with the most similar structure (very low RMSD of 3.1 Å) is the bacterial enzyme BluB (PDB ID: 2ISK). BluB catalyzes yet another type of reaction; the degradation of FMN to 5,6-dimethylbenzimidazole for the synthesis of Vitamin B_{12} (*34*). The organization of the subclasses is further confirmed by the placement of the sequence of the active site lid of BluB and the difference in chemistry from the rest of the superfamily (*31*). BluB and IYD have an insertion in the middle of the αC and

αDE for the active site lid compared to that of NOX and FRP (Figure 5.1). Here, we have enzymes of different function with very similar structure characterized from two very distant organisms, mammals and bacteria.

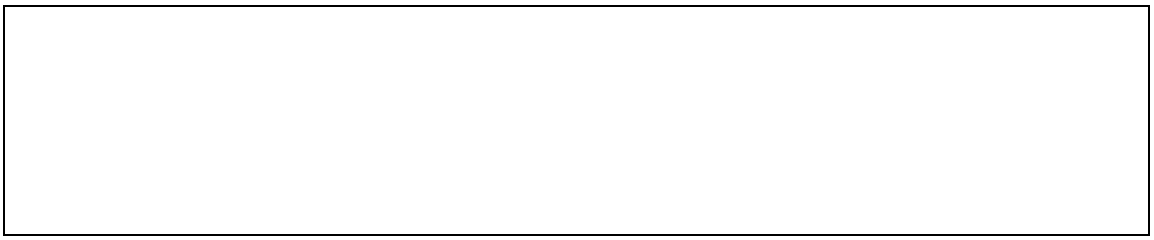


Figure 5.1 Secondary structure alignment of representative enzymes from the NADPH oxidase/flavin reductase superfamily. Protein Data Bank identifiers are as follows: IYD is 3GFD, BluB is 1ISL, NOX is NOX1, and FRP is 2BKJ. Figure adapted from Thomas, et. al. (*31*).

Structural homology of enzymes and their relation to the respective sequences has been a major point of study. The overall structures of bacterial enzymes in the NADPH oxidase/flavin reductase superfamily are very similar despite the sequence location of the active site lid. Enzymes are classified by function through the Enzyme Commission (EC) and by structure through secondary structure classes. Enzyme functions range over a variety of structural classes and enzymes in the same EC family catalyze different reactions, as clearly evident by the NADPH oxidase/flavin reductase superfamily.

The variation between structure and function occurs through two types of evolution: divergent and convergent. Divergent evolution is when a common ancestor yields superfamilies of enzymes that catalyze different reactions. Convergent evolution independently generates unrelated enzymes that catalyze the same type of chemical reaction. Within the structural superfamily, all of the enzymes belong to EC 1, the oxidoreductases, indicating they have a common ancestor thus divergent

structural evolution. However, they are different subclasses within. Both NOX and FRP are in "acting on NADPH" subclass (1.6) and IYD is in a different subclass "acting on a halogen" (1.22) (*64*). As far as classifying enzymes, ID, the other dehalogenase found in thyroid cells, belongs to subclass 1.97, with an alternative function classification of "other" (*64*). ID, although it is a flavoprotein that acts on a halogen, it does not use its flavin directly for dehalogenation. Instead, it performs its deiodination through the use of selenocysteines, so it is not directly related to convergent evolution. ID and IYD dehalogenation reactions are not classified in the same subclass because they catalyze dehalogenation in different manner.

IYD evolved from a structural superfamily of bacterial enzymes that are classified in the same enzyme class of reactions. In order to determine how dehalogenation activity evolved from its superfamily, we must bridge the gap between Eubacteria (bacteria) and Eukaryotes. The availability of DNA libraries allows us to search the sequence for a specific match based on DNA sequence or the respective translated protein sequence. The use of the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (*65*) has greatly enhanced search capabilities. Currently, the BLAST database has sequencing data for about 200 organisms, almost a quarter of the available genome sequences are animals. The sequences deposited in NCBI are annotated by GNOMON to affiliate a structure and a function.

A BLAST search of the mammalian IYD protein sequence results in approximately 30 homologous sequences from different organisms. Many of the aligned sequences are annotated as "iodotyrosine dehalogenase 1" while others are

more generally labeled as “nitroreductase” or “oxidoreductase” from the Conserved Domains Database (*66*). The BLAST output also indicates the putative FMN binding sites based on the sequence alignment. Many non-mammalian animals (Figure 5.2) have a homologous protein including: Vertebrate classes Osteichthyes and Aves are represented by *Danio rerio* (zebrafish) and *Gallus gallus* (chicken), respectively, Cnidaria are represented by *Nematostella vectensis* and *Hydra magnipapillata* (sea anemone and sea squirt, respectively), Echinodermata is represented by *Strongylocentrotus purpurtus* (sea urchin), *Branchiostoma floridae* in subphylum Cephalochordata (lancelet), *Xenopus laevis* or *Xenopus tropicalis* for Amphibia (frog), Arthropoda represented by *Apis mellifera* (honeybee), *Drosophilia melengaster* (fruit fly), and *Tribolium castaneum* (beetle), Nematoda represented by *Caenorhabditis elegans* (worm). The amino acid sequences of predicted IYD from these animals have ~60 % homology to the *Mus musculus* sequence of IYD.

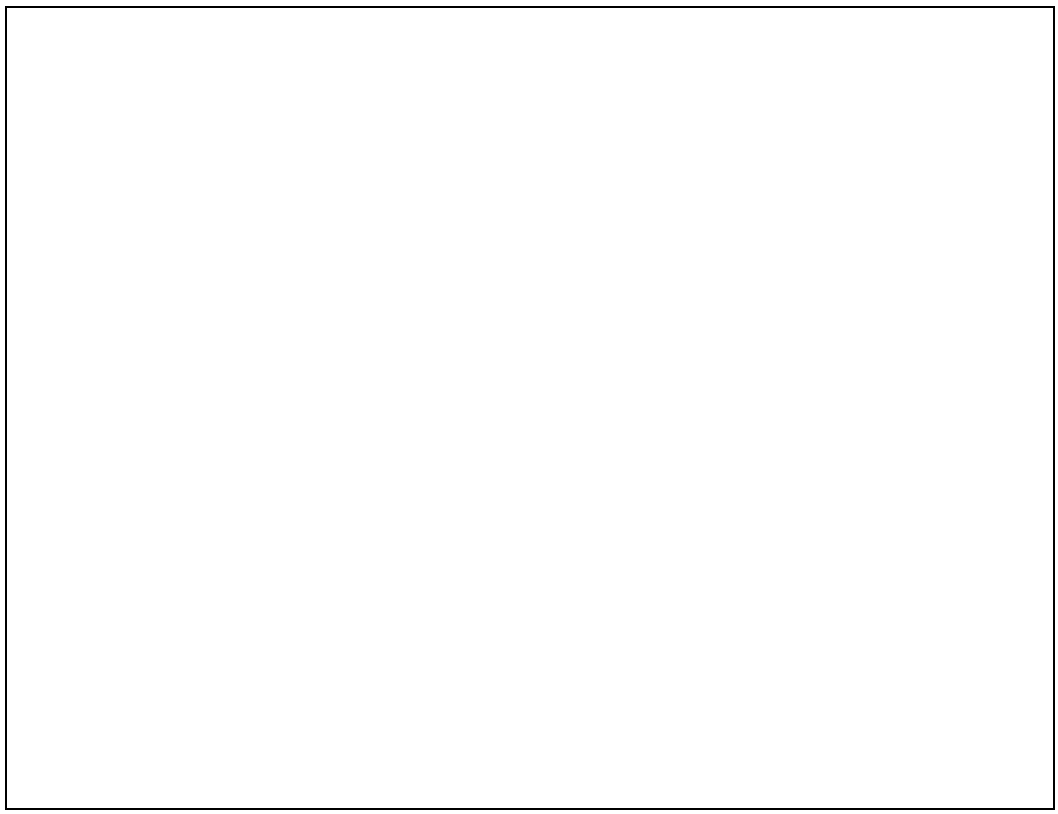


Figure 5.2 Phylogenetic tree of metazoan (animal) life. Classes and Phyla in blue indicate animals where predicted or hypothetical iodotyrosine deiodinases have been identified. Deiodination is primary necessary in glands were thyroid hormones are found, specifically follicular thyroids and endostyles. Branches are not drawn to scale.

Sequence alignments of predicted iodotyrosine deiodinase proteins indicate that eukaryotic organisms in different phyla may possess deiodinase activity (Figure 5.3). There are four criteria to base this prediction: an N-terminal membrane domain, flavin binding residues, an active site lid in the same domain as mammalian IYD, and the substrate binding residues E153, Y157, and K178. Mammalian IYD, in its native form, has a membrane binding region (Figure 5.3, green box), which is known to be necessary for retaining the ability to be reduced by NADPH *in vivo* (*18*). IYD does not bind NADPH directly and prior to catalysis, FMN must be reduced. So this must be a requirement for *in vivo* reduction. The second criterion is the flavin binding

residues (two blue boxes). Two flavin binding regions are located on opposite polypeptide chains, so dimerization is necessary for flavin binding. Specifically, threonine 235 in the mouse sequence directly hydrogen bonds with the N5 of FMN. This contact is constant within the subclass of the superfamily and is presumed to help stabilize the anionic semiquinone (*67*). The third criterion is the presence of sequence for the substrate-induced structure surrounding the active site that defines IYD from the rest of the proteins in the superfamily (red and light blue boxes). The homologous sequence must have sequence in the same region as IYD that the bacterial sequences lack. The third necessity is the presence of the three substrate binding residues, E153, Y157, and K178, defined in Chapter 4 (orange). These residues were found necessary for binding (thus catalysis) so they must also align to the mammalian IYD sequences. Notice, as expected based on the active site lid placement, the bacterial FRP sequence does not adhere to these criteria.

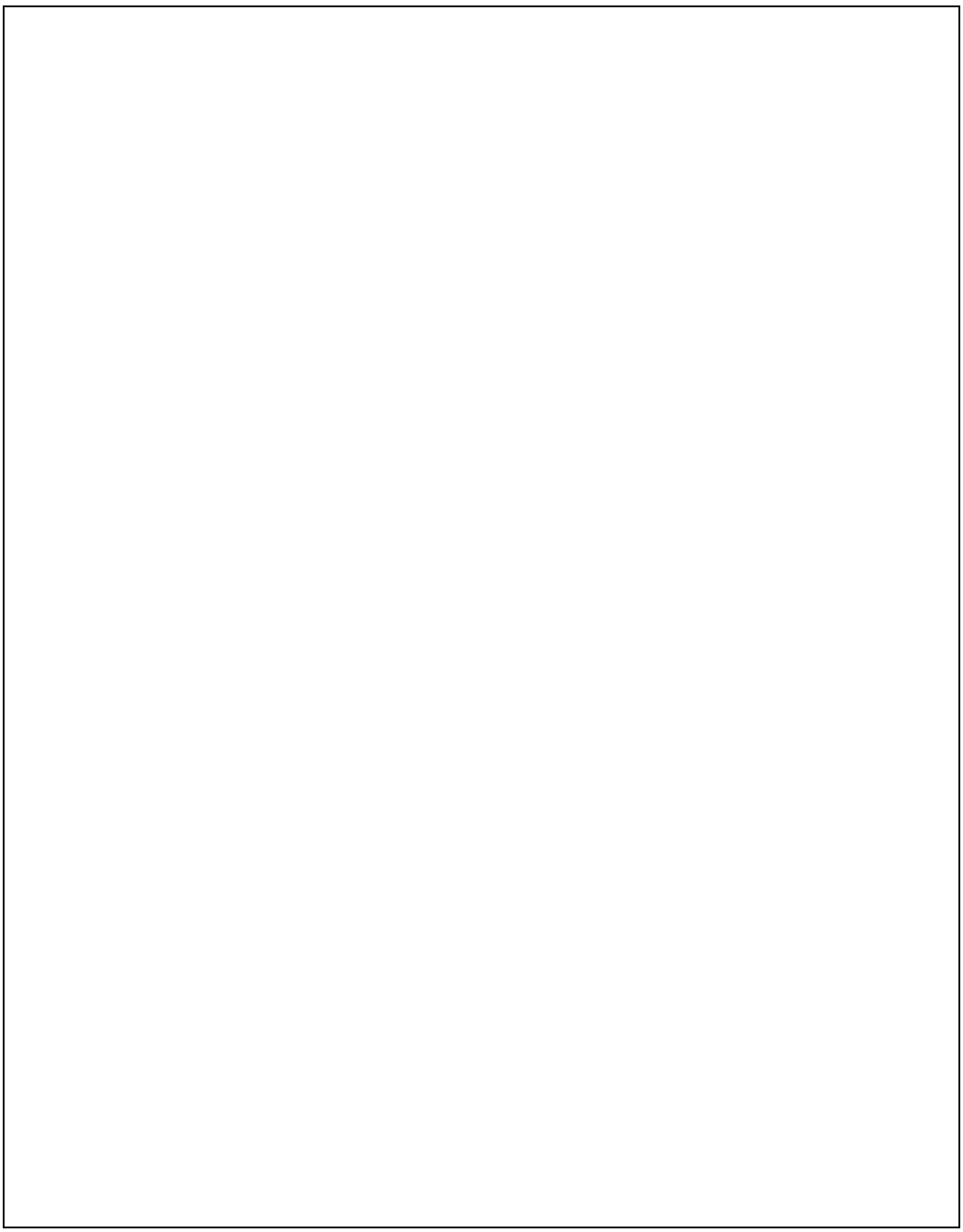


Figure 5.3 Alignment of predicted IYD from organisms in different classes. The key domains from comparison are highlighted. The flavin binding regions are boxed in violet. The membrane binding domain is in green; substrate binding helices: Red; and the substrate induced structure loop is in light blue. The three residues highlighted in orange (E, Y and K) hydrogen bond to the substrate. Helical and β-sheet structural features are shown on the top line for the IYD(Δtm) structure.

Representatives from each Phylum were chosen to test the hypothesis that the defined criteria are descriptive of deiodinase activity. Each of the organisms chosen has been well characterized and is used as model organisms to developmental biology. Starting closest to Mammalia, we chose sequences progressively farther away from known deiodinases. The *Danio rerio* (zebrafish) gene was chosen from the Class Osteichthyes. Zebrafish are known to have a thyroid and thyroid hormone (*35*), so it is an obvious first choice because it needs all the necessary machinery to produce thyroid hormone. The predicted IYD should in fact be iodotyrosine deiodinase. The second sequence chosen is from *Branchiostoma floridae* (lancelet) in the subphylum Cephalochordata. Lancelets have an endostyle that exhibits thyroid-like activity and accumulates iodide (*68*). The next outward branch is another subphylum of Deuterostomia. There was evidence for a homologous sequence in *Strongylocentrotus purpurtus* (sea urchins), the subphylum Echinodermata, but the NCBI accession could not be confirmed. This subphylum was skipped and we proceeded to Protostomia and the subphylum Arthropoda. Four insects were found to have a sequence homologous to IYD. The sequence from *Apis mellifera* (honeybee) was ultimately chosen based on the least amount of excess sequence and a lack of additional cysteines. The final and furthest removed sequence was chosen from the organism *Nematostella vectensis* (sea anemone) in the phylum Cnidaria and the Class Anthosoa. Sea anemones live underwater in what is thought to be iodide rich environment. Iodide is highly water-soluble so it is less likely that an organism would need to harbor iodide in such a manner that would require a specific deiodination recycling mechanism. Additionally, the gene from sea anemone does not have a

sequence in the N-terminus that may bind the membrane like the known mouse sequence. This can serve as a representative slightly deviated from the defined criteria. If some or all of the sequences retain deiodinase activity, we can then hypothesize where the evolution of a dehalogenase arose, from an otherwise NADPH oxidase/flavin reductase superfamily. The sequences in Figure 5.3 from the representative phyla were expressed and assayed for deiodinase activity. The genes were truncated to express the predicted soluble domain of the entire sequence. Furthermore, the reason that these organisms would need a deiodinase is investigated and presented through a developmental biology perspective.

5.2 Experimental

5.2.1 Materials

Genes from zebrafish (Accession: XM_691419), lancelet (Accession: XM_002237516), honeybee (Accession: XM_397179) and sea anemone (Accession: XM_001633119) were ordered from Blue Heron Biotechnology (Bothell, WA) or GenScript (Piscataway, NJ). Rosetta2 (DE3) *Escherichia coli* were purchased from Novagen (San Diego, CA). pET28-SUMO was obtained with permissions from Dr. Christopher Lime. All other reagents were obtained at the highest grade available and used without further purification.

5.2.2 IYD presence and Phylogenetic analyses

The presence/absence of IYD and related proteins was determined by searching gene names in public databases and through species-specific BLAST searches. For the BLAST analysis, sequence similarity greater than 57% similarity and a specific hit to the NADPH oxidase/flavin reductase conserved domain (*66*) was considered significant. Protein sequences were extracted from NCBI and sequence identifiers can be found in Appendix C. Sequence alignments were performed using ClustalW (*69*). Transmembrane regions were predicted by TMHMM (*70*). Structure modeling was performed using SWISS-MODEL (*71*) and was allowed to search for the template or model based on PDB ID 3GFD.

Expression of IYD from zebrafish, lancelet, honeybee, and sea anemone

pET28-SUMO plasmids containing homologous IYD genes were transformed into Rosetta2 *E. coli*. One colony inoculated a 20 mL LB culture containing kanamycin and chloramphenicol, which was in turn used to inoculate 1 L LB with the same antibiotics. Cells were grown at 37 °C with shaking to an OD_{600} of 0.6. Protein expression was induced with 0.4 mM IPTG and allowed to shake at 18 °C for four hours. Cells were harvested by centrifugation for 5 minutes at 5000 x *g* and lysed by three rounds on a French press. Insoluble proteins were removed by centrifugation at 40,000 x *g* for two hours.

Purification of IYD

The soluble proteins were filtered (0.2 μm) and loaded on a HiTrap Ni^{+2} chelating column. The column was washed with three column volumes of wash buffer (500 mM sodium chloride, 50 mM sodium phosphate (pH 8), 10 mM imidazole) and five column volumes of wash buffer with 70 mM imidazole. SUMO-IYD was eluted with 250 mM imidazole in wash buffer. Fractions containing SUMO-IYD were pooled, digested with Ulp1, and dialyzed against wash buffer containing 0.1 mM dithiothreitol (DTT). IYD was re-loaded on a HiTrap Ni^{+2} chelating column and washed with three column volumes of wash buffer. Step wise elution of IYD and SUMO included 10 column volumes of wash buffer with 70 mM imidazole, 10 column volumes of wash buffer with 100 mM imidazole, and 7 column volumes of wash buffer with 250 mM imidazole. Fractions containing IYD without SUMO were pooled and dialyzed against wash buffer containing 0.1 mM DTT. Upon removing the protein from dialysis, if the solution was cloudy with yellow precipitate, DTT was added to a final concentration of 1 mM to restore the protein to solution.

Deiodination activity of IYD

The discontinuous assay described in Chapter 3 was performed on purified IYD. Protein dependence was measured at 20 μM DIT for 30 minutes and time dependence of each protein was measured at 0.08 μM IYD at 20 μM DIT over one hour. Data was collected and fit to Michaelis-Menten equation as reported previously.

5.3 Results and Discussion

IYD homolog's were found in metazoan species in the phyla of Vertebrata, Urochodata, Arthropoda, Nematoda, and Cnidaria. A complete list of NCBI accession numbers based on a BLAST search can be found in the appendix. Of these, select

organisms from each phyla were chose as individual representatives of a general branch of metazoans. Chordata are represented two *Danio rerio* (zebrafish) for Osteichthyes and *Branchiostoma floridae* (lancelet) for Cephalochordata, Arthropoda represented by *Apis mellifera* (honeybee), and Cnidaria are represented by *Nematostella vectensis* (sea anemone).

Sequences were aligned using ClustalW (*69*). TMHMM predicted that the genes from zebrafish, honeybee, and lancelet each have an N-terminal membrane anchor of approximately 25 amino acids. Although the gene from sea anemone has sequence that aligns to mouse IYD, TMHMM did not predict that this region is membrane bound. We had predicted that the N-terminal membrane domain is critical to defining these proteins as a deiodinase. The gene from sea anemone filled the remaining criteria and so it was used as an intermediate between the two subgroups of the superfamily. The proteins from these will be referred to as the first letter of the Genus and Species before IYD; drIYD represents the protein from the homologous sequence found in *Danio rerio*, etc. The abbreviation IYD(Δtm) refers to the truncated mouse protein sequence.

5.3.1 Protein expression and purification

Previously, we had expressed full length mammalian proteins in the HEK 293 cell line and not encountered any folding problems (*18*). So, the full length sequences were ordered according to their native DNA sequence reported in NCBI (Table 2-1). The genes in pcDNA3.1 vector expressed in HEK 293 cells according to Western Blot analysis but were observed in the pellet as inclusion bodies and possessed no

deiodinase activity. Recent advances in our lab suggested that we could express truncated IYD in *E. coli* through the use of an N-terminal SUMO fusion protein.

The truncation length to remove the N-termianl membrane domain was directed by the information from the sequence alignment and the length of the predicted membrane domain from TMHMM. The truncation was made in approximately the same location compared to that of the known soluble expression sequence of mouse protein. The actual amino acid lengths expressed of the specified locus from each organism are listed in Table 2-1.

Table 2-1. Homologous IYD gene accession numbers and specific lengths of amino acids used for expression in *E. coli*.

Species	Protein	NCBI Locus	Amino Acids Expressed[a]
Danio rerio	drIYD(Δtm)	XM_691419	42-297
Branchiostoma floridae	bfIYD(Δtm)	XM_002237516	37-292
Apis mellifera	amIYD(Δtm)	XM_397179	44-296
Nematostella vectensis	nvIYD	XM_001633119	1-264

[a]The numbers refer to the amino acid number of the respective sequence.

Truncated gene sequences were expressed in Rosetta 2 *E. coli* in the pET28-SUMO vector. The fusion protein was found to be soluble and ranged from >10 to 30% of the soluble proteins. The fusion protein was initially purified on a Ni^{+2} chelating column and proteolysed from the SUMO fusion using Ulp1 during dialysis. SUMO and predicted IYDs were separated on a second pass through the Ni^{+2} column (Figure 5.4). IYD eluted as a yellow protein and was confirmed to have one mole of FMN per mole of protein. The homologous proteins bind flavin at the same molar equivalents as the mammalian IYD proteins. The purified proteins required DTT in

the final dialysis. This is most likely due to the additional cysteines in the sequence. DTT would reduce these cysteines, creating an environment conducive to a more soluble protein.

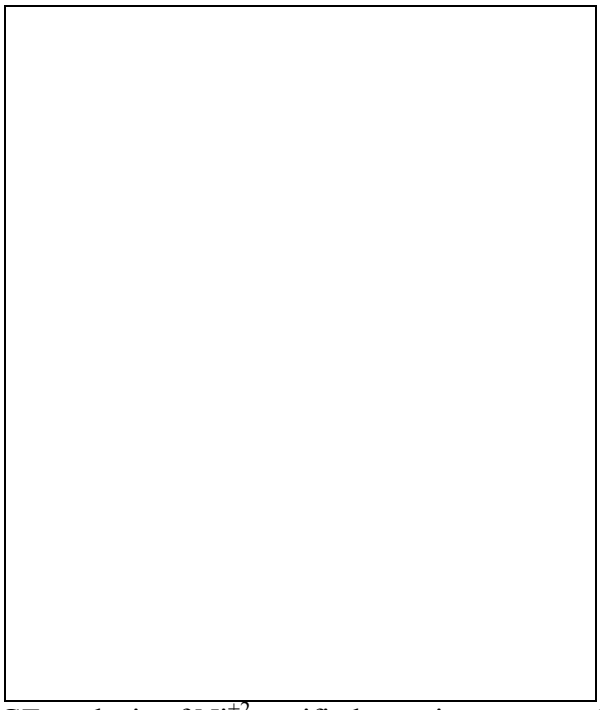


Figure 5.4 SDS-PAGE analysis of Ni^{+2} purified proteins expressed in *E. coli*. Proteins from different organisms are labeled by the first letter of the genus and species. *Danio rerio*: drIYD; *Branchiostoma floridae*: bfIYD; *Apis mellifera*: amIYD; *Nemostella vectininas*: nvIYD.

5.3.2 Homologous sequences are IYD

Purified proteins were tested for deiodinase activity based on the assay previously described. Each protein displayed Michaelis-Menten kinetic properties and was found to possess deiodinase activity! (Figure 5.5) The rates of deiodination of the truncated enzymes are very similar to that of mouse IYD(Δtm) (Table 5-2). The k_{cat}/K_M values for these homologous enzymes are all within error of the mouse enzymes previously studied without the N-terminal membrane anchor (*40*). The similarity in k_{cat}/K_M describes these enzymes as having the same degree of catalytic

efficiency indicating that dehalogenation of a mono or dihalogenated tyrosine could be the native substrate.

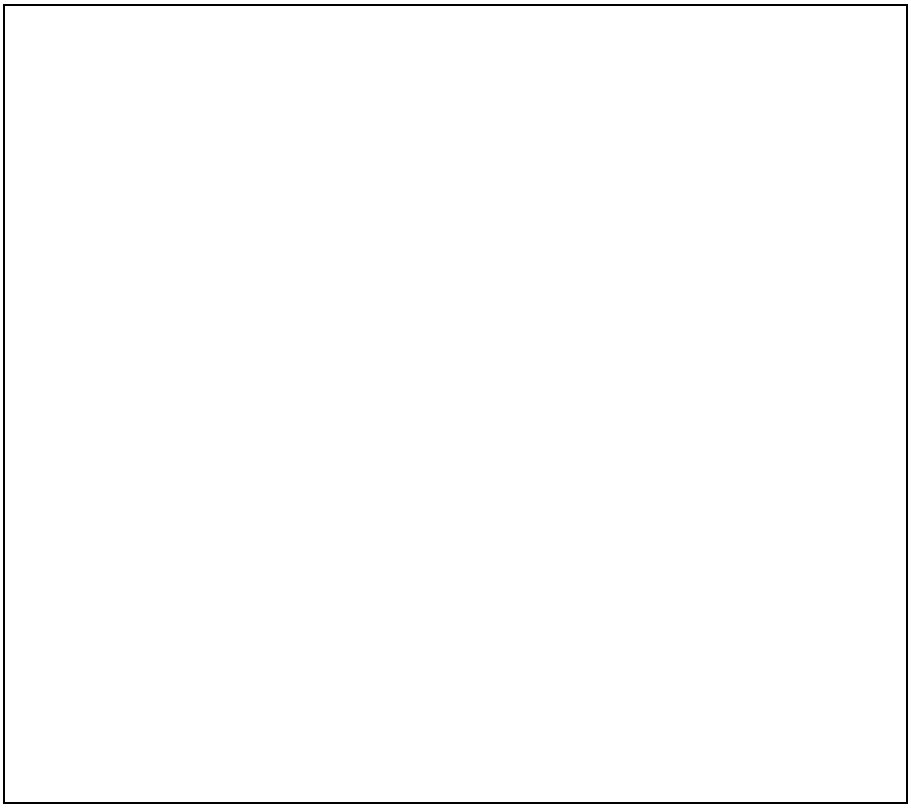


Figure 5.5 Initial rates of DIT deiodination by deiodinases. Data points are an average of three trials and the error bars represent the standard deviation at each concentration. Kinetic constants were obtained by fitting to Michaelis-Menten kinetics using Origin 7.0.

Since all of the predicted proteins possessed deiodinase activity with a similar rate, just one of these homologous proteins was tested for its affinity for the substrate. The zebrafish protein (drIYD(Δtm)) was the simplest to purify so it was used for equilibrium titration studies with MIT. Two independent titrations (Figure 5.6) were averaged and fit to a quadratic binding equation as previously described in Chapter 3 (*21*). drIYD(Δtm) binds MIT with 10 times lower affinity than mouse IYD(Δtm) (0.9 ± 0.07 μM vs 0.09 ± 0.04 μM). This is only 2 times tighter than the affinity of

mouse IYD(Δtm)DM for MIT (Chapter 3). IYD(Δtm)DM was previously used as a model for IYD(Δtm) because of its ability to be expressed as soluble protein in *E. coli* and its similarity to parent protein IYD(Δtm) expressed in Sf9 or *P. pastoris*. The five fold difference in binding affinity is relatively insignificant and we can assume that iodinated tyrosines are the native substrates for these homologous proteins.

Table 5-2. Kinetic parameters of IYD(Δtm) from different species expressed in *E. coli*.

Species	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (min^{-1} μM^{-1})
Mus musculus[a]	6.9 ± 1.3	19 ± 3	0.36 ± 0.09
Nematostella vectensis	6.1 ± 2.3	19 ± 16	0.32 ± 0.31
Danio rerio	4.1 ± 0.2	29 ± 7	0.29 ± 0.08
Branchiostoma floridae	7.2 ± 1.5	8 ± 1	0.53 ± 0.07
Apis mellifera	8.2 ± 1.0	6 ± 3	1.2 ± 0.74

[a]Data from (*40*).

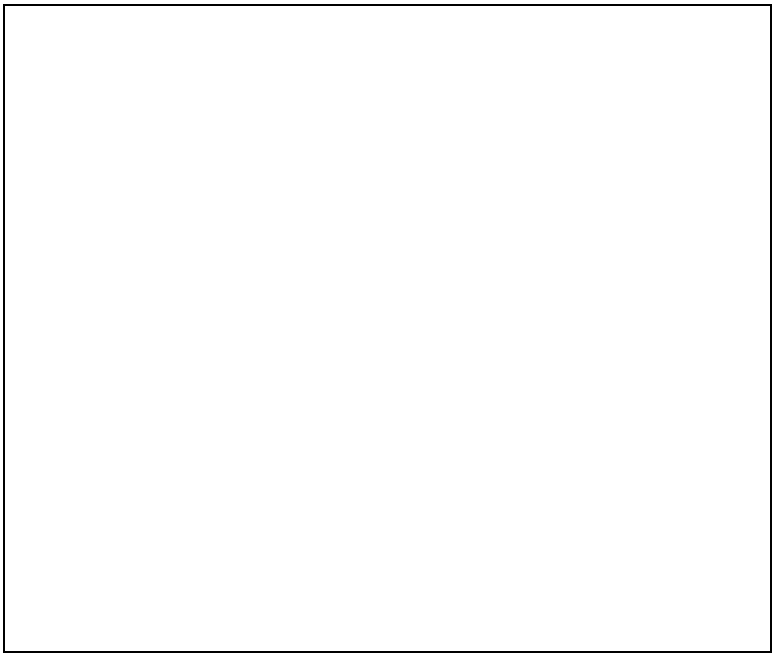


Figure 5.6 Fluorescence quenching of FMN bound to drIYD(Δtm) by MIT. Fluorescence emission (λ_{em} 527 nm; λ_{ex} 450 nm) was monitored during titration of MIT to a solution of drIYD(Δtm) (4.5 μM in 10 mM phosphate, pH 7.4). Data points of two trials and an average (black square) are plotted and error bars represent standard deviation.

5.3.3 Structure modeling of homologous IYD

The structure modeling of homologous proteins was completed to determine how the sequences would fit to the known structure of IYD. From this we can see if the sequence aligns well to the known structure and easily map the sequence differences. Protein sequences were entered in SWISS-MODEL to thread the sequence to a model structure from the Protein Data Bank. This can be done either by a user defined PDB ID or allowing the database to choose a PDB file. Each sequence was submitted using both methods and the outcome was the same. The mouse IYD crystal structure in complex with MIT (PDB ID 3GFD) was chosen as the user-defined structure to thread the homologues sequence. When SWISS-MODEL was

allowed to search for a model PDB, the same PDB ID (3GFD) was the result. Therefore, both the user-defined query and the open ended query resulted in the same theoretical structure. SWISS-MODEL reports a QMEAN z-score as an indication of goodness of fit for the provided sequence to the model. For the four sequences QMEAN z-scores were: -0.73 for sea anemone, -0.979 for zebrafish, -1.844 for lancelet, and -1.795 for honeybee. The human sequence which has 90% sequence identity to the mouse sequence was reported with a QMEAN z-score of -0.83 indicating that these z-scores are appropriate to use the predicted structure as a model. However, when the NOX bacterial sequence was threaded to the mouse IYD 3GFD structure, SWISS-MODEL reported a warning. A low QMEAN z-score (-5.29) was reported indicating that mouse IYD structure is a bad model to fit the NOX sequence. Similarly, when the zebrafish sequence was threaded to the PDB ID 1NOX, a similar warning arose from SWISS-MODEL. When the NOX sequence was used as input to SWISS-MODEL and allowed to search the database for an appropriate structure, the 1NOX PDB was reported with a QMEAN z-score of 0.102. Thus, the program works because it chose the appropriate structure for the known sequence. When this protein from the same structural superfamily was threaded to the mouse IYD PDB a poor structure was constructed. This is a good indication that the sequences chosen are in fact IYD sequences and not artifacts of the structural superfamily.

There are no significant differences in the predicted structures of any of the four homologous proteins to the mouse structure. We can see the overall fold of each of the sequences is very similar (Figure 5.7). Specific differences are mapped and displayed as red side chains in Figure 5.8. Differences are defined as residues that

have different functional groups than the mammalian sequence. For instance, a change from tyrosine to phenylalanine is considered a difference because of the loss of the hydroxyl group. A change from valine to leucine or isoleucine or any variation of those three is not considered a difference. The one carbon difference between glutamate and aspartate is not considered significant because the functional groups remain and in theory can make the same hydrogen bonding contacts.

Figure 5.7 Structural alignment of backbone from SWISS-MODEL predicted structure of homologous IYD sequences and mouse IYD. Zebrafish drIYD is in green, honeybee am IYD is in blue, lancelet bfIYD is pink, sea anemone nvIYD is yellow, and mouse IYD (PDB ID 3GFD) is in gray with FMN and MIT in yellow sticks.

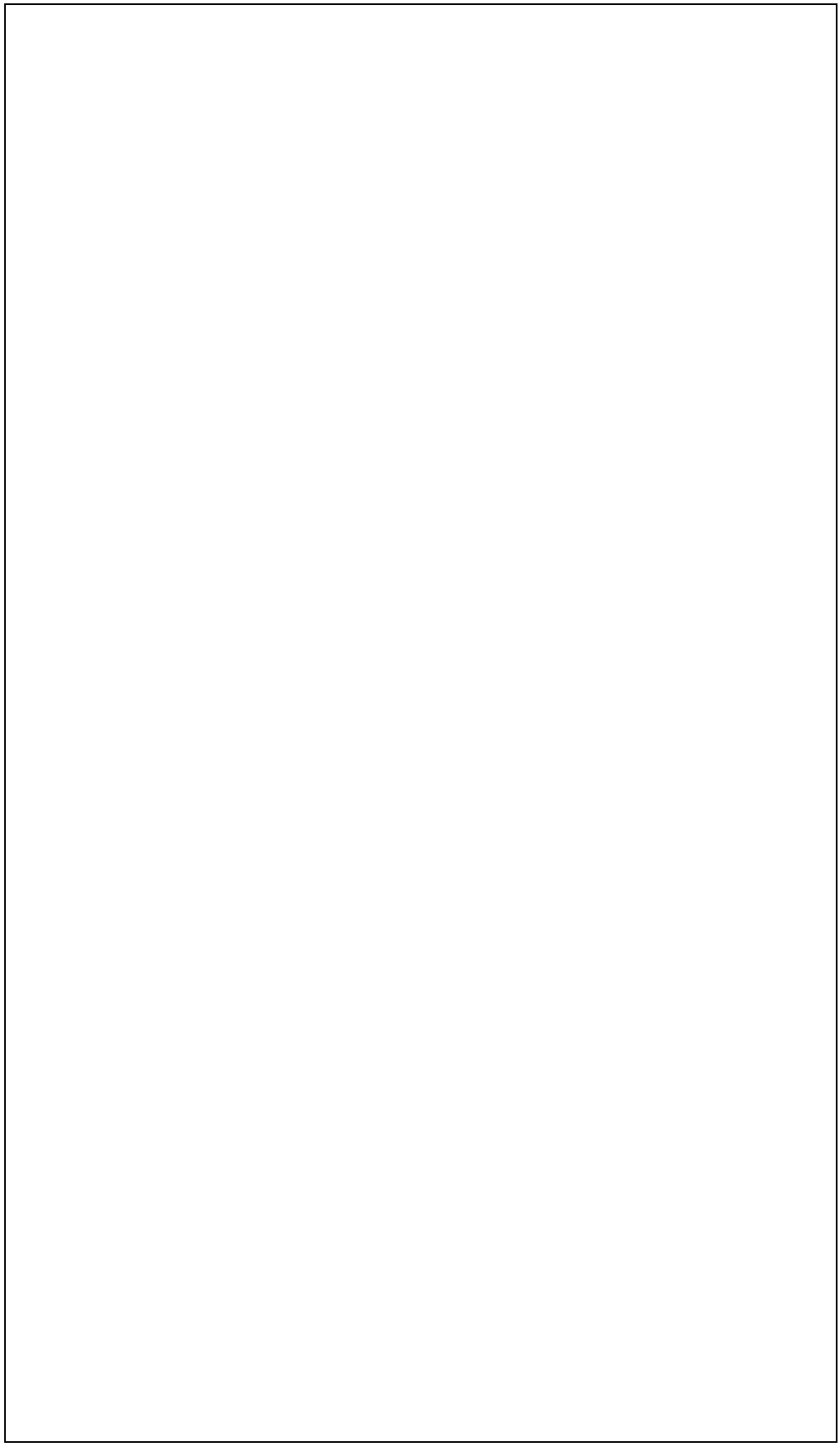


Figure 5.8 Modeling by SWISS-MODEL of homologous sequences to the mouse IYD crystal structure. Side chains in red represent non-conserved residues between the respective sequence and the mammalian sequence. To the right, each active site is highlighted. FMN and MIT are in yellow.

The vast majority of the different amino acids are located on the surface of the protein. When we closely examine the active site we are able to see the pocket does not have any residues highlighted in red indicating that the active site is highly conserved (Figure 5.8, right). This is very important because it tells us that even with only 60% sequence identity, the interior of active site are nearly identical. This is a good comfirmation that deiodination is the native activity.

The network of amino acids that form the active site lid are also conserved, evidenced by the structure we see in green in Figure 5.8 and the red box in the sequence alignment from Figure 5.3. Amino acids R100, N156, M161, R165, W165, D168, L169, K171, W176, L237 from one chain and V205, H206, Y207, and Y208 from the other are highly conserved. These are the residues that form the tight network of hydrogen bonds to sequester the substrate from solution. These interactions are important in defining a protein as being a dehalogenase rather than the bacterial oxidoreductase. Flavin binding location and the core α and β folds are the same across the structural superfamily but their active sites vary vastly. The nitroreductases do not create a more shallow binding pocket for their substrates when compared to the dehalogenases in the same superfamily. Presence of these amino acids should be used in the future as defining criteria for dehalogenation.

Initially, the presence of cysteines in the aligned sequence was used as criteria to discriminate between a deiodinase and a oxidoreductase (*19*). Since it is now known that they are not required for catalysis, the position of the cysteines was

ignored in this study. The lancelet sequence does not have cysteines at either location (mouse residues 217 and 239). Sea anemone and honeybee sequences have an additional cysteine replacing a serine at mouse residue 214. In the mouse (and human) sequence, this is a serine on the criss-crossing helix at the dimer interface. It faces towards the other chain and presumably forms a hydrogen bond with S214 on the opposite chain based on the proximity of the side chain. On the sea anemone and honeybee sequences the cysteines may in theory form a disulfide bond in that close proximity. The two cysteines of the honeybee and sea anemone sequences can be seen at their respective dimer interfaces in Figure 5.8 (amIYD and nvIYD, respectively).

It was originally thought that any sequence with additional cysteines would not be an ideal choice for expression in *E. coli* based on based on poor expression results described in Chapter 2. Upon purification, these two proteins would precipitate at 4 °C within three hours in wash buffer. However, after the addition of more DTT, the proteins would refold in a matter of minutes, as observed by the disappearance of yellow precipitate and the return of a clear rather than opaque solution. The addition of DTT leads us to believe that the cysteines are reduced in solution. Thus, they would not form the predicted disulfide bond at residue 214 and perhaps are forming similar hydrogen bonds to that expected from the mouse crystal structure.

The homologous proteins from zebrafish, lancelet, honeybee, and sea anemone do in fact function as iodotyrosine deiodinases. The defining criteria were met by four selected sequences and our hypothesis was correct. The sequences all

have a region defined as an active site lid, which was predicted by modeling to fold in the same manner as the mouse IYD sequence. Each homologous IYD bound equimolar flavin in a constructive conformation as indicated by deiodinase activity. An N-terminal membrane domain is not necessary for dithionite-dependent catalysis but will be necessary for facilitating reducing equivalents to FMN *in vivo* (*72*). Although we do not know the specific reductase that reduces the FMN bound to IYD, it is apparent that one must exist based on previous work (*18, 38, 41*). More interestingly, many of these organisms have a known requirement for iodide, specifically sea squirt, sea urchin, lancelet, zebrafish, and frog (*5, 7, 35, 73, 74*). Thyroid hormones, T3 and T4, have been discovered during metamorphosis in organisms without a follicular thyroid (*75-77*).

There is good indication that halogenated tyrosines exist in the investigated organisms. All of these organisms go through metamorphosis during their development. Thyroid hormone has been known for years to influence metamorphosis. Almost 100 years ago, frogs were studied with and without a thyroid. Those without a thyroid remained as tadpoles and upon iodinated tyrosines fed in their diet, morphed into frogs (*4*). Frogs, although may seem distant, are not very different from mammals when you look at the broad scale of the animal kingdom. Frogs are still chordates, the same phylum as mammals, but are part of the Amphibia class. The Vertebrata subphylum, including the Class Osteichthyes (e.g. zebrafish), all have a follicular thyroid. The presence of a thyroid indicates that these animals would endogenously produce thyroid hormone and possess regulating machinery similar to mammals. Organisms that produce their own thyroid hormone using a

thyroid would need a pathway to recycle the uncoupled iodinated tyrosines so it isn't so surprising that IYD is present. Homologous IYD sequences were detected in organisms that have an endostyle, specifically the subphyla Urochorata (e.g. sea squirt) and Cephalochordata (e.g. lancelet). An endostyle is a region near the notochord that morphs into a thyroid in the most basic chordate, the lamprey (*78*).

Thyroid hormone signaling pathway has been defined in lancelets and zebrafish (*6, 35*). The presence of thyroid peroxidase, thyroglobulin, a sodium iodide symporter, and thyroid stimulating hormone receptor indicate that lancelets, like the zebrafish chordates, endogenously produce thyroid hormone (*6*). Sea urchins have evidence of thyroid hormone related function, which leads us to believe that even though this sequence was not expressed, it most likely would have the deiodinase (*79*). The alternative to thyroid hormone production is uptake from the environment. Phytoplankton, a part of the diet for marine life, is a large source of iodide in marine environments and implicated in metamorphosis. There is evidence of thyroid hormone in mollusks that would be endogenously synthesized from iodide uptake (*79*). Contrast to using MIT for thyroid hormones, organisms might possess MIT for its antioxidant properties to scavenge radicals (*80*). This property would make iodotyrosines useful signaling molecules between cells (*81*) as they can insert themselves in lipid membranes when bound to other small molecules (*72*).

Growth, morphological, and physiological stages of chordates have been explored with respect to a thyroid gland and its respective properties (*6, 79, 82, 83*). Authors have isolated the protein from bovine and porcine thyroids (*36*), performed thyroidectamies (*4*), and predicted enzyme evolution through bioinformatics (*84*).

This is the first study to express and purify homologous enzymes to test the hypothesis that proteins predicted to be deioidinases indeed function as deiodinases. We biochemically characterized each protein and confirm that the annotated sequence of a predicted or hypothetical protein is in fact a deiodinase. Looking back, it should not be so surprising that each of the sequences chosen are enzymes that retain deiodinase activity. But the question still remains as to where the shift occurred from an oxidoreductase to deiodinase.

Chapter 6: Conclusions

IYD is a flavoprotein essential to a properly functioning thyroid. It forms a homodimer with two identical active sites that each bind FMN and is a member of the NADPH oxidase/flavin reductase structural superfamily. The location of the helical lid that folds to surround the active site separates proteins in this superfamily into three subclasses. This structural difference, in addition to its completely different function, places IYD in a separate subclass of the structural superfamily.

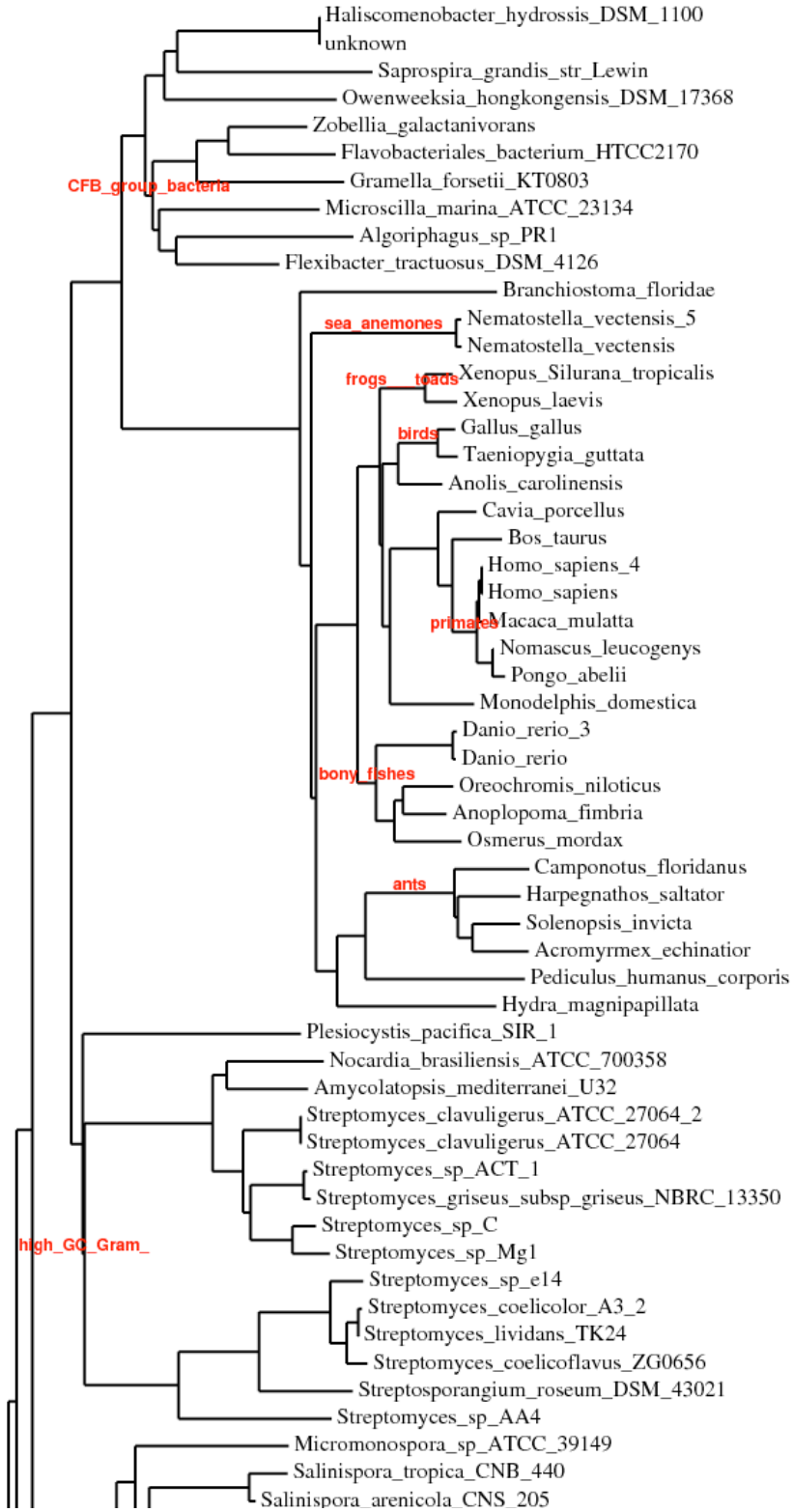
IYD performs reductive dehalogenation to remove iodide from a mono or diiodinated tyrosine which are the byproducts of thyroid hormone synthesis. The protein is primarily found in the thyroid where iodide is harbored to make thyroid hormones. Thyroid hormones help regulate mammalian metabolism. Thyroid hormones are known to aid in development, specifically for organisms that undergo metamorphosis. These organisms have a thyroid or endostyle and may only have expression of IYD as larvae and not as adults. Additionally, the organisms without a thyroid (lancelet and sea anemone) have homologous sequences to the thyroid peroxidase, sodium iodide symporter, and transthreyatin. The honeybee genome only has a sequence homologous to the sodium iodide symporter. This may be a false positive and could be a symporter for other ions.

Four organisms from different levels of the phylogenetic tree of life were chosen to represent respective phyla as we trace the activity of iodotyrosine deiodinase. Each of these proteins displayed deiodinase activity and therefore our

criteria for choosing sequences was successful. However, this has yet to give us any information to dissect where the protein shifts from its structural superfamily of oxidreductases to a dehalogenase.

Criteria were easily defined only after heterologous expression of IYD was successful. IYD was engineered to express a soluble protein in *E. coli* by use of fusion protein, two his-tags and mutation of two cysteines to alanines. This protein was found to be biochemically equivalent to the wild-type truncated protein expressed in HEK 293, Sf9, and *Pichia pastoris (40)*. Expression in *E. coli* was critical to making mutations to study proteins. The cysteine mutant protein can be used as a model of the wild-type protein to create additional mutations to study substrate coordination. In the active conformation, substrate hydrogen bonds to three amino acids (E153, Y157, and K178) on the active site lid. E153 is necessary for proper binding of the substrate in the active site and thus disrupts catalysis. Mutation to Y157 decreased the affinity of the protein for MIT and increased enzymatic turnover implying a decrease in stability of the enzymes active complex. Information could not be reported on K178 due to improper folding *in vivo*. The active site lid, the presence of the three coordinating amino acids, flavin binding residues, and an N-terminal hydrophobic sequence were the criteria chosen to define a deiodinase from the other members of the superfamily. Additionally, the network of amino acids protecting the active site from solvent that form the lid are conserved based on the sequences already known to have deiodinase activity. The presence of these should be considered criteria for future studies.

A recent BLAST search resulted in identification of homologous sequences from CFB bacteria *Haliscomenobacter hydrossis* and the Crusteaca *Daphnia pulex* (water flea). The bacterial genome was published in June 2011 (*85*) and the water flea genome was published only months earlier in February 2011 (*86*). The sequence available from the bacterial genome is annotated as a nitroreductase while the water flea sequence is annotated as "iodotyrosine dehalogenase" on the NCBI. Based on the BLAST results using the predicted IYD from CFB bacteria, a phylogenetic tree was created by TreeDyn (Figure 6.1).


Haliscomenobacter_hydrossis_DSM_1100
unknown
Saprospira_grandis_str_Lewin
Owenweeksia_hongkongensis_DSM_17368
Zobellia_galactanivorans
Flavobacteriales_bacterium_HTCC2170
Gramella_forsetii_KT0803
CFB_group_bacteria
Microscilla_marina_ATCC_23134
Algoriphagus_sp_PR1
Flexibacter_tractuosus_DSM_4126
Branchiostoma_floridae
sea_anemones
Nematostella_vectensis_5
Nematostella_vectensis
frogs
toads
Xenopus_Silurana_tropicalis
Xenopus_laevis
birds
Gallus_gallus
Taeniopygia_guttata
Anolis_carolinensis
Cavia_porcellus
Bos_taurus
Homo_sapiens_4
Homo_sapiens
primates
Macaca_mulatta
Nomascus_leucogenys
Pongo_abelii
Monodelphis_domestica
Danio_rerio_3
Danio_rerio
bony_fishes
Oreochromis_niloticus
Anoplopoma_fimbria
Osmerus_mordax
Camponotus_floridanus
ants
Harpegnathos_saltator
Solenopsis_invicta
Acromyrmex_echinatior
Pediculus_humanus_corporis
Hydra_magnipapillata
Plesiocystis_pacifica_SIR_1
Nocardia_brasiliensis_ATCC_700358
Amycolatopsis_mediterranei_U32
Streptomyces_clavuligerus_ATCC_27064_2
Streptomyces_clavuligerus_ATCC_27064
Streptomyces_sp_ACT_1
Streptomyces_griseus_subsp_griseus_NBRC_13350
Streptomyces_sp_C
high_GC_Gram_
Streptomyces_sp_Mg1
Streptomyces_sp_e14
Streptomyces_coelicolor_A3_2
Streptomyces_lividans_TK24
Streptomyces_coelicoflavus_ZG0656
Streptosporangium_roseum_DSM_43021
Streptomyces_sp_AA4
Micromonospora_sp_ATCC_39149
Salinispora_tropica_CNB_440
Salinispora_arenicola_CNS_205

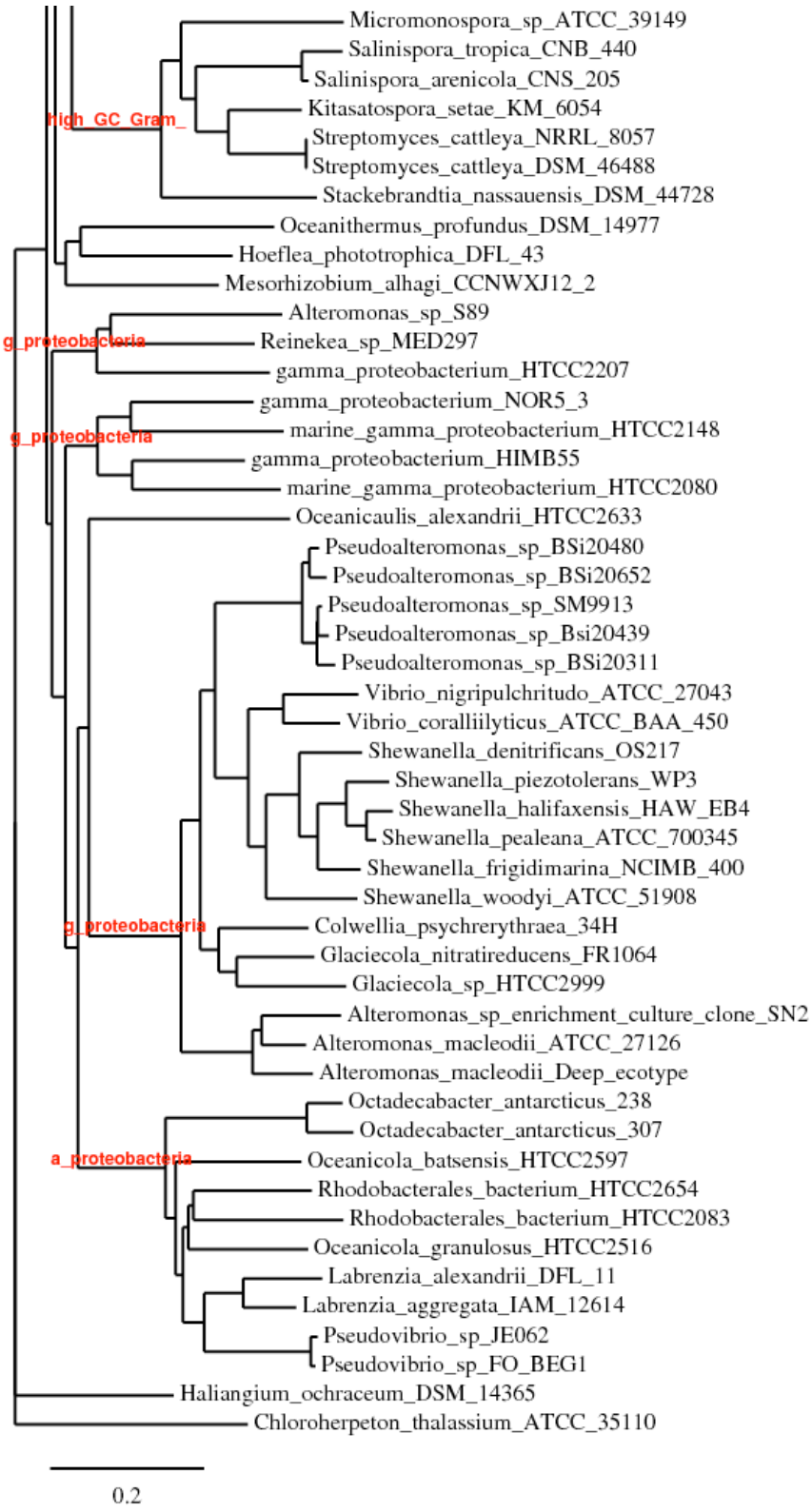


Figure 6.1 The phylogenetic tree based on BLAST results using the hypothetical IYD sequence from *Haliscomenobacter hydrossis* query created by Phylogeny.fr TreeDyn.

From this tree we can see that CFB bacteria are on a clade closest to known IYD sequences and farthest from other bacterial sequences. This is indicative that the protein may in fact be a deiodinase and not a nitroreductase, or at the very least, close to the branch point between the two. These sequences will be purchased and screened for deiodinase activity. Neither organism has sequence information as evidence of the other thyroid-like proteins (thyroid peroxidase, etc). Lacking other thyroid proteins could indicate that these proteins do not have deiodinase activity despite adhering to the specified criteria. It is also plausible that they do function as iodotyrosine deiodinases and serve an unknown physiological role. With this additional information, perhaps we can trace the evolution of iodotyrosine deiodinase.

Chapter 7: Appendices

A.

Name on Tube	Name	Vector	N-terminal Fusion Tags	C-terminal Fusion Tags	Domains	Mutations	expression strain	soluble (S) or insoluble (I)	freezer stock of plasmid	freezer stock of cells
IYD NR	IYD NR	pET102	Trx		NOX		BL21(DE3)	no vis exp		
IYD NR	IYD NR	pET102	Trx		NOX		BL21(DE3)pLysS	no vis exp		
IYD NR	IYD NR	JF1-pGEX	GST		NOX		BL21(DE3)	no vis exp		
IYD NR	IYD NR	pET102	Trx		NOX		Rosetta 2	I		
IYD NR	IYD NR	JF1-pGEX	GST		NOX		Rosetta 2	I		
IYD NR	IYD NR	pMal-c2	MBP		NOX		Rosetta 2	I		
IYD NR	IYD NR	pMal-p2	MBP		NOX		Rosetta 2	I		
IYD NR	IYD NR	pET43.1	NusA		NOX		Rosetta 2	I		
IYD Δtm DM	IYD Δtm DM	pET101/D-TOPO	n/a	V5-epitope His	I, NOX		Rosetta 2	I		
JA3-19	IYD Δtm DM	pET21a	n/a	His	I, NOX		Rosetta 2	I	JA3-19-1,2,4	JA3-21?
JA3-47-2-3	Gst IYD Δtm	pGEX4T-1	GST	His	I, NOX		Rosetta 2	I		
JA3-51-2	Gst IYD Δtm DM	pGEX4T-1	GST	His	I, NOX	C217A, C239A	Rosetta 2	I		
JA3-75	Trx IYD Δtm	pET32a	Trx His S-Tag	His	I, NOX		Rosetta 2	I	JA3-42-2	
JB1-55-4	Trx IYD Δtm DM	pET32a	Trx His S-Tag	His	I, NOX	C217A, C239A	Rosetta 2	S	JB1-55-4	JB1-55-4
	Trx IYD Δtm DM	pET32a	Trx His S-Tag		I, NOX	C217A, C239A	Rosetta 2	I		
JA3-83	Trx (Δhis) IYD Δtm	pET32a (ΔHis$_1$)	Trx S-Tag	His	I, NOX	YNILLL on pET32a C217A, C239A	Rosetta 2	I		
JA3-84	Trx (Δhis) IYD Δtm DM	pET32a (ΔHis$_1$)	Trx S-Tag	His	I, NOX	YNILLL on pET32a C217A, C239A,	Rosetta 2	I		
JA4-7	Trx IYD Δtm DM	pET32a	Trx His S-Tag	His	I, NOX	Y157F	Rosetta 2	S	JA4-7	JA4-7
JA4-8	Trx IYD Δtm DM	pET32a	Trx His S-Tag	His	I, NOX	C217A, C239A, K178Q	Rosetta 2	S	JA4-8	JA4-8
JA4-9	Trx IYD Δtm DM	pET32a	Trx His S-Tag	His	I, NOX	C217A, C239A, E153Q	Rosetta 2	S	JA4-9	JA4-9
	IYD NR	pcDNA3.1/Zeo			NOX		HEK293	no vis exp		
	IYD Δtm	pcDNA3.1/Zeo			I, NOX		HEK293	S		
	IYD NR	pcDNA3.1/Zeo		His	NOX		HEK293	I		
	IYD Δtm	pcDNA3.1/Zeo		His	I, NOX		HEK293	S		
	IYD Δtm	pcDNA3.1/Zeo		His	I, NOX		CHO	S		
JA5-19	IYD	pcDNA3.1/Zeo			MD, I, NOX					
	IYD Δtm	baculovirus		His	I, NOX		Sf9	S	x, virus	
	IYD Δtm	pPICZα		His	I, NOX	codon usage	p. pastoris	S	x	
	hIYD	pET32a	Trx		MD, I, NOX			I		
	hIYD1 (hIYD Δtm)	pET32a	Trx		I, NOX			S*		
	hIYD1	pGEX-6P-1	GST		I, NOX			I		
	hIYD1	pET43.1	NusA		I, NOX			I		
	hIYD1	pTrcHis	His		I, NOX			I		

*partial solubility at 20uM IPTG. At 50uM soluble protein is lost. According to JC progress report

B. Propagation of error

Error propagation for the k_{cat}/K_M values was calculated using the following equation.

C.

organism	common name	gene sequence	protein sequence
Homo sapien	human	NM_203395.2	NP_981932.1
Rattus norvegicus	norway rat	NM_001025000	NP_001020171
Sus scrofa	pig	NM_214416.1	NP_999581.1
Mus musculus	mouse	NM_027391	NP_081667
Canis lupus familiaris	dog	XM_856900	XP_533449
Pan troglodyte	chimpanzee	XM_001135488	XP_001135488
Macaca mulatta	monkey	XM_001098926	XP_001098926
Ornithorhynchus anatinus	platypus	XP_001505946	XM_001505896
Bos taurus	wild boar	NM_001102165.1	NP_001095635.1
Gallus gallus	red jungle fowl	XP_419670	XP_419670
Xenopus laevis	frog	NM_001093860	NP_001087329
Gadus morhua	atlantic cod	ES470904	
Oncorhynchus mykiss	rainbow trout	BX302875	
Danio rerio	zebrafish	XM_691419	XP_696511
Branchiostoma floridae	lancelet	XM_002237516	XP_002237552.1
Hydra magnipapillata	hydra	XM_002164492	XP_002164528
Nematostella vectensis	starlet sea anemone	XP_001633169	XM_001633119
Apis mellifera	honey bee	XM_397179	XM_397179
Anopheles gambiae	malaria mosquito	XM_315442	XM_315442.4
Drosophilia melengaster	fruit fly	NP_648433	NP_648433
Aedes aegypti	mosquito	XP_001660302	XP_001660302
Tribolium castaneum	beetle	XM_961277	XM_961277
Caenorhabditis elegans	worm	NP_498712	NM_066311.2
Haliscomenobacter hydrossis	CFB bacteria	NC_015510.1	YP_004447048.1
Daphnia pulex	water flea	GL732523.1	EFX90111.1

Bibliography

1. Zimmermann, M. B., Jooset, P.L., Pandav, C.S. (2008) Iodine-deficiency disorders., *Lancet 372*, 1251-1262.
2. Johnson, K. S., Coale, K.H., Jannasch, J.W. (1992) Analytical chemistry in oceanography, *Anal Chem 64*, 1065-1075.
3. Verhaeghe, E. F., Fraysse, A., Guerquin-Kern, J.L., Wu, T.D., Deves, G., Mioksowski, C., Leblanc, C., Ortega, R., Ambroise, Y., Potin, P. (2008) Microchemical imaging of iodine distribution in the brown alga Laminaria digitata suggests a new mechanism for its accumulation, *J Biol Inorg Chem 13*, 257-269.
4. Swingle, W. W. (1923) Iodine and amphibian metamorphosis, *Biological Bulletin 45*, 229-253.
5. Paris, M., Escriva, H., Schubert, M., Brunet, F., Brtko, j., Ciesielski, F., Roecklin, D., Vivat-Hannah, V., Jamin, E.L., Cravedi, J.P., Scanlan, T.S., Renaud, J.P., Holland, N.D., Laude, V. (2008) Amphioxus postembryonic development reveals the homology of chordate metamorphosis, *Curr Biol 18*.
6. Paris, M., Brunet, F., Markvo, G.V., Schubert, M., Laudet, V. (2008) The amphioxus genome enlightens the evolution of the thyroid hormone signaling pathway, *Developmental Genes Evolution 218*, 667-680.
7. Gaupale, T. C., Mathi, A.A., Ravikumar, A., Bhargava, S.Y. (2009) Localization and enzyme activity of iodotyrosine dehalogenase 1 during metamorphosis of frog *Microhyla ornata*, *Trends in Comparative Endocrinology and Neurobiology 1163*, 402-406.
8. Ruppert, E. E. (2005) Key characters uniting hemichordates and chordates: homologies or homoplasties?, *Can. J. Zool. 83*, 8-23.
9. Heyland, A., Hodin, J. (2004) Heterochronic developmental shift caused by thyroid hormone in larval sand dollars and its implications for phenotypic plasticity and the evolution of nonfeeding development, *Evolution 58*, 524-538.
10. Ventrui, S., Donati, F.M. Ventrui, A., Ventrui, M. (2000) Environmental iodine deficiency: A challenge to the evolution of terrestrial life?, *Thyroid 10*, 727-729.
11. De La Vieja, A., Dohan, O., Levy, O., Carrasco, N. (2000) Molecular analysis of the sodium/iodide symporter: impact on theyroid and extrathyroid pathophysiology, *Physiol Rev 80*, 1083-1105.
12. deGroot, L. J., Laarse, PR., Refetoff, S., Stanbury, J.B. (1984) *The Thyroid and its Diseases*, Wiley and Sons, New York.
13. Rokita, S. E., Adler, J.M., McTamney, P.M, Watson, J.A. (2010) Efficient use and recycling of the micronutrient iodide in mammals, *Biochimie 92*, 1227-1235.
14. Moreno, J. C., Klootwijk, W. . (2008) Mutations in the iodotyrosine deiodinase gene and hypothyroidism, *N. Engl. J. Med. 358*, 1811-1818.

15. Bianco, A. C., Salvatore, D., Gereben, B., Berry, M.J., Larsen, P.R.,. (2002) Biochemistry, cellular and molecular biology and physiological roles of the iodothyronine seleodeiodinases, *Endocr Rev. 23*, 38-89.
16. Larsen, P. R. (1997) Update on the human iodothyronine seleodeiodinases, the enzymes regulation activation and inactivation of thyroid hormone, *Biochem. Soc. Trans 25*, 588-592.
17. Kunishima, M., Friedman, J. E., and Rokita, S. E. (1999) Transition-State Stabilization by a Mammalian Reductive Dehalogenase, *J. Am. Chem. Soc. 121*, 4722-4723.
18. Watson, J. A., McTamney, P. M., Adler, J. M. (2008) Flavoprotein Iodotyrosine Deiodinase Functions without Cysteine Residues, *ChemBioChem 9*, 504-506.
19. Friedman, J. E., Watson, J. A., Jr., Lam, D. W. H., and Rokita, S. E. (2006) Iodotyrosine Deiodinase is the First Mammalian Member of the NADH Oxidase/Flavin Reductase Superfamily, *J. Biol. Chem. 281*, 2812.
20. Callebaut, I., ,Curcio-Morelli, C., Mornon, J.-P., Gereben, B., Buettner, C., Huang, S., Castro, B., Fonseca, T.L., Harney, J.W., Larsen, P.R., Bianco, A.C. (2003) The iodothyronine selenodeiodinases are thioredoxin-fold family proteins containing a glycoside hyrolase clan GH-A-like structure, *J. Biol. Chem. 278*, 36887.
21. Warner, J. R., Copley, S. D. (2007) Pre-steady-state kinetic studies of the reductive dehalogenation catalyzed by tetrachlorohydroquinone dehalogenase, *Biochemistry 46*, 13211-13222.
22. McCarthy, D. L., Navarrete, S., Willett, W.S., Babbitt, P.C., Copley, S.D. (1996) Exploration of the relationship between tetrachlorohydroquinone dehalogenase and the glutathione S-transferase superfamily, *Biochemistry 35*, 14634-14642.
23. Goswami, A., and Rosenberg, I. N. (1979) Characterization of a flavoprotein iodotyrosine deiodinase from bovine thyroid. Flavin nucleotide binding and oxidation-reduction properties, *J. Biol. Chem. 254*, 12326.
24. Rosenberg, I. N., and Goswami, A. (1979) Purification and Characterization of a Flavoprotein from Bovine Thyroid with Iodotyrosine Deiodinase Activity, *J. Biol. Chem. 254*, 12318.
25. Goswami, A., Rosenberg, I.N. (1977) Sudies on a soluble thyroid iodotyrosine deiodinase: activation by NADPH and electron carriers, *Endocrinol 101*, 331-341.
26. McTamney, P. M., Roktia, S.E. (2009) A mammalian reductive deiodinase has broad power to dehalogenate chlorinated and brominated substrates., *J Am Chem Soc 141*, 14212-14213.
27. Mani, A. R., Ippolito, S., Moreno, J.C., Visser, T.J., Moore, K.P. (2007) The metabolism and dechlorination of chlorotyrosine in vivo, *J Biol Chem 282*, 29114-29121.
28. Wu, W., Samoszuk, M.K., Comhair, S.A., Thomassen, M.J., Farver, C.F., Dweik, R.A., Kavuru, M.S., Erzurum, S.C., Hazen, S.L. (2000) Eosinophils generate brominating oxidants in allergen-induced asthma., *J Clin Invest 105*, 1455-1463.

29. Stanbury, J. B. (1957) The requirement of monoiodotyrosine deiodinase for triphosphopyridine nucleotide, *J Biol Chem 228*, 801-811.
30. Zhou, T., Taylor, M.M., DeVito, M.J., Crofton, K.M. (2002) Developmental exposure to brominated diphenyl ethers results in thyroid hormone disruption, *Toxicol Sci 66*, 105-116.
31. Thomas, S. R., McTamney, P. M., Adler, J. M., LaRonde-LaBlanc, N., Rokita, S. E. (2009) Crystal structure of iodotyrosine deiodinase, a novel flavoprotein responsible for iodide salvage in thryriod glands, *J Biol Chem 284*, 19659-19667.
32. Hecht, H. J., Erdmann, H., Park, H.J., Sprinzl, M., Schmid, R.D. (1995) Crystal structure of NADH oxidase from *Thermus thermophilus*, *Nat Struct Biol 2*, 1109-1114.
33. Tanner, J. J. L., B.; Tu, S.; Krause, K.L. (1996) Flavin Reductase P: Structure of a Dimeric Enzyme That Reduces Flavin, *Biochemistry 35*, 13531-13539.
34. Taga, M. E., Larsen, N.A., Howard-Jones, A.R., Walsh, C.T., Walker, G.C. (2007) BluB cannibalizes flavin to form the lower ligand of vitamin B12, *Nature 446*.
35. Opitz, R., Maquet, E., Zoenen, M., Dadhich, R., Costagliola, S. (2011) TSH receptor function is required for normal thyroid differentiation in zebrafish, *Mol Endocr 25*, 1579-1599.
36. Rosenberg, I. N. (1970) Purification Of Iodotyrosine Deiodinase From Bovine Thyroid, *Metabolism 19*, 785-&.
37. Gnidehou, S., Caillou, B., Talbot, M., Ohayon, R., Kaniewski, J., Noel-Hudson, M.-S., Morand, S., Agnangji, D., Sezan, A., Courtin, F., Virion, A., and Dupuy, C. (2004) Iodotyrosine Dehalogenase 1 (DEHAL1) is a Transmembrane Protein Involved in the Recycling of Iodide Close to the Thyroglobulin Iodination Site, *The FASEB Journal*, 04-2023fje.
38. Watson, J. A. (2006) Insight into the Structure and Mechanism of Iodotyrosine Deiodinase, the First Mammalian Member of the NADH Oxidase/ Flavin Reductase Superfamily. Ph.D. Dissertation, in *Department of Chemistry and Biochemistry*, University of Maryland, College Park.
39. Esposito, D., Chatterjee, D.K. (2006) Enhancement of soluble protein expression through the use of fusion tags, *Curr Opin Biotechnol 17*, 353-358.
40. Buss, J. M., McTamney, P. M., Rokita, S. E. (2012) Expression of a soluble form of iodotyrosine deiodinase for active site characterization by engineering the native membrane protein from *Mus musculus*, *Protein Sci*, 351-361.
41. McTamney, P. M. (2009) Catalytic Features of the Iodine Salvaging Enzyme Iodotyrosine Deiodinase. Ph.D. Dissertation, in *Department of Chemistry and Biochemistry*, University of Maryland, College Park.
42. Baneyx, F., Mujacic, M. (2004) Recombinant protein folding and misfolding in *Escherichia coli*, *Nat Biotechnol 22*, 1399-1408.
43. Terpe, K. (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commerical systems, *Appl Michrobiol Biotechnol 71*, 211-222.

44. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. (2002) *Short protocols in molecular biology*, 5th ed. ed., Wiley.
45. Roskams, J., Rodgers, L. (2002) *Lab Ref: A handbook of recipes, reagents, and other reference tools for use at the bench*, Cold Spring Laboratory Press, Cold Spring Harbor, NY.
46. Blackwell, J., Horgan, R. (1991) A novel strategy for production of a highly expressed recombinant protein in an active form, *FEBS 295*, 10-12.
47. deMarco, A., Vigh, L., Diamant, S., Goloubinoff, P. (2005) Native foliding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol- overexpressed molecular chaperones, *Cell Stress and Chaperones 10*, 329-339.
48. Oganesyan, N., Ankoudinova, I., Kim. S., Kim, R. (2007) Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization, *Protein Expr Purif 52*, 280-285.
49. Koziol, J. (1971) Fluorometric analyses of riboflavin and its coenzymes, *Methods Enzymol 18*, 253-285.
50. Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, in *The Proteomics Protocols Handbook* (Walker, J. M., Ed.), pp 571-607, Human Press.
51. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server, *John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press*, 571-607.
52. Warner, J. R., and Copley, S. D. (2007) Pre-Steady-State Kinetic Studies of the Reductive Dehalongenation Catalyzed by Tetrachlorohydroquinone Dehalogenase, *Biochemistry 46*, 13211-13222.
53. Rosenberg, I. N., Goswami, A. (1984) Iodotyrosine deiodinase from bovine thyroid, *Methods Enzymol. 107*, 488-500.
54. Perez-Irgtxeta, C., and Andrade-Navarro, M. (2008) K2D2: Estimation of protein secondary structure from circular dichroism spectra, *BMC Structr Biol 8*.
55. Munro, A. W., Kelly, S. M., Price, N. C. (1999) Methods in Molecular Biology, pp 111-123.
56. Edmondson, D. E., Tollin, G. (1971) Circular dichroism stuides of the flavin chromophore and of the relation between redox properties and flavin environment in oxidases and dehydrogenases, *Biochemistry 10*, 113-123.
57. Vrielink, A., Lloyd, L. F., Blow, D.M. (1991) Crystal structure of cholestreol oxidase from *Brevibacterium sterolicum* refined at 1.8 Å resolution, *J Mol Biol 219*, 533-554.
58. Fisher, A., Thompson, T.B., Thoden, J., Baldwin, T.O., Rayment, I. (1996) The 1.5 Å resolution crystal structure of bacterial luciferase in low salt conditions, *J Biol Chem 271*, 21956-21968.
59. Tu, S.-C. (2001) Reduced flavin: donor and acceptor enzymes and mechanisms of channeling, *Antioxid Redox Signal 3*, 881-897.

60. Tanner, J. J., Tu, S.C., Barbour, L.J., Barnes, C.L., Krause, K.L. (1999) Unusual folded conformation of nicotinamide adenine dinucloetide bound to flavin reducatse P, *Protein Sci 8*, 1725-1732.
61. The PyMOL Molecular Graphics System, Version 1.5.0.1 Schrödinger, LLC.
62. Solis-S., J. C., Villalobos, P., Orozco, A., valverde-R, C. (2004) Comparative kinetic characterization of rat thyroid iodotyrosine dehalogenase and iodothyronine deiodinase type 1, *J Endocrinol 181*, 385-392.
63. Kelly, S. M., and Price, N. C. (2000) The use of circular dichroism in the investigation of protein structure and function, *Curr Protein Pept Sci 1*, 349-384.
64. Bairoch, A. (2000) The ENZYME database in 2000, *Nucleic Acids Res 28*, 304-305.
65. Altschul, S. F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool, *J. Mol. Biol. 215*, 403-410.
66. Marchler-Bauer, A. (2011) CCD: a conserved domain database for the functional annotation of proteins, *Nucleic Acids Res 39*, 225-229.
67. Yang, K. Y., Swenson, R.P. . (2007) Modulation of the redox properties of the flavin cofactor through hydrogen-bonding interactions with the N(5) atom: role of alphaSer254 in the electron-transfer flavoprotein from the methylotrophic bacterium W3A1., *Biochemistry 46*, 2289-2297.
68. Fredriksson, G., Ericson, L.E. Olsson, R. (1984) Iodine binding in the endostyle of larval *Branchiostoma lanceolatum* (Cephalochordata), *Gen Comp Endocrinol 56*, 177-184.
69. Larkin, M. A. B. G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R.*, Thompson J.D., Gibson T.J. and Higgins D.G. . (2007) ClustalW and ClustalX version 2, *Bioinformatics 23*, 2947-2948.
70. Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E. L. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes, *J Mol Biol 305*, 567-580.
71. Arnold, K., Bordoli, L., Kopp, J., Schwede, T. (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homoly modelling, *Bioinformatics 22*, 195-201.
72. Harshman, S. (1979) Action of staphylococcal alpha-toxin on membranes: some recent advances., *Mol Cell Biochem 23*, 143-152.
73. Hiruta, J., Mazet, F., Ogasawara, M. (2006) Restricted expression of NADPH oxidase/peroxidase gene (*Duox*) in zone VII of the ascidian endostyle, *Cell Tissue Res 326*, 835-841.
74. Heyland, A., Moroz, L.L. (2005) Cross-kingdom hormonal signaling: an insight from thyroid hormone functions in marine larvae, *J Exp Biol 208*, 4355-4361.
75. Wheeler, B. M. (1950) Halogen metabolism of Drosophila Gibberosa, *J Exp Zool 115*, 83-107.
76. Heyland, A., Reitzel, A. M., and Hodin, J. (2004) Thyroid hormones determine developmental mode in sand dollars (Echinodermata: Echinoidea), *Evol. Dev. 6*, 382-392.

77. Wright, G. M., Youson, J.H.,. (1976) Transformation of the endostyle of the anadromous sea lamprey, *Petromyzon marinus* L., during metamorphosis. I. Light microscopy and autoradiography with I125, *Gen Comp Endocrinol 30*, 243-257.
78. Kluge, B., Renault, N., Rohr, K.B. (2004) Anatomical and molecular reinvestigation of lamprey endostyle development provides new insight into thyroid gland evolution, *Deve Genes Evol 215*, 32-40.
79. Heyland, A., Price, D.A., Bodnarova-Buganova, M., Moroz, L. (2006) Thyroid hormone metabolism and peroxidase function in two non-chordate animals, *J Exp Zool 306*, 551-566.
80. Oziol, L., Faure, P., Vergely, C., Rochette, L., Artur, Y/, Chomard, P. (2001) *In vitro* free radical scavenging capactiy of thyroid hormones and structural analogues, *J Endocr 170*, 197-206.
81. Crockford, S. J. (2009) Evolutionary roots of iodine and thyroid hormones in cell-cell signaling, *Integr Comp Biol 49*, 155-166.
82. Eales, J. G. (1997) Iodine metabolism and thyroid-related functions in organisms lacking thryoid follicles: are thyroid hormones also vitamins?, *PSEMB 214*, 302-317.
83. Heyland, A., Reitzel, A.M., Hodin, J. (2004) Thyroid hormones determine developmental mode in sand dollars (Echinodermata: Echinodidea), *Evol Dev 6*, 382-392.
84. Kawahara, T., Quinn, M., Lambeth, D. (2007) Molecular evolution of the reactive oxygen-generating NADPH oxidase (Nox/Duox) family of enzymes, *BMC Evol Biol 7*, 109-130.
85. Hajnalka Daligault, A. L., Ahmet Zeytun, Matt Nolan, Susan Lucas, Tijana Glavina Del Rio, Hope Tice, Jan-Fang Cheng, Roxanne Tapia, Cliff Han, Lynne Goodwin, Sam Pitluck, Konstantinos Liolios, Ioanna Pagani, Natalia Ivanova, Marcel Huntemann, Konstantinos Mavromatis, Natalia Mikhailova, Amrita Pati, Amy Chen, Krishna Palaniappan, Miriam Land, Loren Hauser, Evelyne-Marie Brambilla, Manfred Rohde, Susanne Verbarg, Markus Göker, James Bristow, Jonathan A. Eisen, Victor Markowitz, Philip Hugenholtz, Nikos C. Kyrpides, Hans-Peter Klenk, and Tanja Woyke. (2011) Complete genome sequence of *Haliscomenobacter hydrossis* type strain (OT), *Stand Genomic Sci. 4*, 352-360.
86. Colbourne, J. K., Pfrender,M.E., Gilbert,D., Thomas,W.K., Tucker,A.,, Oakley, T. H., Tokishita,S., Aerts,A., Arnold,G.J., Basu,M.K.,, Bauer, D. J., Caceres,C.E., Carmel,L., Casola,C., Choi,J.H.,, Detter, J. C., Dong,Q., Dusheyko,S., Eads,B.D., Frohlich,T.,, Geiler-Samerotte, K. A., Gerlach,D., Hatcher,P., Jogdeo,S.,, Krijgsveld, J., Kriventseva,E.V., Kultz,D., Laforsch,C.,, Lindquist, E., Lopez,J., Manak,J.R., Muller,J., Pangilinan,J.,, Patwardhan, R. P., Pitluck,S., Pritham,E.J., Rechtsteiner,A., Rho,M.,, Rogozin, I. B., Sakarya,O., Salamov,A., Schaack,S., Shapiro,H.,, Shiga, Y., Skalitzky,C., Smith,Z., Souvorov,A., Sung,W., Tang,Z.,, Tsuchiya, D., Tu,H., Vos,H., Wang,M., Wolf,Y.I., Yamagata,H.,, Yamada, T., Ye,Y., Shaw,J.R., Andrews,J., Crease,T.J., Tang,H.,, Lucas, S. M., Robertson,H.M., Bork,P.,

Koonin,E.V., Zdobnov,E.M.,, and Grigoriev, I. V., Lynch,M. and Boore,J.L. (2011) The ecoresponsive genome of *Daphnia pulex*, *Science 331*, 555-561.