Spring 4-21-2011

Truncated dUTPase of Dictyostelium discoideum is active and likely to form trimers

Phuoc V. Nguyen
University of Nebraska-Lincoln, gotrice1980@yahoo.com

TRUNCATED dUTPASE OF *DICTYOSTELIUM DISCOIDEUM* IS ACTIVE AND LIKELY TO FORM TRIMERS

by

Phuoc Nguyen

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Biological Sciences

Under the Supervision of Professor Catherine Chia

Lincoln, Nebraska

May, 2011

TRUNCATED dUTPASE OF *DICTYOSTELIUM DISCOIDEUM* IS ACTIVE AND LIKELY TO FORM TRIMERS

Phuoc Nguyen, M.S.

University of Nebraska, 2011

Adviser: Catherine Chia

Enzymes mediating nucleotide metabolism are potential chemotherapy targets of tumor cells. Deoxyuridine triphosphatase (dUTPase) hydrolyzes dUTP into dUMP and pyrophosphate, playing a crucial role in limiting the misincorporation of uracil into DNA. Uracil incorporated into DNA is subjected to excision, but repeated excision repair leads to DNA strand breaks and cell death. Most dUTPases are homotrimers having three active sites with each active site consisting of motif III of the first subunit, motifs I, II, and IV from the second subunit and motif V from the third subunit. Because the cell biology of the eukaryote *Dictyostelium discoideum* parallels that of mammalian cells, a study of the enzymology of *D. discoideum* dUTPase will provide insights into its regulation. The dUTPase gene was synthesized and the recombinant protein was expressed in *Escherichia coli*. A trimer of 9-kDa segment, lacking motifs I and II, was purified and confirmed by mass spectroscopic analyses and N-terminal sequencing. The enzyme has dUTPase activity with a Km of 0.566 $\pm$ 0.02 μM, a specific activity of 0.0259 μmol dUMP/min/μg at pH 8.0 and 25 ˚C, and optimal activity at pH 8.0 and 50 ˚C.

Table of Contents

Chapter 1

Introduction

DNA integrity is one of the most important if not the most important function of cells. This is because if DNA integrity is compromised it could lead to problems such as cancer. One of the problems cells have to keep in check is the incorporation of uracil instead of thymine into DNA. Uracil can be incorporated into DNA because DNA polymerases cannot distinguish between thymine and uracil (Mustafi *et al*, 2003). A low amount of uracil insertion is able to be repaired by cell, but high amounts could lead to cell death. Uracils in the DNA are subject to excision repair, but when uracil is abundant then uracil will reincorporate into DNA during repair synthesis. This will lead to double-stranded DNA breaks, and subsequently cell death. This is referred to as thymineless cell death (Vertessy and Toth, 2009).

The enzyme deoxyuridine triphosphatase (dUTPase), encoded by the *dut* gene, plays a crucial role in preventing thymineless cell death. dUTPase catalyzes the hydrolysis of deoxyuridine triphosphate (dUTP) into deoxyuridine monophosphate (dUMP) and pyrophosphate (Figure 1A). In carrying out this function, the enzyme fulfills two roles. The first is the production of dUMP, which is the precursor for the biosynthesis of deoxythymidine triphosphate (dTTP). The second role is to keep the concentration of dUTP low; therefore limiting the incorporation of uracil into DNA. Figure 1B diagrams the role of dUTPase in the nucleotide biosynthesis pathway. Briefly, ribonucleotide reductase acts on UDP to produce dUDP, which takes on another phosphate group to become dUTP. dUTPase acts on dUTP to produce dUMP. This

activity of dUTPase provides the substrate for thymidylate synthase for the conversion of uridylate to thymidylate (McIntosh and Haynes, 1997).

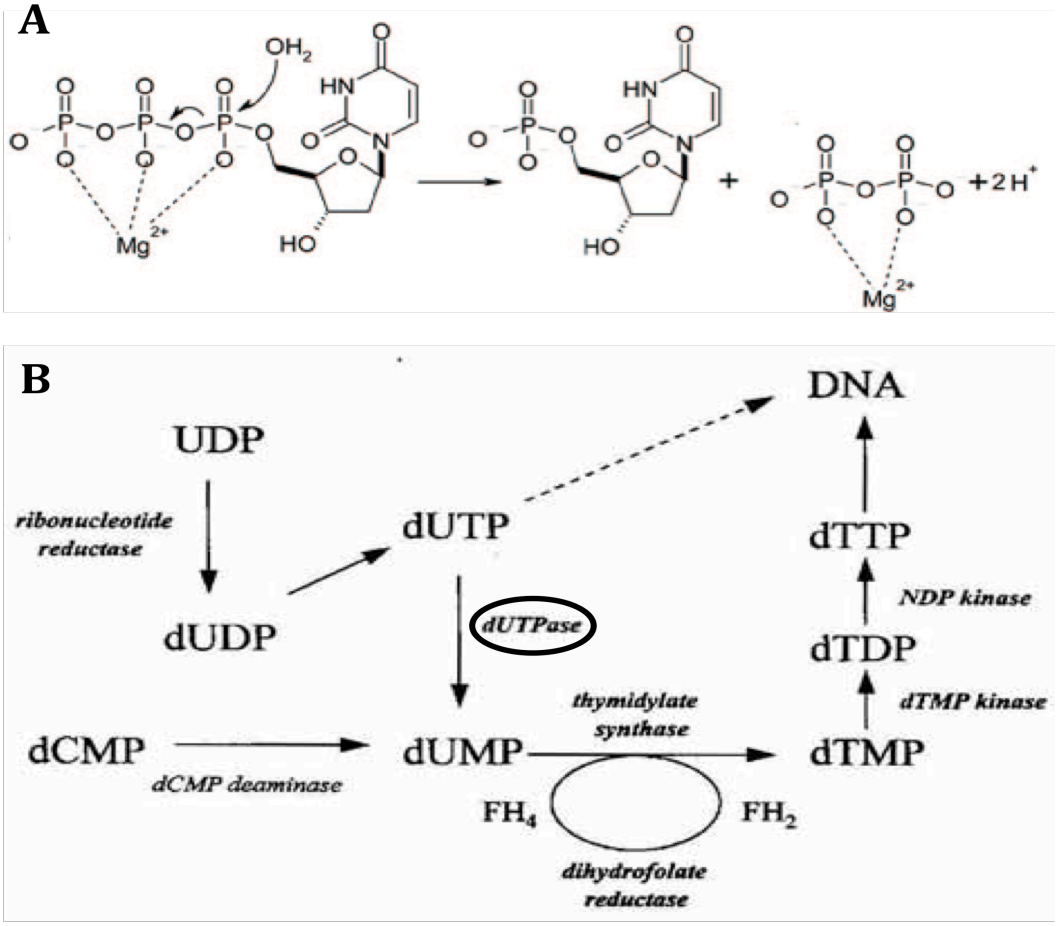


Figure 1: **dUTPase catalyzes the hydrolysis of dUTP into dUMP**. ***A***. Magnesium interacts with the phosphate groups of dUTP enabling a nucleophilic attack by a water molecule. This causes the loss of pyrophosphate releasing two protons in the process. ***B***. dUTPase serves two roles in the deoxynucleotide biosynthesis pathway: To supply dUMP for thymidylate synthase, and to minimize dUTP levels (dashed arrow). A second source of dUMP is the deamination of dCMP by dCMP deaminase. In the thymidylate synthase reaction, the methyl donor is N^5N^{10}-methylene tetrahydrofolate synthesized from tetrahydrofolate (FH_4). Dihydrofolate reductase reduces dihydrofolate (FH_2) to FH_4. Panel A was published by Vertessy and Toth (2009), and panel B was published by McIntosh and Haynes (1997).

The ratio of dTTP to dUTP plays a crucial role in thymineless cell death prevention. The availability of dTTP alone does not prevent the incorporation of uracil

into DNA because DNA polymerases cannot distinguish between the two (Mustafi *et al*, 2003). It is the relative levels between dTTP and dUTP that will prevent the misincorporation of uracil; thus, preventing cell death. The activity of dUTPase enables the cell to keep the dTTP/dUTP concentrations at the physiological ratio of 24 to 1 (Toth *et al*, 2007). As mentioned above, uracil in DNA is tolerated to a certain level because of base excision DNA repair mechanisms, but too much will lead to the breakage of double-stranded DNA and subsequently cell death (Vertessy *et al*, 1998). Studies on *Escherichia coli* and S*accharomyces cerevisiae* dUTPases showed that the activity of the enzymes is essential for the viability of the organisms (Larsson *et al*, 1996). DNA integrity and ultimately cell viability are largely depend on the activity of dUTPase.

dUTPase has been shown to be a target for anticancer and antimicrobial therapies (McIntosh and Haynes, 1997; Wilson *et al*, 2008; Webley *et al*, 2000). It has been shown that dUTPase is up-regulated in human tumor cells (Vertessy and Toth, 2009). For many years fluoropyrimidine 5-fluorouracil (5-FU) has been used as the main agent for cancer therapies, especially colorectal cancer (Longley *et al*, 2003). 5-FU inhibits to thymidylate synthase. This inhibition led to the increase concentration of dUTP within cells, and subsequently cell death or at least it decreased the tumor growth (Wilson *et al*, 2008). Since the whole purpose of 5-FU or any thymidylate synthase inhibitor is the misincorportation of uracil into DNA, the expression of dUTPase is therefore an important determinant of how effective the inhibition is (Wilson *et al*, 2008). This is because dUTPase maintains low intracellular dUTP thus preventing DNA damage associated with uracil misincorportation (Ladner *et al* 2000). dUTPase protects the cells from thymineless death (Wilson *et al,* 2008; Webley *et al*, 2000) since it catalyzes the

hydrolysis of dUTP. One study showed that the elevated expression of dUTPase prevented the expansion of intracellular dUTP pool following thymidylate synthase inhibition in breast cancer cells (Wilson *et al*, 2008). On the other hand, a second study showed that the decreased dUTPase expression enhanced dUTP pool expansion after thymidylate synthase inhibition in yeast cells (Tinkelenberg *et al*, 2002). Similarly, a third study showed that the suppression of dUTPase expression in breast and colon cancer cell lines resulted in high dUTP levels (Koehler *et al*, 2004). These evidences suggest that inhibiting thymidylate synthase alone is not fully effective, and that dUTPase is a potential therapeutic co-target with thymidylate synthase. Studying of this enzyme will give one a better understanding of how this enzyme works and the knowledge will give one the ability to come up with ways to treat cancer patients effectively. For antimicrobial therapies, the enzyme dUTPase has an even greater potential as a target. The enzyme plays a critical role in the pathogens *Plasmodium falciparum* and *Mycobacterium tuberculosis* where dUMP is solely derived from the activity of dUTPase (Vertessy and Toth, 2009). In other organisms about 60% of the dUMP pool comes from the activity of dUTPase and the remaining 40% come from dCMP deaminase (Guillet *et al*, 2006) (Figure 1B). Though a minor amount of dUMP does not come from dUTPase, but the conversion of dUTP is carried out by dUTPase. This activity of dUTPase is what prevents the misincorporation of uracil into DNA. Thus, the loss of dUTPase activity means the death of pathogens.

The *dut* gene is found in eukaryotes, prokaryotes and many viruses (Baldo *et* al, 1999). It is logical to reason that since this enzyme involved in DNA integrity, it is in the nucleus of eukaryotes. dUTPase isoforms have been found in some eukaryotic

organisms. Human cells contain nuclear and mitochondrial isoforms (Ladner *et al*, 1996). Studies showed that the two isoforms were similar in amino acid sequence and catalytically similar, having the same Km. The different between the isoforms was that the nuclear isoform was smaller in size. The smaller nuclear isoform was 30-fold more abundant than the larger mitochondrial isoform (Ladner *et al,* 1996; Merenyi *et al*, 2010). It is correct to say that the function of dUTPase is to keep the integrity of DNA and prevent cell death but what is the actual mechanism of this enzyme?

Most dUTPases are homotrimers, and the hallmark of dUTPases is five conserved motifs referred to as motifs I to V (Barabas *et al*, 2004; Vertessy & Toth, 2009). Monomeric but not dimeric dUTPases also are found to have these five motifs (Tarbouriech et al, 2005; Hidalgo-Zarco *et al*, 2001) that were shown to work cooperatively to form the active site of the enzyme (Barabas *et al*, 2004). The enzyme requires Mg^{2+} as a cofactor; studies showed that the catalytic activity of dUTPase was highest in the presence of Mg^{2+} (Mustafi *et al*, 2003). Each active site consists of motif III of the first subunit, motifs I, II, and IV from the second subunit and motif V from the third subunit. Motifs I, II, and IV coordinate the metal ion and the phosphate chain of dUTP. The initiation of the catalytic activity of the enzyme is by a nucleophilic attack on the alpha phosphate by a water molecule that is coordinated by aspartate80 of motif III (Barabas *et al*, 2004). The interactions between Mg^{2+} and the phosphate chain bend the phosphate chain into a catalytically competent conformation that enables the attack by the water molecule (Barabas *et al*, 2004). The attack of the water molecule leads to the hydrolysis of the second and third phosphate groups. The products dUMP and pyrophosphate are released from the active site. As shown in figure 2, the products

releasing was done by the action of motif V through the stacking interaction between histidine 145 and the uracil ring of dUTP (Mol *et al*, 1996; Vertessy *et al*, 1998). The active site model of figure 2 is an altered version of the *M. tuberculosis* dUTPase published by Vertessy and Toth (2009) where the normal sequence would have phenylalanine instead of histidine at the position 145 of motif V. The interaction between phenylalanine 145 of the normal sequence and the uracil ring leads to the pulling of dUMP out of the active site pocket.

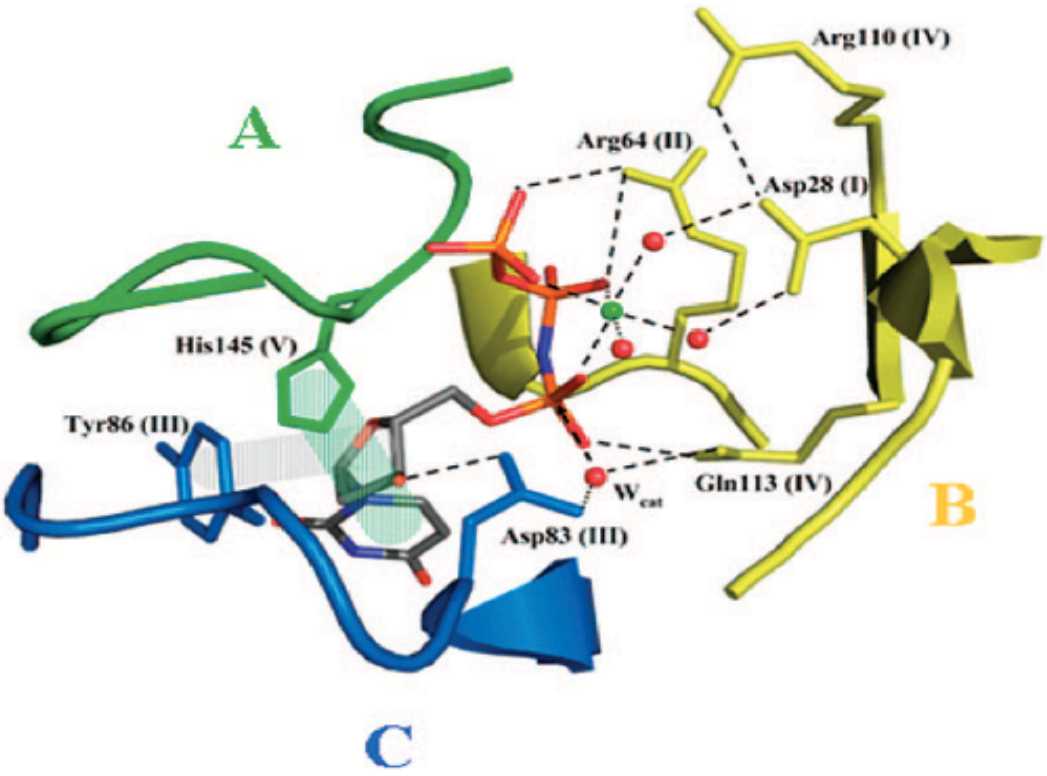


Figure 2. dUTPase activity of each active site involves all three subunits. Homotrimeric dUTPases have three active sites, each of which involves all three subunits. Each active site consists of motifs I, II, and IV from one subunit (yellow, B), motif III from another (blue, C), and motif V from the third subunit (green, A). Each catalysis involves all three subunits. The dUTPase active site model in this figure is an altered version of the M. tuberculosis enzyme published by Vertessy and Toth (2009) where the normal sequence

would have phenylalanine instead of histidine at the position 145. Residues in the normal that participate in the catalysis are: Asp28 of motif I, Arg64 of motif II, Asp83 and Tyr86 of motif III, Arg 110 and Gln113 of motif IV, and Phe145 of motif V. The substrate, dUTP, is in the middle of the figure. The red dots (spheres) are water molecules that participate to form a binding network between the substrate and the enzyme. The hydrolysis reaction is initiated by a nucleophilic attack by the catalytic water molecule (W_{cat}) coordinated by Asp83 of motif III. The green dot is Mg^{2+}, a cofactor, and dashed lines are hydrogen bonds (Vertessy and Toth, 2009).

The dUTPase in this study is from the social amoeba *Dictyostelium discoideum*. The organism is a single-cell eukaryote in which many of its genes are homologous to higher eukaryotes that are absent in other systems such as *Saccharomyces cerevisiae* (http://dictybase.org). *D. discoideum* grows as single cells but, upon stress such as lack of nutrition, up to 100,000 cells aggregate together to form a fruiting body consisting of a ball of spore cells held up by a stalk. Because of this organism's unique characteristics, it has become a great system for studying cell mobility, cell differentiation, phagocytosis, chemotaxis, and signal transduction (Bazzaro & Ponte, 1995; Gerisch & Weber, 2000; Buenemann et al, 2010). These processes are applicable to higher eukaryotes such as humans.

There had been no studies of *D. discoideum* dUTPase; therefore, all characteristics of this dUTPase are inferred through similarity. The predicted dUTPase monomer has 179 amino acids with the size of 19.5 kDa (gene ID: DDB_G0293374; http://dictybase.org). *D. discoideum* dUTPase is predicted to be a homotrimer and function to catalyze the hydrolysis of dUTP into dUMP and pyrophosphate. dUTPase has been studied extensively in human, rat and other organisms for the purposes of a deeper understanding of this important enzyme. The information from the studies of the *D. discoideum* dUTPase potentially could be used to human. The five conserved motifs

have been determined in human and rat dUTPases. *D. discoideum* dUTPase do contain the five conserved motifs that are the hallmark of dUTPases as shown by the sequence alignment with human and rat dUTPases in Figure 3. There is a 63.2% identity among the human nuclear, rat and *D. discoideum* dUTPases; therefore, knowledge about *D. discoideum* dUTPase has potential toward biomedical usage especially in humans.

Recombinant *D. discoideum* dUTPase was expressed in *E. coli* with pGEX-2T as the expression vector. The dUTPase was expressed in-frame with glutathione s-transferase (GST), enabling the purification of the GST-dUTPase fusion protein using glutathione affinity chromatography. The recombinant dUTPase lacks the first 37 residues at the N-terminal end and has an expected size of 15.19 kDa. The purpose of not including the first 37 amino acids was for structural studies. This peptide has been found to be quite flexible and it could prevent the protein from crystallizing (Nemeth-Pongracz et al, 2007; Bekesi et al, 2004; Bajaj and Moriyama, 2007). This region of dUTPase is not involved in the enzymatic activity of the enzyme.

```
H. sapiens      MPCSEETPAISPSKRARPAEVGG-----------MQLRF 28
R. norvegicus   MPCSED-PAVSVSKRARAEDD-------------ASLRF 25
D. discoideum   MPIEQKYFSLFSNLFKRLTTNNNNNNYLKMAPPNFETFKV 40
                            Motif I
H. sapiens      ARLSEHATAPTRGSARAAGYDLYSAYDYTIPPMEKAVVKT 68
R. norvegicus   VRLSEHATAPTRGSARAAGYDLYSAYDYTIPSMEKALVKT 65
D. discoideum   KKLSDKAIIPQRGSKGAAGYDLSSAHELVVPAHGKALAMT 80
                          Motif II          Motif III
H. sapiens      DIQIALPSGCYGRVAPRSGLAAKHFIDVGAGVIDEDYRGN 108
R. norvegicus   DIQIAVPSGCYGRVAPRSGLAVKHFIDVGAGVIDEDYRGN 105
D. discoideum   DLQIAIPDGTYGRIAPRSGLAWKNFIDCGAGVIDSDYRGN 120
                             Motif IV
H. sapiens      VGVVLFNFGKEKFEVKKGDRIAQLICERIFYPEIEEVQAL 148
R. norvegicus   VGVVLFNFGKEKFEVKKGDRIAQLICERILYPDLEEVQTL 145
D. discoideum   VGVVLFNHSDVDFKVAVGDRVAQLIFERIVTPEPLEVDEI 160
                     Motif V
H. sapiens      DDTERGSGGFGSTGKN------------------------ 164
R. norvegicus   DNTERGSGGFGSTGKN------------------------ 161
D. discoideum   DETQRGAGGFGSTGVK------VQN--------------- 179
```

Figure 3: **The five conserved motifs of dUTPases are found in the predicted dUTPase of *D. discoideum***. The five highly conserved motifs found in all dUTPases are highlighted for *H. sapiens* (human) (2HQU; www.pdb.org), *R. norvegicus* (rat) (P70583; www.uniprot.org) and *D. discoideum* (soil amoeba) (Q54BW5; www.dictybase.org).

The expression system was engineered with a thrombin cleavage site between GST and dUTPase, but the dUTPase amino acid sequence also had another thrombin cleavage site within its sequence. This second cleavage cut the recombinant dUTPase into two segments (Gasteiger *et al*, 2003). The shorter segment contained the front 60 residues of the protein (residues 38-97; Figure 3) with an expected size of 6.33 kDa; it is referred to as the 6-kDa segment. The longer segment contained the remaining 82 residues (residues 98-179; Figure 3) with an expected size of 8.85 kDa. This segment is referred to as the 9-kDa segment. The focus of this study is on the 9-kDa segment, which lacks motifs I and II.

Using the 9-kDa segment, I determined if the 82 residues still formed trimers. Crystal structures of trimeric dUTPases from various organisms had been solved, and the analyses identified three different types of subunit interface interactions. First, pairwise interactions are contacts between neighboring subunits at the subunit interfaces. Second, arm crossing interactions are when C-terminal arms cross over and contact the neighbor subunit. Third, contacts of the residues within the threefold axis of the trimer are called threefold interactions. The residues responsible for these interactions are scattered along the sequence, with the exception of the arm-crossings which are by the C-termini. The threefold interactions come from the interactions of residues within the active site, which are the residues of the five conserved motifs. (Cedergren-Zeppezauer *et al*, 1992;

Quesada-Soriano *et al*, 2007; Mol *et al*, 1996; Fiser & Vertessy, 2000). The findings from my studies showed that the 9-kDa segment, missing motifs I and II, forms trimers.

Studies show that dUTPase activity involves all five conserved motifs (Toth *et al*, 2007; Barabas *et al*, 2004; Takacs *et al*, 2010). From the evolutionary point of view, the five motifs are highly conserved among organisms including viruses that encoded the *dut* gene; therefore, these five motifs must be functionally important. I asked if the 9-kDa segment, without motifs I and II, has dUTPase catalytic activity. A single aspartic acid149/asparagine change in motif I drastically decreased the catalytic efficiency of human and mycobacterial dUTPases (Takacs *et al*, 2010). The mutation of serine72 to alanine in motif II resulted in a catalytically compromised enzyme in *E. coli* (Palmen *et al*, 2008). A third study showed that the deletion of the C-terminus carrying motif V resulted in an Epstein-Barr virus protein without enzymatic activity (Freeman *et al*, 2009). The 9-kDa segment of my studies showed dUTPase activity.

My experimental results show that the 9-kDa segment of *D. discoideum* dUTPase forms trimers and, in the absence of motif I and II, had catalytic activity. The measured Km of 0.566 ± 0.02 μM indicated a high substrate affinity, and the optimal pH and temperature for activity were comparable to the catalytic activities of dUTPases of other organisms.

Chapter 2

Materials and Methods

The diagrams below outline the strategies taken for the growth of *E. coli* BL21 that contained the expression vector pGEX-2T with the inserted dUTPase gene (Figure 4) and the purification of the 9-kDa segment of the recombinant *D. discoideum* dUTPase (Figure 5).

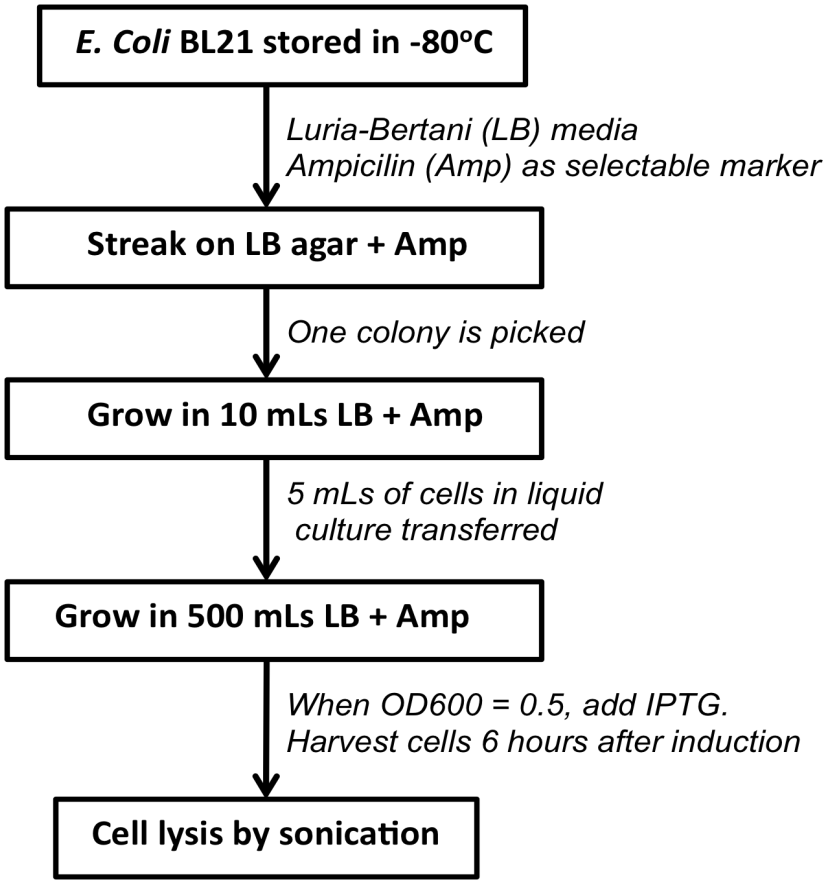


Figure 4: **Expression of recombinant *D. discoideum* dUTPase**. A whole cell lysate was prepared from *E. coli* BL21 transformed with an inducible expression plasmid containing the inserted dUTPase gene.

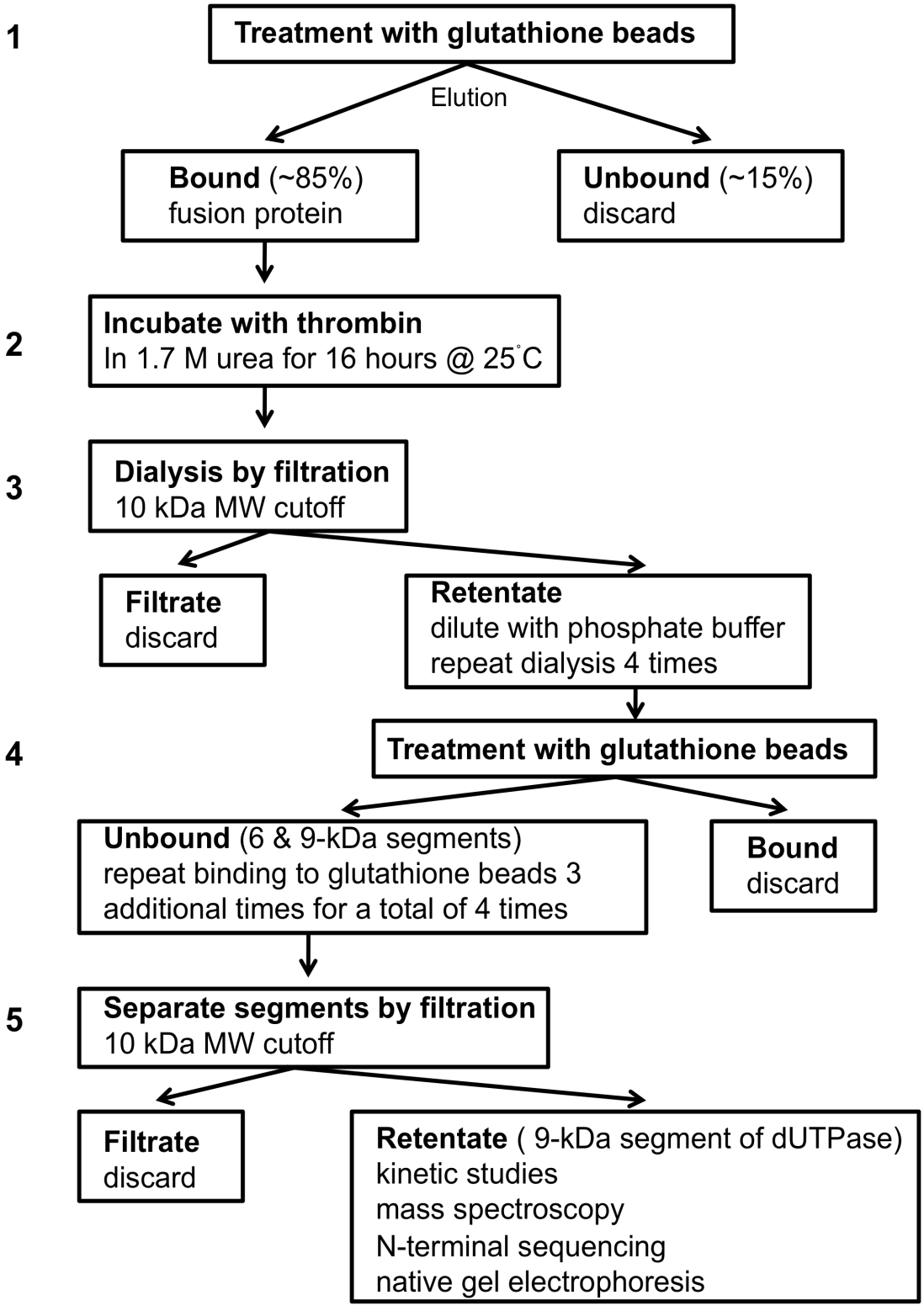

1
Treatment with glutathione beads
Elution
Bound (~85%)
fusion protein
Unbound (~15%)
discard
2
Incubate with thrombin
In 1.7 M urea for 16 hours @ 25˚C
3
Dialysis by filtration
10 kDa MW cutoff
Filtrate
discard
Retentate
dilute with phosphate buffer
repeat dialysis 4 times
4
Treatment with glutathione beads
Unbound (6 & 9-kDa segments)
repeat binding to glutathione beads 3
additional times for a total of 4 times
Bound
discard
5
Separate segments by filtration
10 kDa MW cutoff
Filtrate
discard
Retentate (9-kDa segment of dUTPase)
kinetic studies
mass spectroscopy
N-terminal sequencing
native gel electrophoresis

Figure 5: **Steps of the purification of the 9-kDa segment of the recombinant *D. discoideum* dUTPase**. Step numbers are on the left. The purification process started with the isolation of the fusion protein from lysed cells (Figure 4). The 85% bound and 15% unbound of step 1 were the estimated percentages of GST-dUTPase fusion protein recovered and lost, respectively, from a typical cell lysate, derived from 5 g cells, after using a glutathione affinity matrix. Samples were run on SDS-gels, and after Coomassie Blue Staining, analyzed by ImageJ. This program used the gel image where pixel was calculated from the gel lane and each bands within that lane, enabling estimations on protein recovered and lost. Step 1 = treatment with glutathione beads; Step 2 = incubate with thrombin; Step 3 = dialysis; Step 4 = treatment with glutathione beads; Step 5 = separate segments by filtration.

2.1 dUTPase Cloning

The *Dictyostelium discoideum* dUTPase gene was synthesized by GenScript (GenScript Corporation, Piscataway, NJ) and inserted into vector pUC57 using restriction sites EcoRI and BamHI. The plasmid was transformed into *E. coli* Top 10 cells (Invitrogen, Carlsbad, CA) of the genotype F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG. Plasmid DNA was purified using a Maxiprep kit (Qiagen Inc., Valencia, CA), digested with EcoRI and BamHI, and run on a 0.8% low melt agarose gel with TE (10 mM Tris-Cl, pH 7.5, 1 mM EDTA) buffer. The band corresponding to the insert was excised and transferred to a sterile centrifuge tube. The DNA was extracted from the agarose gel (Figure 1 Appendix) using a Gel Cleanup Kit (Eppendorf, Hamburg, Germany).

The purified insert was ligated into the expression vector pGEX-2T (GE Healthcare, Uppsala, Sweden) that would generate a GST-fusion protein. The pGEX-2T expression vector was a system under the control of the *tac* promoter. This promoter was

to be induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) when ready for the high-level protein expression. Before ligation, the pGEX-2T plasmid was collected from *E. coli* Top10 cells using a Plasmid Maxi Kit (Qiagen, Valencia, CA) and linearized by digestion with EcoRI and BamHI. The expression vector and insert were quantified using Hoechst dye with calf thymus DNA as standards according the standard protocol (Sambrook and Russell, 2001). At this point the insert and the linearized vector were combined in different ratios. The molar ratios of vector to insert used were: 1:1, 1:3, 1:10 and 1:25. Ligase reaction buffer and ligase were added and incubated for 1 hour at room temperature (measured at 23˚C).

Ligation mixtures were added to *E. coli* Top 10 cells (Invitrogen). The process of transformation was followed according to standard protocol (Sambrook and Russell, 2001). Briefly, 2 µL from each of the different ligation reactions were added to vials each containing 50 µL of competent cells. The mixtures were incubated on ice for 30 minutes, and transferred to a 42˚C water bath for 30 seconds. Vials containing the mixtures were then put back onto ice, and 250 µL of Super Optimal broth with Catabolite repression (SOC) medium that had been warmed to 37˚C were added to each vial. The vials were then put into a 37˚C shaking water bath for 1 hour. When one hour was up, the contents of each vial were plated on LB agar with ampicillin (100 µg/mL). The plates were incubated at 37˚C overnight. The next day, each plate was checked for colonies.

Colonies were grown in flasks containing 5 mL of liquid LB with ampicillin (100 µg/mL) overnight at 37˚C. Plasmid DNA was extracted from cells using a FastPlasmid Mini kit (Eppendorf), digested with EcoRI and BamHI, and run on a 0.8% agarose gel with TE buffer. With the positive results, the extracted plasmid was transformed into *E.*

coli BL21 cells (as shown in Figures 2 and 3 Appendix) according to the standard protocol described above.

2.2 Cell Growth and Collection

E. coli BL21 cells transformed with the pGEX-2T-dUTPase plasmid were stored at -80^{o}C. When ready to grow, cells were streaked onto a LB agar plate with ampicillin (100 µg/mL). This process was done in a sterile environment, preferably the hood. The plate was incubated overnight at 37 ˚C. A single colony was picked with a sterile toothpick and placed into a flask containing 10 mL of liquid LB medium with ampicillin (100 µg/mL). This was the seed culture. The flask was incubated at 37 ˚C with shaking at 200 rpm for 4 hours or when the liquid culture started to get cloudy. Five milliliters of the seed culture was transferred into 500 mL of liquid LB plus ampicillin (100 µg/mL). This large culture was incubated at 37 ˚C with shaking at 200 rpm and the absorbance at 600 nm was read every 30 minutes. Absorbance readings were plotted against time to project when the absorbance would reach 0.5. Absorbance was read again at the predicted time for comparison. At this point the cells were ready for the induction of protein expression (refer to Figure 4).

Solid IPTG (Invitrogen, Carlsbad, CA), which served to induce protein expression, was dissolved with sterile water to a concentration of 1 M. The 1M IPTG was added to the large culture to a final concentration of 1 mM. After the addition of IPTG, the large culture was grown for an additional of 6 hours before cells were harvested by centrifugation at 10,000 x g for 15 minutes at 4 ˚C (Figure 4). Cells were stored as pellets at -80 ˚C.

2.3 Protein Extraction

Cell pellets were combined with binding buffer (20 mM potassium phosphate (Kpi), 1.5 mM β -mercaptol ethanol, pH 7.4) at a ratio of 10 mL of binding buffer per gram of cell mass. For example, 5 grams of cells and 50 mL of binding buffer were added into a stainless steel container. To inhibit proteolysis, phenylmethylsulfonyl fluoride (PMSF; 750 µL of 100 mM) and ethylenediaminetetraacetic acid (EDTA; 150 µL of 0.5 M) were added to a final concentration of 1.5 mM each. The container was put on ice water to prevent heating during sonication. The cells were lysed by sonication for 5 minutes (30 Watt with 10 seconds sonication and 20 seconds cool) using a Sonicator 3000 (Misonix Incorporated, Farmingdale, NY). At this point, it was estimated by examining the remaining cell pellet that 95% of the cells were lysed, and protein had been released into the solution. Centrifugation (15,000 x g for 60 minutes at 4 ˚C) was carried out to separate liquid from cell debris. The supernatant (45 mL) was collected and filtered using a Whatman 0.2 µm polysulfone (PSU) with glass microfilter (GMF). The cell-free extract or lysate was always stored on ice or at 4˚C, and used immediately for protein purification.

2.4 Protein Purification

The lysate was treated with glutathione beads (GE Healthcare, Uppsala, Sweden) (Step 1, Figure 5). When passed through the column containing glutathione beads, the GST-fused dUTPase (fusion protein) will bind to the beads and all other contaminants will pass through. To make sure no contaminant remained in the column, the column was washed thoroughly with binding buffer. The bed volume of the glutathione beads in the column was 7 mL, so the volume of binding buffer used for washing was 20 times the

bed volume, which was 140 mL. One thing to remember was to make sure the glutathione bead column was washed off of all debris and equilibrated with binding buffer before use.

Fractions were collected with a fraction collector (150 drops per 13 X 100 mm tube). After the column had been washed thoroughly, elution buffer (5 mM Tris-HCl, 1.5 mM β -mercaptol ethanol, pH 8.2) was added to the column with a volume of 10 times the bed volume, which was 70 mL, to release the fusion protein from the glutathione beads.

The absorbance at 280 nm of fractions was measured to determine the protein elution profile. The fractions collected from the addition of elution buffer were where the fusion protein resided. This was determined by the peak absorbance readings. In this case, fractions 32-37 were collected. The total volume collected was 35 mL with absorbance at 280 nm of 2.6483. This was equal to 92.7 mg of fusion protein based on OD_{280} of 1 equals 1 mg of protein per mL.

To prepare the fusion protein for thrombin cleavage, urea was added (Step 2, Figure 5). It is important to point out that the GST fused dUTPase protein has two thrombin cleavage sites (Figure 6; Gasteiger *et al*, 2003). One site cleaves GST from dUTPase, and the other site is within motif II of the dUTPase amino acid sequence, that results in the 6-kDa and 9-kDa segments of dUTPase. Seven milliliters of 10 M urea (in water) were added to make a final concentration of 1.7 M. Thrombin (36432 unit/mg; Haematologic Technologies Inc., Essex Junction, VT) was added to a final concentration of 10 unit/mg. In this case 25.44 μL was added and incubated at room temperature (measured at 21˚C) for 16 hours with shaking.

After the incubation of the GST-dUTPase fusion protein with thrombin, 1 M Kpi buffer, pH 7.4 was added to 20 mM. The preparation was then dialyzed to reduce the urea concentration (Step 3, Figure 5) using a 10 kDa cutoff Centricon filter (Amicon Inc. Beverly, MA; 1,500 x g for 1 hour at 4 ˚C). The filtrate was discarded. The retentate was diluted with binding buffer to its original volume, and the filtration step was repeated four more times for a total of five times. The final urea concentration was less than 1 mM based on the total volume of binding buffer (37 mL) added after each filtration for the five filtrations. The pH of the preparation was checked to make sure it was at pH 7.4 before continuing to the next step.

The preparation then passed through the glutathione bead column, which had been washed and equilibrated (Step 4, Figure 5). The unbound fractions (75 drops per fraction) containing the 9-kDa and 6-kDa segments of dUTPase were collected and pooled, and the bound fractions eluted with elution buffer containing GST and fusion protein were discarded. This step was repeated three more times for a total of four times. Protein elution profiles were monitored by measuring the absorbance of fractions at 280 nm.

The preparation that contained both the 9-kDa and 6-kDa segments was further purified by filtration (Step 5, Figure 5). A 10 kDa cutoff filter (Amicon Inc, Beverly, MA) was used to separate the 9-kDa segment from the 6-kDa segment (1,500 x g for 1 hour at 4 ˚C). The filtrate was discarded and the retentate containing the 9-kDa segment was collected and used for kinetic studies, SDS-PAGE and native gel electrophoresis, mass spectroscopic analyses and N-terminal sequencing. From 92.7 mg of fusion

protein, 1.1 mg of the 9-kDa segment of dUTPase was obtained, a yield of 1.2% (see Table 1).

2.5 BCA Protein Assay

Protein concentrations were measured using the BCA protein assay (Pierce, Rockford, IL) according to the BCA protein assay manual. Briefly, protein standards of known concentration were prepared by diluting a stock standard solution of bovine serum albumin (BSA) with binding buffer. Working Reagent was prepared in the ratio of 50 parts Reagent A: 1 parts Reagent B. To run the assay, 0.1 mL of each standard or sample were added into labeled tubes. For the blank, 0.1 mL of binding buffer was used. Next, 2 mL of Working Reagent was added to each tube and mixed well. The tubes were incubated at room temperature for 2 hours. The absorbance at 562 nm of the assay mixtures, with water as reference, was measured. The blank was subtracted from standards and samples. A standard curve was a plot of the (blank corrected) net absorbance at 562 nm vs. protein concentration. Using the equation generated from the best fit trendline, the protein concentration of the sample was determined. In this case the concentration of the 9-kDa segment dUTPase was 80 µg/mL. The protein concentration estimated from absorbance at 280 nm values with the assumed 1 absorbance unit equals 1 mg of protein per milliliter was fairly accurate. The difference between the protein concentration using the BCA protein assay and absorbance at 280 nm was small. For the eluted GST-dUTPase fusion protein (Step 1, Figure 5) the BCA protein assay produced a concentration of 1.28 mg/mL and the absorbance at 280 nm was 1.21, a difference of 5.5%. For the 9-kDa segment the results were 0.08 mg/mL from the BCA assay and 0.09 mg/mL from absorbance at 280 nm, a difference of 11%.

2.6 Kinetic Analysis

2.6a Temperature Optima

For kinetic studies, an HPLC method measured the product (dUMP) produced by dUTPase. First a standard was determined for the product (dUMP). Both the product dUMP and the substrate dUTP were available commercially from Jena Bioscience (Jena, Germany). dUTP (100 mM in sodium salt) was diluted with 10 mM Bicine buffer, 2 mM MgCl, pH 8.0 to a final concentration of 1 μM. To run a standard assay, 990 μL of 1 μM dUTP were added to a 13 x 100 mm test tube kept in a 25˚C water bath for 5 minutes before the addition of dUTPase. If the experiment was to be carried out at different temperatures, then tubes containing 1 μM dUTP were incubated for 5 minutes at the particular temperature before the addition of enzyme. In this case the temperatures were 20, 30, 40, 50, 60 65, 70 and 80˚C. After 5 minutes of incubation, 10 μL of dUTPase (0.0466 μg) were added. The mixture was incubated for 10 minutes and the reaction was quenched with the addition of 20 μL of 1 M HCl. Two hundred microliters of the reaction solution were chromatographed on a Mono Q 5/50 GL anion exchange column (1 mL bed volume; GE Healthcare, Giles, UK) on BioCAD Vision Workstation HPLC (GMI Inc., Ramsey, MN). The column was washed with a gradient of 10 to 500 mM KCl, 5 mM Kpi, pH 8 with a flow rate of 2 mL per minute. The deoxynucleotides were eluted as the salt concentration increased with dUMP first and then dUTP (see Figure 7 Appendix for an example chromatogram). The peak areas of the HPLC profile (chromatogram) was integrated using Igor Pro Software (WaveMetrics, Lake Oswego, OR) and compared to the standard curve to determine the amount of dUMP produced.

2.6b pH Optima

To measure the effect of pH on dUTPase activity at 25˚C, the same HPLC assay described was used with different buffers. In this case, the buffers were citrate buffer for pH 1-5, MES for pH 6, HEPES for pH 7, TrisHCl for pH 8-9, and CAPS for pH 10-11.

2.6c Km Determination

To determine Km of the 9-kDa segment, a method to measure the amount of protons produced by the reaction was monitored indirectly by changes in the absorbance of cresol red (Freeman *et al*, 2009). This slope was considered as the rate of the enzyme activity because the change in the absorbance decreased as protons are produced (Freeman *et al*, 2009). dUTP solutions at the final concentrations of 0.1, 0.3, 0.5, 0.7, 1.0, 3.0, 5.0, and 7.0 μM were prepared by diluting the 100 mM stock dUTP with 10 mM Bicine buffer, 2 mM MgCl, pH 8.0 in the 1 mL reaction volume.

Cresol Red (Sigma-Aldrich Corp., St. Louis, MO) was dissolved with 10 mM Bicine buffer, 2 mM MgCl, pH 8.0 to a concentration of 25 μM. To run the assay, 985 μL of 25 μM cresol red was added first into a cuvette, followed by 5 μL of the appropriate dUTP solution, and lastly, 10 μL of dUTPase (0.0466 μg) was added. The absorbance change was monitored at 573 nm using the Kinetics program version 3.0.0.0 and a Cary 50 spectrophotometer (Varian Inc., Santa Clara, CA) over one minute with 0.1 second increments. The progress curve data were recorded, and the rate of absorbance change in the linear portion was determined. The time points used for rate measurements were the same among replicates.

2.7 Gel Electrophoresis

Sodium dodecyl sulfate polyacrylamide (SDS-PAGE) and non-denaturing (native) gel electrophoresis were carried out to examine the size of the 9-kDa segment. The 16.25% acrylamide; 1.5 mm thick SDS gel was prepared using the Hoefer Dual Gel Caster (Pharmacia Biotech, San Francisco, CA). The running conditions were 100 Volts for 2 hours with 1X Tris Glycine Running Buffer (TGRB) containing 0.1% SDS. The 1X TGRB was diluted from a 4X TGRB stock pH 8.3 (0.10 M Tris and 0.77 M glycine in water) with deionized water. The gel was than stained with Coomassie Blue R-250 (0.25 g/L Coomassie R-250, 50% methanol, 10% acetic acid, 40% deionized water). The non-denaturing 15% acrylamide gel was prepared without SDS. The standards were prepared from four different proteins: bovine serum albumin (66 kDA), ovalbumin (45 kDa), trypsin inhibitor (soybean; 20 kDa), and β-lactoglobulin-A (18 kDa). The non-denaturing gel was run at 100 Volts for 2 hours with 1X Tris Glycine Running Buffer without SDS. The gel was then stained with Coomassie Blue R-250 overnight and destained with destaining solution (40% methanol, 10% acetic acid, 50% deionized water). *Arabidopsis* dUTPase, having a molecular weight of 53.6 kDa (PDB ID: 2P9O; Bajaj and Moriyama, 2007) in its trimer form, was used as positive control in running the native gel.

2.8 Sample preparation for Mass Spectroscopic analyses and N-terminal Sequencing

SDS-PAGE was done to prepare samples for mass spectroscopic analyses and N-terminal sequencing. The 16.25% acrylamide SDS gels were prepared 1.5 mm thick using the Hoefer Dual Gel Caster. After electrophoresis the gel then was stained with Coomassie Blue R-250 (as described) for 1 hour and destained with destaining solution. The bands (2.4 μg per lane) corresponding to the 9-kDa segment dUTPase were excised

and put into a microcentrifuge tube containing 10% methanol, 90% deionized water. The sample was sent to Nebraska Center for Mass Spectrometry (University of Nebraska-Lincoln) for MS/MS analysis.

The gel for N-terminal sequencing analyses was used to transfer the protein onto Polyvinylidene fluoride (PVDF). After the run, the gel was soaked in transfer buffer (10 mM Tris base, 100 mM glycine, 10% methanol, pH 8.3) for 3 minutes. High retention PVDF (Millipore, Bedford, MA 01730) was pre-wet in methanol, and the blotting paper, both thin and medium, were also soaked in transfer buffer. The protein transfer apparatus used was the semi-dry called Semi-Phor (Hoefer Scientific, San Francisco, CA). The sandwich was set up with 2 medium and 1 thin blotting papers, PVDF, gel, and 1 thin and 2 medium blotting papers in order from anode to cathode. The running conditions were 40 mA for 3 hours. The PVDF was then removed, rinsed with deionized water, and immediately stained with Coomassie Blue (0.2 mg/L Coomassie R-250, 20% methanol, 0.5% acetic acid) for 5 minutes. The blot was then destained in 30% methanol, 70% deionized water, until the protein bands became visible, and then washed three times with deionized water, 5 minutes each. The blot was dried in a dust-free location. The bands corresponding to the 9-kDa segment dUTPase (2.4 µg per lane) was excised and transferred into the microcentrifuge tube. The sample was sent to the Protein Structure Core Facility of the University of Nebraska Medical Center (Omaha, NE) for N-terminal sequencing.

2.9 Calculations carried out for the analysis of the data.

Protein Amount: from gram to mole and molarity

GST-dUTPase Fusion Protein (42 kDa)

42 kDa = 42000 g/mole

(measured protein, gram) x (1 mole / 42000 gram) = moles of GST-dUTPase fusion protein.

(mole of fusion/mL of solution) X (1000 mL / 1 L) = Moles per liter GST-dUTPase fusion protein.

9-kDa segment (8.85 kDa)

8.85 kDa = 8850 g/mole

(measured 9-kDa segment, gram) x (1 mole / 8850 gram) = moles of 9-kDa segment.

(mole of 9-kDa segment/mL of solution) X (1000 mL / 1 L) = Moles per liter 9-kDa segment.

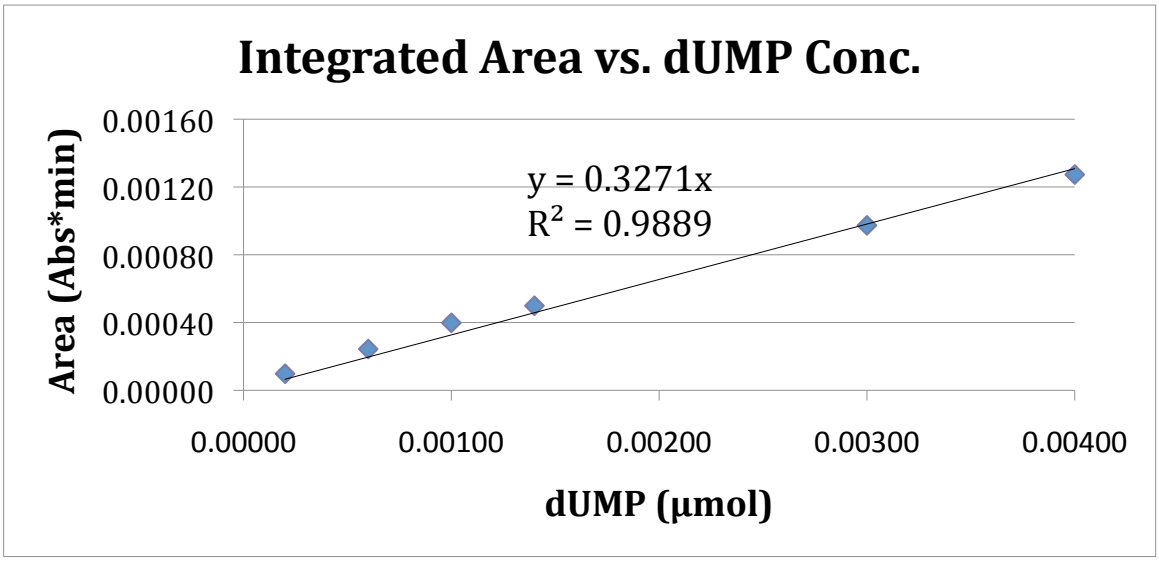
Enzyme Activity

HPLC method

From the chromatogram from each reaction, the area of the peak corresponding to dUMP was integrated using Igor software. The value of the area was used to calculate the dUMP concentration using the equation from the dUMP standard curve.

Equation from standard curve:
y (Abs*min) = x (µmol dUMP) + (Abs*min)
Using this equation to plug in the measured area for y value to obtain the x value, which is the amount of dUMP in micromoles.

Below is an example dUMP standard curve.



Since the injection volume was 200 µL and the total enzyme reaction volume was 1 mL:
Calculated µmol dUMP X 5 = µmol dUMP in 1 mL Reaction

Specific Activity:
Total µmol dUMP / 10 minutes / µg 9-kDa segment = Specific Activity of the 9-kDa segment of dUTPase (µmol dUMP/min/ µg 9-kDa segment).

Cresol Red method (Absorbance at 573 nm):
Enzyme activity was observed on the progress curve. Point A and B were picked when the curve was linear. This represented the initial velocity of enzyme.

(Δ Abs) / (Δ time) = 9-kDa segment arbitrary velocity (Δ Abs/s)

Km Determination was calculated from using Michaelis-Menten equation using the Visual Enzymatics software :
$v=(V_{max} \text{ X } [S])/(K_m + [S])$

This equation describes the rate of substrate conversion. In this equation, V is the rate of conversion, V_{max} is the maximum rate of conversion, [S] is the substrate concentration, and K_m is the Michaelis constant.

Chapter 3

Results

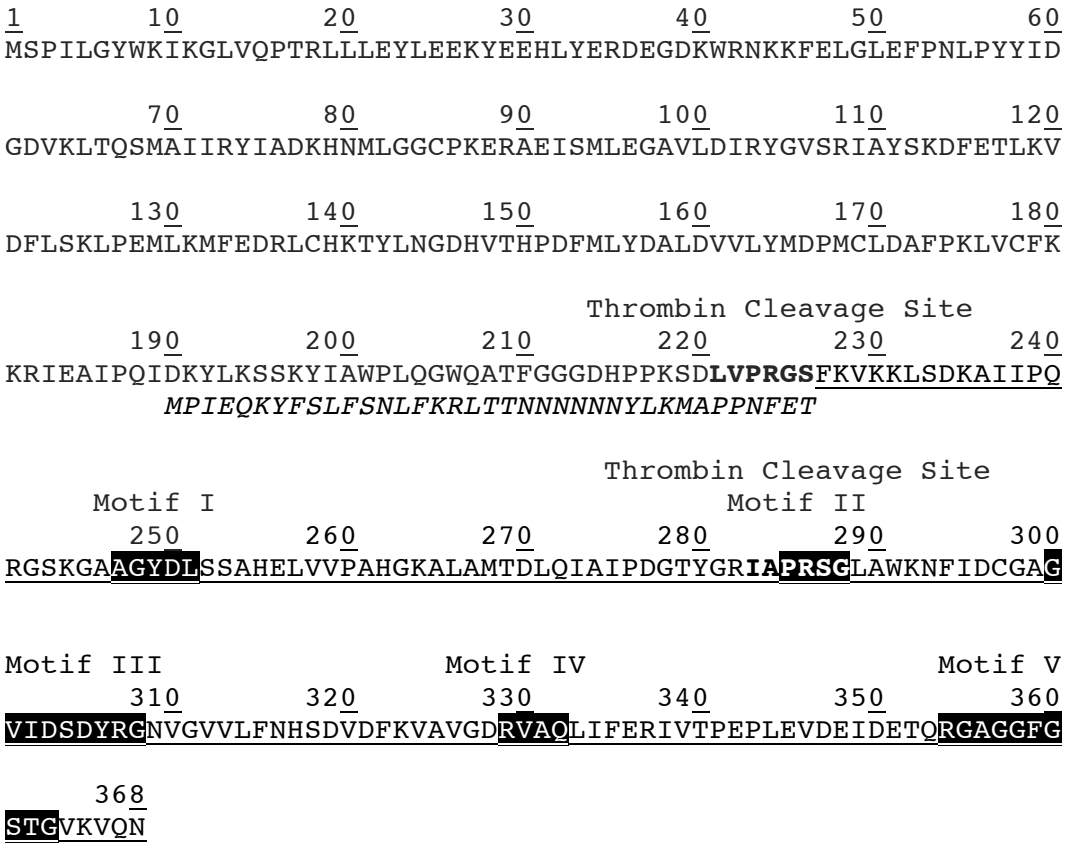


Figure 6: **Predicted amino acid sequence of the GST-dUTPase fusion protein and thrombin cleavage sites.** GST (residues 1-220) is linked to recombinant dUTPase (residues 227-368; underlined) by one thrombin cleavage site **LVPRGS**. A second thrombin cleavage site (**IAPRSG**), in motif II, splits the dUTPase into two polypeptides referred to as the 6-kDa and 9-kDa segments. The two thrombin recognition sites have different amino acid sequences because thrombin recognizes more than one sequence; though characteristically they follow the same general pattern (Gasteiger *et al*, 2003). The five conserved motifs are shaded black. The full length (native) *D. discoideum* dUTPase is predicted to have an additional 37 residues at the N-terminus (italicized; DDB_G0293374; www.dictybase.org).

The synthesized *D. discoideum* dUTPase protein was smaller than the full length. The reason for omitting the 37 residues at the N-terminal was because this is a flexible segment (Nemeth-Pongracz *et al*, 2007; Bekesi *et al*, 2004; Bajaj and Moriyama, 2007) of the protein (Figure 6) and potentially could interfere with protein crystallization, a long-term goal of this project.

The system for dUTPase expression and purification was the GST fused system. This system produced dUTPase fused to GST with an engineered thrombin cleavage site between GST and dUTPase. Protein purification was done using affinity chromatography that takes advantage of the binding of GST to glutathione. The cleavage of dUTPase from GST was done by the serine protease thrombin. The predicted size of the GST-dUTPase fusion protein is 42 kDa, and the recombinant dUTPase is 15 kDa.

3.1 Protein Purification

3.1a Inclusion of 1.7 M urea greatly increased the efficiency of thrombin cleavage

After cells were induced to express protein, they were lysed and the lysate was treated with glutathione beads (Glutathione Sepharose) to bind the GST-dUTPase fusion protein (Figure 7). The fusion protein was cleaved with thrombin to separate dUTPase from GST (Figure 7). Thrombin cleavage efficiency was low therefore different treatments were tested to increase the cleavage efficiency. It was roughly estimated that 50% of all the fusion protein were cleaved by thrombin as shown in Figure 7 (Image J) (Abramoff *et al*, 2004). One important point was that when thrombin cleavage efficiency was low, the subsequent protein purification steps were unable to detect the presence of dUTPase because a small amount of protein was lost in each preparation step. Also, when high percentage of fusion protein was left un-cleaved, the subsequent step always

contained some fusion protein contaminant. The inclusion of 1 and 2 M NaCl, and 1.7 and 2.9 M urea were tested. The addition of NaCl did not increase the cleavage. On the contrary, the addition of NaCl resulted in no to faint detection of cleaved GST or dUTPase as shown in Figure 4 Appendix.

Urea was shown to increase the thrombin cleavage efficiency greatly (Figure 4 Appendix). Thrombin cleavage efficiency of the fusion protein was low; therefore, urea, a protein denaturant, was added to serve as a possibility to solve this problem. By carrying out thrombin digestion of the recombinant protein in the presence of protein denaturant will help to expose the cleavage site to the enzyme that might have gotten covered up (Novagen Thrombin Kit; Novagen, Madison, WI). Both the 1.7 M and 2.9 M urea increased the efficiency; thus, 1.7 M was chosen because the 2.9 M was a high amount of salt in the solution and unnecessary when 1.7 M did the job. A comparison of thrombin-cleaved fusion protein with and without urea is shown in Figure 7. It was estimated that 95% of all fusion protein were cleaved by thrombin in the presence of 1.7 M urea as compared to 50% without urea (Image J) (Abramoff *et al*, 2004). The band corresponds to dUTPase ran faster than predicted is because dUTPase has been cleaved into two smaller segments that were not separated on a 15% acrylamide SDS gel. The segments are shown when ran on a 16.25% acrylamide gel as shown in Figure 8.

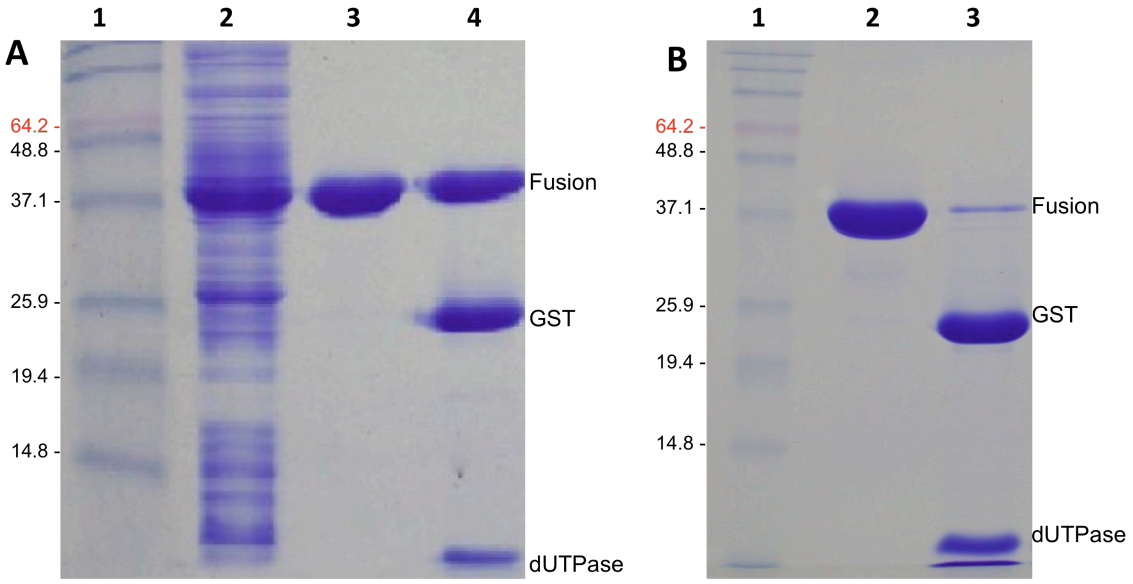


Figure 7: **Efficient cleavage of GST-dUTPase fusion protein by thrombin in the presence of urea (15% acrylamide SDS-gels)**. ***A***. Lane 1: molecular markers in kDa. Lane 2: lysate (12.4 µg) showing the major species was the GST-dUTPase fusion protein (fusion). Lane 3: GST-dUTPase fusion protein (10.1 µg) before the addition of thrombin. Lane 4: Fusion protein (10.1 µg) cleaved with thrombin; the species are GST-dUTPase fusion protein, GST and dUTPase from top to bottom. ***B***. Lane 1: molecular markers in kDa. Lane 2: GST-dUTPase fusion protein (9.8 µg) before thrombin cleavage. Lane 3: Fusion protein (9.7 µg) cleaved with thrombin in the presence of 1.7 M urea; the species are GST-dUTPase fusion protein, GST, dUTPase and dye-front from top to bottom.

3.1b 9-kDa segment of dUTPase was obtained

Thrombin cleaved GST from dUTPase from the first thrombin cleavage site and dUTPase into two smaller segments from the second cleavage site. It was observed that after the fusion protein was incubated with thrombin there were two smaller bands appeared below the predicted size of the recombinant dUTPase as shown in Figures 8 and 5 Appendix. Using Expasy tool (Gasteiger *et al*, 2003), it was determined that there was a secondary thrombin cleavage site within the dUTPase sequence. Thrombin cleaved the

recombinant dUTPase into two segments, referred to as the 6-kDa and the 9-kDa segments of *D. discoideum* dUTPase. The actual sizes of the 6-kDa and 9-kDa segments, based on the location of the thrombin cleavage sites, were predicted to be 6.34 and 8.85 kDa, respectively. After the fusion protein was cleaved with thrombin, the 6-kDa and the 9-kDa segments were obtained through treatments with glutathione beads (Figure 8; Figure 5 Appendix).

A 10-kDa cutoff filter effectively separated the 9-kDa from the 6-kDa segment, suggesting that the 9-kDa segment was actually in its trimeric form (predicted size of 26.55 kDa). The 9-kDa segment was purified and was used for kinetic analysis (Figure 8). A BCA protein assay determined the concentration of the 9-kDa segment of dUTPase was 80 µg/mL.

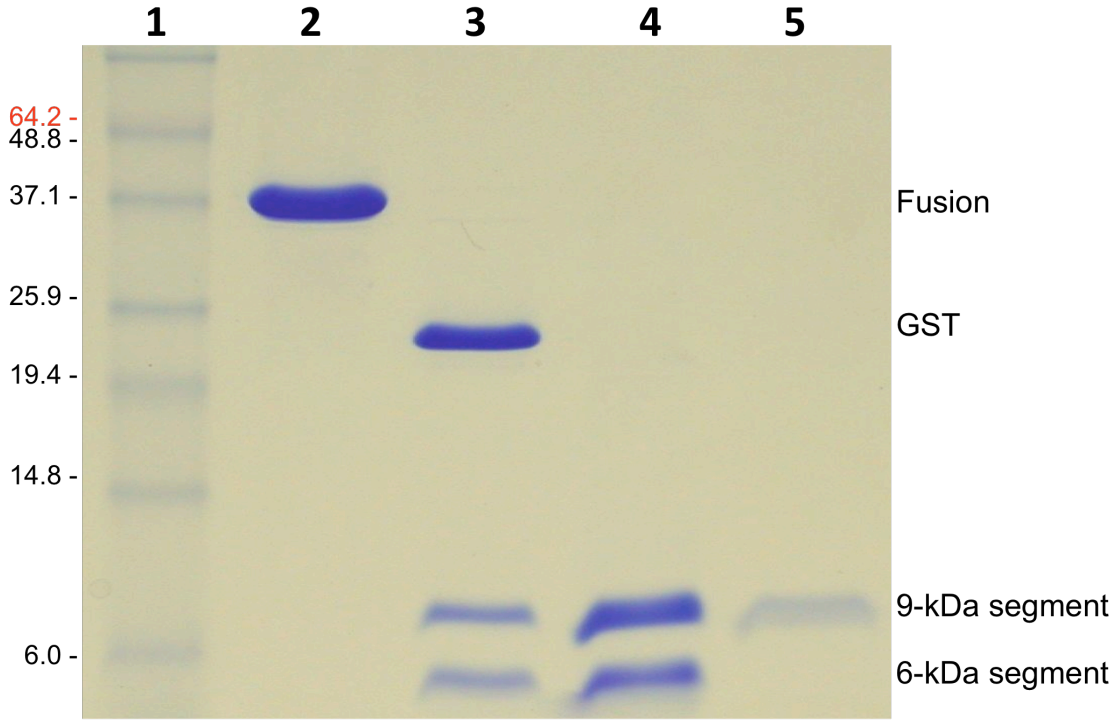


Figure 8: **9-kDa segment of dUTPase was obtained (16.25% acrylamide SDS-gel)**. Lane 1: molecular markers (kDa). Lane 2: eluted GST-dUTPase fusion protein (Step 1, Figure 5). Lane 3: fusion protein after cleavage with thrombin in the presence of 1.7 M urea (Step 2, Figure 5). Lane 4: unbound 6-kDa and 9-kDa segments after last (fourth) pass through a glutathione matrix (Step 4, Figure 5). Lane 5: 9-kDa segment of dUTPase collected by filtration (Step 5, Figure 5).

At each step of the protein purification, the starting and the recovered protein amounts were determined. Also, the dUTPase activity of the fusion protein and the 9-kDa segment were measured to make sure the protein was active from beginning to the end of the purification process. Table 1 shows the measured protein amount of each step of the purification and the measured protein specific activity of GST-dUTPase fusion and the 9-kDa segment. For some steps the protein amount was measured by taking absorbance at 280 nm, and the purified 9-kDa segment amount was measured both by absorbance and BCA protein assay. The extinction coefficient of the fusion protein and the 9-kDa segment were calculated from the ProtParam server (http://ca.expasy.org/tools/protparam.html; Gasteiger *et al*, 2005) to be 53080 and 6990 M^{-1} cm^{-1}, respectively. Based on these extinction coefficients, the assumption that OD_{280} of 1 equals to 1 mg/mL is 21% underestimated for the fusion protein and 27% overestimated for the 9-kDa segment because absorbance of 1 would equals to 0.79 mg/mL for fusion protein and 1.27 mg/mL for the 9-kDa segment. The important point thing is dUTPase activity was present throughout the purification steps and that the protein concentration value used for kinetic studies was from BCA protein assay. As

shown in Table 1, the starting fusion protein amount as 92.7 mg and the purified 9-kDa segment was 1.1 mg. Specific activity was measured at 0.0028 and 0.0256 μmol dUMP per minute per μg of protein for fusion protein and 9-kDa segment, respectively. Note that the size of the fusion protein had taken into account in the measurement of its specific activity. Log of molecular weight vs. migration were also carried out, and the calculated sizes for fusion, GST, 9-kDa and 6-kDa segments were similar to the predicted sizes (Figure 6 Appendix). Three gel images were used to calculate for the different protein species. For GST-dUTPase fusion protein calculated sizes were: 49.8, 40.3 and 42.0 kDa with an average size of 44.0 kDa. For GST the sizes were: 26.8, 24.5 and 23.8 kDa with an average size of 25.0 kDa. The resulted were 8.3 and 8.8 and 8.2 kDa for the 9-kDa segment with an average size of 8.4 kDa. Both the 6- and 9-kDa segments were observed when subjected to higher acrylamide (16.25% acrylamide) SDS gel electrophoresis in panel C of Figure 6 Appendix. Only one measurement for the 6-kDa segment with a calculated size of 5.6 kDa.

Table 1: Purification Summary of 9-kDa segment of *D. discoideum* dUTPase

Step	Starting Protein mg[1]	Recovered Protein mg[1]	% Recovery	Specific Activity (µmol dUMP/min*µg protein)
Fusion Protein	92.7	--	--	0.0028
Dialysis by filtration (after thrombin digestion)	92.7	78.5	85.3%	not measured
Post glutathione sepharose cycling	78.5	6.0	7.6%	not measured
Purified 9-kDa segment	6.0	1.3 1.1[2]	18%	0.0256

[1]Using A_{280} where 1 absorbance unit equals 1 mg/mL.
[2]Using BCA protein assay (see Materials and Methods).

6-kDa Segment of *Dictyostelium* dUTPase

```
                          Motif I                                   Motif II
  230        240        250        260        270        280
FKVKKLSDKAIIPQRGSKGAAGYDLSSAHELVVPAHGKALAMTDLQIAIPDGTYGRIAPR
```

9-kDa Segment of *Dictyostelium* dUTPase

```
                  Motif III                 Motif IV                  Motif V
  290        300        310        320        330        340        350        360      368
SGLAWKNFIDCGAGVIDSDYRGNVGVVLFNHSDVDFKVAVGDRVAQLIFERIVTPEPLEVDEIDETQRGAGGFGSTGVKVQN
     KNFIDCGAGVIDSDYRG
                    RGNVGVVLFNHSDVDFKV     VAQLIFERI
                                  KVAVGDRVAQLIFERI
                                              RIVTPEPLEVDEIDETQRG
                                                                  GAGGFGSTGVKVQN
                                                                  GAGGFGSTGVKV
```

Figure 9: **Mass spectroscopic analyses and N-terminal sequencing detect the 9-kDa segment**. The predicted 6- and 9-kDa segments of the expressed *D. discoideum* dUTPase are shown with the five conserved motifs underlined. Residue numbers correspond to the dUTPase sequence presented in Figure 6. Peptides detected by mass spectroscopic analyses are in red, and amino acid sequence from N-terminal sequencing is in blue.

3.2 The 9-kDa segment of dUTPase was confirmed by Mass Spectroscopic analyses and N-terminal Sequencing

It was important to determine if the protein obtained was actually the predicted 9-kDa segment of *D. discoideum* dUTPase. The 9-kDa segment protein sample was analyzed by mass spectroscopy and N-terminal sequencing. Results from mass spectroscopic analyses showed that all the detected peptides were from the 9-kDa segment of dUTPase. More importantly, residues of the 6-kDa segment were not detected by mass spectroscopic analyses (Figure 9). The 9-kDa segment was further confirmed by N-terminal sequencing in which the sequence of the first five residues were: SGLAW (Figure 9). This amino acid sequence was the same as the predicted N-terminus of the 9-kDa segment of *D. discoideum* dUTPase (Figure 6, residues 287-291).

3.3 9-kDa segment of dUTPase forms trimers

Once the protein obtained was confirmed to be the 9-kDa segment of *D. discoideum* dUTPase, it was necessary to determine if the protein formed trimers before carrying out kinetic studies. Because the 9-kDa segment was retained by a 10 kDa cutoff Centricon, it seemed plausible that it behaved as a species larger than 8.85 kDa. Nondenaturing gel electrophoresis was done and the results indicated that the 9-kDa segment was likely in trimer form (Figure 10). Standards were prepared from four different proteins: bovine serum albumin, ovalbumin, trypsin inhibitor (soybean), β-lactoglobulin-A . The 9-kDa segment dUTPase would have a molecular weight of 26.55 kDa if trimeric. On a native (non-denaturing) 15% acrylamide gel the 9-kDa segment migrated to a position between 20 and 45 kDa. This indicated that the 9-kDa segment formed trimers. *Arabidopsis* dUTPase as the positive control, having a molecular weight

of 53.6 kDa (PDB ID: 2P9O; Bajaj and Moriyama, 2007) in its trimer form migrated to a position between 45 and 66 kDa as well.

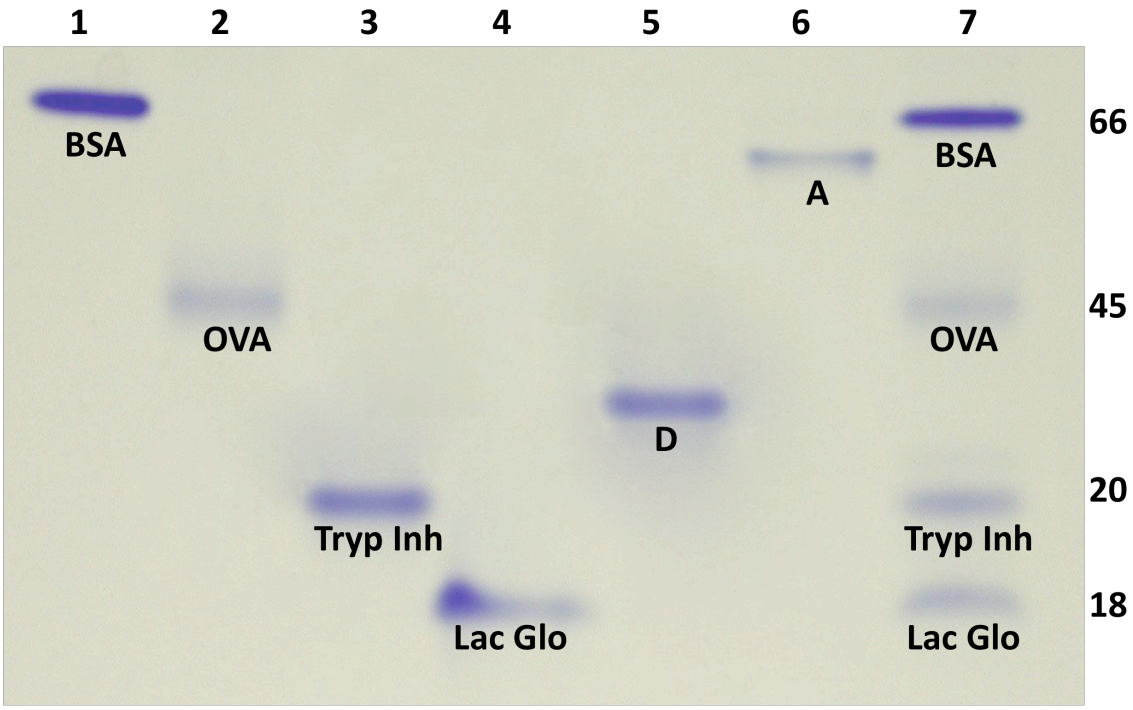


Figure 10: **9-kDa segment of dUTPase forms trimers (15% acrylamide native gel)**. Lane 1: BSA, bovine serum albumin, 66 kDa. Lane 2: OVA, ovalbumin, 45 kDa. Lane 3: Tryp Inh, trypsin inhibitor (soybean), 20 kDa. Lane 4: Lac Glo, β-lactoglobulin-A, 18 kDa. Lane 5: D, 9-kDa segment of *D. discoideum* dUTPase. Lane 6: A, *Arabidopsis* dUTPase, which served as a control. Lane 7: the four standard proteins run in the same lane.

3.4 Michaelis constant of 9-kDa segment of dUTPase

The cresol red method was used for determining the Km of the 9-kDa segment. This method indirectly measures the enzyme activity where the decreasing rate of absorbance at 573 nm is inversely proportion to the rate of enzyme activity. In this method the progress curve was observed where the enzyme activity rate was determined

during the linear portion (Figure 8 Appendix). Since the change in absorbance is correlated to enzyme activity, it was necessary to use the amount of enzyme that was within the linear range of the rate of absorbance changes. To determine this, dUTPase activity was measured in an assay where substrate concentration was constant and enzyme concentration was varied. The results showed the amount of enzyme used for the Km determination was within the linear range (Figure 11).

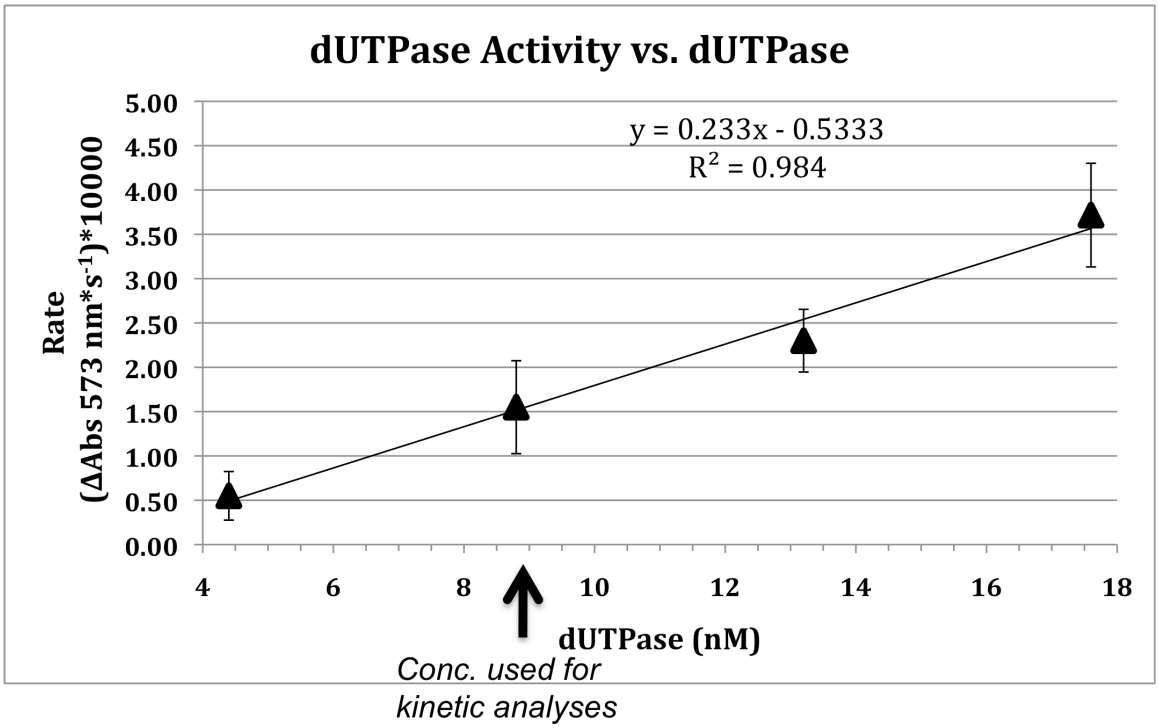


Figure 11: **Linear range of enzyme activity of the 9-kDa segment of dUTPase**. The dUTPase concentration used for kinetic studies was 8.9 nM (arrow), which was within the linear range of activity in assays where dUTP was 10 µM. The error bars are the standard deviation of changes in cresol red absorbance (see Materials and Methods).

The enzyme concentration used for the Km determination was 8.9 nM with substrates concentrations of 0.1, 0.3, 0.5, 0.7, 1.0, 3.0, 5.0 and 7.0 µM. The Km for the 9-kDa segment of *D. discoideum* dUTPase was 0.566 $\pm$ 0.02 µM (Figure 12). An

example progress curve and the data calculations using the cresol red method are shown in Figure 8 Appendix. A Lineweaver-Burk plot was constructed using the average values of the enzyme activity from the Km determination assays. The trendline was linear with a R squared value of 0.99853 (Figure 13), which is an indication that the determined Km is not an under or overestimation.

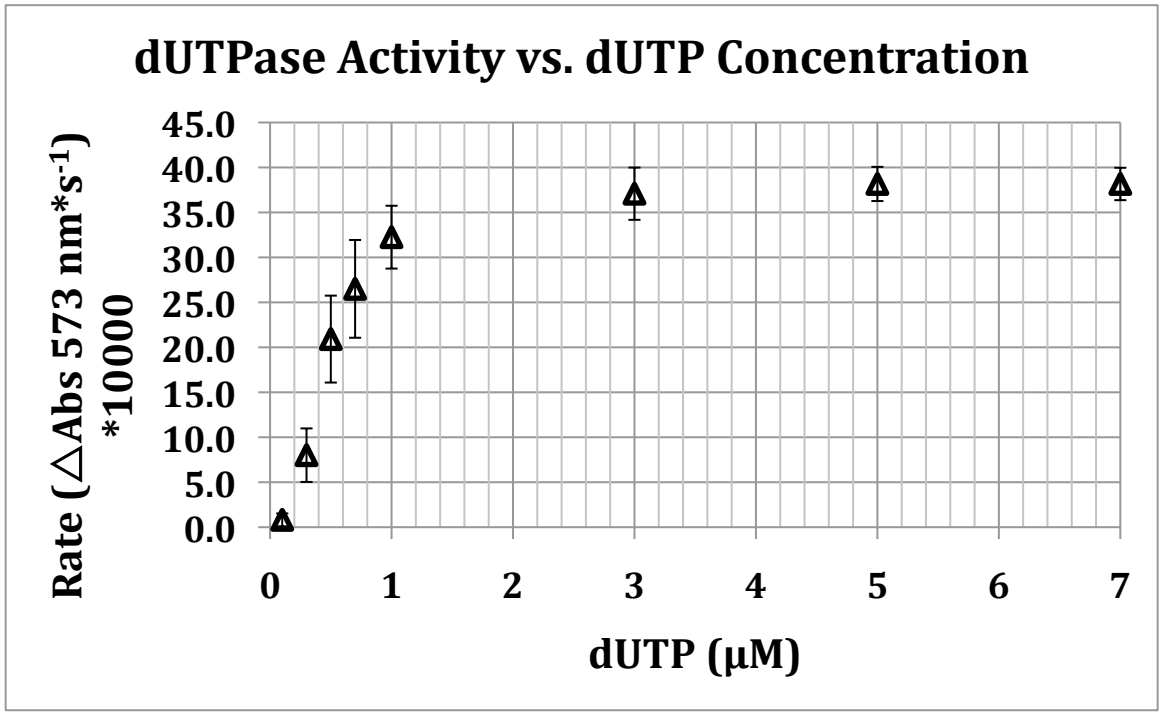


Figure 12: **dUTPase activity of the 9-kDa segment as a function of dUTP concentration**. Error bars are the standard deviation. Enzyme activity is determined by measuring the decrease in absorbance at 573 nm of cresol red over time. The slope between two points on the linear portion of progress curve was recorded and used as the rate of activity.

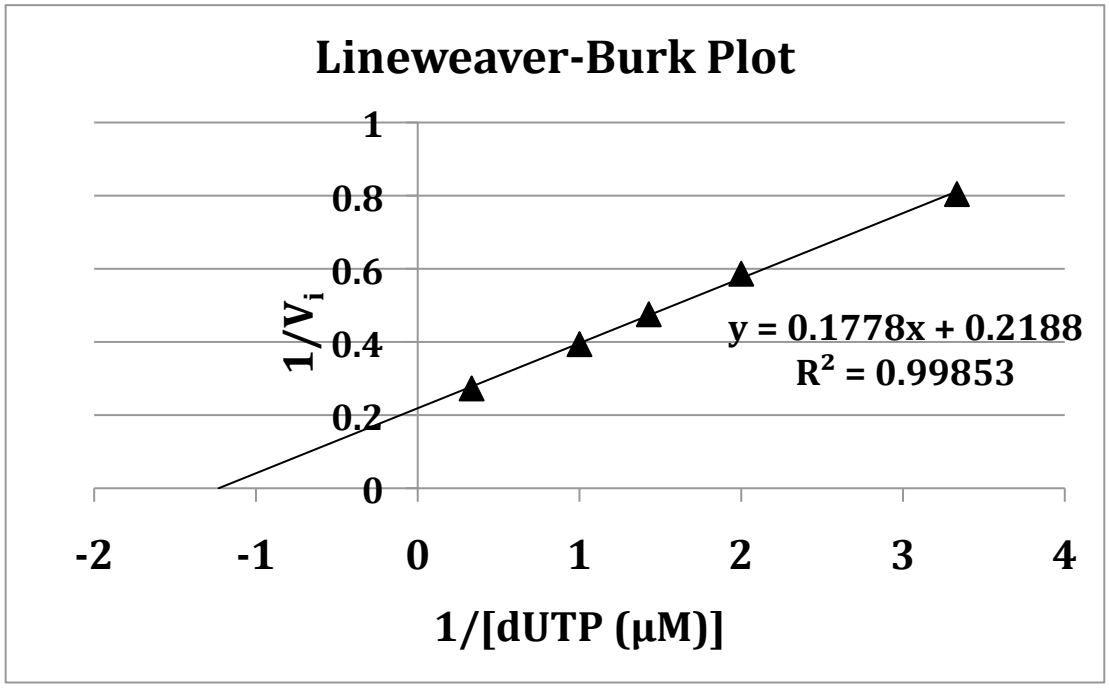


Figure 13: **Lineweaver-Burk plot of activity**. The calculated Km was 0.566 $\pm$ 0.02 µM.

3.5 pH and Temperature Optima

The method used was the HPLC measurement of dUMP produced, which is correlated to the activity of the enzyme. Examples of the data analyses and a chromatogram are in Figure 7 Appendix. The optimum pH of the 9-kDa segment dUTPase was 8.0 at 25˚C (Figure 14). At this pH the enzyme had a specific activity of 0.0259 µmol dUMP per min per µg of enzyme. Different buffers were used for different pH ranges. Table 2 shows using citrate buffer and MES at pH 6, the average specific activity were 15.4 and 17.2 µmol dUMP per minute per µg of protein, respectively. Tris buffer and HEPES were used at pH 8, and the average specific activities were 17.8 and 18.2 µmol dUMP per minute per µg of protein, respectively. The data enable one to

confidently state that the different buffers used had relatively small effects effect on the enzyme activity.

Differences in dUTPase Activity Due to Different Buffers				
Buffer	**pH**	**Average Spec. Act. (µmol dUMP/min*µg dUTPase**	**Standard Deviation**	**% Deviation**
Citrate Buffer	6	15.4	0.56	3.6%
MES	6	17.17	1.13	6.6%
Tris Buffer	8	17.79	1.67	9.3%
HEPES	8	18.2	1.62	8.9%

Table 2: **dUTPase activity of the 9-kDa segment was affected by different buffers**. Citrate buffers and MES were used to test the enzyme activity at pH 6, and Tris and HEPES buffers were used to test activity at pH 8.

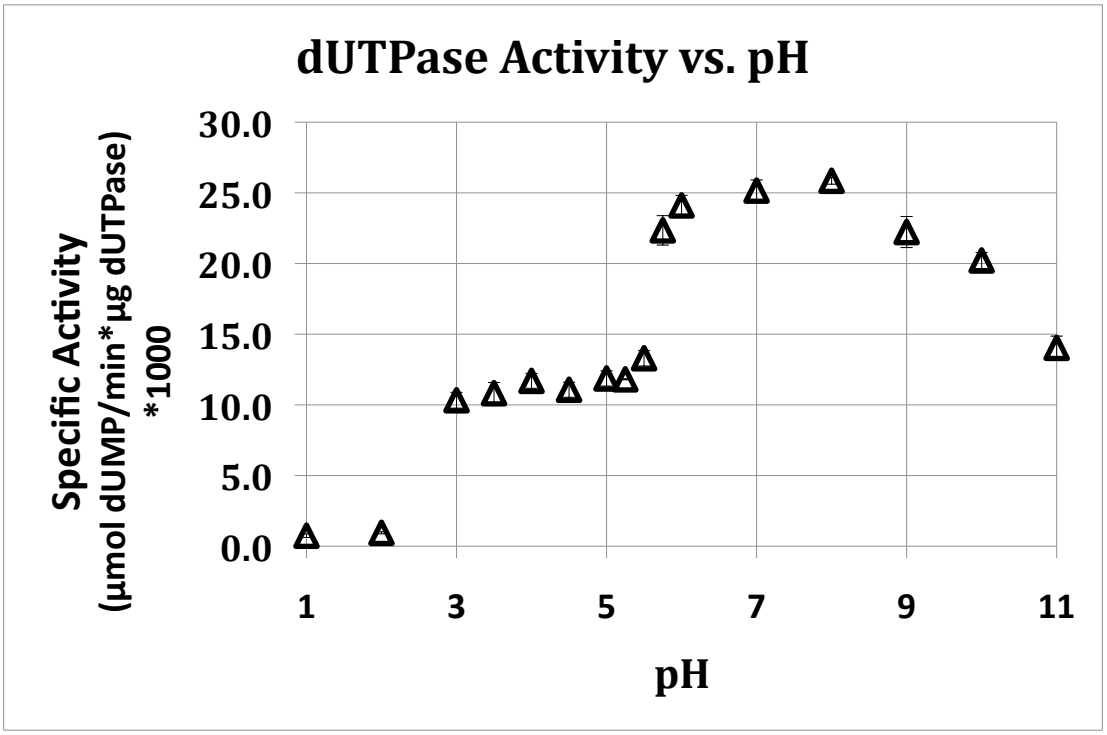


Figure 14: **dUTPase activity of the 9-kDa segment as a function of pH**. Optimum catalytic activity was at pH 8 based on the dUMP produced in µmol per min per µg**.** Enzyme activity was determined by the production of dUMP as measured by HPLC. Error bars are the standard deviation of the activity measurements.

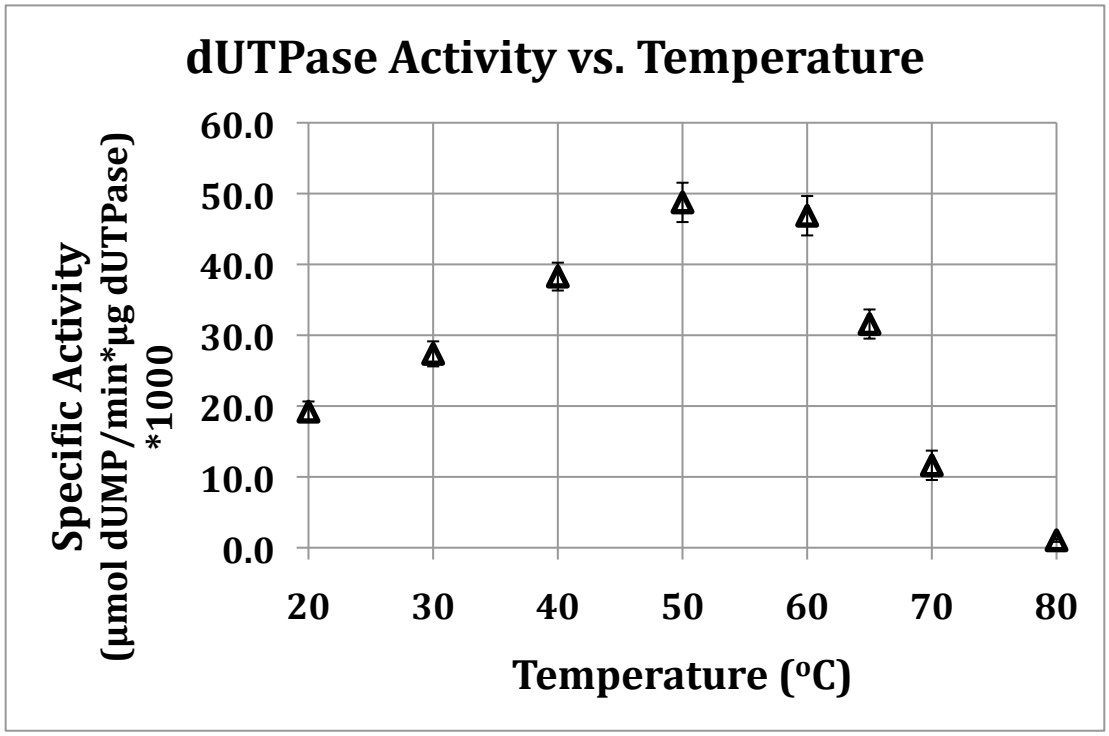


Figure 15: **dUTPase activity of the 9-kDa segment as a function of temperature.** Optimum temperature for catalytic activity was 50˚C based on the dUMP produced in µmol per min per µg protein. Enzyme activity was determined by the production of dUMP as measured by HPLC. Error bars are the standard deviation of the activity measurements.

The optimum temperature of catalysis by the 9-kDa segment dUTPase was 50˚C (Figure 15). At this temperature the enzyme had a specific activity of 0.0488 µmol dUMP per min per µg of enzyme.

In summary, the 9-kDa segment, which contained the remaining 82 residues on the C-terminal end of *D. discoideum* dUTPase, was obtained by the protein purification process, and was confirmed by mass spectroscopic analyses and N-terminal sequencing. This 9-kDa segment was likely to form trimer as shown from the nondenaturing gel. Additionally, the 9-kDa segment, even though lacking motifs I and II, was functional.

Chapter 4

Discussion

The dUTPase's function is to catalyze the hydrolysis of dUTP into dUMP. In carrying out this function, the enzyme fulfills two roles. The first is the production of dUMP, which is the precursor for the biosynthesis of dTTP. The second role is to keep the concentration of dUTP low; therefore limiting the incorporation of uracil into DNA. Most dUTPases are homotrimers, and the hallmark of dUTPases is five conserved motifs referred to as motifs I to V (Barabas *et al*, 2004; Vertessy & Toth, 2009). These five motifs were shown to work cooperatively to form the active site of the enzyme (Barabas *et al*, 2004). There have been no studies of *D. discoideum* dUTPase; therefore, all characteristics of this dUTPase are inferred through similarity.

The *D. discoideum* dUTPase has an amino acid sequence of 179 residues predicted from the information on the genome database of *D. discoideum* (*www.dictybase.org*). The gene ID is DDB_G0293374, and the gene location is on chromosome 6 with the coordinates 2804871 to 2805532. The recombinant protein lacked 37 residues from the N-terminal end. The reason for this was this part of the protein was flexible and the inclusion of this part will make protein crystallography difficult (Nemeth-Pongracz *et al*, 2007; Bekesi *et al*, 2004; Bajaj and Moriyama, 2007). The synthesized protein consisted of 141 residues and contained the five conserved motifs of dUTPases (Figures 3 and 6).

The protein purification process was followed the standard protocols with some modifications. One modification was the addition of 1.7 M urea. Thrombin cleavage

efficiency was low so that the separation of the cleaved proteins from the fusion was greatly difficult. Because the percentage of fusion protein that got cleaved was low the subsequent steps of the purification almost all recombinant proteins were either lost or was at an undetectable level (Figure 7; Figure 4 Appendix). The hypothesis was that since dUTPase was fused to GST and because of this the thrombin cleavage site possibly could have gotten covered up by GST making it hard for thrombin to access its site. The fact that with the addition of urea, a protein denaturant, cleavage efficiency greatly increased was an indication that the site was likely gotten covered by GST. Thrombin cleavage efficiency, in the presence of 1.7 M urea, increased from 25% of fusion protein got cleaved to 95% (Figure 7). Enzyme activity was detected after the fusion was incubated with urea, in which the urea was ridden off by filtration (data not shown).

It was observed that the recombinant protein had a secondary thrombin cleavage site (Gasteiger *et al*, 2003). This led to the production of 6-kDa and 9-kDa segments of dUTPase. The main focus was on the 9-kDa segment, which had only 82 remaining amino acids *D. discoideum* dUTPase (Figure 8). After the collection of the 9-kDa segment, the focus was to investigate to see if this segment of the protein would form trimers.

It was important to determine the reason the recombinant dUTPase got cut into 2 segments was not due to any non-specific proteolysis, and that the 9-kDa segment was actually the predicted peptide. To do this mass spectroscopic and N-terminal sequencing analysis were carried on the 9-kDa segment collected. The mass spectroscopic analyses showed that all the peptides detected were of the 9-kDa segment, and that there was no detection of peptides within the 6-kDa segment (Figure 9). This confirmed that the

obtained sample was of the 9-kDa segment of *D. discoideum* dUTPase. To further confirm this, N-terminal sequencing was carried of the protein sample. The results showed that the first five amino acids on the N-terminus were SGLAW as predicted. The results from mass spectroscopic analyses and N-terminal sequencing confirm that the protein obtained was the 9-kDa segment of *D. discoideum* dUTPase produced by thrombin cleavage. The yield of the purified 9-kDa segment (1.1 mg) from the mixture of 6-kDa and 9-kDa segments (6.0 mg) was 18%, and from the fusion protein (92.7 mg) was 1.2%. These protein amounts are from Table 1, and the percentages are calculated using these numbers. On the side note, the protein quantification based on the assumption that OD_{280} of 1 equals to 1 mg/mL is 21% underestimated for the fusion protein and 27% overestimated for the 9-kDa segment because absorbance of 1 would equals to 0.79 mg/mL for fusion protein and 1.27 mg/mL for the 9-kDa segment. The focus here is dUTPase activity was present throughout the purification steps and that the protein concentration value used for kinetic studies was from BCA protein assay.

Once it was shown that the protein being used was the 9-kDa segment the focus was back to the determination if the segment would form trimers. The retention of the protein by the 10-kDa cutoff filter is the first indication that it is acting as a larger species . This shows that the 9-kDa segment was not a monomer and a multimer. The results from running a nondenaturing gel showed that the protein band was at the area of predicted size (Figure 10). The 9-kDa segment trimer would have a molecular weight of 26.6 kDa. The gel showed that the 9-kDa segment band lied between the ovalbumin and trypsin inhibitor (soybean) having the sizes of 45 kDa and 20 kDa, respectively. From this it was certain that the 9-kDa was not a monomer, but what about a dimer? If the 9-

kDa segments dimerized it would have a size of 17.7 kDa, but the protein band lied above the trypsin inhibitor which has the size of 20 kDa. There was always the possibility that the 9-kDa segment could be a tetramer and pentamer, but this was unlikely since there had been no indication of the existence of dUTPases of tetramer or above. Overall, the evidence from filtration and especially from nondenaturing gel were good indications the 9-kDa segment of *D. discoideum* dUTPase, with only 82 residues from the C-terminal end, formed trimer. The results from nondenaturing gel analyses is a reasonable indication that the 9-kDa segment forms trimers, but this method alone is not enough to say for certain of the trimerization of the 9-kDa segment. Gel chromatography such as gel filtration is needed for this confirmation.

Most dUTPases are homotrimer consisting of three subunits and three active sites. There are also monomeric and dimeric dUTPases. Monomeric dUTPases occur in herpesviruses, such as the Epstein-Barr virus (Tarbouriech *et a*l, 2005). Dimeric dUTPases are found in organisms such as the gastric pathogen *Campylobacter jejuni* and the parasitic protozoan *Leishmania major* (Moroza *et al*, 2004; Hidalgo-Zarco *et al*, 2000). Dimeric dUTPases are found to not contain the five conserved motifs as in the trimeric and monomeric dUTPases. The trimeric dUTPase's active site is a combination of the five motifs from all three subunits, in which each active site involves the cooperation of motifs I, II and IV from one subunit, motif III from another and motif V from the third subunit. The 9-kDa segment of *D. discoideum* dUTPase lacked motif I and motif II. The question was asked whether the 9-kDa segment of dUTPase still carries out its function. The results from kinetic studies showed that the 9-kDa segment has activity, which was an indication that the participations of motifs I and II were not essential for the

enzyme's function. The activity possibly did get affected by the absence of the first two motifs. Further investigations would be needed for this hypothesis. This could be done by obtaining a full length *D. discoideum* dUTPase to compare activity with the 9-kDa segment. One thing for certain was the *D. discoideum* dUTPase catalyzed the hydrolysis of dUTP into dUMP, and the activity was not lost in the absence of motif I and motif II.

The 9-kDa segment has dUTPase activity, but how does its activity compare to dUTPases of other organisms? The 9-kDa segment of *D. discoideum* dUTPase had a Km of 0.566 $\pm$ 0.02 μM (Figure 12). The Lineweaver-Burk plot showed a linear pattern indicating that the determined Km value was not under or over estimated (Figure 13). The Km of dUTPases is different among different organisms. The values vary from *Escherichia coli* (Larsson, Nyman *et a*l., 1996) of 0.21 μM to 28 μM of the dUTPase of the mouse mammary tumor virus *gag-pro* transframe protein (p30) (Koppe *et al*., 1994). The reported Km of dUTPases from human (Toth *et al*, 2007) and *Drosophila melanogaster* (Kovari *et al*., 2004) are 3.6 μM and 0.4 μM respectively. For *Chlorella* virus PBCV-1 (Zhang *et al*., 2005) the reported Km is 11.7 μM. For the 9-kDa segment of *D. discoideum* dUTPase the Km value was on the lower end of the spectrum. First, this confirmed that the binding of substrate to the active site was functional. Secondly, the association rate between substrate and the 9-kDa segment was quite high from the point of view where the Km value lies within range of different known Km from dUTPases of other organisms. It shows that only a small amount of substrate was required for the enzyme to become saturated. It is known from studies that motifs I, II and IV coordinate the triphosphate of the substrate in the active site; therefore the Km of

the 9-kDa segment did get affected. One possibility was that, without motifs I and II, the affinity would be even higher. One possible way to test this is by comparing to the full length *D. discoideum* dUTPase. One the side note, the Km measured is not a representation of the substrate affinity. It is the substrate concentration where the enzyme velocity is half Vmax. Kcat was not able to obtain because the cresol red method was used to compare enzyme velocity at different substrate concentrations, not for the measurement of product formation. It is because of this Kcat and subsequently Kcat/Km were unable to obtain.

The results from the optimum catalysis pH of *D. discoideum* 9-kDa segment of dUTPase showed that the optimum pH was 8.0 at 25 ˚C (Figure 14). The pattern was as the pH level increased the activity also increased though not linearly. From pH 3 to 5 the level of activity was quite similar. When the pH was at 6 the activity jumped more than two fold and stayed fairly constant until after pH 9. It was not a clear cut at which specific pH has the highest activity, but at pH 8 the activity was a little higher than pH 6, 7 and 9. The reason for this pattern was not fully understood. One possible explanation was that upon substrate binding the flexible C-terminus arm closed upon the active site; therefore, the active site was fairly insensitive to the minor pH changes in the surrounding environment (Vertessy *et al*., 1998; Larsson *et al*., 1996). Similar pattern was also observed in dUTPase from *E. coli* where the enzyme's function was not affected by pH in the range of 6-9 (Larson *et al*., 1996). Usually, different type of enzyme has its own pH range for the enzyme's activity to be optimal. In the case of the 9-kDa segment of *D. discoideum* dUTPase it was optimal at the pH range from 6 to 9. The reported optimum pH for dUTPase from *Chlorella* virus (PBCV-1) is at 8.5 (Zhang *et al*., 2005).

In *D. melanogaster* the dUTPase optimum pH range is 6-10 (Giroir and Duetsch, 1987). It is usually the case that when an enzyme is put into an environment of pH outside its range its 3-D structure will get affected, and this will affect the enzyme's function. Overall, the 9-kDa segment, though in the absence of motifs I and II, has similar pH optimum catalysis to dUTPases from organisms such as the bacteria *E. coli*, the multicellular eukaryote *Drosophila,* and *Chorella* virus.

The results from the optimum temperature catalysis of *D. discoideum* 9-kDa segment of dUTPase showed the activity was highest at 50 ˚C at pH 8.0 (Figure 15). The pattern was that the activity increased as the temperature increased quite linearly, but then the activity greatly decreased beyond 50 ˚C. As the temperature increased the enzyme was able to bind to its substrate at a faster rate because molecules were moving around faster in the reaction mixture, but only to a point. This was where the tradeoff between higher activity and the risk of denaturation occurs. Beyond the temperature 50˚C the enzyme started to denature and its activity dropped really fast. The reported optimum catalysis temperature for *Chlorella* virus (PBCV-1) is 50 ˚C (Zhang *et al*., 2005). In *D. melanogaster* and *P. falciparum* the optimum temperatures are 56.3 ˚C and 50 ˚C, respectively (Kovari *et al*., 2004; Quesada-Soriano *et al*., 2008). The 9-kDa segment of *D. discoideum* dUTPase exhibited similar optimum temperature catalysis of dUTPases of organisms mentioned above. *D. discoideum* normally grows in the temperature of 20 ˚C. Interestingly, the temperature optimum was much higher, which could be an indication that the enzyme was quite stabled.

The data showed that dUTPase activity was present even without a significant portion of the protein. The question was what were the possible functions of the missing

portion? Firstly, the missing portion contained motif I and II that participate, though not essential as shown by the present study, in the catalytic activity of the enzyme. Secondarily, studies have shown that NLS locates on the N-terminal end of dUTPase (Merenyi *et al.*, 2010). This is logical because dUTPase involves in keeping the DNA integrity. In eukaryotic organisms, proteins that are involved in nuclear functions must pass through the nuclear envelope and into the nucleus after synthesis. In this process, proteins greater than 35 kDa, requires the presence of a NLS for a direct and specific localization to the nucleus (Merenyi *et al.*, 2010). Different NLS signals had been identified on dUTPases. Even though the NLS signals are varied in the amino acid sequence, they do share common characteristics. The NLS signal motifs are generally high amount of basic residues such as lysine and arginine , and the presence of proline (Christophe *et al*, 2000; Fries *et al*, 2007). An example is that studies have shown in *D. melanogaster* cells contain two dUTPase isoforms. NLS is found on the nuclear isoform, and the sequence is determined to be PAAKKMKID (Merenyi *et al*., 2010). Possible future studies on *D. discoideum* dUTPase to answer such questions as: Does *D. discoideum* dUTPase contain the nuclear and mitochondrial isoforms? Do they contain localization signals? What are the signals?

In summary, we have shown that the 9-kDa segment of *D. discoideum* dUTPase, with only 82 remaining residues on the C-terminal end, was likely a trimer. Also the 9-kDa segment, in the absence of motif I and II, carried out its catalysis activity. The results from Km, optimum pH and temperature of the 9-kDa segment showed that it had similar functional characteristics as compared to dUTPases of other organisms. The only

exception was there is a wide range of Km among dUTPases, and the 9-kDa segment's Km fell toward the lower end of spectrum.

References

Abramoff, M. D., Magelhaes, P. J., Ram, S. J. (2004). Image Processing with ImageJ. *Biophotonics International*, volume 11, issue 11, pp. 36-42.

Bajaj, M., Moriyama, H. (2007). Purification, crystallization and preliminary crystallographic analysis of deoxyuridine triphosphate nucleotiddohydrolase from *Arabidopsis thaliana*. *Acta Crystallographica Section F* **63**, 409-411.

Barabas, O., Pongracz, V., Kovari, J., Wilmanns, M., Vertessy, B. G. (2004). Structural Insights into the Catalytic Mechanism of Phosphate Ester Hydrolysis by dUTPase. *The Journal of Biological Chemistry* **279**, 42907-42915.

Bekesi, A., Zagyva, I., Hunyadi-Gulyas, E., Pongracz, V., Kovari, J., Nagy, A. O., Erdei, A., Medzihradszky, K. F., Vertessy, B. G. (2004). Developmental regulation of dUTPase in *Drosophila melanogaster*. *Journal of Biological Chemistry* **21**, 22362-22370.
Bozzaro, S., Ponte, E. (1995). Cell adhesion in the life cycle of *Dictyostelium*. *Experientia* **51**, 1175-1188.

Buenemann, M., Levine, H., Rappel, W., Sander, L. M. (2010). The Role of Cell Contraction and Adhesion in *Dictyostelium* Motility. *Biophysical Journal* **99**, 50-58.

Cedergren-Zeppezauer, E. S., Larsson, G., Nyman, P. O., Dauter, Z., Wilson, K. S. (1992). Crystal structure of a dUTPase. *Nature* **355**.

Christophe D., Christophe-Hobertus, C., Pichon, B. (2000). Nuclear targeting of proteins: how many different signals? *Cell Signal* **12**, 337-341.

Fiser, A., Vertessy, B. G. (2000). Altered Subunit Communication in Subfamilies of Trimeric dUTPases. *Biochemical and Biophysical Research Communications* **279**, 534-542.

Freeman, L., Buisson, M., Tarbouriech, N., Van der Heyden, A., Labbe, P., Burmeister, W. P. (2009). The Flexible Motif V of Epstein-Barr Virus Deoxyuridine 5'-Triphosphate Pyrophosphatase Is Essential for Catalysis. *The Journal of Biological Chemistry* **284**, 25280-25289.

Fries, T., Betz, C., Sohn, K., Ceasar, S., Schlenstedt, G., Bailer, S. M. (2007). A novel conserved nuclear localization signal is recognized by a group of yeast importins. *The Journal of Biological Chemistry* **282**, 19292-19301.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M. R., et al. (2005). Protein Identification and Analysis Tools on the ExPASy Server. John M Walker (ed): The Proteomics Protocols Handbook, Humana Press: 571-607.

Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R. D., Bairoch A. (2003). ExPASY: the proteomics server for in-depth protein knowledge and anlysis. *Nucleic Acids Res*. **31**, 3784-3788.

Gerisch, G., Weber, I. (2000). Cytokinesis without myosin II. *Current Opinion Cell Biology* **12**, 126-132.

Giroir, L. E., Deutsch, W. A. (1987). Drosophila Deoxyuridine Triphosphatase Purification and Characterization. *The Journal of Biological Chemistry* **262**, 130-134.

Guillet, M., Van Der Kemp, P. A., Boiteux, S. (2006). dUTPase activity is critical to maintain genetic stability in *Saccharomyces cerevisiae*. *Nucleic Acids Research* **34**, 2056-2066.

Hidalgo-Zarco, F., Camacho, A. G., Bernier-Villamor, V., Nord, J., Ruiz-Perez, L. M., Gonzalez-Pacanowska, D. (2001). Kinetic properties and inhibition of the dimeric dUTPase-dUDPase from *Leishmania major*. *Protein Science* **10**, 1426-1433.

Koehler S. E., Ladner R. D. (2004). Small interfering RNA-mediated suppression of dUTPase sensitizes cancer cell lines to thymidylate synthase inhibition. *Molecular Pharmacology* **66**, 620-626.

Koppe, B., Menendez-Arias, L., Oroszlan, S. (1994). Expression and Purification of the Mouse Mammary Tumor Virus *gag-pro* Transmembrane Protein p30 and Characterization of Its dUTPase Activity. *Journal of Virology* **68**, 2313-2319.

Kovari, J., Barabas, O., Takacs, E., Bekesi, A., Dunbrovay, Z., Pongracz, V., Zagyva, I., Imre, T. Szabo, P., Vertessy, B. G. (2004). Altered Active Site Flexibility and a Structural Metal-binding Site in Eukaryotic dUTPase. *The Journal of Biological Chemistry* **279**, 17932-17944.

Ladner, R. D., Lynch F. J., Groshen S., Xiong Y. P., Sherrod A., Caradonna S. J., Stoehlmacher J., Lenz H. J. (2000). dUTP nucleotidohydrolase isoform expression in normal and neoplastic tissues: association with survival and response to 5-fluorouracil in colorectal cancer. *Cancer Res.* **60**, 3493-3503.

Ladner, R., McNulty, D. E., Carr, S. A., Roberts, G. D., Caradonna, S. J. (1996). Characterization of Distinct Nuclear and Mitochondrial Forms of Human Deoxyuridine Triphosphate Nucleotidohydrolase. *The Journal of Biological Chemistry* **13**, 7745-7751.

Larsson, G., Nyman, P. O., Kvassman, J. (1996). Kinetic Characterization of dUTPase from *Escherichia coli*. *The Journal of Biological Chemistry* **271**, 24010-24016.

Longley D. B., Harkin D. P., Johnston P. G. (2003). 5-fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **3**, 330-338.

Mustafi, D., Bekesi, A., Vertessy, B. G., Makinen, M. W. (2003). Catalytic and structural role of the metal ion in dUTP pyrophosphatase. *Proceedings of the National Academy of Sciences* **100**, 5670-5675.

McIntosh, E., Haynes, R. H. (1997). dUTP pyrophosphatase as a potential target for chemotherapeutic drug development. *Acta Biochimica Polonica* **44**, 159-172.

Merenyi, G., Konya, E., Vertessy, B. G. (2010). Drosophila proteins involved in metabolism of uracil-DNA possess different types of nuclear localization signals. *The FEBS Journal* **277**, 2142-2156.

Mol, C. D., Harris, J. M., McIntosh E. M., Tainer, J. A. (1996). Human dUTP pyrophosphatase: uracil recognition by a beta hairpin and active sites formed by three separate subunits. *Structure* **4**, 1077-1092.

Nemeth-Pongracz, V., Barabas, O., Fuxreiter, M., Simon, I., Pichova, I., Rumlova, M., Zabranska, H., Svergun, D., Harmat, V., Klement, E., Hunyadi-Gulyas, E., Medzihradszky, K. F., Konya, E., Vertessy, B. G. (2007). Flexible segments modulate co-folding of dUTPase and nucleocapsid proteins. *Nucleic Acids Res*. **35**, 495-505.

Nord, J., Kiefer, M., Adolph, H., Zeppezauer, M. M., Nyman, P. O. (2000). Transient kinetics of ligand binding and role of the C-terminus in the dUTPase from equine infectious anemia virus. *The FEBS Letters* 472, 312-316.

Palmen, L. G., Becker, K., Bulow, L., Kvassman, J. (2008). A Double Role for a Strictly Conserved Serine: Further Insights into the dUTPase Catalytic Mechanism. *Biochemistry* **47**, 7863-7874.

PDB ID: 2P9O
Bajaj, M., Moriyama, H. Structure of dUTPase from *Arabidopsis thaliana*.

Pri-Hadash, A., Hareven, D., Lifschitz, E. (1992). A Meristem-Related Gene from Tomato Encodes a dUTPase: Analysis of Expression in Vegetative and Floral Meristems. *The Plant Cell* **4**, 149-159.

Quesada-Soriano, I., Leal, I., Casas-Solvas, J. M., Vargas-Berenguel, A., Baron, C., Ruiz-Perez, L. M., Gonzalez-Pacanowska, D., Garcia-Fuentes, L. (2008). Kinetic and thermodynamic characterization of dUTP hydrolysis by *Plasmodium falciparum* dUTPase. *Biochimica et Biophysica Acta* **1784**, 1347-1355.

Quesada-Soriano, I., Musso-Buendia, J. A., Tellez-Sanz, R., Ruiz-Perez, L. M., Baron, C., Gonzalez-Pacanowska, D., Garcia-Fuentes, L. (2007). *Plasmodium falciparum* dUTPase: Studies on protein stability and binding of deoxyuridine derivatives. *Biochimica et Biophysica Acta* **1774**, 936-945.

Sambrook, J., Russell, D. W. Molecular Cloning: A Laboratory Manual. Third edition. Cold Spring Harbor, New York. Cold Spring Harbor Laboratory Press. 2001.

Takacs, E., Nagy, G., Leveles, I., Harmat, V., Lopata, A., Toth, J., Vertessy, B. G. (2010). Direct contacts between conserved motifs of different subunits provide major contribution to active site organization in human and mycobacterial dUTPases. *FEBS Letter* **584**, 3047-3054.

Tarbouriech, N., Buisson, M., Seigneurin, J., Cusack, S., Burmeister, W. P. (2005). The Monomeric dUTPase from Epstein-Barr Virus Mimics Trimeric dUTPases. *Structure* **13**, 1299-1310.

Tinkelenberg B. A., Hansbury M. J., Ladner R. D. (2002). dUTPase and uracil-DNA glycosylase are central modulators of antifolate toxicity in *Saccharomyces cerevisiae*. *Cancer Res*. **62**, 4909-4915.

Toth, J., Varga, B., Kovacs, M., Malnasi-Csizmadia, A., Vertessy, B. G. (2007). Kinetic Mechanism of Human dUTPase, an Essential Nucleotide Pyrophosphatase Enzyme. *The Journal of Biological Chemistry* **282**, 33572-33582.

Vertessy, B. G., Larsson, G., Persson, T., Bergman, A., Persson, R., Nyman, P. O. (1998). The complete triphosphate moiety of non-hydrolyzable substrate analogues is required for a conformational shift of the flexible C-terminus in *E. coli* dUTP pyrophosphatase. *The FEBS Letters* **421**, 83-88.

Vertessy, B. G., Toth, J. (2009). Keeping Uracil Out of DNA: Physiological Role, Structure and Catalytic Mechanism of dUTPases. *Accounts of Chemical Reseach* **42**, 97-106.

Webley, S. D., Hardcastle, A., Ladner, R. D., Jackman, A. L., Aherne, G. W. (2000). Deoxyuridine triphosphatase (dUTPase) expression and sensitivity to the thymidylate synthase (TS) inhibitor ZD9331. *British Journal of Cancer* **83**, 792-799.

Wilson, P. M., Fazzone, W., LaBonte, M. J., Deng, J., Neamati, N., Ladner, R. D. (2008). Novel opportunities for thymidylate metabolism as a therapeutic target. *Molecular Cancer Therapeutics* **7**, 3029-3037.

Zhang, Y., Moriyama, H., Homma, K., Van Etten, J L. (2005). Chlorella Virus-Encoded Deoxyuridine Triphosphatases Exhibit Different Temperature Optima. *Journal of Virology* **79**, 9945-9953.

Appendix A

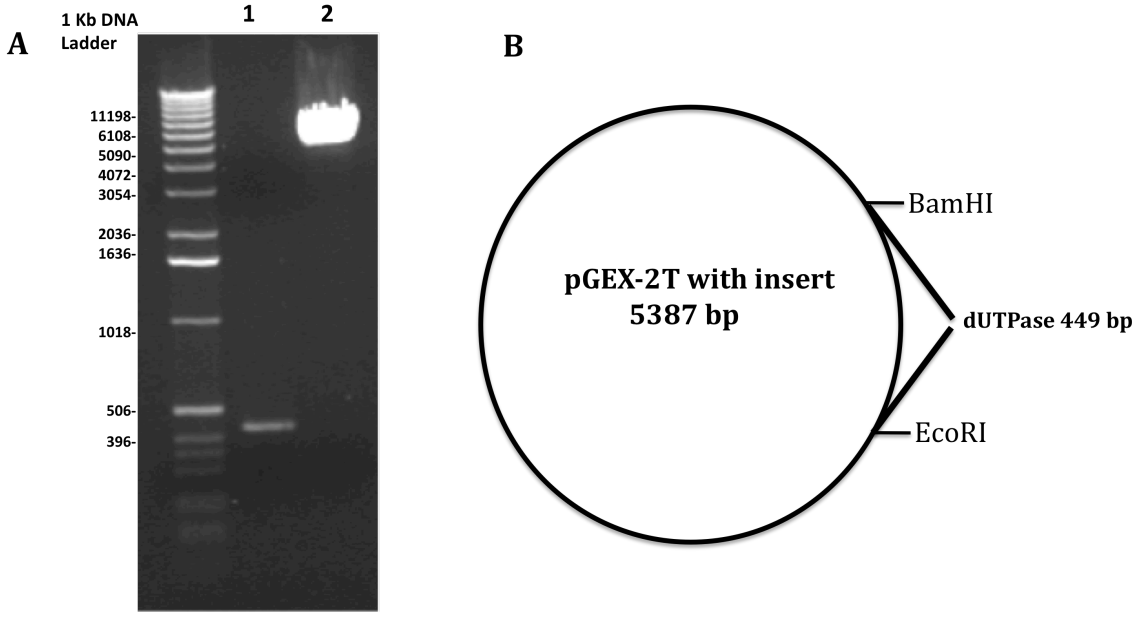


Figure 1 Appendix: **Determination of the sizes of both digested D. *discoideum* dUTPase gene and expression vector pGEX-2T**. ***A***. Inserts and pGEX-2T after digestions by EcoRI and BamHI and purified (0.8% agarose in TE). Lane 1*: D. discoideum* dUTPase DNA (449 bp) from pUC57 after digestion and purification. Lane 2: Expression vector pGEX-2T (4938 bp) after digestion and purification. The vector band appears to be running slower than 4938 bp due to the heavy load. ***B.*** The physical map of the predicted size of the expression plasmid (pGEX-2T) with the *D. discoideum* dUTPase gene after ligation.

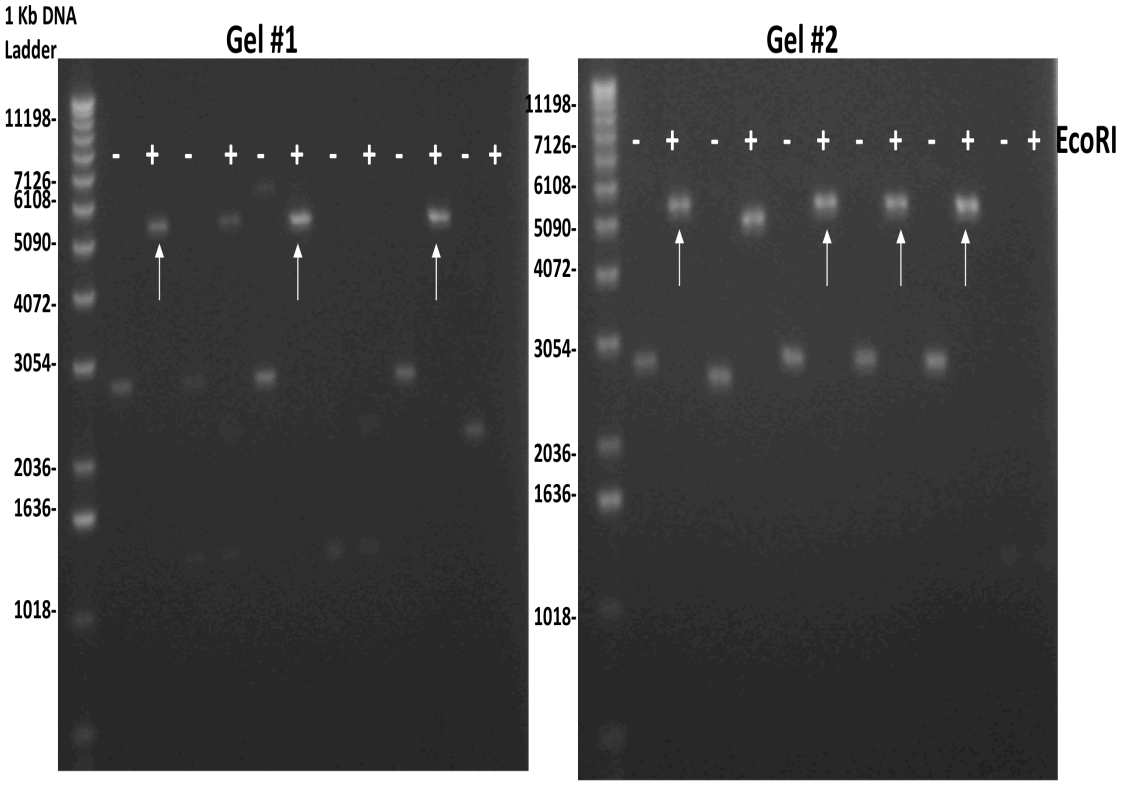


Figure 2 Appendix: **Identification of transformed Top 10 *E. coli* (0.8% agarose in TE)**. EcoRI-digested plasmid DNA isolated from presumptive transformants identified as containing the expression vector with insert (indicated by arrows). The predicted size of the vector with insert was 5387 bp. The faster migrating bands were undigested plasmids.

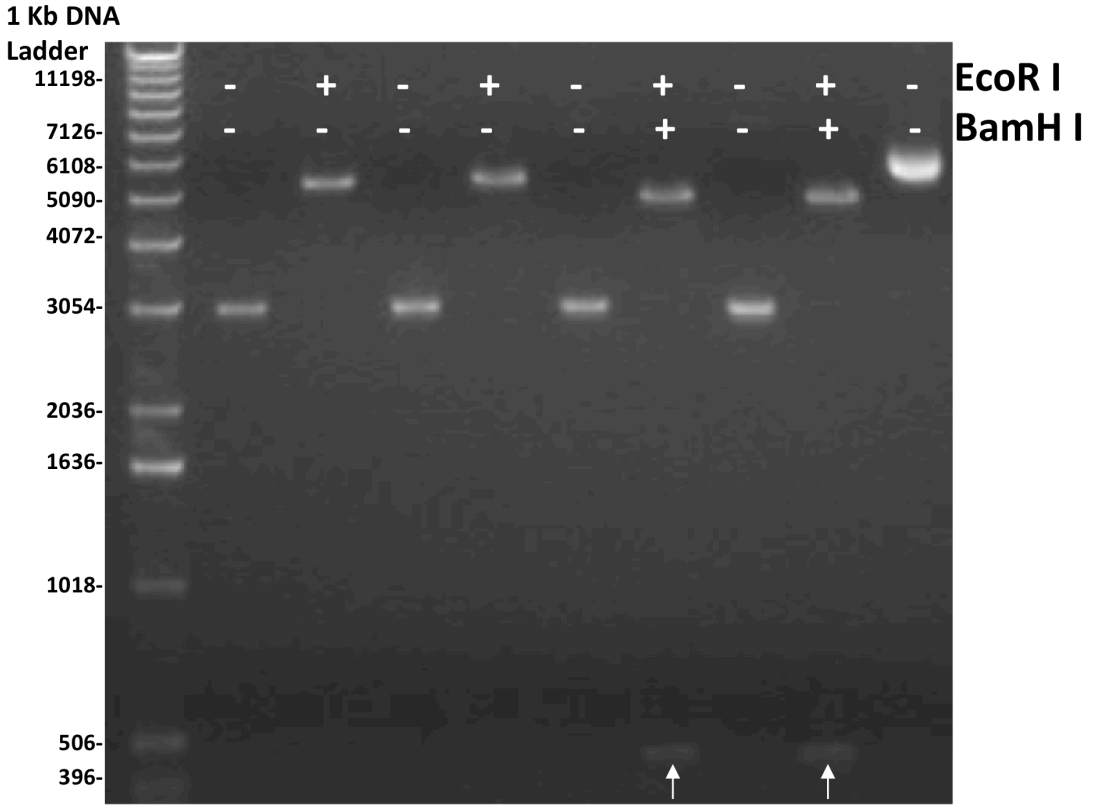


Figure 3 Appendix: **The *D. discoideum* dUTPase gene was present when plasmid DNA from Top 10 *E. coli* was digested with both EcoRI and BamHI (0.8% agarose in TE)**. The gene was at the predicted size (449 bp) (arrows). The bands running close to 3054 bp are undigested plasmid DNA. Single (EcoRI) digested plasmids were at the predicted size (5387 bp).

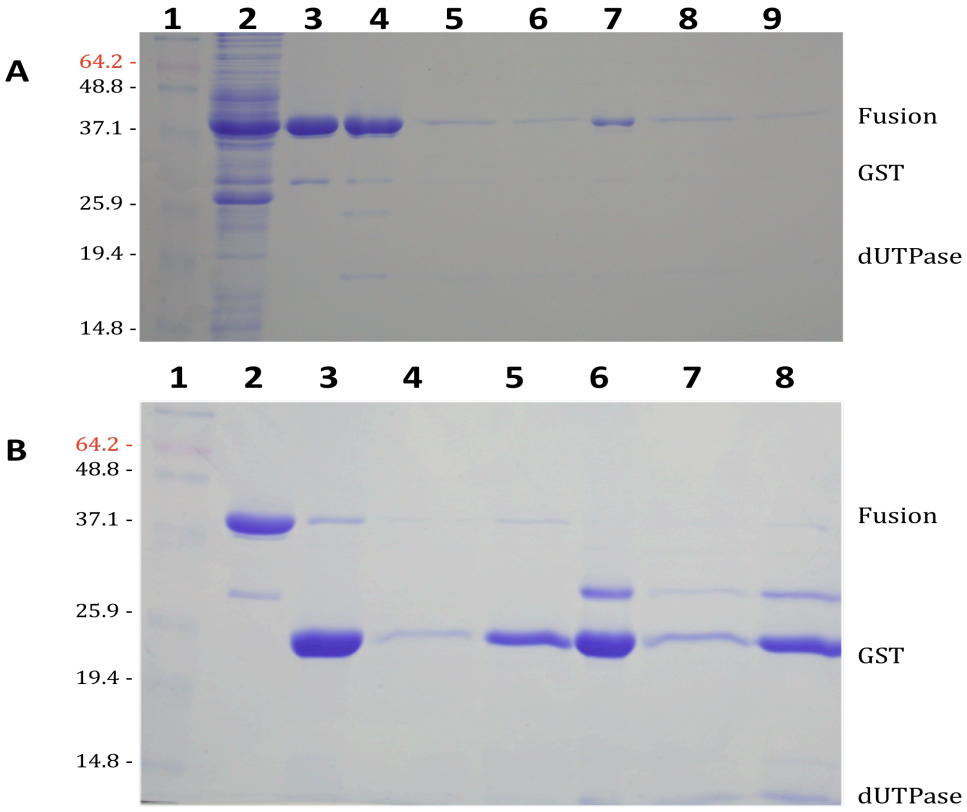


Figure 4 Appendix: **Thrombin cleavage efficiency did not increase in the presence of NaCl but greatly increased in the presence of urea (15% acrylamide SDS-gels)**. ***A***. Lane 1: molecular markers (kDa). Lane 2: cell lysate. Lane 3: eluted GST-dUTPase fusion protein (10.1 µg) (Step 1, Figure 5). Lane 4: fusion protein after thrombin cleavage (10.1 µg) in the presence of 1 M NaCl. Lanes 5 and 6: unbound peak fractions after passage through a glutathione beads matrix (Step 4, Figure 5). Lane 7: fusion protein (10.1 µg) after thrombin cleavage in the presence of 2 M NaCl. Lanes 8 and 9: unbound peak fractions after passage through a glutathione beads matrix (Step 4, Figure 5). There was protein precipitation when 2 M NaCl was used, and this is why lane 7 shows lower protein amount. ***B***. Lane 1: molecular markers (kDa). Lane 2: eluted GST-dUTPase fusion protein (10.1 µg) (Step 1, Figure 5). Lane 3: fusion protein after thrombin cleavage (10.1 µg) in the presence of 1.7 M urea (Step 2, Figure 5). Lanes 4 and 5: unbound peak fractions after passage through a glutathione beads matrix (Step 4, Figure 5). Lane 6: fusion protein (10.1 µg) after thrombin cleavage in the presence of 2.9 M urea. Lanes 7 and 8: unbound fractions after passage through a glutathione beads matrix (Step 4, Figure 5). Peak fractions were fractions with relative high OD_{280} readings compared to other fractions indicating the presence of protein. The protein species that is bigger than GST but smaller than fusion in panel B is likely the GST fused 6-kDa segment based on the location of the band and an measured size of 30 kDa (Image J) (refer to Figure 5 Appendix and the Background section for information about the segments of dUTPase).

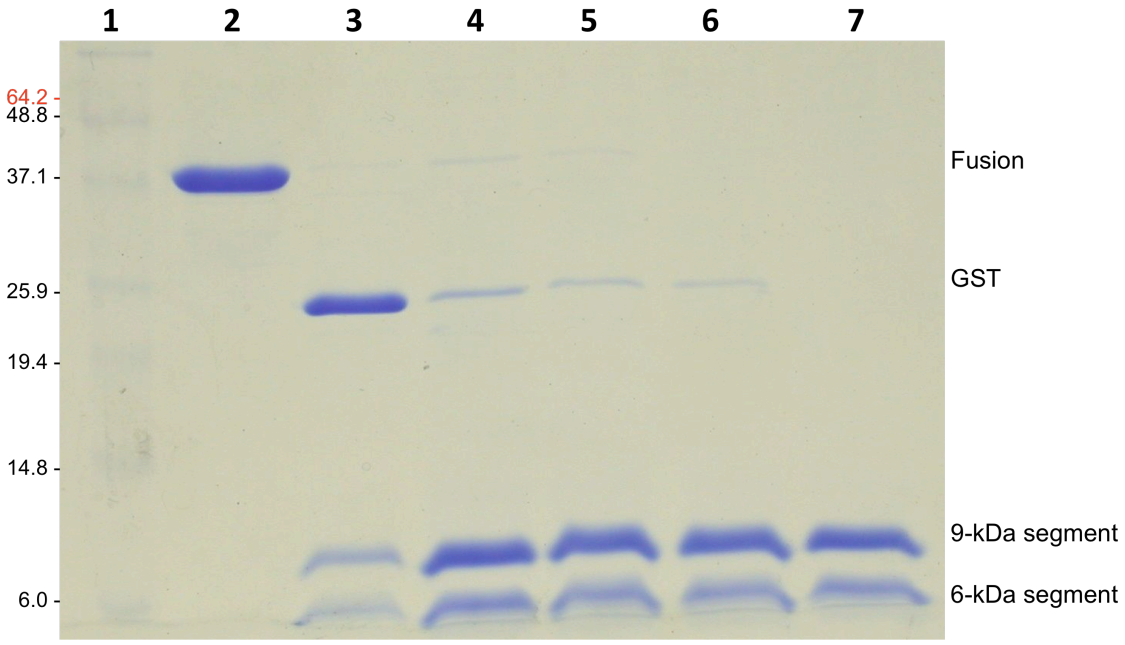
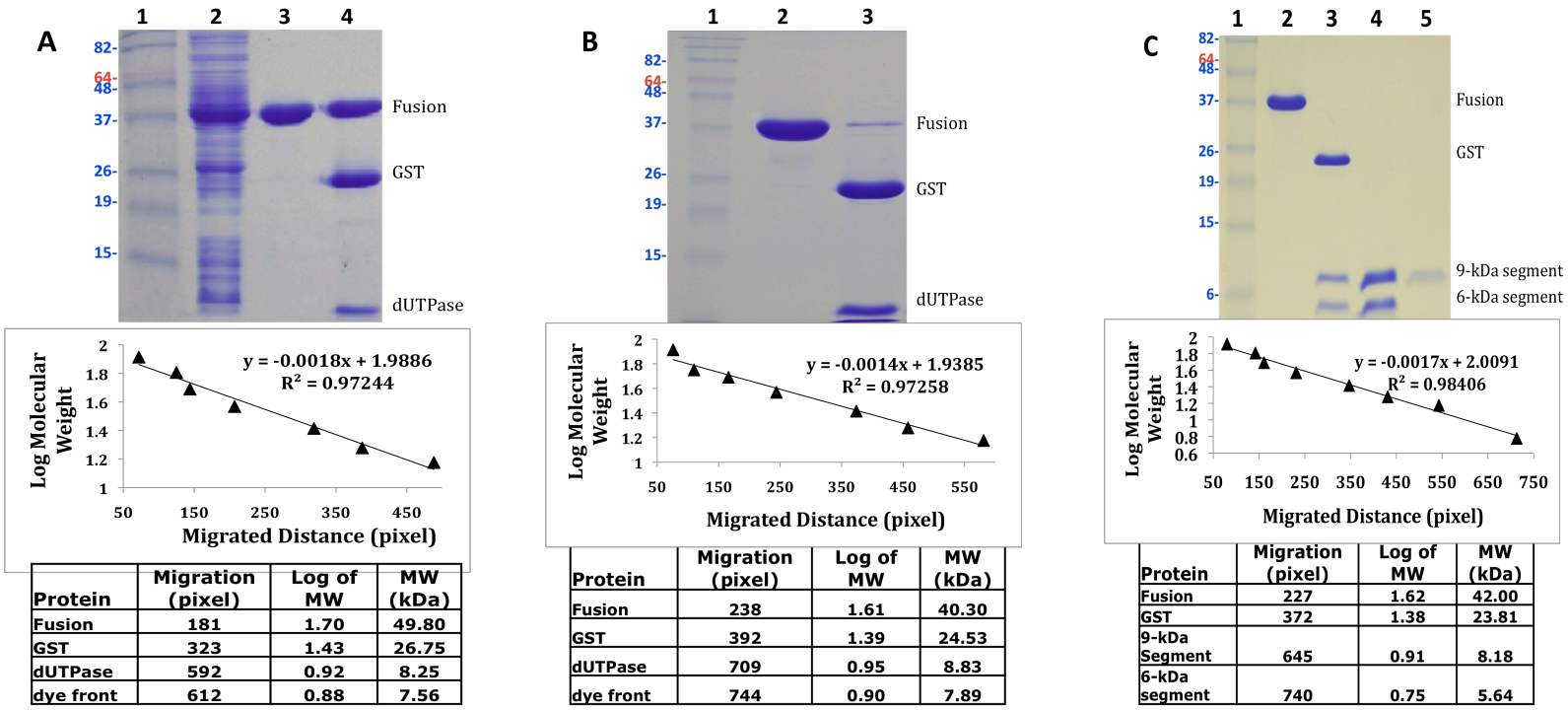


Figure 5 Appendix: **Highly pure dUTPase was obtained (16.25% acrylamide SDS-gel)**. Lane 1: molecular markers (kDa). Lane 2: GST-dUTPase fusion protein (Step 1, Figure 5). Lane 3: fusion protein cleaved by thrombin in the presence of 1.7 M urea (Step 2, Figure 5). Lanes 4-7: thrombin cleaved preparation after each of four additional treatments with glutathione beads (Step 4, Figure 5). The collected unbound protein after the last treatment with glutathione beads contain the 9-kDa and 6-kDa segments of dUTPase.



Protein	Migration (pixel)	Log of MW	MW (kDa)
Fusion	181	1.70	49.80
GST	323	1.43	26.75
dUTPase	592	0.92	8.25
dye front	612	0.88	7.56

Protein	Migration (pixel)	Log of MW	MW (kDa)
Fusion	238	1.61	40.30
GST	392	1.39	24.53
dUTPase	709	0.95	8.83
dye front	744	0.90	7.89

Protein	Migration (pixel)	Log of MW	MW (kDa)
Fusion	227	1.62	42.00
GST	372	1.38	23.81
9-kDa Segment	645	0.91	8.18
6-kDa segment	740	0.75	5.64

Figure 6 Appendix: **Molecular weights of GST-dUTPase fusion protein, GST, dUTPase, 9-kDa and 6-kDa segments are similar to the predicted sizes.** Graphs (log of molecular weight vs. migration distance) of the standard markers of each gel were constructed using ImageJ (Abramoff *et al*, 2004). The molecular weights of the different recombinant protein species were calculated using the best fit line for each graph. ***A***. 15% acrylamide SDS-gel. Lane 1: Molecular markers in kDa. Lane 2: Lysate. Lane 3: GST-dUTPase fusion protein. Lane 4: Fusion protein cleaved by thrombin without urea. ***B***. 15% acrylamide SDS-gel. Lane 1: Molecular markers in kDa. Lane 2: GST-dUTPase fusion protein. Lane 3: Fusion protein cleaved by thrombin in the presence of 2 M urea. ***C.*** 16.25% acrylamide SDS gel. Lane 1: Molecular markers in kDa. Lane 2: GST-dUTPase fusion protein. Lane 3: Fusion protein cleaved by thrombin in the presence of 1.7 M urea. Lane 4: 9-kDa and 6-kDa segments of dUTPase. Lane 5: 9-kDa segment of dUTPase. Panels A and B are used for Figure 7, and pancel C is used for Figure 8.

HPLC quantification of dUMP produced by dUTPase activity assays

Title: dUTPase Activity Assay-Temperature Dependent Date:15-17Jun2010

Sample ID	dUTP Conc. (µM)	dUMP Area (Abs*min)	dUMP (µmol)	Total dUMP (µmol)	Enzyme Amt. (µg)	Temperature (°C)	Incubation (min)	Specific Activity*1000 (µmol dUMP/min*µg dUTPase)	Average Spec. Act. (µmol dUMP/min*µg dUTPase)	Standard Deviation
11	5	4.29E-04	1.31E-03	6.56E-03	0.0466	30	5	28.16		
12	5	4.42E-04	1.35E-03	6.75E-03	0.0466	30	5	28.99		
13	5	4.58E-04	1.40E-03	7.00E-03	0.0466	30	5	30.03		
14	5	4.20E-04	1.28E-03	6.41E-03	0.0466	30	5	27.52		
15	5	4.48E-04	1.37E-03	6.85E-03	0.0466	30	5	29.41		
16	5	4.08E-04	1.25E-03	6.24E-03	0.0466	30	5	26.78		
17	5	4.09E-04	1.25E-03	6.25E-03	0.0466	30	5	26.84		
18	5	3.86E-04	1.18E-03	5.90E-03	0.0466	30	5	25.31		
19	5	3.80E-04	1.16E-03	5.81E-03	0.0466	30	5	24.95		
20	5	3.91E-04	1.20E-03	5.98E-03	0.0466	30	5	25.67	27.4	1.77

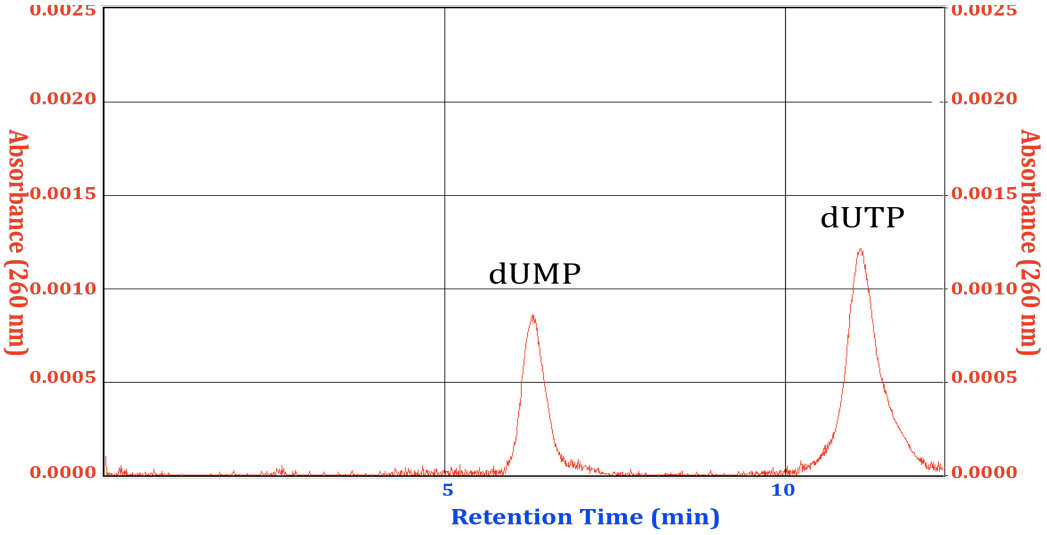


Figure 7 Appendix: **Measurement by HPLC of dUMP produced in dUTPase activity assays for the determination of optimal temperature and pH**. The area under the peak, corresponding to dUMP of the HPLC chromatogram, was integrated and the dUMP amount was determined by comparison to peaks of known dUMP standards (see Materials and Methods). The table shows how the average specific activity (27.4 µmol dUMP/min*µg dUTPase) was calculated from two trials each having five replicates (Sample IDs 11 through 20), and the calculated standard deviation (1.77) of the ten measurements. Shown below the table is an example HPLC chromatogram from an activity assay carried out at 30 ˚C.

Example of Data Analyses from Using Cresol Red Method

dUTPase Activity-Substrate Dependent (Cresol Red)

Sample #	dUTP (μM)	Rate (△Abs 573 nm*s^{-1})	Rate(Blank Cor.) (△Abs 573 nm*s^{-1})*10000	Average Rate (△Abs 573 nm*s^{-1})	Std Dev
1	0.1	0.00002	1		
2	0.1	0.00002	1		
3	0.1	0.000018	0.8		
4	0.1	0.00004	0		
5	0.1	0.00004	0		
6	0.1	0.00005	1		
7	0.1	0.00008	0		
8	0.1	0.00009	1		
9	0.1	0.0001	2		
10	0.1	0.00003	1		
11	0.1	0.00004	2		
12	0.1	0.00002	0	0.8	0.716
61	3	0.0004	39		
62	3	0.0004	39		
63	3	0.0004	39		
64	3	0.0004	36		
65	3	0.00035	31		
66	3	0.0004	36		
67	3	0.00045	37		
68	3	0.00045	37		
69	3	0.0005	42		
70	3	0.0004	38		
71	3	0.0004	38		
72	3	0.00035	33	37.1	2.906

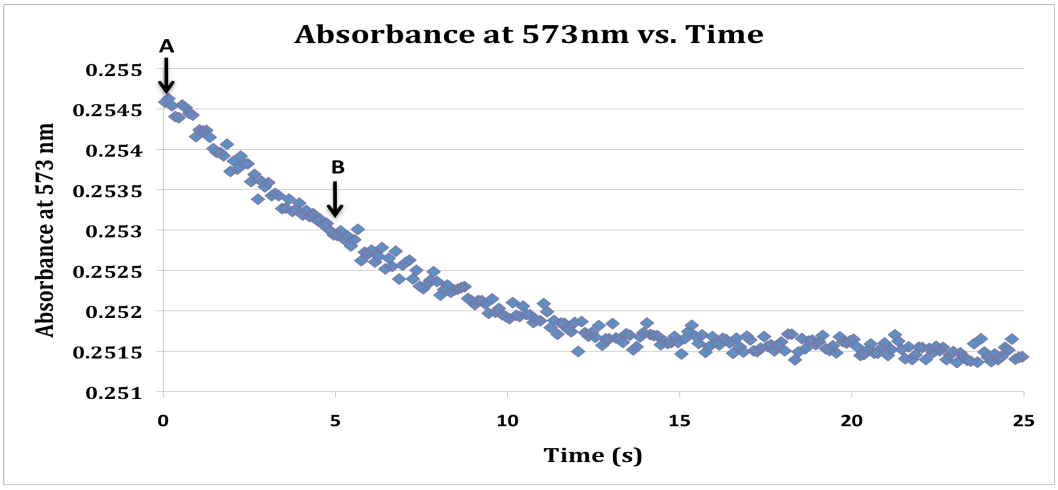


Figure 8 Appendix: **Cresol Red method is used to determine Km**. Enzyme activity is determined by measuring the decrease in absorbance at 573 nm over time. The slope between two points (labeled A and B) on the linear portion of the shown progress curve was recorded and used as the rate of activity (dUTP concentration = 3 μM). The rate is corrected by subtracting the rate without enzyme. An average rate was calculated from three trials each having four replicates. The standard deviation of the twelve measurements was determined.

The synthesized *D. discoideum* dUTPase gene was inserted into pUC57 and used for transformation. The isolated plasmid was digested with EcoRI and BamHI and the purified dUTPase gene is shown in lane 1 (Figure 1 Appendix A). The expression vector was pGEX-2T with the size of 4938 bp after linearized by EcoRI and BamHI and purified as shown in lane 2 (Figure 1 Appendix A). Panel B of Figure 1 Appendix shows a simple physical map of the gene inserted into the expression plasmid resulting in a predicted size of 5387 bp. Figure 2 Appendix shows the identified Top 10 cell colonies that contained the expression plasmid of the predicted size (5387 bp). The EcoRI-linearized plasmids have a size of 5387 bp indicated by the arrows. To show the presence of the dUTPase gene, the plasmids were digested with both EcoRI and BamHI. When digested with both restriction enzymes, the insert will be cut from the plasmid. This is shown by the presence of the bands at the 449 bp area, as shown by the arrows in Figure 3 Appendix. It was established that the insert was successfully put into the expression vector.

The GST-tagged dUTPase protein was expressed and purified using affinity chromatography. IPTG was used to induce expression of the GST-dUTPase fusion protein, and the protein was purified by treatment with glutathione beads (see Materials and Methods). A problem was the low efficiency of thrombin cleavage of GST-dUTPase fusion protein making the purification process problematic; therefore, denaturants were tested to increase the efficiency. One such denaturant was NaCl, and that the addition of NaCl did not increase the cleavage efficiency (Figure 4 Appendix A). Also, when there was protein precipitation when 2 M NaCl was used, and this is why lane 7 of Figure 4 Appendix shows lower protein amount compared to lane 4 though the detected protein

was only the fusion indicating cleavage efficiency was very low. The other denaturant used was urea. With the addition of 1.7 M urea, the thrombin cleavage efficiency greatly increased (Figure 4 Appendix B). Thrombin not only cleaved dUTPase from GST, but it also cleaved dUTPase into two peptides known as 6-kDa and 9-kDa segments. The segments were successfully purified through treatments with glutathione beads (see Materials and Methods) as shown in lane 7 of Figure 5 Appendix.

I then wanted to see if the bands corresponding to the GST-dUTPase fusion protein, GST, 9-kDa and 6-kDa segments of *D. discoideum* dUTPase are similar to the predicted sizes. The method used was log of molecular weights versus migration distances of the standard molecular markers to estimate to the size of the unknown protein species. Figure 6 Appendix shows the estimated molecular sizes of GST-dUTPase fusion protein, GST, 9-kDa and 6-kDa segments were similar to the predicted sizes. The average size for the 9-kDa segment is 8.4 kDa. The 6-kDa segment was observed when subjected to higher acrylamide (16.25%) gel electrophoresis with the only measured size of 5.6 kDa.

Figure 7 and Figure 8 of the Appendix were examples of data collection and analyses from kinetic studies using HPLC and cresol red (see Materials and Methods for more details).