Structural Analysis of Thymidylate Synthase Complimenting Protein from *Thermatoga Maritima* in the Presence of dTMP

Deepa Mathew

Stanford Linear Accelerator Center (SLAC)

Stanford Synchrotron Radiation Laboratory (SSRL)

Summer Internship

August 20, 2003

STRUCTURAL ANALYSIS OF THYMIDYLATE SYNTHASE COMPLIMENTING PROTEIN FROM *THERMATOGA MARITIMA* IN THE PRESENCE OF DEOXYTHYMIDINE 5'-MONOPHOSPHATE

Deepa Mathew
Summer Internship in the Stanford Synchrotron Radiation Lab, Stanford Linear Accelerator Center, Palo Alto, CA
June 11-August 22, 2003

Abstract:

A recent finding shows that thymidylate synthase-complimenting proteins (TSCPs) are responsible for the same mechanism of dUMP to dTMP conversion in prokaryotes as thimidylate synthase (TS) in eukaryotes. The existence of TSCPs, prevalent in prokaryotes which most often lack TS proteins, yields new opportunities for antimicrobial drugs. A detailed study of the structural and functional aspects of TSCPs in *Thermatoga Maritima* (TM0449) through protein crystallography yields further insight into the catalytic reaction, enzymatic activity, and inhibitor design of the proteins. Inhibition of TSCPs will provide a promising mechanism to halt the DNA production and replication of microbes and various forms of bacteria.

I. Introduction:

Recent developments in bioinformatics and structural genomics have led the way to the discovery of thimidylate synthase-complimenting protein of the thy1 gene of *Thermotoga maritma*. The protein exists almost exclusively in prokaryotes, and replaces the mechanism of 2'-deoxyuridine 5'-monophosphate (dUMP) to 5'-deoxythymidine 5'-monophosphate (dTMP) conversion by thymidylate synthase (TS) in eukaryotes.

Thymidylate synthesis is the sole de nova catalytic reaction that converts dUMP to dTMP in eukaryotes when additional sources of thymidylate are not present. Thymidylate synthesis involves the methylation of dUMP to produce dTMP, which comprises the DNA nitrogen base thymine. During thymidylate synthesis, CH_2H_4folate serves as a carbon donor to produce H_4Folate, while $FADH_2$ acts as a reducing agent to form FAD. Since dTMP lies in the path of DNA production, the inhibition of TS has been the at the center of anti-cancer drugs for several years (See Figure 1).

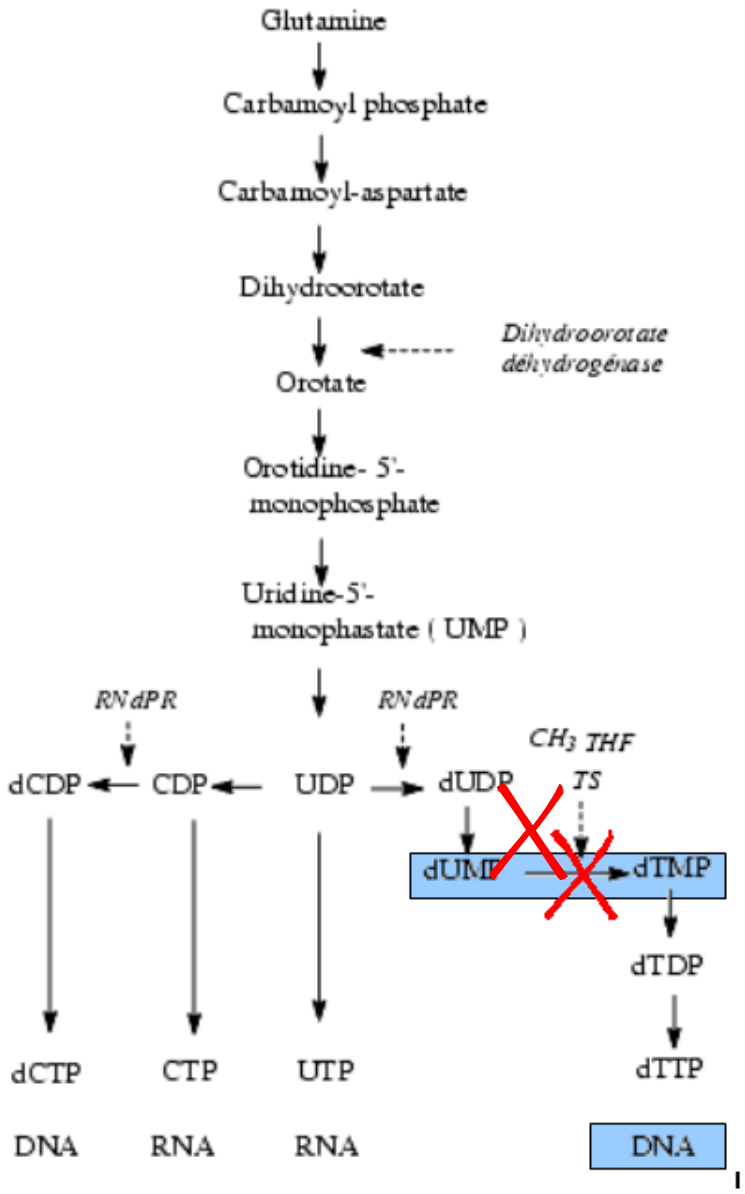

Figure 1: Thymidylate synthase complimenting protein (TSCP) in prokaryotes performs the identical conversion of dUMP to dTMP as thymidylate synthase (TS) in eukaryotes. An analysis of TSCP reveals that the newly discovered protein differs significantly in structure from TS. Further insight into the catalytic reaction and properties of TSCPs allows for the development of inhibitors and shed light on new antimicrobial compounds.

A structural and functional analysis of TSCPs reveals that the structure of the thy1 enzyme shares little similarity to the structure of TS proteins. The catalytic reaction also of dUMP to dTMP conversion in TSCP also differs from that of TS. Although TS has previously been portrayed as the sole de nova catalytic reaction for synthesizing dTMP in all species, a recent breakthrough reveals that the TS, or thyA gene is absent in most prokaryotes. Instead, TS is substituted by TSCP in most prokaryotes. The successful inhibition of TSCP proteins would lead to "thymineless" cell death, and the study of substrate compounds paves the way to the development of antimicrobial drugs.

II. Background Information:

Project Description:

The project consisted of three parts: the growth and optimization of TSCPs protein crystals of TM0449 containing dTMP substrate, the diffraction and data collection procedure, and the refinement and analysis of the protein structure. To understand the interaction and bonding mechanism between the protein and substrate, we crystallized and determined the structure of a complex of TM0449 in the presence of substrate dTMP.

Protein Structure

Amino acids are the "building blocks," or monomers of proteins. Amino acids are composed of an amino group (basic), a carboxyl group (acidic) , a hydrogen atom, and an R group attached to an α carbon atom. Mirror-image configurations of the four groups around the α carbon atom form the L isomer and D isomer, and proteins are composed of L isomers exclusively. A range of twenty amino acids act as structural units for proteins.

Amino acids are characterized by three primary subgroups: hydrophobic, polar (charged), and polar (uncharged). The hydrophobic, or nonpolar amino acids include alanine (Ala, A), valine (Val, V), leucine (Leu, L), isoleucine (Ile, I), praline (Pro, P), methionine (Met, M), phenylalanine (Phe, F), and tryptophan (Trp, W). The polar, charged amino acids are composed of aspartic acid (Asp, D), glutamic acid (Glu, E), lysine (Lys, K), Arginine (Arg, R), and histidine (His, H). The polar, uncharged

amino acids consists of glycine (Gly, G), serine (Ser, S), threonine (Thr, T), cysteine (Cys, C), tyrosine (Tyr, Y), asparagines (Asn, N), and glutamine (Glx, Z).

Amino acids are linked by peptide bonds, in which the α-carboxylic group and α-amino group of adjacent amino acids are joined. The formation of a dipeptide from two amino acids results in the loss of a water molecule, and a series of peptide bonds linking amino acids forms a polypeptide chain.

Structure in Understanding Function

Protein crystallography provides an insight into the structure and bonding of proteins by revealing the three-dimensional atomic configuration of the atoms and amino acids that make up proteins. The structure of proteins unveils the atomic detail of how proteins fold, bond to other molecules, and operate as enzymes. The three-dimensional configuration is also vital in providing molecular detail, such as locating a protein's active site, opening fields for antibiotics and pharmaceutical drug compounds that target specific proteins.

Common Proteins: Hemoglobin

The molecular and structural representations of protein revealed using x-ray crystallography are important in determining and understanding the function and interactive behavior of familiar proteins. For instance, the three-dimensional structure of hemoglobin unveiled the mechanisms by which hemoglobin functions within the body.

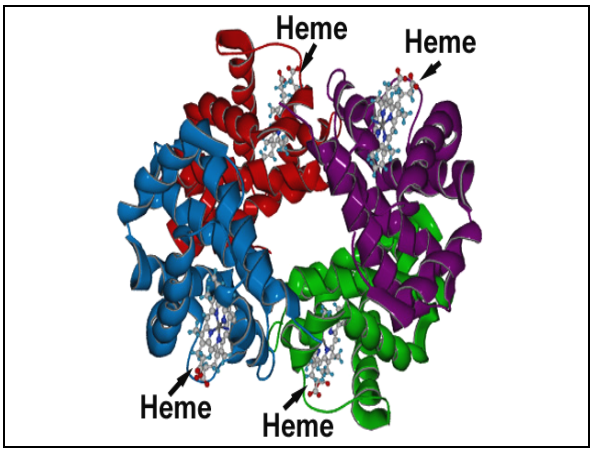

Figure 2: Molecular Model of Hemoglobin with Heme (Ball and stick) subgroups

Abundant in red blood cells, hemoglobin is the oxygen carrier in blood. Hemoglobin binds with oxygen molecules in the lungs, transports and releases the oxygen to various body tissues, binds with carbon dioxide from tissues, and carries the carbon dioxide back to the lungs. X-ray crystallography reveals the tetrameric structure of hemoglobin, as well as four heme prosthetic groups to which oxygen binds, enabling hemoglobin to bind with up to four oxygen molecules.

III. Crystallography Information:

Unit Cells, Crystals, X-rays, and Diffraction:

Solid crystal structures assume an array of atoms and ions. The basic unit of the lattice structure is referred to as the unit cell. The x, y, and z edges of the unit cell are designated as a, b, and c, with angles of α, β, and γ respectively. Varying dimensions and levels of symmetry give rise to seven crystal classes: triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, and cubic.

Solid crystal structures assume an array of atoms and ions. The basic unit of the lattice structure is referred to as the unit cell. The x, y, and z edges of the unit cell are designated as a, b, and c, with angles of α, β, and γ respectively. Varying dimensions and levels of symmetry give rise to seven crystal classes: triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, and cubic.

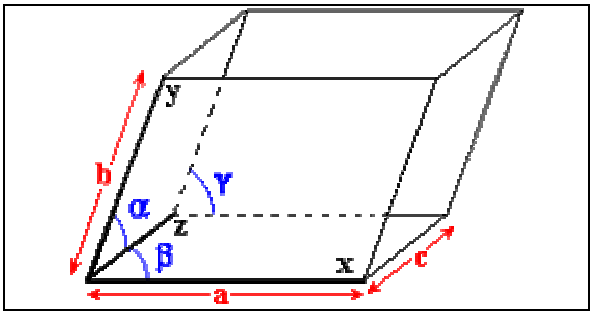

Figure 3: Unit Cell Dimensions of Triclinic Crystal Class

X-rays are vital in protein crystallography and structure determination because x-rays have approximately the same wavelengths as the wavelengths between the individual atoms of the crystal. Upon striking the individual atoms within the crystalline structure, the high-energy x-rays from the synchrotron source scatter to create a diffraction pattern. Synchrotron sources allow the x-ray wavelengths to be adjusted in correlation to the space groups and size of the unit cell. The diffraction pattern is recorded on a diffraction film, appearing as spots with varying intensities. A series of calculations and computer programs are then used to interpret the collected data to create a 3-d electron density map based on the intensity and position of the diffraction pattern.

When x-ray beams strike an atom, the beams are diffracted by the atom's electrons at an angle of 2θ. Although the diffraction angles and intensities of the x-ray waves can be located by the detector, the phase cannot be determined by

intensities alone. X-ray waves in phase are reinforced, while x-rays out of phase are cancelled. Phase and amplitude information are vital in solving the protein structure, and the missing information is often referred to as the "phase problem."

Three methods are commonly used to recover vital phase information: molecular replacement (MR), multiple isomorphous replacement (MIR), and multiwavelength anomalous diffraction phasing (MAD).

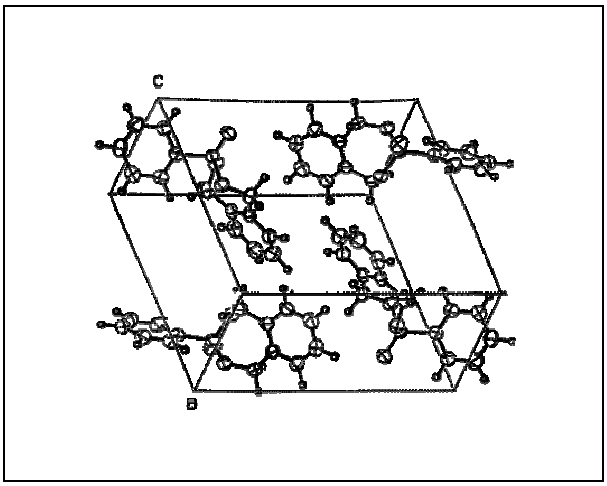

Figure 4: Protein aligned within a unit cell of a crystal structure.

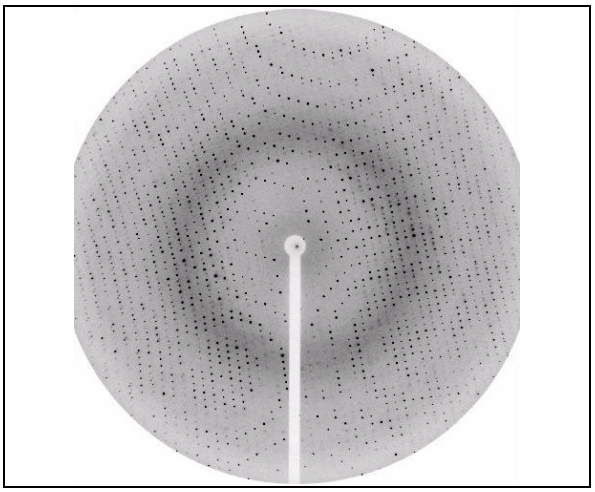

Figure 5: Diffraction Pattern resulting from the scattering of x-rays from atoms within the crystal

Structure Solution Techniques:

Three methods are commonly used to recover vital phase information: molecular replacement (MR), multiple isomorphous replacement (MIR), and multiwavelength anomalous diffraction phasing (MAD).

Molecular Replacement (MR)

Molecular replacement is a technique that can be used if a homologous structure of the crystal is already known. Molecular replacement involves determining the translation and rotational angle that positions the analogous model structure in the unit cell of the sample with the aim of a resulting diffraction measurements that correspond to the diffraction measurements of the sample.

Multiple Isomorphous Replacement (MIR)

During multiple isomorphous replacement, protein crystals are soaked in an aqueous buffer solution containing heavy atom. The heavy atom crystals and protein crystals must be isomorphous for the procedure. The heavy atoms allow the crystals to be treated as larger molecules, and the diffraction pattern is derived from the Fourier transform, which uses the vectors of the protein and heavy atom samples.

Figure 3: Multiple Isomorphous Replacement (MIR)

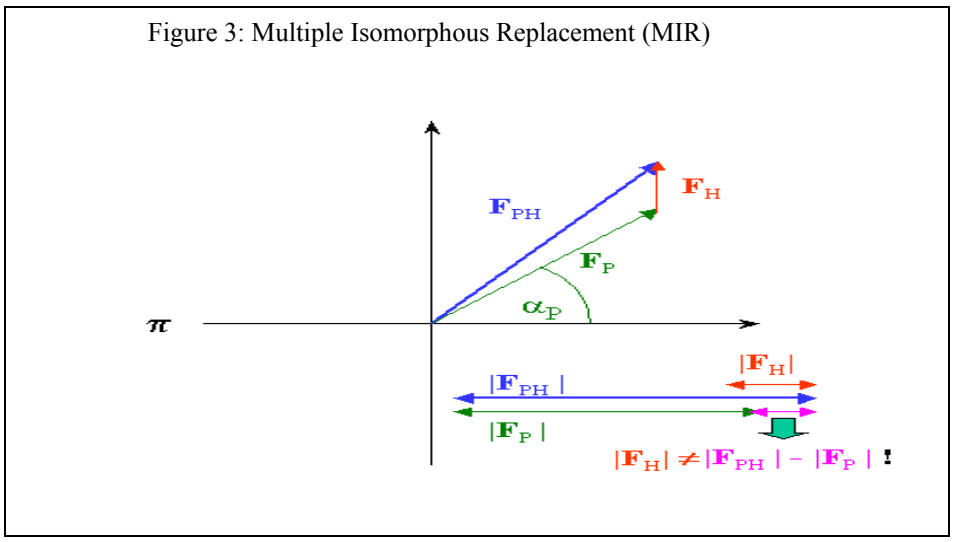

Figure 6: MIR Vector Calculations

Multiwavelength Anomalous Diffraction Phasing (MAD)

In the multiwavelength method, the wavelengths of x-ray beams striking the crystal are varied around the absorption edges of atoms in the crystalline structure. The X-ray absorption patterns of these atoms create differing diffraction patterns that are relevant to the phase information. In multiwavelength method, the wavelengths of x-ray beams striking the crystal are varied around the absorption edges of atoms in the crystalline structure.

IV. Crystallography of TM0449 with dTMP:

Crystallization Techniques:

Hanging Drop Vapor Diffusion Method:

A drop containing a mixture of concentrated protein, buffer, and precipitant is placed on a coverslip over a reservoir of buffer and precipitant mixture in a vacuum sealed well. The drop maintains a lower concentration than that of the reservoir solution.

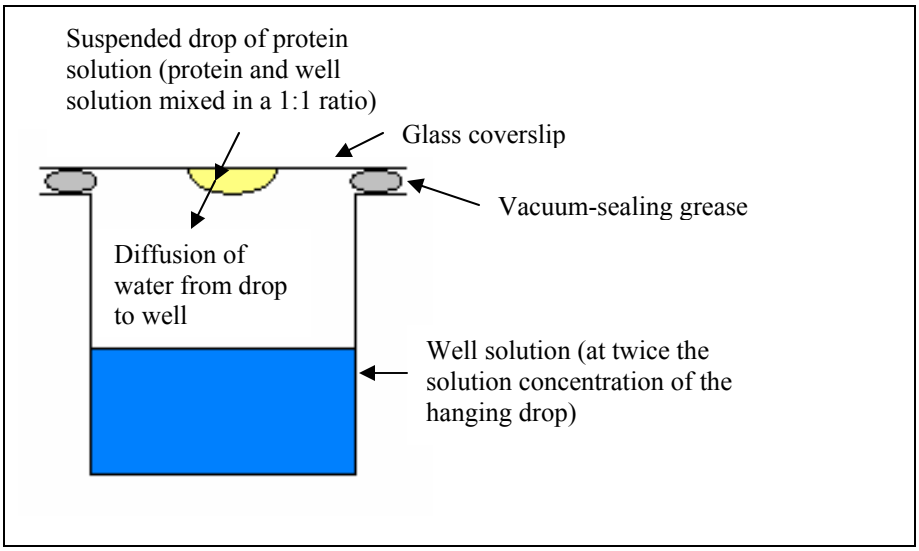

Figure 7: Hanging Drop Vapor Diffusion Well Setup

Solid crystal structures assume an array of atoms and ions. The basic unit of the lattice structure is referred to as the unit cell. The x, y, and z edges of the unit cell are designated as a, b, and c, with angles of α, β, and γ respectively. Varying dimensions and levels of symmetry give rise to seven crystal classes: triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, and cubic.

Sitting Drop Vapor Diffusion Method:

Similar to the setup of the hanging drop, a drop containing a mixture of concentrated protein is placed on a mount higher than the reservoir solution over a reservoir of buffer and precipitant mixture in a vacuum sealed cell. The drop maintains a lower concentration than that of the reservoir solution. As the system reaches equilibrium, water is transferred from the protein drop to the reservoir

solution until the precipitant concentration in the protein drop equals the precipitant concentration in the reservoir solution.

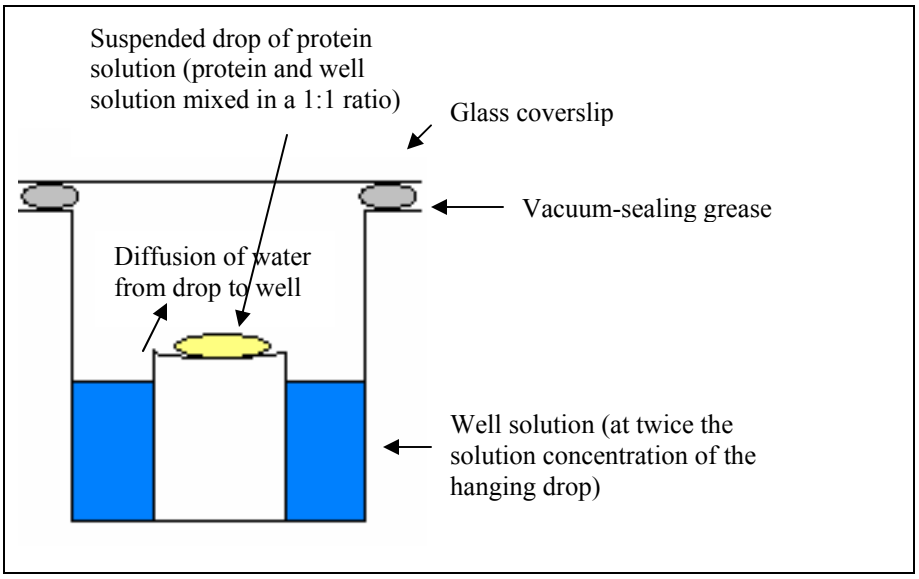

Figure 8: Sitting Drop Vapor Diffusion Well Setup

Dialysis Method:

Protein and solution and Precipitant are placed in adjoining compartments separated by a semi-permeable membrane. The buffer and precipitant concentrations can be altered, while the protein concentration remains constant. Additionally, the water and precipitant are exchanged through the membrane layer, while the protein solution remains set. A common method of dialysis involves placing the protein-containing solution into a dialysis button that is sealed with a membrane layer.

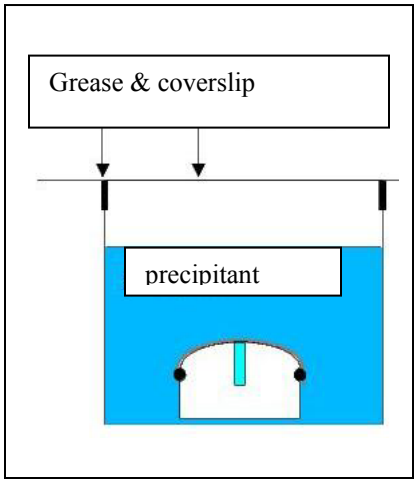

Figure 9: Dialysis Well Setup

Complex Preparation

Solution	Desired pH	Mol. Weight
1 M Tris	8.0	157.6 $^{g}/_{mol}$
1 M Hepes	7.5	238.3 $^{g}/_{mol}$

Tris Calculations: (157.6 $^{g}/_{mol}$ ÷ 1000 mL) * 40 mL = **6.30 g** Tris.. Dissolve 6.30 g of Tris in H_2O to make 40 mL of solution. Add NaOH to solution to reach pH 8.0.
Hepes Calculations: (238.3 $^{g}/_{mol}$ ÷ 1000 mL) * 40 mL = **9.53 g** Hepes. Dissolve 9.53 g of Hepes in H_2O to make 40 mL of solution. Add NaOH to solution to reach pH 7.5.

Protein Concentration & Tests to Check Protein Quality

The Bradford Protein Assay is a procedure for determining protein concentration from the change in absorbance caused by Coomassie Blue G-250 binding to protein, as the solution turns blue when the dye binds with the protein. A blank solution is made of a mixture containing 200 µL 10mm Tris (7.5 pH) buffer solution and 200 µL BioRad solution. A sample contains 200 µL 10mm Tris (7.5 pH), 200 µL BioRad solution, and 1 µL of protein. The solutions are measured at 595 nm with an absorbance range of 300nm to 700 nm on a single reference wavelength spectrometer.

Well Setup

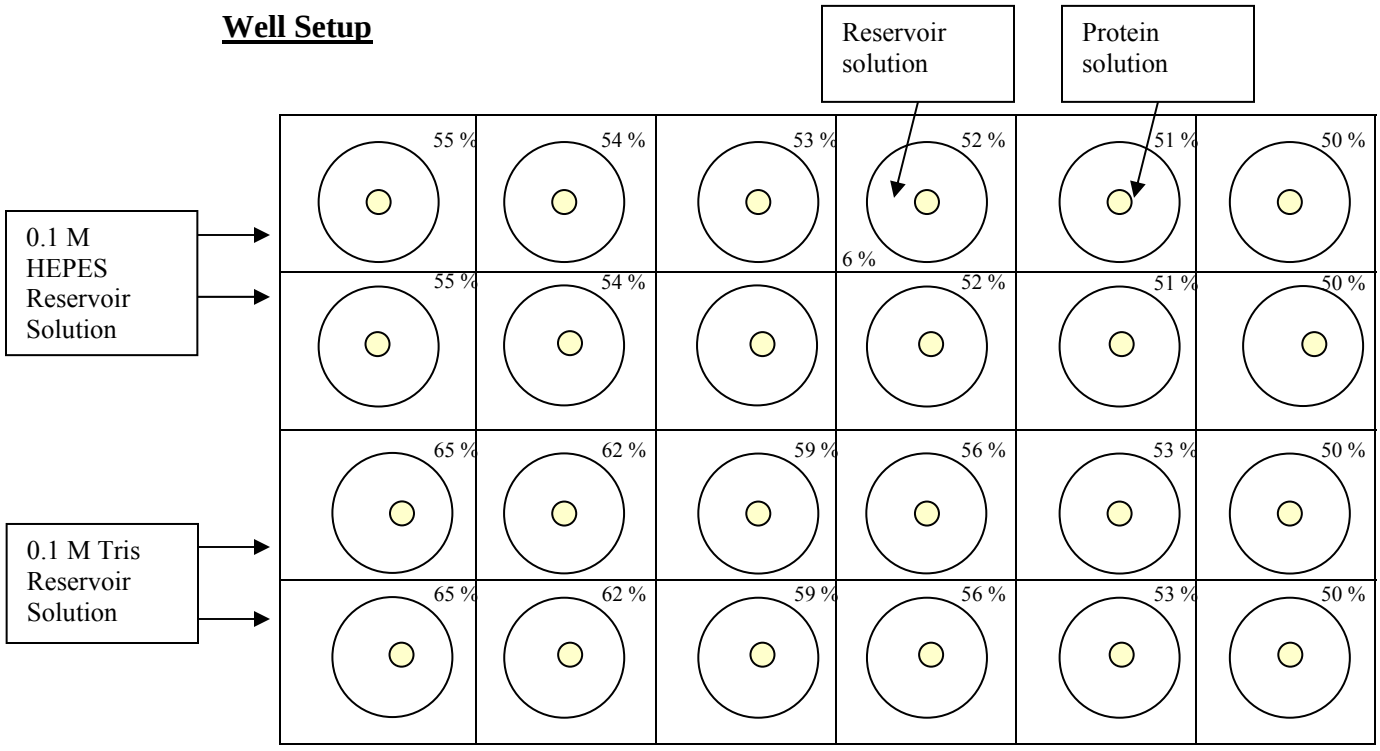

Figure 10: TSCP Well Grid

Key:

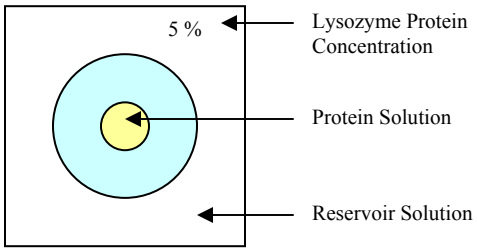

Figure 11: TSCP Grid Key

Lysozyme Crystallization:

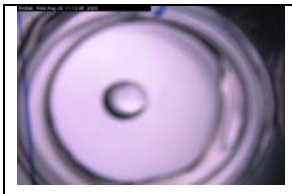	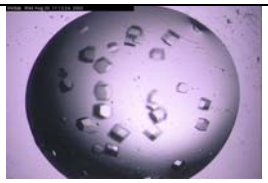	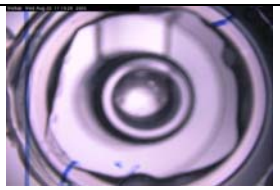	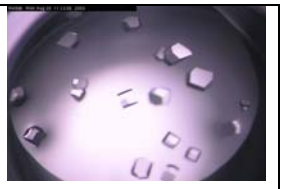
Figure 12: Lysozyme Hanging Drop	Figure 13: Lysozyme Hanging Drop Crystals	Figure 14: Lysozyme Sitting Drop Setup	Figure 15: Lysozyme Sitting Drop Crystals

A) Lysozyme Solution Preparation
   a. Prepare 0.1 M Sodium Acetate
      i. Dissolve 13.608 g Sodium Acetate into 100 mL of distilled H_2O.
      ii. Set H_2O pH to 4.20 by adding concentrated HCl solution.
      iii. Prepare 10 mL of 20 mM (0.02M) solution of Sodium Acetate from stock solution. For 2 mL divisions, remove 2 mL of stock solution and dilute up to 10 mL with H_2O for 0.02 Mconcentration.
      iv. If a large quantity of HCl is added and the volume rises significantly, recalculate the final concentration. For example, if the final volume is 120 mL, the final concentration is $^{0.1\ mol}/_{0.12\ L}$, or 0.833 M.

B) Crystallization with Hanging Drops
   a. Prepare 5% NaCl solution
      i. To prepare a 500 µL solution, add 250 µL NaCL to 250 µL Sodium Acetate.
   b. Prepare 6% NaCl solution
      i. To prepare a 500 µL solution, add 300 µL NaCL to 200 µL Sodium Acetate.
   c. Prepare 7% NaCl solution
      i. To prepare a 500 µL solution, add 350 µL NaCL to 150 µL Sodium Acetate.

C) Lysozyme Dillution
   a. Initial Concentration: 40 $^{mg}/_{mL}$
   b. Make 30 $^{mg}/_{mL}$ solution.

i. Take 75 µL of concentrated solution and bring to 100 µL with H_2O.

ii. Calculations: 40 mg/mL * ? = 30 mg/mL * 100 mL.

iii. ? = 75 mL

c. Make 20 $^{mg}/_{mL}$ solution.

i. Take 50 µL of concentrated solution and bring to 100 µL with H_2O.

d. Make 15 $^{mg}/_{mL}$ solution.

i. Take 37.5 µL of concentrated solution and bring to 100 µL with H_2O.

D) Lysozyme Solution Grid

Key:

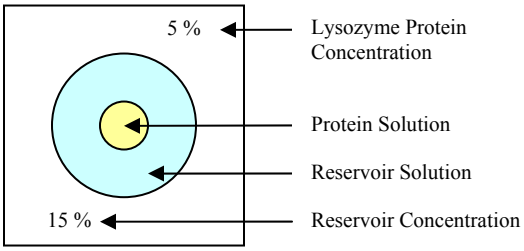

Figure 16: Lysozome Grid Key

Grid Setup:

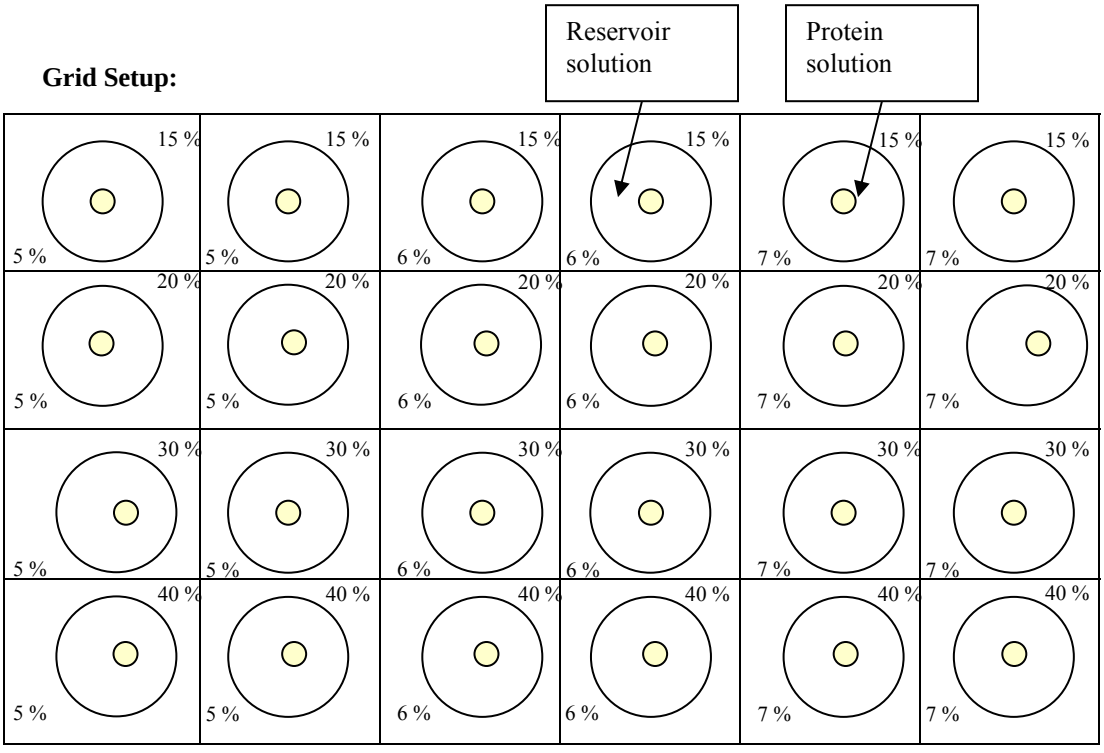

Figure 17: Lysozyme Well Grid

dTMP Crystallization

Figure 18: TSCP with dTMP crystals in a hanging drop. The solution is tinted yellow due to the presence of the protein..

Figure 19: TSCP with dTMP crystals in a hanging drop. Gel Precipitate has formed in the solution.

Well Observations

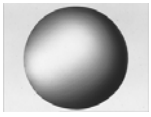

Fig 20: Clear Drops signal Soluble Protein

Fig 21: Precipitate

Fig 22: Gel Precipitate

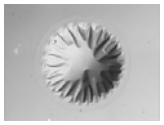

Fig 23: Skin Layer

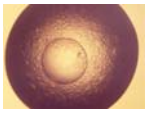

Fig 24: Phase Separation

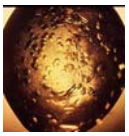

Fig 25: Oil Droplets

Fig 26: Microcrystalline Precipitate

Fig 27: 1D-Needles

Fig 28: 2D-Plates

Fig 29: 3D-Boulders

V. Visualization Programs

Graphics Programs and Crystallography:

Data collected along the x-ray beamline consists of photographs of a circular dimension of spots with varying intensities, which are stored as pixels for computer analysis. Computer programs are then used to interpret the data. Graphics programs are first used to interpret the intensities and sizes of the pixels to create a rough 3-d model of the electron density clouds.

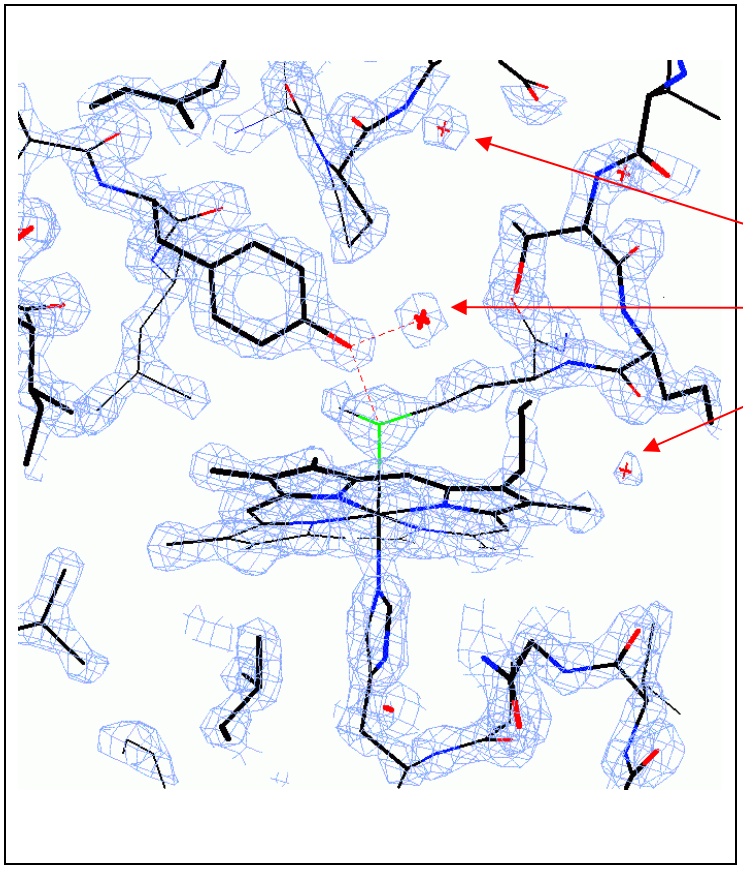

Figure 30: Graphics Programs interpret diffraction spots and output rough electron density maps that are then used to create a structural model of the protein.

Connecting Diffraction Map to Protein Model:

Subsequent steps in building a model require the use of specific programs. The "O Graphics" program is used to generate a mainchain trace of the alpha-carbon atoms in the molecule, determine where the sequence matches the density, build a rough model, optimize the fit of the model to the density, and evaluate the model.

VI. Results

Structural Representation of TM0449 with dTMP

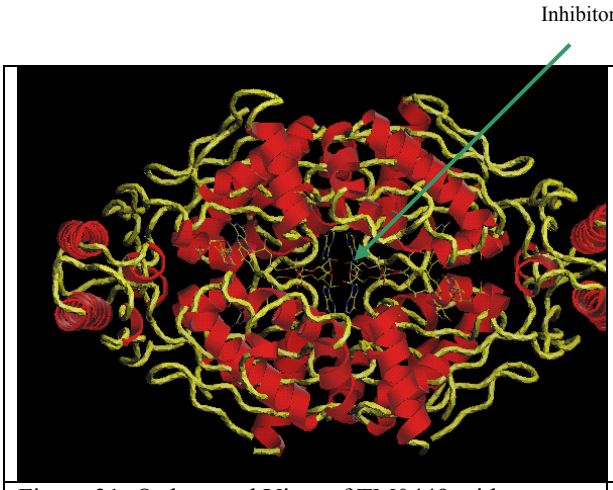

Figure 31: Orthogonal View of TM0449 with inhibitor.

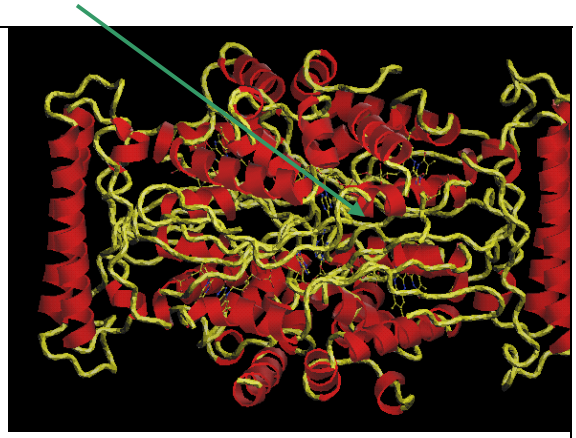

Figure 32: View of TM0449 perpendicular to the view of Figure 20

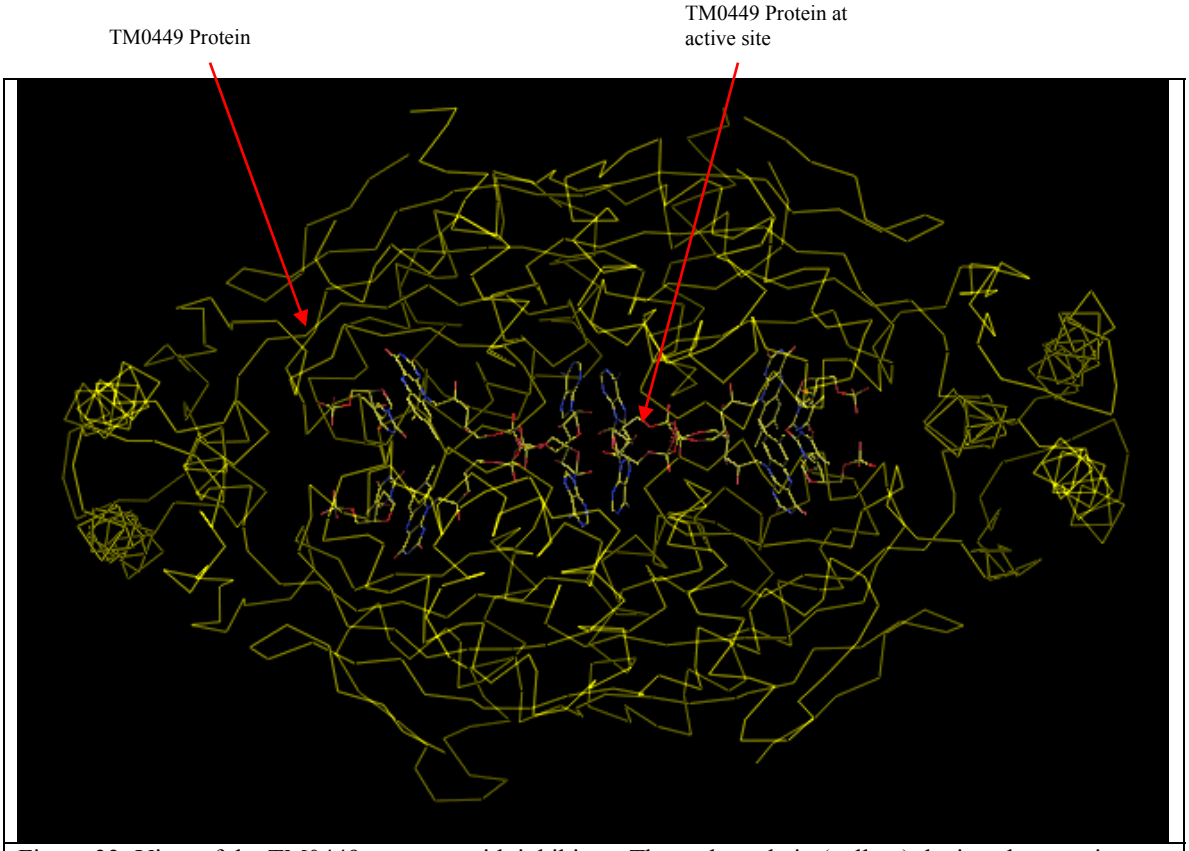

Figure 33: View of the TM0449 tetramer with inhibitor. The carbon chain (yellow) depicts the protein. The inhibitor is positioned at the active site and displayed in a ball-and-stick representation colored by atom.

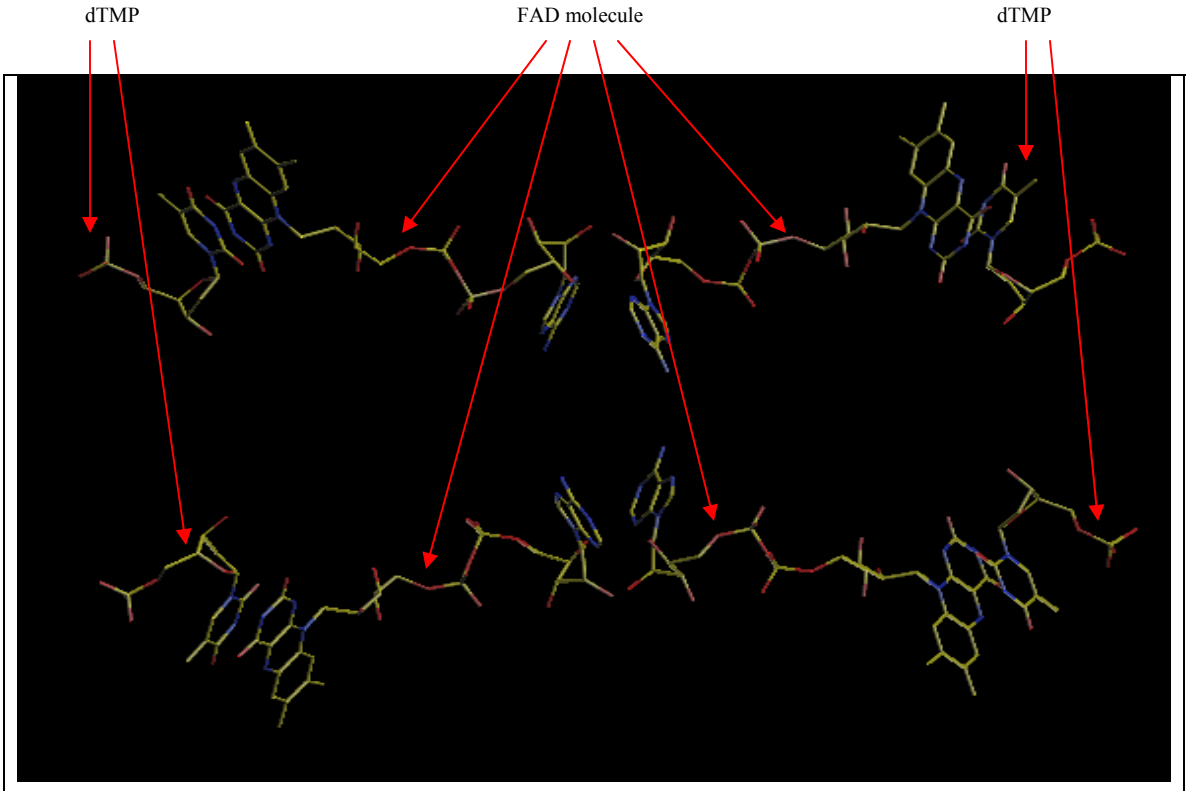

Figure 34: Inhibitor showing FAD and dTMP

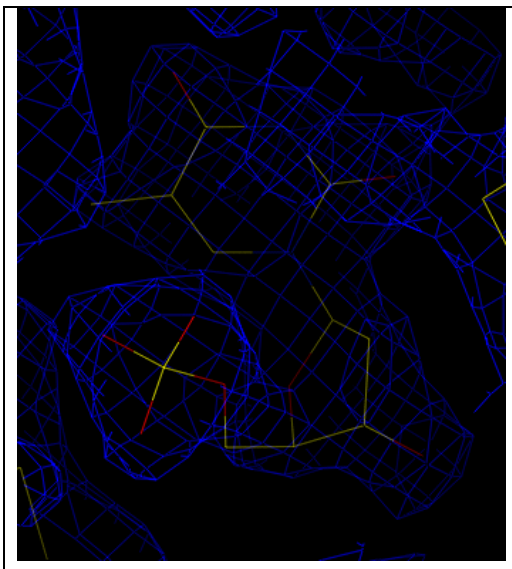

Figure 35: dTMP structure with electron density (blue).

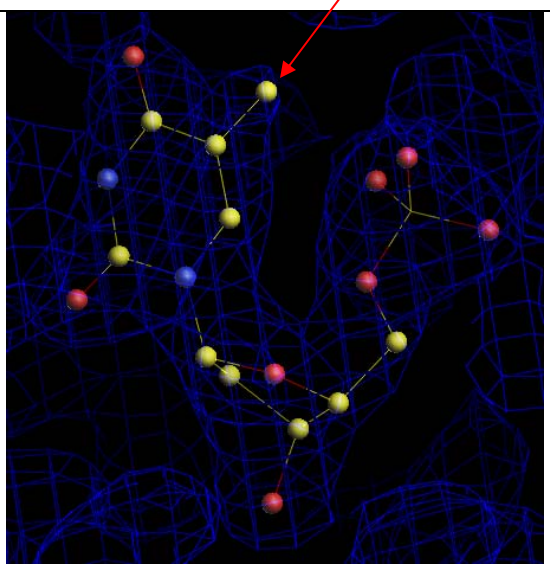

Figure 36: Ball and Stick representation of dTMP with surrounding electron density shows the methyl group donated by tetrahydrofolate.

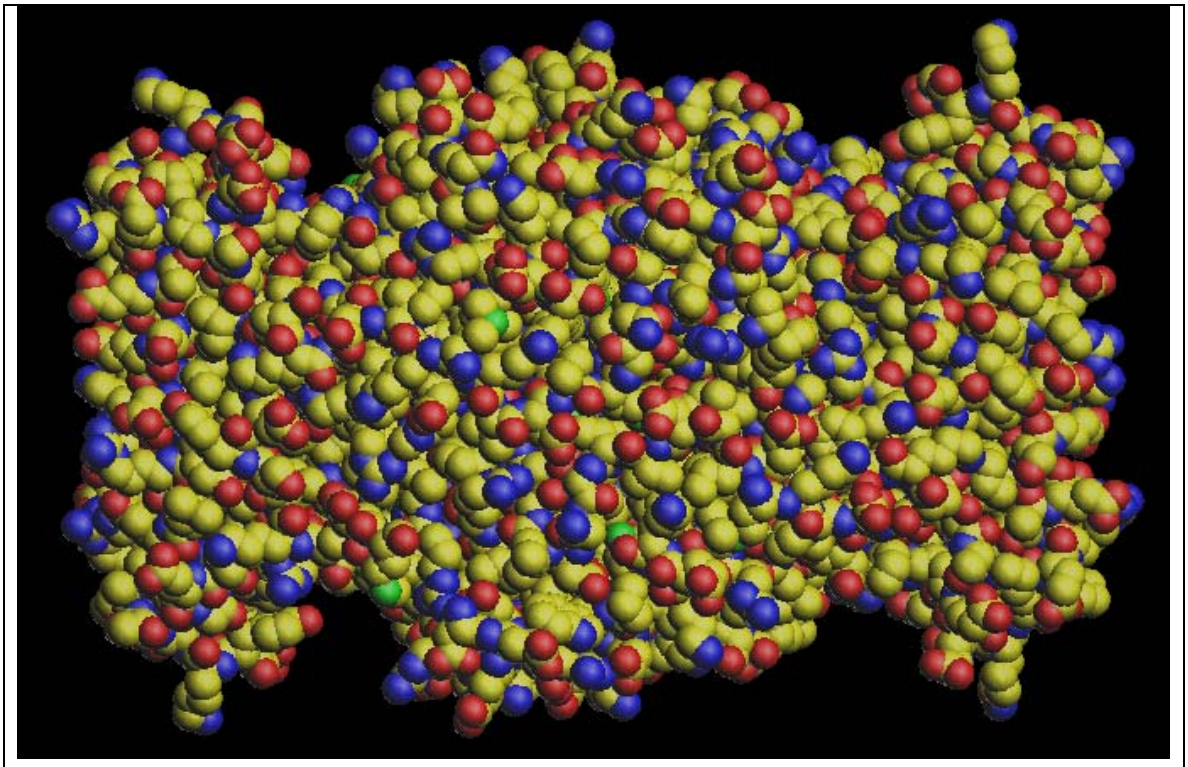

Figure 37: Space-Filling Model of TM0449 with dTMP. Yellow atoms depict Carbon, red atoms depict Oxygen, blue atoms depict Nitrogen, and green atoms depict Sulfur.

VII. Discussion:

An analysis of the TM0449 crystal structure revealed the presence of dTMP product bound to the active site. The dTMP inhibitor consisted of four FAD molecules with four smaller dTMP molecules affixed at the end of each FAD molecule. The dTMP molecule differs from the dUMP reactant due to the presence of a methyl group affixed at the end of the dTMP molecule.

The catalytic reaction of dUMP to dTMP conversion of TSCP also differs significantly from the catalytic reaction of TS. Like TS, TSCP uses cofactor 5, 10-methylenetetrahydrofolate (MTF) for carbon exchange, but TSCP also uses flavin adenine dinucleotide (FAD) and a pyridine nucleotide, NADH or NADPH, for its catalytic reaction.

The presence of TSCP exclusively in prokaryotes and its rare occurrence in eukaryotes provides significant promise for the development of a successful inhibitor to target pathogenic bacteria. The development of a TSCP inhibitor with little crossreactivity with TS will lead to "thymineless" cell death of target bacteria and yield promising designs for antimicrobial drugs. Because FAD substrate mimics would affect many vital reactions that also utilize FAD, an inhibitor that modeled the product of the catalytic reaction, dTMP, would adapt to the active site.

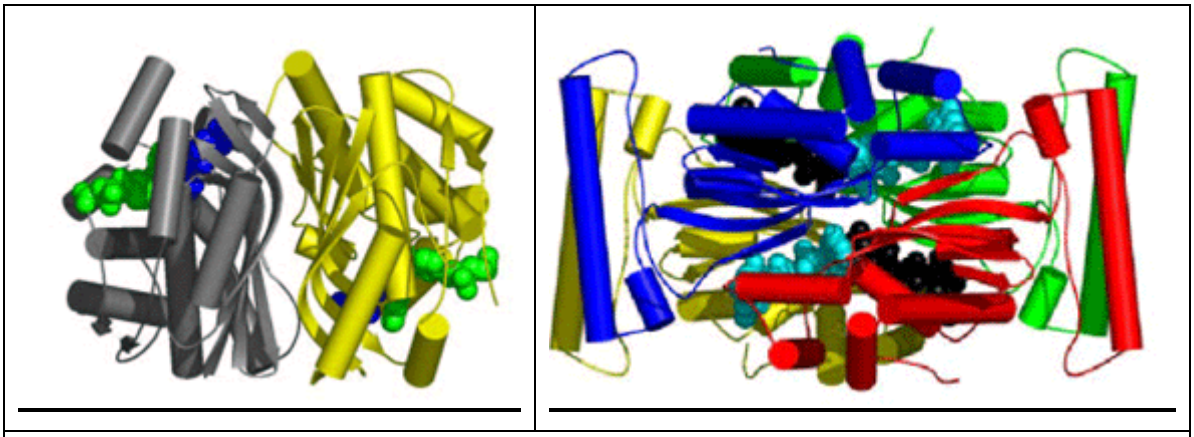

Figure 38: Comparison of TS protein on the left to TSCP protein on the right reveal little structural similarity. Spherical molecules indicate substrates bound at the active site.

Thymidylate synthase complimenting proteins (TSCPs) are responsible for dUMP to dTMP conversion in prokaryotes as thymidylate synthase (TS) in eukaryotes. In most prokaryotes, TSCP substitutes the function of TS. The inhibition of TS can halt the production of dTMP, which comprises the DNA nitrogen base thymine. The successful inhibition of TSCP leads to "thymineless" cell death, and the search for a TSCP inhibitor with little crossreactivity yields a promising technique to combat pathogenic bacteria.

VIII. Conclusion:

TM0449 protein crystals in the presence of dTMP were successfully grown in optimal conditions. The crystals were then diffracted at the Advanced Light Source at the Berkeley National Laboratory and produced refined diffraction patterns. The collected data was then interpreted using Nuclear Magnetic Resonance (NMR) systems to construct the protein structure. A previous analysis of TSCP in the

presence of the dUMP reactant displayed that dUMP bound to the active site. The dTMP product differs from dUMP due to the presence of a methyl group donated by tetrahydrofolate in the catalytic reaction. The structural analysis of TM0449 TSCP in the presence of dTMP revealed that in the catalytic reaction converting dUMP to dTMP, the product, dTMP, also successfully bound to the protein at the active site.

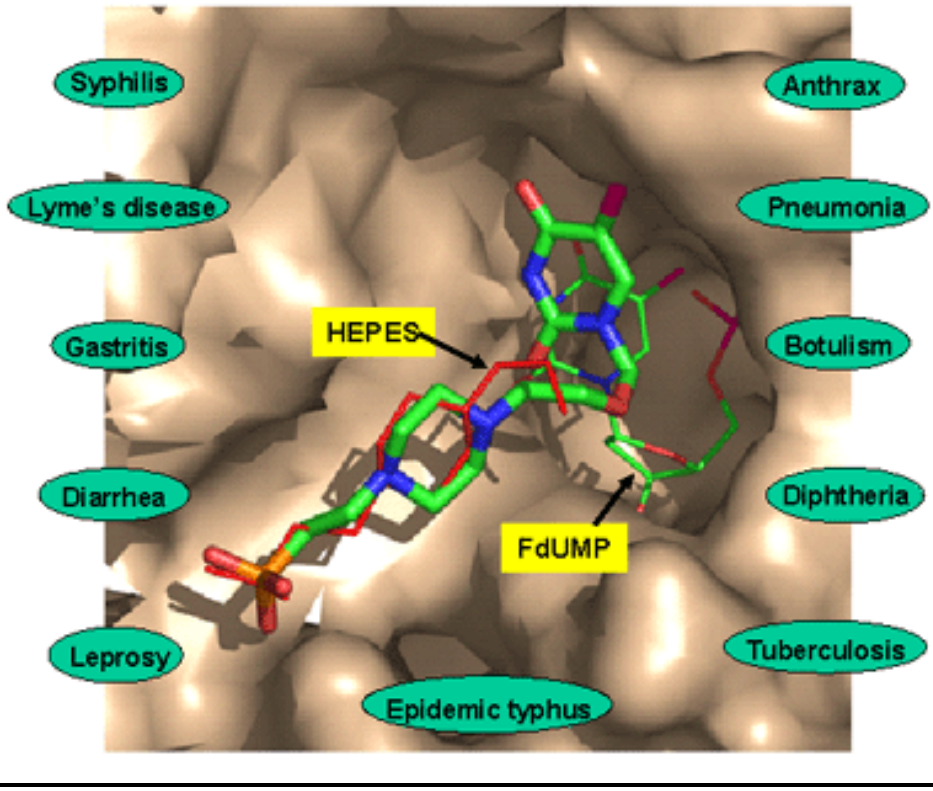

Figure 39: Surface representation of TSCP with dUMP inhibitor (center in color) and diseases that a successful inhibitor could suppress.

VIX. Acknowledgements:

I would like to take this opportunity to thank Mike Soltis for providing this internship opportunity and for arranging an x-ray diffraction visit to the Lawrence Berkeley National Laboratory for data collection and Dr. I. I. Mathews for his invaluable guidance and inspiration. Additionally, I also thank Amanda Prado, Ashley Deacon, Chu-Young Kim for their project support, and the Structural Molecular Biology Group for making my SSRL work experience an enjoyable and memorable one.

References

Jones, C., Mulloy, B., and Sanderson M.R., *Crystallographic Methods and Protocol*, New York: Humana Press, 1979.

Jones, T.A., Kjeldgaard, M., *Electron-Density Map Interpretation*, Methods in Enzymology, Academic Press, 1997.

Mathews, I. I., Deacon, A.M., Canaves, J.M., McMullan, D., Lesley, S.A., Agarwalla, S., and Kuhn, P., (2003) Structure *11*, 677-690.

McPherson, Alexander, *Preparation and Analysis of Protein Crystals*, New Jersey: Krieger Publishing Company, 1982.

Stryer, Lubert, *Biochemistry*, New York: W.H. Freeman and Company, 1995.