UNIVERSITY OF CALIFORNIA, SAN DIEGO

Targeting Translation in Human Thymidylate Synthase and Hepatitis C

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Chemistry

by

Nicholas Brunn

Committee in charge:

Professor Thomas Hermann, Chair
Professor Rommie Amaro
Professor John Johnson
Professor Marianne Manchester
Professor Susan Taylor

2013

The Dissertation of Nicholas Brunn is approved, and it is acceptable

in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2013

TABLE OF CONTENTS

Signature Page……iii

Table of Contents……iv

List of Abbreviations……vi

List of Figures……viii

List of Tables……xii

Acknowledgements……xiii

Vita and Publications……xiv

Abstract of the Dissertation……xv

Chapter 1. Ribosomes, Translation, and Regulation: Introduction……1

Inhibiting Translation……2

Assays……3

In Vitro Translation Assay……4

Thiouracil Crosslinking Assay……5

2AP Assay……6

Gel Shift Assay……7

Chapter 2.Targeting Thymidylate Synthase: Introduction……9

Target Validation……11

Stabilizing Mutations as Surrogates for Ligands……18

Compound Screening……20

Peptide Screening……25

Conclusions……30

Acknowledgements……31

Chapter 3. Analysis of Thymidylate Synthase Protein and RNA Interactions: Introduction……32

TS Site 1 Truncation Mutant Analysis……32

TS Protein-RNA Crosslinking and PAGE Analysis……35

Mass Spectrometry Analysis.....38
Structural Examination of Protein-RNA Binding Regions…….....40
RNA Binding Activity of Peptides.....47
Applications of 2AP Screen.....49
Conclusions.....53
Acknowledgements…….56
Chapter 4. Targeting Translation in the HCV IRES: Introduction.....57
Compound Testing.....59
Structure Activity Relationships……..63
Insights into IRES driven translation……..……....64
Conclusions.....67
Chapter 5. Crystallography: Introduction…….…....68
Protein-RNA Complex.....68
RNA Crystallography.....73
Conclusions…….……....75
Chapter 6. Materials, Methods, and Supplemental Figures.....77
References.....93

LIST OF ABBREVIATIONS

HCV: Hepatitis C Virus

hTS: Human thymidylate synthase

IRES: Internal Ribosome Entry Site

NTR: Non-translated Region

UTR: Untranslated Region

RNA: Ribonucleic acid

DNA: Deoxyribonucleic acid

mg: milligram

tU: thiouracil

elF: Elongation Factor

DTT: Dithiothreitol

EDTA Ethylenediamenetetracetic acid

mL: milliliter

2AP: 2-aminopurine

Nt: nucleotide(s)

UV: Ultraviolet

PAGE Polyacrylamide Gel Electrophoresis

WT: Wild-type

dTMP: 2'-deoxythymidine-5'-monophosphate

dUMP: 2'-deoxyuridine-5'-monophosphate

mTHF: N^5, N^{10}-methylenetetrahydrofolate

THF: Tetrahydrofolic acid

5-FU: 5-fluorouracil

dTTP: 2'-deoxythymidine-5'-triphosphate

CHAPS: 3-[(3-Cholamidopropyl)-dimethylammonio] 1-propane sulfonate

SDS: Sodium dodecyl sulfate

LIST OF FIGURES

Figure 1.1 Flowchart of the IVT assay……...……4

Figure 1.2 Predicted secondary structure of thiouracil labeled (red) RNAs……...6

Figure 1.3 Structures of the 2-aminopurine RNA……..7

Figure 1.4 Diagram of gel shift assay……...…..8

Figure 2.1 Thymidylate synthase pathway……..…9

Figure 2.2 Human TS RNA with detailed secondary structure of the site 1 RNA……..10

Figure 2.3 TS Site 1 mutants (Mt1-Mt4) ……11

Figure 2.4 Bicistronic reporter constructs used for IVT assay……..……..13

Figure 2.5 Monocistronic reporter constructs…….……14

Figure 2.6 Dose response of huTS protein in IVT assay ……..……..16

Figure 2.7 Relative expression levels of IRES driven renilla luciferase in TS1 & ΔTS1 constructs…….....16

Figure 2.8 Relative firefly translation of the bicistronic reporter constructs and mutants…..17

Figure 2.9 IRES driven renilla luciferase translation of bicistronic constructs…….....17

Figure 2.10 Firefly luciferase expression of bicistronic constructs conducted in TS immunodepleted whole cell extract……..…18

Figure 2.11 Correlation of free energy stabilization (kcal/mol) introduced by mutations into TS site 1 structure in full cell extract……...……19

Figure 2.12 Correlation of free energy stabilization (kcal/mol) introduced by mutations into TS site 1 structure in immunodepleted cell extract……...……20

Figure 2.13 Subset of exploratory screen of RNA friendly molecules in IVT assay……...21

Figure 2.14 Dose response of ES-291 and ES-ME in TS1 bicistronic construct……..…22

Figure 2.15 Dose responses of ES-ME and ES-291 in full reticulocyte lysate in ΔTS1 construct……..……22

Figure 2.16 Structures of ES-291, ES-ME, and tetrahydrofolic acid (THF)……...……..23

Figure 2.17 THF dose response in TS1 construct, in both full and depleted (dep) extracts….24

Figure 2.18 Inhibition of TS enzymatic activity by ES-ME, ES-291, and 5-FU……..25

Figure 2.19 Native PAGE analysis of hexapeptides binding to 33mer TS site 1 motif…..….26

Figure 2.20 Dose responses of peptides (B) & (C) with 33mer TS1 RNA in native PAG….27

Figure 2.21 Presence of hexapeptides B, C, and D and their effect on translation as determined by reporter expression in bicistronic constructs……...27

Figure 2.22 IVT assay in bicistronic TS1 and ΔTS1 constructs of additional hexapeptides, and pentapeptides inspired by the lead peptide ……..….29

Figure 2.23 Titration of the RLRRLR peptide in the IVT assay……..……29

Figure 3.1 Names and predicted secondary structures for the TS1 truncation mutants……33

Figure 3.2 IVT assay reporter expression for the TS site 1 truncation mutants……34

Figure 3.3 Predicted secondary structure of thiouracil labeled RNAs……..…35

Figure 3.4 SDS-PAGE denaturing and native gel of TS protein protein-RNA crosslink complexes with either a 33mer or a 16mer RNA……..…37

Figure 3.5 Native PAGE kinetic analysis of TS protein and R33 crosslinking……38

Figure 3.6 Color mapping of the mass spectrometry sequence analysis ……..39

Figure 3.7 Elution profiles……..…40

Figure 3.8 Structure of the human thymidylate synthase solved at low salt (active) conditions.....42

Figure 3.9 Peptide mapping from figure 3.6 overlaid onto thymidylate synthase structure….43

Figure 3.10 Thymidylate synthase structure with electrostatic potential surface mapping.....44

Figure 3.11 20 nucleotide RNA overlay onto TS sequence……..…..……..45

Figure 3.12 TS crosslinking activity map with a 14mer terminal loop RNA……..46

Figure 3.13 2AP assay examining various TS1 RNA ligands…….……48

Figure 3.14 Gel shift of peptides TS82-89 and TS84-92……..49

Figure 3.15 2AP assay of polymyxin B, E, and B-nonapeptide (BNP)……50

Figure 3.16 IVT assay dose response of polymyxin B and E ……..……51

Figure 3.17 IVT assay of two TS peptides (82-89 and 84-92) implicated in the binding interactions of the TS1 RNA…….....51

Figure 3.18 Proposed model for the hTS protein binding to the TS site 1 RNA…….52

Figure 4.1 Secondary structure of the HCV IRES, and the crystal structure of subdomain IIa.58

Figure 4.2 HCV IRES Subdomain IIa crystallized in either the absence or presence of a benzimidazole translation inhibitor…….59

Figure 4.3 Dose response IVT assay on benzimidazole compound……60

Figure 4.4 Compound screen……...60

Figure 4.5 Result of IVT assay and scaffolds……..61

Figure 4.6 Dose response curves for compounds (2) and (3) in in the IVT assay…….61

Figure 4.7 Dose response curves of compounds (189), (191), and (192……..….62

Figure 4.8 Structure activity relationship analysis of a small subset of benzoxazoles……..…63

Figure 4.9 Dose response of benzimidazole in bicistronic reporter constructs…….65

Figure 4.10 Bicistronic construct ΔdII ……66

Figure 4.11 Normalized translation levels of firefly luciferase in WT and ΔdII…..……..….66

Figure 5.1 TS protein crystallized in the presence of 4 fold excess 33mer TS1 RNA in high salt and low salt conditions……..69

Figure 5.2 TS protein crystalized in the absence of additives or RNA in high salt and low Salt conditions……70

Figure 5.3 Unit cells of high salt (A) and low salt (B) crystal structure……71

Figure 5.4 Overlay of the TS protein in high salt conditions and low salt conditions……..…73

Figure 5.5 T7 RNA transcription scheme……74

Figure 5.6 Secondary structure of the TS site 1 RNA crystallization scheme……..……75

Figure 6.1 SDS-PAGE gel of Ni2+ affinity purification……..78

Figure 6.2 Time course of the bicistronic reporter constructs in the IVT assay…….....99

LIST OF TABLES

Table 3.1 Mass spectrometry analysis of trypsin digestion……55

Table 6.1 Crystallographic data and refinement statistics……89

Table 6.2 Compound name, compound number and luciferase expression……90

Table 6.3 Oligo sequences for TS site 1 mutants……92

Table 6.4 Oligo sequences for cloning HCV IRES into IVT reporter……93

Table 6.5 IVT assay results of the TS mRNA truncation mutants……94

Table 6.6 IVT assay results of the TS protein titration……94

Table 6.7 IVT assay results of the bicistronic reporter containing the stabilizing mutation....95

Table 6.8 IVT assay results of the stabilizing mutations within the TS depleted extract……95

Table 6.9 2AP assay: peptides and the fluorescence dose response……96

Table 6.10 2AP assay: peptides……97

Table 6.11 IVT assay results of the lead……98

ACKNOWLEDGEMENTS

I would like to acknowledge several individuals who played integral roles in shaping my college and graduate experience. Prof. Galen George of Santa Rosa Junior College, Dr. Tom Huxford of San Diego State University, Dr. Anette Schneemann of the Scripps Research Institute, and Dr. Thomas Hermann of UCSD all are responsible for making me the individual I am today. I will be deeply indebted to them, for all that they have done. I would like to thank the many current and former members of Hermann lab, including Kevin Rynearson, Kejia Ding, Maia Carnevali, Jerod Parsons and Emily Garcia-Sega. Additionally, I would like to acknowledge the lab of Dr. Pat Jennings and Dr. Ulrich Mueller, here at UCSD, for providing the use of various sonication equipment and centrifuges.

Portions of the text from Chapter 2 and Figures 2.1-2.21 were adapted from a published article, Brunn et al., (2012) Targeting a Regulatory Element in Human Thymidylate Synthase mRNA. *ChemBioChem*. The dissertation author was the primary author, Emily Garcia Sega and Melody B. Kao were co-authors and Thomas Hermann was the corresponding author of this article.

Portions of the text from Chapter 3 and Figures 3.1-3.18 have been submitted for publication entitled, “Molecular recognition of mRNA by human thymidylate synthase” The dissertation author is the primary author. Co-authors include Sergey Dibrov, Melody B. Kao, and Thomas Hermann was the corresponding author of this article. The UCSD Bimolecular/Proteomics Mass Spectrometry Facility and Majid Ghassemian was responsible for the mass spectrometry analysis.

VITA AND PUBLICATIONS

EDUCATION

University of California, San Diego, CA **2009-2013**
PhD. Chemistry
Thesis Dissertation: "Targeting Translation in Human Thymidylate Synthase and Hepatitis C"

University of California, San Diego, CA **2009-2011**
M.S. Chemistry

San Diego State University, San Diego, CA **2005-2007**
B.S. Chemistry

RESEARCH EXPERIENCE

University of California, San Diego, CA **2009-2013**

The Scripps Research Institute, La Jolla, CA **2007-2009**

San Diego State University, San Diego, CA **2005-2007**

Analytical Sciences, Petaluma, CA **2004-2005**

Publications:

"Hepatitis C virus translation inhibitors targeting the internal ribosome entry site" (2013) Dibrov SM, Parsons J, Carnevali M, Zhou S, Rynearson KD, Ding K, Garcia Sega E, Brunn ND, Boerneke MA, Castaldi MP & Hermann T. *J. Med. Chem.*, in press.

"Screening for inhibitors of the hepatitis C virus internal ribosome entry site RNA" Zhou S, Rynearson K, Brunn ND, Hermann T June, 2013

"Targeting a Regulatory Element in Human Thymidylate Synthase mRNA" Brunn ND, Garcia Sega E, Kao MB, Hermann T. *Chembiochem,* November, 2012

"Structure of a hepatitis C virus RNA domain in complex with a translation inhibitor reveals a binding mode reminiscent of riboswitches" Dibrov SM, Ding K, Brunn ND, Parker MA, Bergdahl BM, Wyles DL, Hermann T. *Proc. Natl. Acad. Sci. USA,* Apr, 2012

"The 2.3-angstrom structure of porcine circovirus 2" Khayat R, Brunn N, Speir JA, Hardham JM, Ankenbauer RG, Schneemann A, Johnson JE. *J. Virol.* Aug, 2011

ABSTRACT OF THE DISSERTATION

Targeting Translation in Human Thymidylate Synthase and Hepatitis C

by

Nicholas Brunn

Doctor of Philosophy in Chemistry

University of California, San Diego, 2013

Professor Thomas Hermann, Chair

The goal of this study was to investigate selective translational inhibition of the human thymidylate synthase (TS), and the hepatitis C virus (HCV) internal ribosome entry site (IRES). This was done through the targeting of critical RNA structures within the TS messenger RNA (mRNA) and the HCV IRES, by small molecules or peptides. The TS protein-TS mRNA interaction was also explored through the use of chemical crosslinking, and mass spectrometry. The validity of targeting the TS mRNA was examined through the

exploration of various TS mRNA stabilized mutants in the context of an *in vitro* transcription translation assay. Additionally, select compounds and peptides were tested in biological assays for their ability to ablate TS mRNA regulated translation and HCV IRES driven translation. The mechanism of HCV IRES driven translation and inhibition was also explored through the use of *in-vitro* transcription-translation assays and IRES mutants.

Compounds were chosen for testing based on their potential to form multiple hydrogen bonds within an RNA fold, and thereby biased to target structured RNA. Peptides were chosen based on their similarity to known RNA binding peptides, or based on the primary sequence of the TS protein. Compounds that target the TS mRNA were further examined using a 2-aminopurine fluorescence assay and gel-shift assays.

To further investigate the respective targets' structures and interactions, macromolecular x-ray crystallography was employed. The TS mRNA is likely dynamic in solution, and less amenable to crystallography than the HCV IRES. To overcome this obstacle, various TS RNA sequences were synthesized, and set in high-throughput crystal screens. The TS protein was also crystallized in the presence of its RNA under two unique conditions in an attempt to capture the TS-mRNA bound state.

Chapter 1. Translation, Regulation, and Inhibition

Introduction

Translation is the process whereby ribosomes process a messenger RNA (mRNA) into the respective amino acids, and ultimately the protein that it encodes. Selectively targeting this protein biosynthesis pathway for therapeutic purposes can drastically affect the relative amounts of key proteins being expressed within a cell. Therefore, this is very amenable to targeting cells or viruses growing out of control and arresting their development. The approach of targeting translation, rather than targeting the protein, could potentially alleviate nonspecific drug-protein interactions observed in many therapeutic agents today. In order to effectively and efficiently target the translation of specific proteins, understanding their mode of regulation within a cell is critical. Many proteins have multiple levels of regulation, and by exploring these mechanisms on a structural and biochemical level, it may possible to design ways of selectively inhibiting their translation.

Ribosomes, Translation, and Regulation

Traditional cap-dependent eukaryotic translation involves 40S ribosomal subunits recognizing the 7-methylguanosine (m7G) capped 5′ end of an mRNA, scanning to find the initiation codon, and then joining with a 60S subunit into translational competent 80S ribosomes (40). At least nine initiation factors play a role in the assembly of the 43S preinitiation complex. Once assembled, this complex can recognize and attach to the capped 5' region of the target mRNA. Subsequently, it scans downstream until it locates the first AUG start codon. After recognizing the initiation codon, the complex attaches to the 60S ribosomal subunit to form translation competent 80S complexes (40, 41, 42, 43). The 43S

ribosomal complexes move along the mRNA relatively easily to scan unstructured RNA without the need for additional cofactors. However if an RNA contains a 5' UTR with even weak secondary structures, the 43S complex requires a number of additional factors, including ATP and elongation factors (elF) 4A, 4B, and 4G (43). In addition, these factor requirements are proportional to the degree of secondary structure in the 5' UTR (44,45).

The cap-independent mechanism of ribosomal recruitment, also called internal translation initiation, utilizes structured 5' RNA sequences called internal ribosome entry sites (IRES), typically in the 5' UTR, to position its mRNA on the ribosomal subunits, and recruit the necessary protein synthesis factors (46, 47, 48). This translational mechanism is most often employed by viruses, although there are some examples of cellular mRNAs utilizing IRES like translation (49).

Traditionally, RNA translational regulation falls into one of three following categories; sequence specific RNA binding proteins, microRNAs, or the availability of phosphorylated elF2 and elF4F (50, 51, 52). The regulation of translation by protein-RNA binding is rare (54). The two better studied examples of this type of interaction are the ferritin mRNA, and the thymidylate synthase mRNA. These RNA binding proteins that lead to translational regulation is, do so in a concentration dependent fashion. The higher the concentration of active protein, the greater the translational inhibition. Additionally, the protein-RNA interaction occurs near the 5' region of the RNA, and prevents the loading of the 43S complex onto the mRNA (53).

Inhibiting Translation

Selectively inhibiting the translation of a specific protein is quite difficult, as the major targets such as ribosomes or elongation factors are essentially identical in architecture throughout the body. Current translational inhibition drugs, such as antibiotics, are extremely effective at shutting down protein synthesis in their prokaryotic pathways. However shutting down the translation of a specific eukaryotic protein is much more difficult. Translation inhibitors are generally not selective in their binding, compounds like cycloheximide, or certain bacterial toxins, will inhibit all protein synthesis processes in all eukaryotic cells, making them quite toxic. Antisense RNAs and siRNA can be quite specific in their effect, but unless the phosphate backbone of the RNA is chemically modified to be resistant to RNase activity, their half-life within a cell is very short, and methods of cellular delivery are severely limited as well (55). Therefore it is critical to discover ways to target the translation of specific proteins, while allowing native protein synthesis to continue minimally affected, and this can often be done by exploiting unique regulatory pathways, and RNA structures.

Assays

Several assays were used in this study to quantify ligand-RNA interactions and their overall effects on translation. These assays can measure several factors, including conformational changes of RNA induced by the ligand, ability to bind to the RNA, and the *in vitro* activity of a ligand in the presence of whole-cell lysate. The assays that will be discussed in this work are the *in vitro* transcription-translation assay (IVT), gel retardation or gel shift assay, thiouracil crosslinking assay, and 2-aminopurine (2AP) assay.

In Vitro Translation Assay

In vitro transcription-translation (IVT) systems can be harnessed to potentially examine DNA, RNA, drug, and protein interactions, as well as their effects on translation. The IVT system for screening drug-RNA interactions relies on having the gene of a reporter protein (typically a firefly or renilla luciferase) cloned downstream of a phage polymerase promoter, but after the RNA target of interest. The DNA can be either linearized or circular, so long as the concentration is kept constant. The DNA is then mixed with eukaryotic cell lysate, for example rabbit reticulocyte lysate, SP6 RNA polymerase, and a mix of salts and amino acids (56). After incubation at 30 ºC (typically for 45 minutes) the reaction is analyzed for the respective levels of reporter proteins. If analyzing a bicistronic construct, this involves the initial addition of firefly luciferase reagent, measurement of luminescence, followed by the subsequent quenching of the firefly signal, and simultaneous addition of renilla luciferase reagent (Figure 1.1). This can be done with either Dual Glo or Stop-N-Glo (Promega) assay reagents.

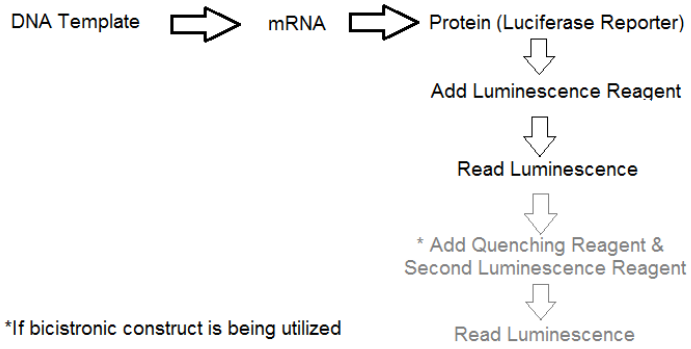


Figure 1.1 Flowchart of the IVT assay.

When the IVT assay is run in the presence of a compound targeting the RNA structure of interest, positive “hits” can be identified by comparing the respective levels of luminescence in the drug free control, with the experimental results. All luminescence values

are normalized with respect to the negative control. Any "hits" are determined by successful translational inhibition by a factor greater than three standard deviations (3σ). If this criterion is met, the compound is then tested in the monocistronic ΔTS1 construct (Figure 2.5), as well as on already synthesized luciferase enzyme. This eliminates the possibility that a compound indiscriminately binds to the ribosomes, thereby reducing all translation, or that it acts on the luciferase enzyme.

Thiouracil Crosslinking Assay

Thiouracil labeled RNA is a powerful tool in quantifying and analyzing RNA synthesis, decay rates, as well potentially purifying away from non-labeled RNA, or for mapping RNA-protein interactions through chemical crosslinking. The presence of the thio-modified uracil has been shown to have little effect on the structure of RNA species, or on cellular physiology. (5). To determine whether the TS protein binds to the RNA as a monomer or dimer, two RNAs containing either the majority of the TS site 1 sequence (R33), or simply the terminal loop and transcription initiation site (R14) were commercially synthesized. In order to detect any transient binding activity, a thiouracil modification was inserted in the transcription initiation site (Figure 1.2). These RNAs were incubated with the protein for set durations, and then separated and visualized on both denaturing, and native gels to determine stoichiometry, relative abundance, and separation of crosslinking products for mass spectrometry analysis.

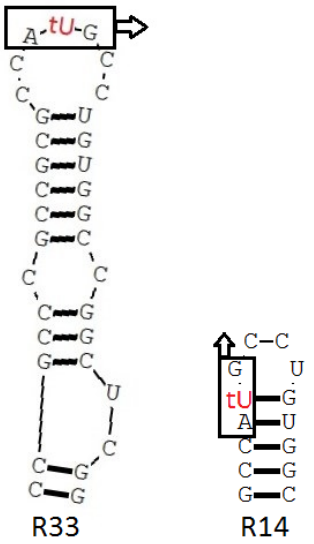


Figure 1.2 Predicted secondary structure of thiouracil labeled (red) RNAs. Box indicates translation initiation site.

2AP Assay

Ultraviolet (UV) absorption and emission, combined with fluorescent labeling allow investigation of the interactions between a RNA and a ligand of interest (6,8). In the 2-aminopurine assay (2AP) a RNA of interest is synthesized with a 2AP substituted adenosine base in the targeted region. The fluorescence of 2AP has been shown to be affected by its chemical environment. Depending on whether the 2AP base is stacked (fluorescence is quenched), or un-stacked (fluorescence is increased), one can monitor the environment of the RNA-ligand interactions. Subsequently, one can monitor the fluorescence excitation and emission spectrums of the 2AP labeled RNA in question, and ultimately characterize the RNA in the absence and presence of a ligand (8). The presence of 2AP does not hinder the formation of Watson-Crick type base pairs with dT/dU, and has been shown to result in minimal modification of the nucleic acid tertiary structure (9).

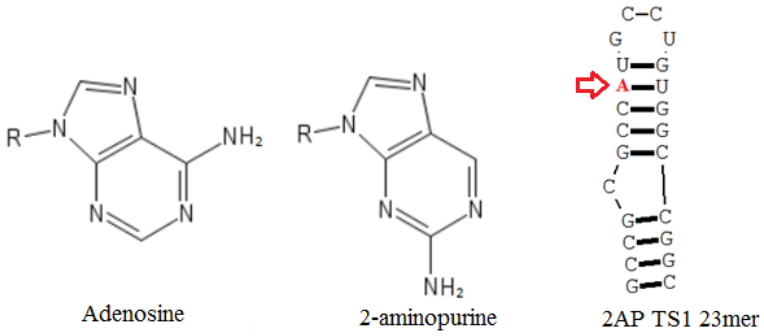


Figure 1.3 Structures of adenosine, the 2-aminopurine analog, and the 2AP (red) labeled TS1 23mer RNA used in the assay.

In this study, the adenine at position 94 in the terminal loop of the thymidylate synthase site 1 hairpin (TS1) was replaced by 2AP (Figure 1.6) to probe the binding of small molecules, peptides, and proteins. When the RNA is in a ligand-free environment, it has a minimal change in its fluorescence excitation and emission profile. The addition of ligands that bind to the TS1 RNA lead to large changes in relative fluorescence output, which can be quantified to determine binding affinities (6).

Gel Shift Assay

The gel shift, or gel retardation assay utilizes native polyacrylamide gels as a means to investigate the interaction of biomacromolecules under near physiological conditions. The gels are typically made of a polyacrylamide concentration of 7%, contain 2mM $MgCl_2$, and are run at a physiological pH of 7.5. The assay allows us to observe the changes in migration of our target RNA in the presence of our ligand(s) of interest (57). After incubation of the RNA with our ligand of interest, they are loaded on the gel, and an electric field is applied. Due to its negative charge, the RNA will migrate towards the anode at the bottom of the gel. If there is a significant “shift” or slowing of the rate of migration of the RNA, it is typically

due to the ligand binding to, and increasing the charge-to-mass ratio and hydrodynamic radius of our target RNA.

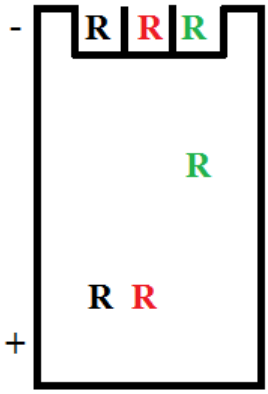


Figure 1.4 Diagram of gel shift assay. In the presence of an electric field, the negatively charged RNA will migrate towards the anode, and the free RNA (black) and RNA in the presence of a non-binding compound (red) will migrate unabated, whereas active bound ligand (green) will slow the rate of migration of the RNA.

Chapter 2. Targeting Thymidylate Synthase Translation

Introduction

The human thymidylate synthase (hTS) protein is a uniquely situated target, as it provides the cells sole *de novo* source of 2'-deoxythymidine-5'-monophosphate (dTMP), an essential component for DNA biosynthesis, via reductive methylation of 2'-deoxyuridine-5'-monophosphate (dUMP) by the reduced folate N^5, N^{10}-methylenetetrahydrofolate (mTHF) Figure 2.1 (30). After dTMP is synthesized, it is subsequently phosphorylated to the DNA synthesis component dTTP.

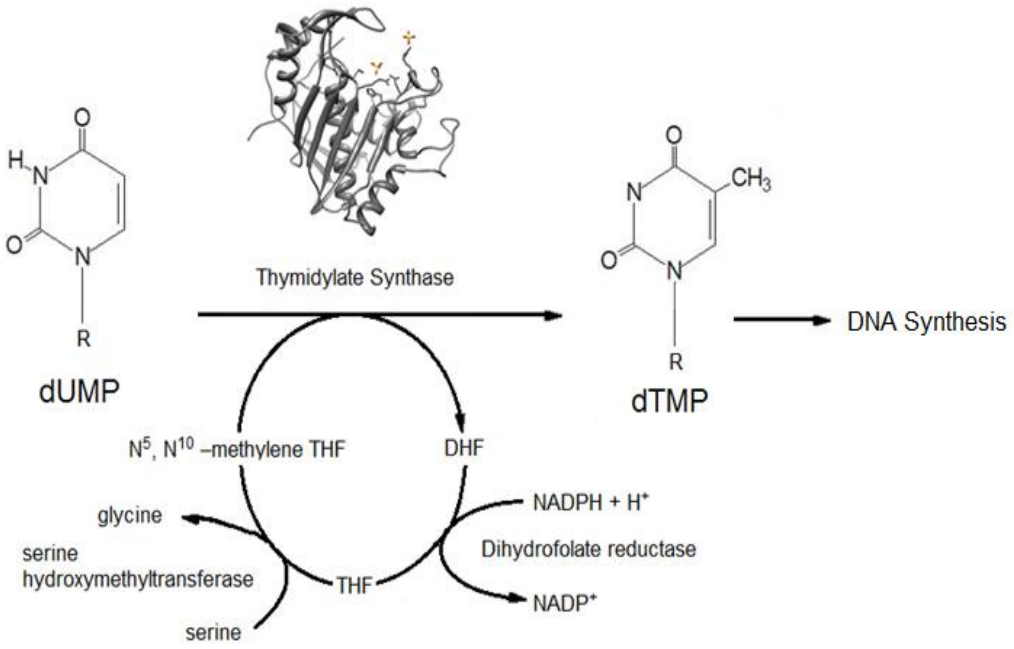


Figure 2.1 Thymidylate synthase pathway, and thymidylate biosynthesis pathways.

Since it is the sole source of dTMP within a cell, TS is an attractive target for cancer chemotherapy, and TS inhibitors are some of the oldest anti-cancer agents available (31, 32). Unfortunately the use of TS inhibitors, such as 5-fluorouracil (5-FU), leads to an eventual

resistance of the drug within the targeted cell, generally through upregulation of TS expression (31, 33). It has been demonstrated that 5-FU induced TS overexpression is brought about by interfering with the autoregulatory mechanism of TS translational control. In its apo-form TS binds to its cognate mRNA at two locations, one of which is located in the 5'-UTR and is predicted to form a stem-loop structure that sequesters the translation initiation site (Figure 2.2) (34, 35), thereby stabilizing the hairpin loop, blocking ribosomal

Figure 2.2 Human TS mRNA with detailed secondary structure of the site 1 RNA.

complex formation (34, 35). Substrate binding to TS abolishes this inhibition, and leads to an increase in TS synthesis (36), thereby negating much of the therapeutic benefit of the TS inhibitor. Compounding these problems, the degradation pathway of the TS protein is

ubiquitin independent, thereby resistant to degradation control. To overcome this drug resistance mechanism, a ligand could be envisaged to target and stabilize the 5'-UTR, thereby mimicking the binding effect of the TS protein, and blocking translation. In this chapter, we explore the validity of this approach through the use of *in vitro* transcription-translation assays, gel shift assays, and enzymatic activity assays.

Target Validation

To validate the potential stabilizing effect that a small molecule ligand could have on the TS RNA hairpin, a series of mutants of the site 1 RNA were created whereby the hairpin was systematically stabilized by closing budges and mismatched bases, thereby acting as a surrogate for ligand-stabilization effects (Figure 2.3).

Figure 2.3 TS Site 1 mutants (Mt1-Mt4) arrows and letters indicate which nucleotide was substituted.

Both the wild type TS site 1 regulatory element and the mutants were cloned upstream of a reporter gene to be used in a coupled *in vitro* transcription-translation (IVT) assay. The TS site 1 hairpin and mutants were constructed in such a way that the TS hairpin sequestered the start codon of a firefly luciferase reporter, under cap-dependent translation, followed by a *Renilla* luciferase reporter, driven by a HCV IRES to provide an internal translation control (Figure 2.4). Additionally, a negative control (ΔTS1) was created, where the 5' end is simply a cap-driven firefly reporter, replacing the TS site 1 hairpin. Utilizing a bicistronic reporter system allows the exclusion of compounds that affect the ribosome indiscriminately, as well as obviates the requirement of normalizing absolute RNA concentrations in the assay. Additionally, in the case that a promiscuous ligand binds to both the TS site 1 hairpin, as well at the HCV IRES, two monocistronic constructs were created to test the relative expression levels of cap driven TS1 site regulated RNA, compared to a cap driven ΔTS1 RNA (Figure 2.5)

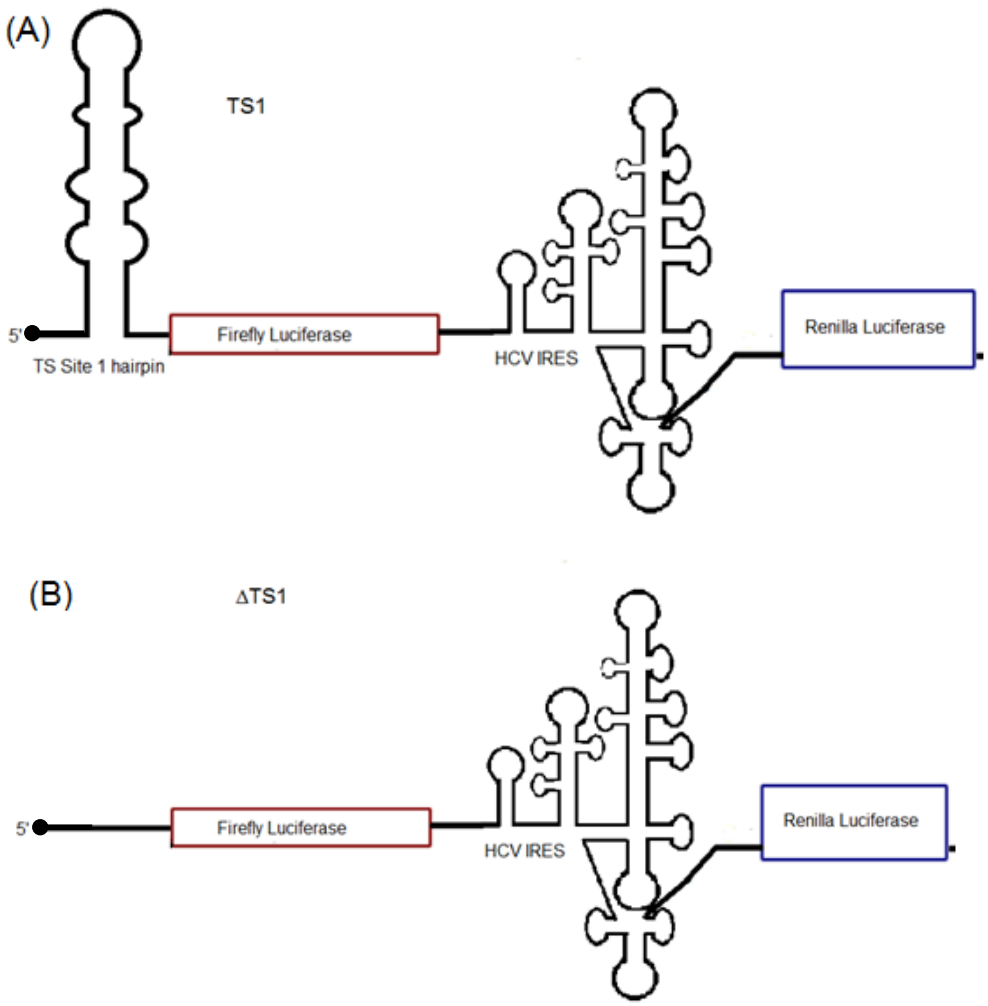


Figure 2.4 Bicistronic reporter constructs used for IVT assay, either containing the TS regulatory motif (A) or a basic cap driven firefly luciferase at the 5' end (B), with both containing an internal HCV IRES driven renilla luciferase.

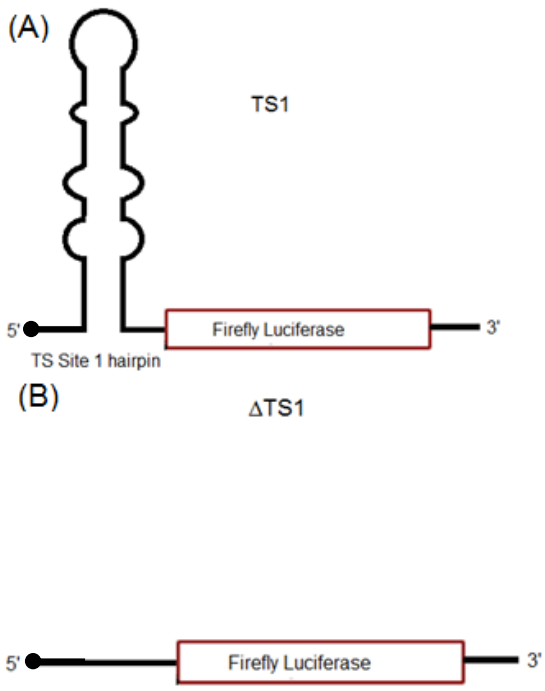


Figure 2.5 Monocistronic reporter constructs TS1 (A), and ΔTS1 (B) used for IVT assays.

The capacity of the TS site 1 to function as a regulatory motif was tested in the context of the bicistronic reporter by the titration of human TS protein into the IVT assay (Figure 2.6). As was shown, the addition of the enzyme suppressed relative translation of the firefly luciferase in a dose dependent fashion, completely suppressing translation at 25μM, but only in the construct containing the TS site 1. The IRES driven renilla luciferase expression was not affected in the presence of the huTS enzyme (Figure 2.7) in either the TS1 or ΔTS1 bicistronic constructs. This result demonstrates that we can reproduce the effect of native TS acting on its own native mRNA (37, 38). When comparing the relative translational efficiencies of the TS1 with the ΔTS1, we observe approximately a 40% decrease (Figure 2.8) in the TS1 construct, but only under the cap driven TS hairpin

regulated translation, while IRES driven translation remained unaffected (Figure 2.9). While structured 5′UTRs provide an energetic barrier for ribosomal read-through (43-45), their stabilizing effect is not sufficient to explain this large of a translational decrease. The logical explanation for the discrepancy in translation would be the presence of native TS protein in the reticulocyte lysate used for the IVT assay. To verify this hypothesis, a sample of the reticulocyte lysate was immunodepleted of TS protein, and the assay was re-run using this depleted extract (Figure 2.10). The results demonstrate that the relative expression levels of the TS1 and ΔTS1 constructs are essentially equal in the absence of TS protein. This finding also confirms that the TS1 secondary structure by itself does not provide sufficient stabilization to prevent the ribosome from accessing the translation initiation site. Rather it requires the presence of the TS protein bound to the RNA to inhibit translation.

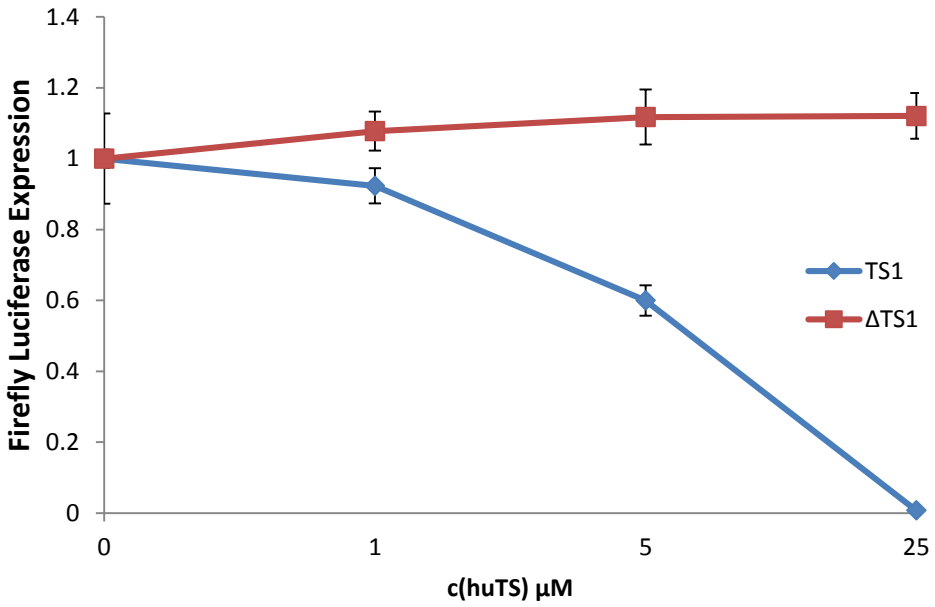


Figure 2.6 Dose response of huTS protein in IVT assay utilizing bicistronic reporter constructs containing the TS site 1 regulatory element (TS1) or the cap driven firefly control (ΔTS1). Results are averages of triplicate experiments. Error bars correlate to 1σ.

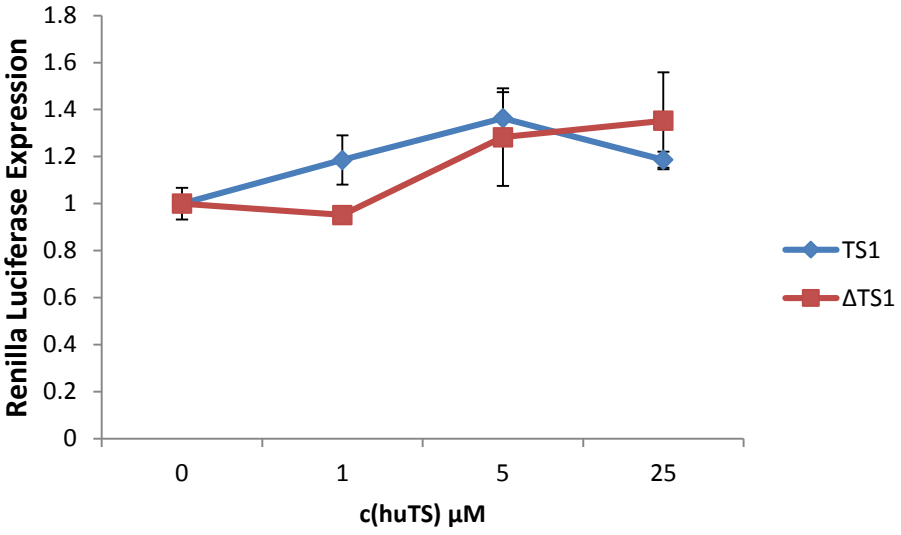


Figure 2.7 Relative expression levels of IRES driven renilla luciferase in TS1 and ΔTS1 constructs in the presence of increasing concentration of huTS protein.

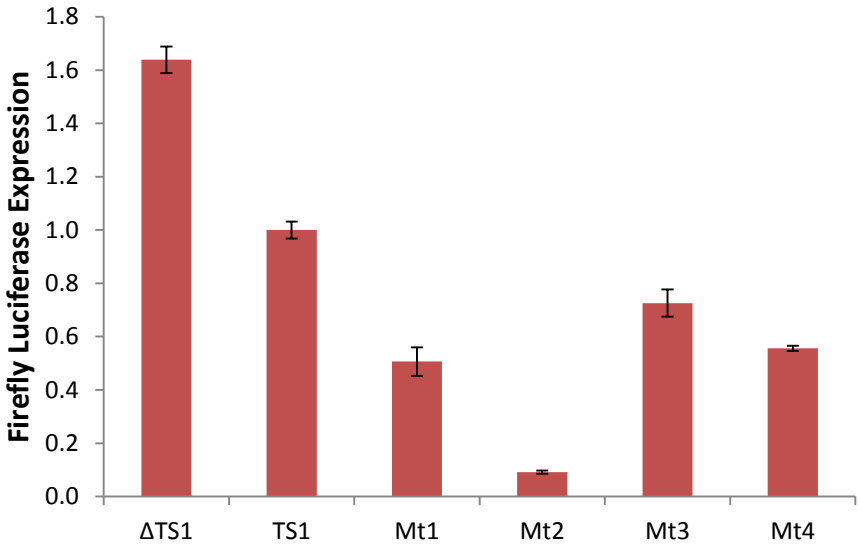


Figure 2.8 Relative firefly translation levels of the bicistronic reporter constructs and mutants. Values are normalized to TS1. Results are averages of triplicate experiments. Error bars correlate to 1σ.

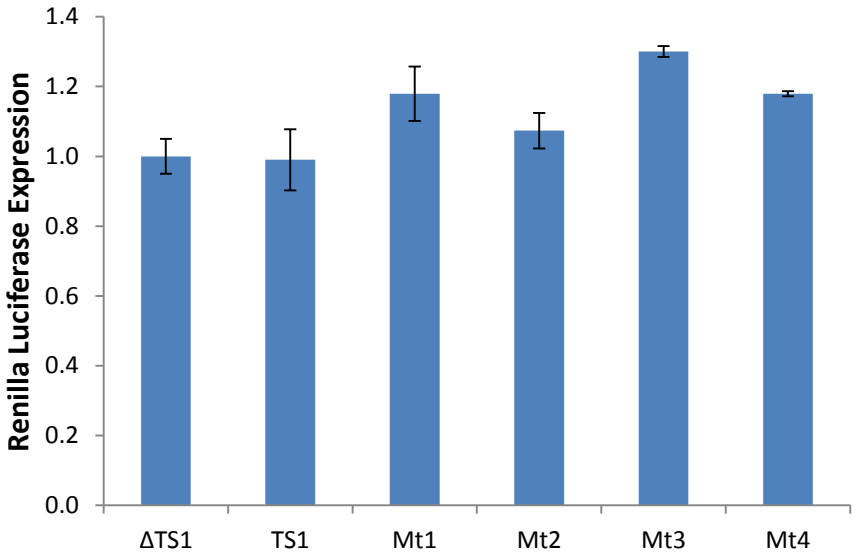


Figure 2.9 IRES driven renilla luciferase translation of bicistronic constructs. Values are normalized to TS1. Results are averages of triplicate experiments. Error bars correlate to 1σ.

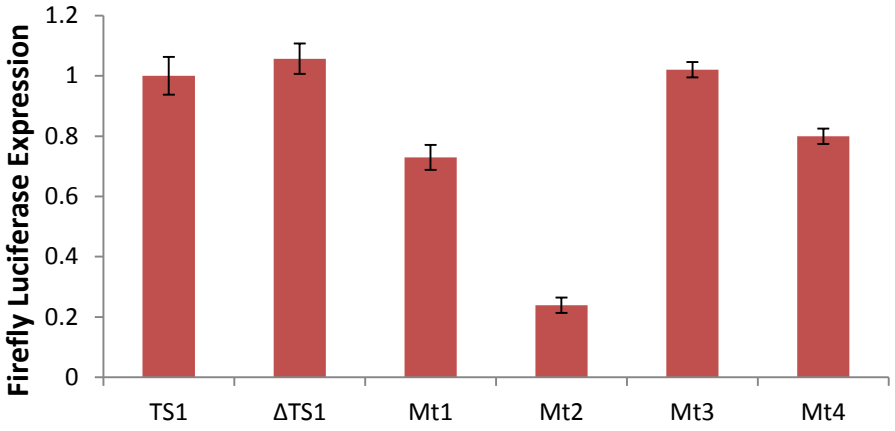


Figure 2.10 Firefly luciferase expression of bicistronic constructs conducted in TS immunodepleted whole cell extract. Values are normalized to TS1. Results are averages of triplicate experiments. Error bars correlate to 1σ.

Stabilizing Mutations as Surrogates for Ligands

The testing of the four mutant hairpin constructs (Figure 2.3) which introduce stabilizing base pairs into the TS site 1 hairpin, showed that there was a direct correlation between the number of Watson-Crick base pairs formed, and reporter repression (Figure 2.8). The reduction of firefly reporter translation was specifically due to the point mutations, as demonstrated by the relatively unaffected levels of IRES translation (Figure 2.9). When comparing the IVT assay conducted in the whole cell extract versus the immunodepleted extract, we notice that the reporter expression levels are partially increased, with the mutant Mt3 displaying an almost wild type expression level. This taken with our finding that the TS site 1 motif offers insufficient stabilizing forces without the presence of TS protein, is suggestive of an evolutionary balance where the wild type hairpin is labile enough to allow ribosomal access to the translation initiation site, while offering a stable secondary structure for the TS protein to recognize and bind to. The RNA folding stabilization energies resulting

from the point mutations was calculated for the TS site 1 motif. A linear correlation between decreased translation efficiency and RNA stabilization is evident in the full cell extract (Figure 2.11).

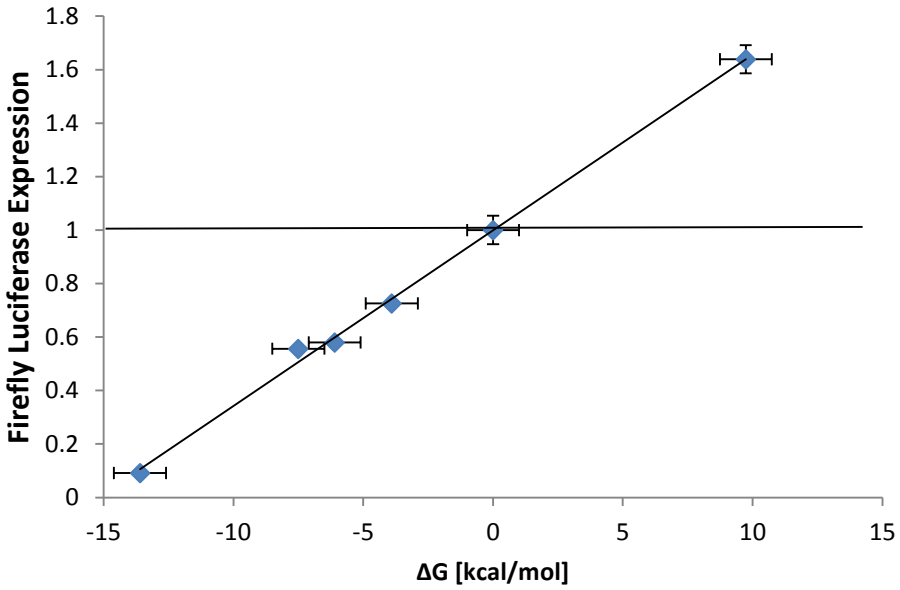


Figure 2.11 Correlation of free energy stabilization (kcal/mol) introduced by mutations into TS site 1 structure in full cell extract.

This correlation is likely due to the additional energy requirements of the ribosome to access the translation initiation site within the hairpin. To examine how much stabilization energy is required to prevent efficient ribosome translation, the same calculations were conducted in the standard extract (Figure 2.11) and the TS depleted extract (Figure 2.12). We can see that in Mt3, an additional 3-4 kcal/mol of free energy is obtained from mutating the C-C mismatch to a G-C base pair, yet this is not sufficient to significantly decrease translation in the absence of TS protein. By averaging the stabilization energy of Mt1 and Mt2, we estimate that a stabilizing energy of ~8kcal/mol will be sufficient to reduce translation by ~50%. This could be easily accomplished by a ligand forming 2 to 3 hydrogen bonds within the TS site 1 RNA, thereby potentially reducing any 5-FU associated TS

enzyme overexpression. This finding validates our approach to discover ligands that suppress TS expression through the stabilization of the TS site 1 RNA motif.

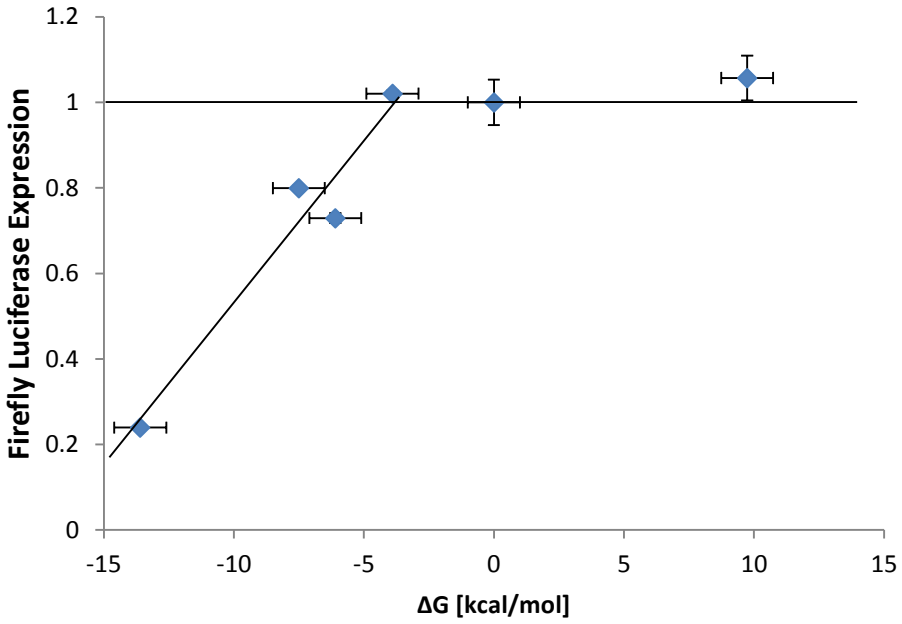


Figure 2.12 Correlation of free energy stabilization (kcal/mol) introduced by mutations into TS site 1 structure in immunodepleted cell extract.

Compound Screening

To selectively target protein translation under the control of the TS site 1 RNA motif, a small screen of exploratory compounds was tested in the IVT assay, utilizing the TS1 bicistronic construct. Compounds were synthesized and chosen based on structural features that provide a bias for RNA binding, and were tested at concentrations of 10μM and 100μM. The initial screen did not yield any compounds that suppressed TS site 1 regulated translation. However, several compounds were identified that increased the reporter translation to that of the ΔTS1 construct (Figure 2.13). Compounds ES-ME and ES-291 are

derivatives of diaminopropionic acid, and dihydropyrimidine respectively (Figure 2.16). Further examination revealed that these compounds selectively increase TS1 regulated translation (Figure 2.14), while not affecting reporter expression in the ΔTS1 construct (Figure 2.15)

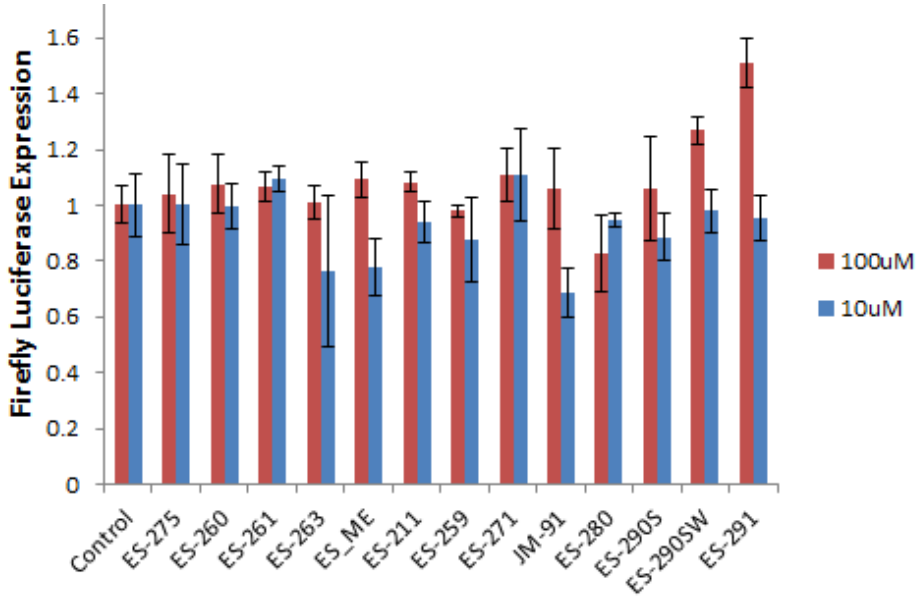


Figure 2.13 Subset of exploratory screen of RNA friendly molecules in IVT assay. Results are averages of triplicate experiments. Error bars correlate to 1σ.

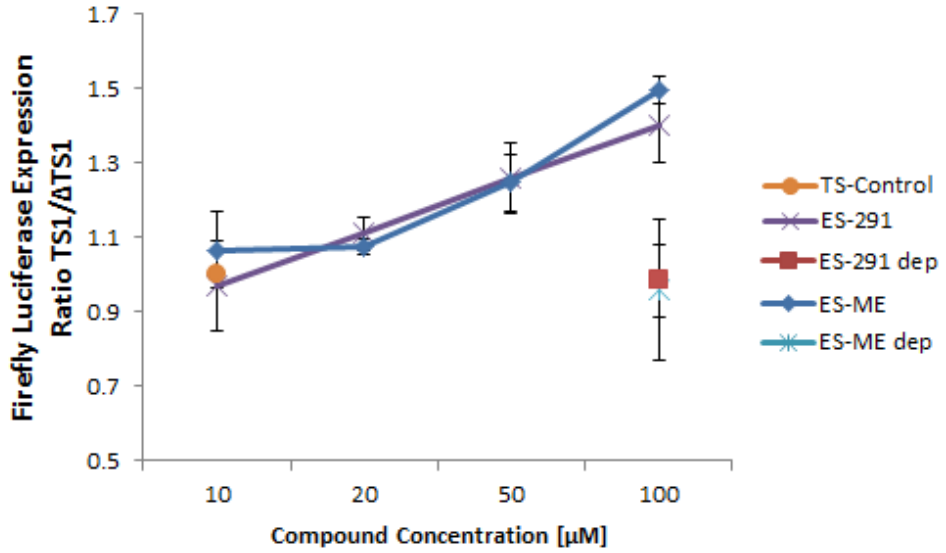


Figure 2.14 Dose response of ES-291 and ES-ME in TS1 bicistronic construct in either full reticulocyte lysate (ES-291 and ES-ME respectively) or depleted extract (dep). Results are averages of triplicate experiments. Error bars correlate to 1σ.

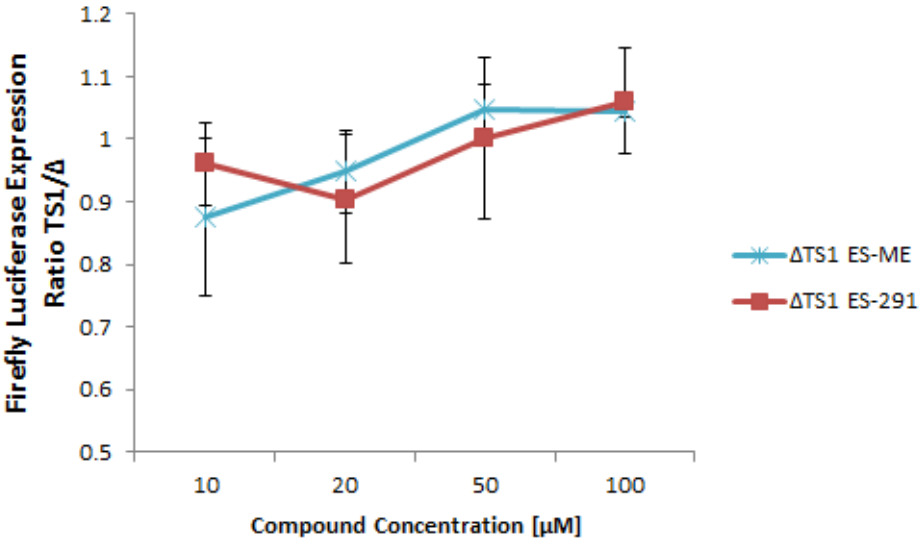


Figure 2.15 Dose responses of ES-ME and ES-291 in full reticulocyte lysate in ΔTS1 construct. Results are averages of triplicate experiments. Error bars correlate to 1σ.

The increase in translation observed in the full lysate, but not in the TS depleted lysate suggests that these compounds ablate the TS protein binding to the RNA, thereby removing its translational suppression. Since the phenomenon of TS protein being displaced from its RNA is observed with chemotherapy enzyme inhibitor agents, as well as the presence of the enzyme substrates, we hypothesize that these compounds are targeting the substrate binding site of the TS enzyme. To test this hypothesis we repeated the same dose response with tetrahydrofolic acid (THF) a structurally similar molecule to ES-ME, as well as a precursor to the TS enzyme co-substrate 5,10-methylenetetrahydrofolic acid (Figure 2.17). The results indicate that like compounds ES-ME and ES-291 (Figure 2.16), THF cause an increase in reporter expression as well.

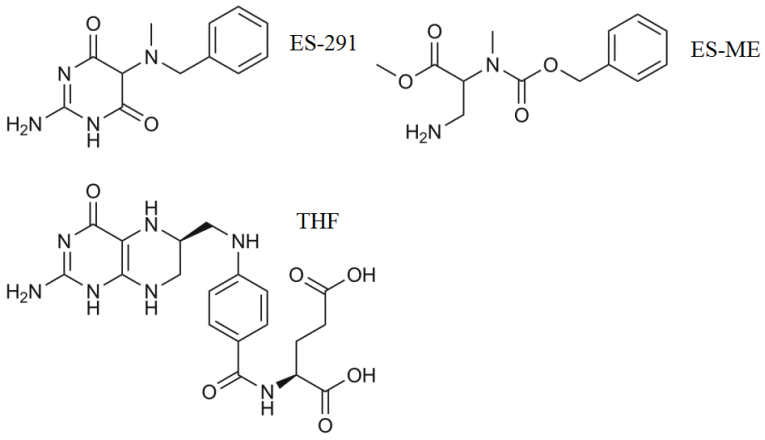


Figure 2.16 Structures of ES-291, ES-ME, and tetrahydrofolic acid (THF).

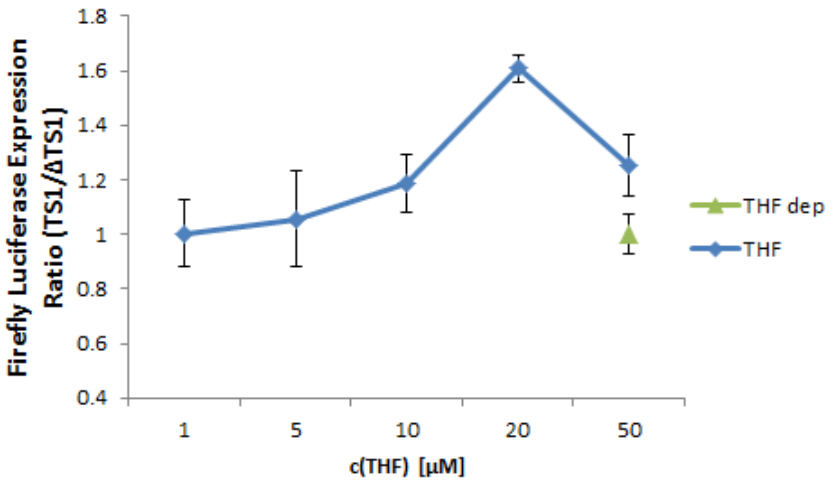


Figure 2.17 THF dose response in TS1 construct, in both full and depleted (dep) extracts. This increase in translation supports the theory that ligand binding to the TS enzyme substrate site ablates the translational repression previously observed for TS protein binding to the TS site 1 RNA motif.

To test whether these compounds were binding to the TS enzyme substrate site, rather than non-specifically interfering with the protein-RNA interaction, ES-ME and ES-291 were tested in their capacity to inhibit the TS enzymatic activity. The addition of ES-ME and ES-291 to a TS *in vitro* enzymatic activity assay demonstrated their ability to inhibit enzymatic activity in a dose dependent fashion (Figure 2.18). The demonstrated enzymatic inhibition in addition to the ablation of TS transcriptional repression suggests that ES-ME, ES-291, and other similarly acting compounds, are all targeting the TS protein, thereby disrupting the autoregulatory feedback pathway.

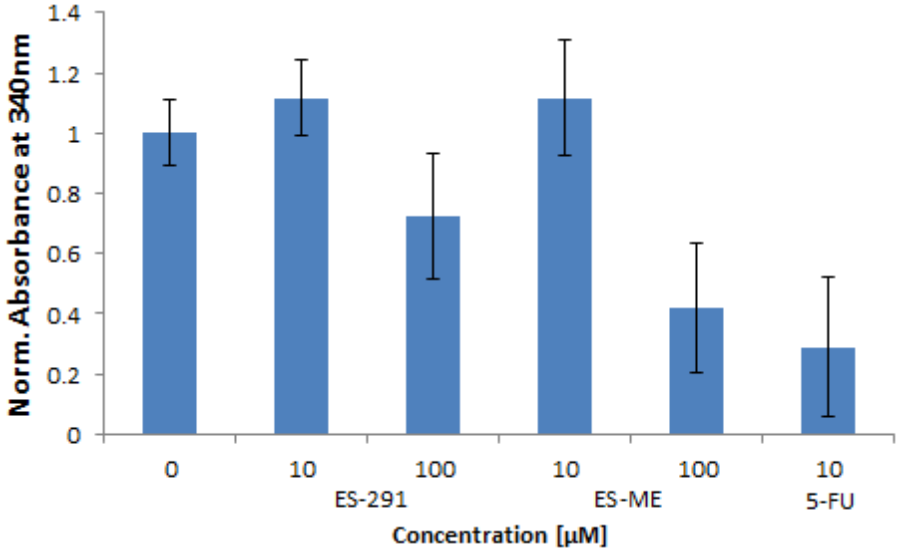


Figure 2.18 Inhibition of TS enzymatic activity by ES-ME, ES-291, and 5-FU as a positive control. TS ligands likely function as TS substrate binding site inhibitors. Results are averages of triplicate experiments. Error bars correlate to 1σ.

Peptide Screening

Due to the lack of promising hits from the small molecule screen, we turned our focus towards short arginine-rich peptides which have been previously shown to have a propensity for RNA binding (39). The initial set of peptides was symmetrically designed in such a way that two neutral residues were flanked by two sets of arginine dyads. These peptides were then incubated with a synthetic 33mer RNA encompassing the majority of the TS site 1 motif (corresponding to nucleotide 80-112 in Figure 2.3), and then analyzed by native polyacrylamide gel electrophoresis (PAGE). Peptides that contained tyrosine or tryptophan amino acid pairs in the center (A, and B respectively) showed nonspecific binding, and at higher concentrations, RNA precipitation and aggregation (Figure 2.19).

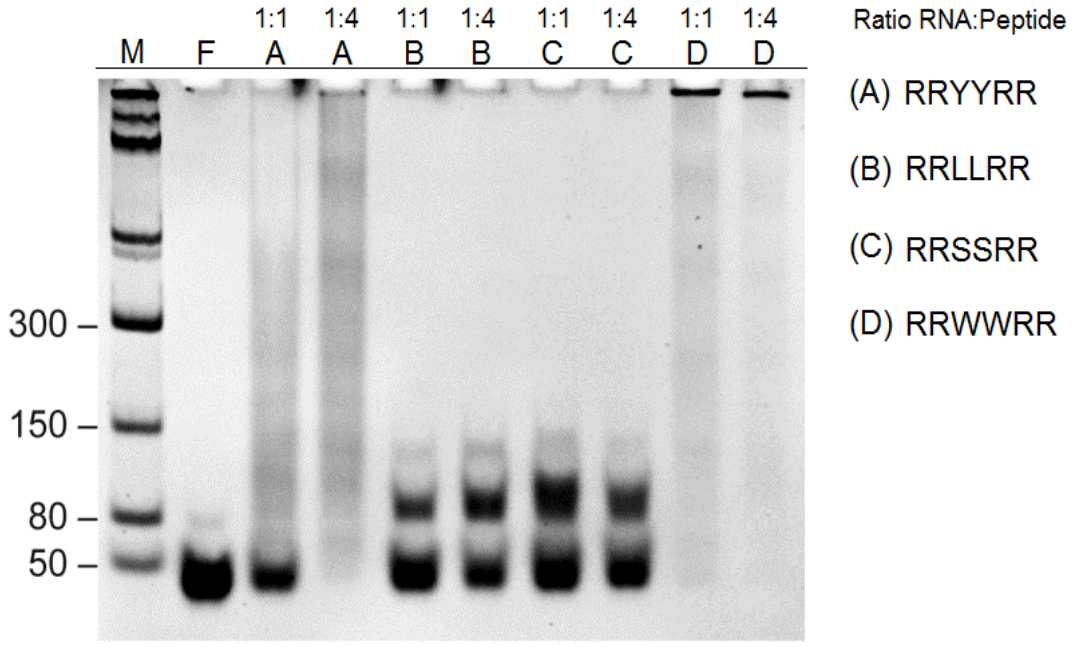


Figure 2.19 Native PAGE analysis of hexapeptides binding to 33mer TS site 1 motif. "M" indicates low range ssRNA marker, "F" indicates peptide free RNA, and stoichiometric ratios are indicated as either 1:1 or 1:4 RNA:Peptide.

Two peptides that showed promising binding activity were those either containing a pair of leucine or serine pairs (peptides B and C respectively). The titration of both peptides B and C demonstrate TS1 RNA binding in a dose dependent fashion with an affinity of ~50μM (Figure 2.20). In order to determine whether the binding of the peptide to the RNA would impact translation, we tested the B, C, and D peptides in the IVT assay. The reporter expression was determined in the presence of 150μM peptide (Figure 2.21) in the TS1 and ΔTS1 constructs.

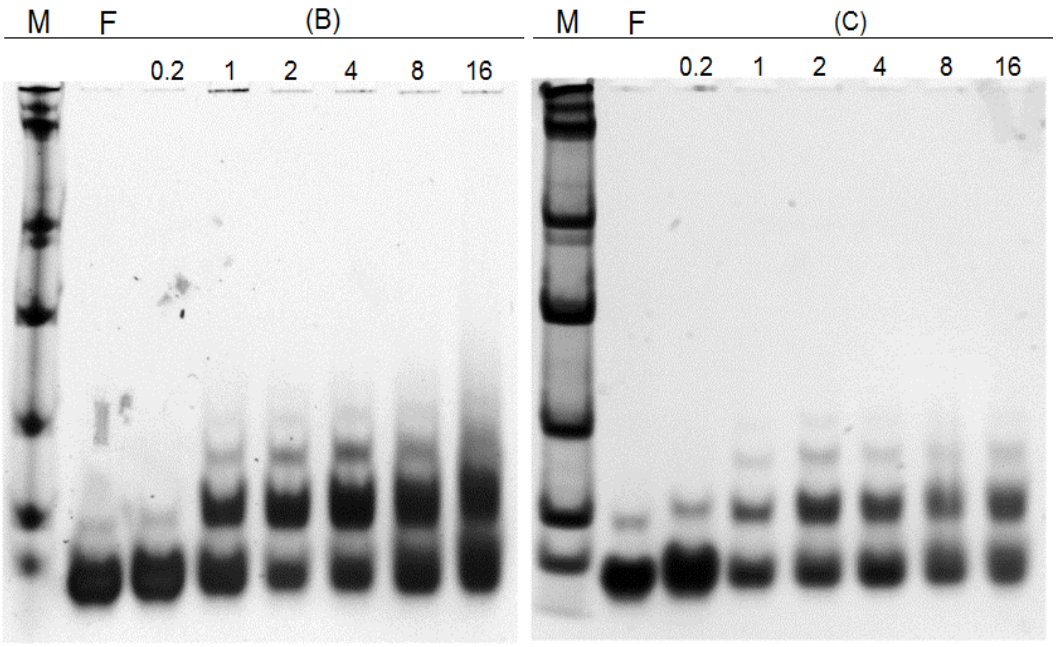


Figure 2.20 Dose responses of peptides (B) and (C) with 33mer TS1 RNA in native PAGE assay. Numbers correlate to the respective ratio of peptide to RNA.

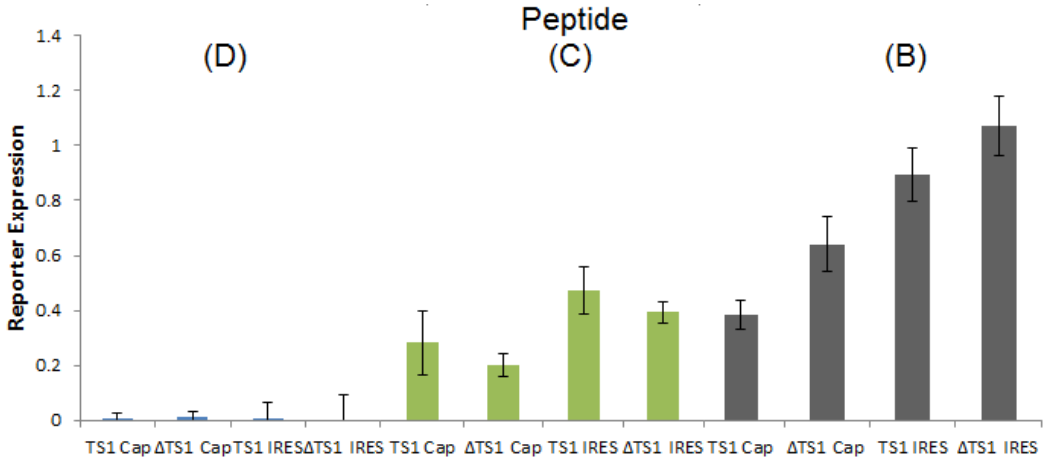


Figure 2.21 Presence of hexapeptides B, C, and D and their effect on translation as determined by reporter expression in bicistronic constructs. Expression was normalized to a peptide-free negative control. Peptide D appears to introduce aggregation, which is consistent with Figure 2.19. Results are averages of triplicate experiments. Error bars correlate to 1σ.

The tryptophan containing peptide (D) exhibited a profile of strong indiscriminant inhibition of translation, consistent with RNA aggregation (observed in Figure 2.19). The serine containing peptide (C) which formed stable complexes in the native gel (Figure 2.20) inhibited translation in an indiscriminant fashion. Levels of both cap driven and IRES driven translation were decreased in the presence of peptide (C). However, in the presence of the leucine containing peptide (B) we observed a marked decrease in translation (60%) within the TS1 specific cap driven construct. IRES reporter expression did not decrease in the presence of peptide (B), which suggests that the leucine peptide minimally affects the ribosomes and the translational machinery, while simultaneously binding to the TS1 target. Further examination of the binding interactions of the peptides, coupled with their behavior in the IVT assay, suggests that optimization of the lead peptide (B) could yield an even more potent TS site 1 inhibitor.

Additional peptides whose sequences were inspired by the lead peptide (B) were tested in the IVT assay, and several promising peptides were discovered that appear to preferentially inhibit TS site 1 regulated translation (Figure 2.22). Other peptides exhibited the previously observed RNA precipitation, while others indiscriminately target the translational machinery of the ribosomes (Figure 2.23) while not precipitating RNA.

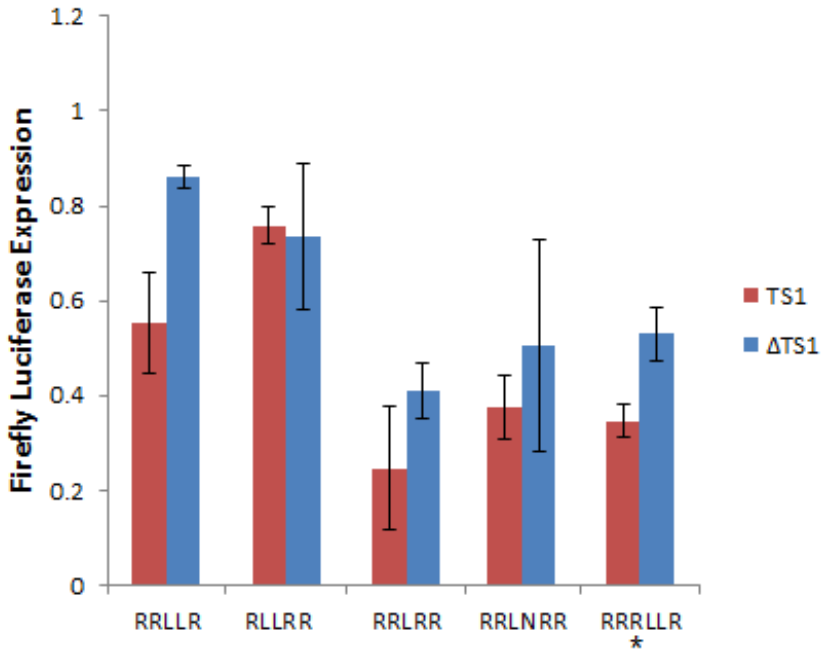


Figure 2.22 IVT assay in bicistronic TS1 and ΔTS1 constructs of additional hexapeptides, and pentapeptides inspired by the lead peptide (B). Values are normalized to a peptide free control, and all peptides are at 150μM *except RRRLLR, which is at 100μM. Results are averages of triplicate experiments. Error bars correlate to 1σ.

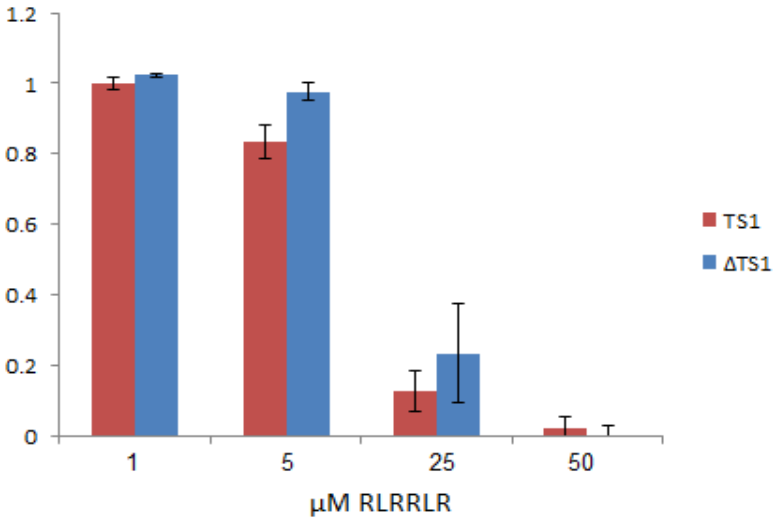


Figure 2.23 Titration of the RLRRLR peptide in the IVT assay. Native PAGE gel shift assay does not indicate precipitation, yet the indiscriminant translation inhibition indicates it is likely targeting the ribosomes and the HCV-IRES. Values are normalized to a compound free control Results are averages of triplicate experiments. Error bars correlate to 1σ.

Conclusions:

This study has demonstrated that the TS site 1 hairpin is a critical component of TS protein autoregulation, and can be examined in the context of mechanistic reporter assays for the purpose of developing a specific inhibitor of TS site 1 regulated translation. As our investigation of stabilizing mutants in TS depleted extract demonstrate, the secondary structure of the TS 5' UTR provides a negligible roadblock in the process of downstream translation. Rather it takes the presence of this TS site 1 hairpin in addition to bound TS protein to decrease translation. Additionally, stabilization of the RNA secondary structure by a small molecule or a peptide does provide sufficient energy to suppress translation, and could be utilized as a method to overcome chemotherapy-induced resistance to TS enzyme inhibitors.

Acknowledgements

Portions of the text and Figures 2.1-2.21 were adapted from a published article, Brunn et al., (2012) Targeting a Regulatory Element in Human Thymidylate Synthase mRNA. *ChemBioChem*. The dissertation author was the primary author, Emily Garcia Sega and Melody B. Kao were co-authors and Thomas Hermann was the corresponding author of this article.

Chapter 3. Thymidylate Synthase Protein and RNA Interactions

Introduction

While many publications examine the TS protein's interactions with its cognate mRNA, the regions of the protein involved in the RNA interactions are unknown. The identification of these sites is critical for gaining insight into the hTS protein-RNA recognition and interaction, and designing future TS inhibitors that do not affect the RNA binding site, or that target the RNA by mimicking the proteins structural features. It has been previously demonstrated that the human TS protein is a dimer in its enzymatically active form, and that it can adopt both "active" and "inactive" conformations (1, 2, 3). The presence of anionic ligands have been shown to drive the TS protein towards the inactive conformation, while enzymatic substrates drive the formation of the active conformation (3). The TS monomer-dimer dissociation constant (K_D) has been reported as 200nM (4). Our study seeks to better define the interaction of the hTS protein with its cognate RNA, specifically the TS site 1 regulatory motif, previously examined in Chapter 2. This was accomplished through the analysis of multiple *in vitro* transcription-translation assay constructs utilizing various truncation mutants of the TS site 1 RNA hairpin, and the use of thiouridine modified TS-RNA hairpins, utilized in a 360nm UV protein-crosslinking assay, described previously in Chapter 1.

TS Site 1 Truncation Mutant Analysis

Since we have previously demonstrated the translational suppression of TS site 1 regulated translation in the presence of the thymidylate synthase enzyme (63), we set out to investigate whether this effect was located on any particular region of the TS site 1 RNA. To

identify whether the protein-RNA interaction was near the base of the secondary structure stem loop, or closer to the apical loop, three TS site 1 RNA truncation mutants were cloned into the bicistronic reporter constructs (Figure 1.5). The truncations were selected to maintain the pyrimidine-rich RNA sequences surrounding the initiation codon, while also maintaining the correct reading frame of the firefly luciferase reporter. These constructs had either 23 (TS1 23), 20 (TS1 20), or 15 (TS1 15) nucleotides composing the TS site 1 motif (Figure 3.1). We examined the truncation mutants in the IVT assay utilized in Chapter 2, in either the presence or absence of 5μM hTS enzyme.

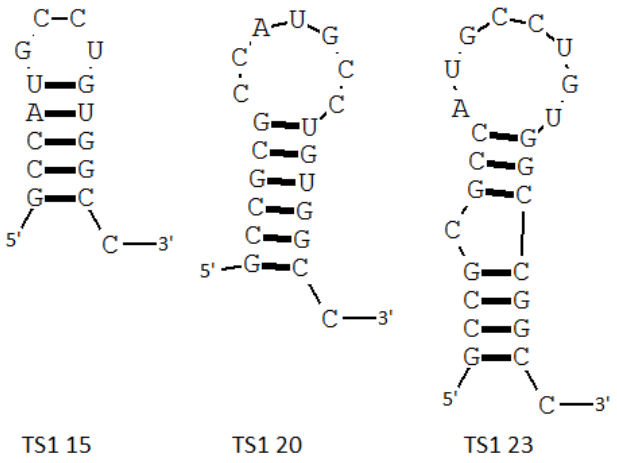


Figure 3.1 Names and predicted secondary structures for the TS1 truncation mutants.

A TS enzyme concentration of 5μM in the IVT assay had been shown to roughly decrease translation by 50% in the wild type TS site 1 sequence reporter (Figure 2.1). The TS1 23 mutant appears to behave very similar to the wild type sequence, both in the relative levels of translation, but also in the susceptibility to translation inhibition in the presence of TS enzyme (Figure 3.2). The TS1 20 mutant exhibited higher levels of translation overall, about 150% higher than the wild type sequence. However upon addition of the TS enzyme,

the respective rates of translation suppression are very similar to that of the wild type sequence. This trend continues for the TS1 15 mutant, which has minimal secondary structure and an even higher basal rate of translation, yet in the presence of 5μM TS enzyme, it still has its respective levels of translation decreased by approximately 50% (Figure 3.2). This data seems to indicate that while the secondary structure of the 5' UTR may play a role in the respective basal rates of translation in the IVT assay, even a 15 nucleotide stem loop containing the translation initiation site is sufficient to observe TS enzyme mediated translation suppression. This indicates that the majority of the TS protein-RNA binding interaction is likely occurring near the apical loop of the RNA.

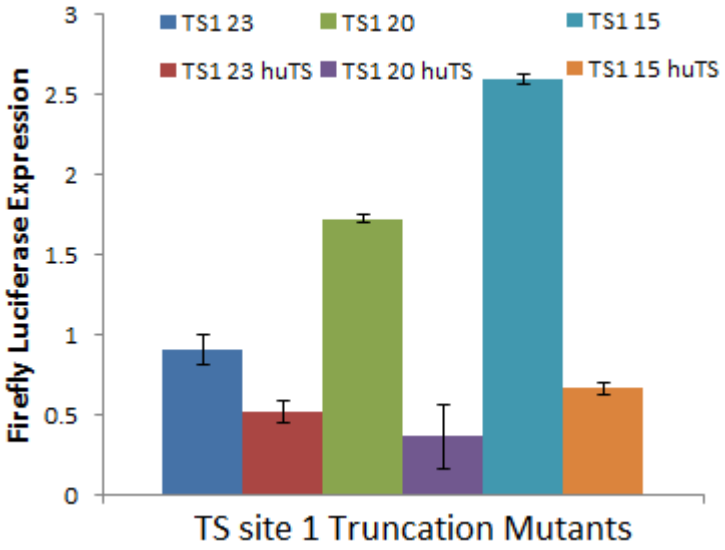


Figure 3.2 IVT assay reporter expression for the TS site 1 truncation mutants in either the absence, or presence of 5μM TS enzyme (huTS). TS1 23, 20, and 15 values are normalized to wild type TS1 reporter translation, whereas the 5μM huTS dosed samples (huTS) are normalized to their enzyme free counterpart. Results are averages of triplicate experiments. Error bars correlate to 1σ.

TS Protein-RNA Crosslinking and PAGE Analysis

To identify the regions of the TS enzyme responsible for the RNA interaction, thiouridine modified TS-RNA hairpins, previously described in Chapter 1, were utilized in a biochemical crosslinking assay (Figure 1.2). Thiouracil modified RNA will crosslink to neighboring atoms when exposed to 360nm wavelength radiation, avoiding problems associated with non-specific crosslinking observed with shorter wavelength radiation. Additionally, as previously mentioned, thiouracil modified RNA maintains the same structure, and functionality as its unmodified counterpart.

Based on the results of the truncation mutant IVT assay (Figure 3.2), using a RNA containing a thiouridine at position U95 of the terminal loop of the TS site 1 secondary structure, should provide the necessary proximity to the TS protein to allow crosslinking to occur.

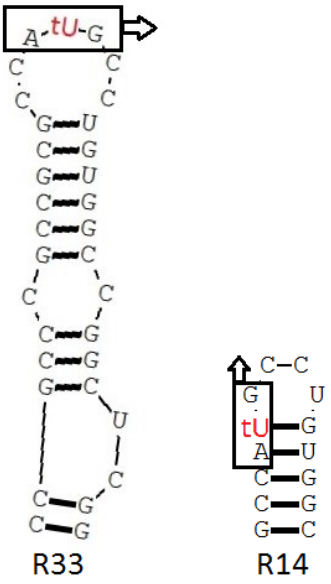


Figure 3.3 Predicted secondary structure of thiouracil labeled (red) RNAs. Box indicates translation initiation site.

Upon incubation with a 1:1 ratio of protein to RNA, the mixture was irradiated with long range UV-light at a wavelength of approximately 360nm. The radiation provides sufficient energy to crosslink the thiouracil to any region of the protein within 1-2 angstroms. After crosslinking, the sample was run on either a native PAGE gel (Figure 3.4 A) or a denaturing sodium dodecyl sulfate PAGE gel (SDS-PAGE) (Figure 3.4 B). Analysis of the SDS-PAGE gel (A) indicates that the R33 construct crosslinks to the TS protein, and does so in a single defined complex. Conversely the R14 RNA crosslinks to multiple places on the TS protein, indicative of potentially nonspecific RNA interactions, which become captured upon UV irradiation. The detection of a single crosslinked complex in the R33 RNA suggests the hTS binding of one RNA precludes the adjacent subunit from binding to another RNA. The RNA binding event might alter the hTS dimer's conformation, disfavoring additional RNA binding in the adjacent subunit. Examination of the native PAGE gel demonstrates that the vast majority of TS protein runs as a dimer in the gel, which is consistent with the dissociation constant of the native complex. In addition, the crosslinked samples indicate that the TS protein-RNA complex can exist as both a dimer, and a monomer, although the dimer is present in excess. It is unclear whether the presence of RNA on a TS protein monomer precludes it from dimerizing, however we can say that the RNA presence on the TS dimer does not force its dissociation. We hypothesized that the TS protein interacts with the RNA, and crosslinks as a dimer, but can eventually disassociate into its respective monomeric components. A time course experiment was conducted to determine the order of binding events. The respective concentrations of TS protein and RNA were lowered, to 25uM and 10uM respectively. The assay was irradiated with 360nm UV light for 50 minutes, with samples of the complex being taken at periodic intervals (Figure 3.5).

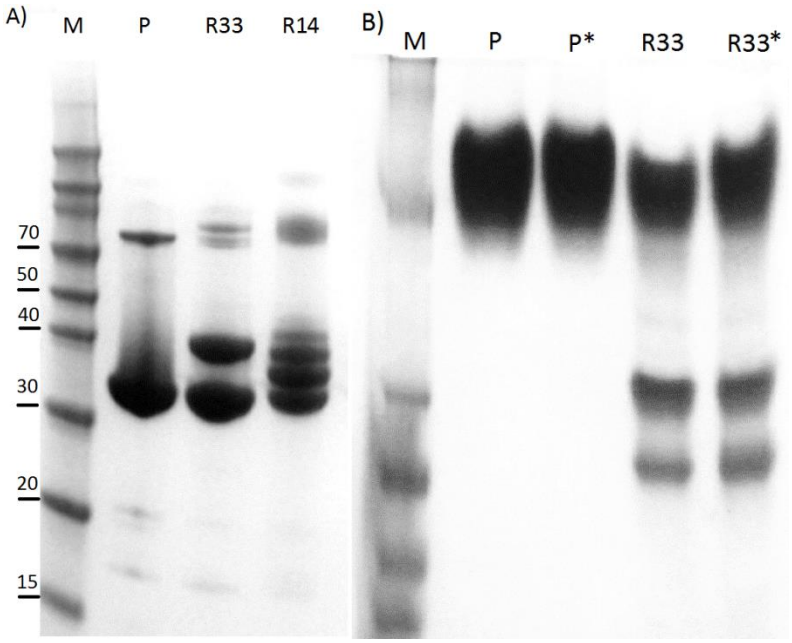


Figure 3.4 SDS-PAGE denaturing gel (A) and native gel (B) of TS protein (P), protein-RNA crosslink complexes with either a TS-RNA 33mer (R33) or a TS-RNA 14mer (R14). Protein and RNA concentrations are 200μM. P* and R* indicated the addition of 1% zwitterionic detergent (CHAPS).

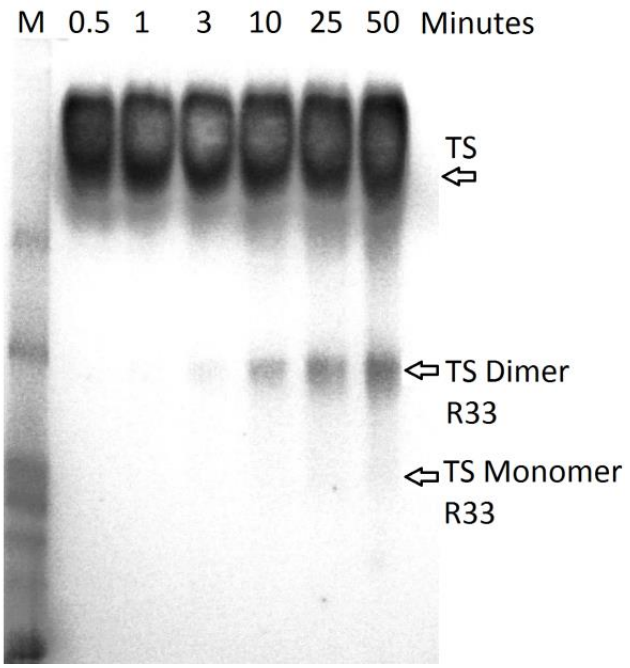


Figure 3.5 Native PAGE kinetic analysis of TS protein (P) and R33 crosslinking over 50 minutes.

Based on the visualization of two bands that shift in the native gel, and that the dominant one is the less mobile, it appears that the majority of the TS protein crosslinks in a time dependent fashion as a dimer, and is stable (Figure 3.5). The presence of both species albeit with the dimer in excess, seems to suggest that the monomer is capable of binding to, and subsequently crosslinking with the TS protein. However this complex is disfavored, possibly due to the low levels of free monomeric TS existing in solution.

Mass Spectrometry Analysis

With the knowledge that we can successfully isolate the TS protein-RNA complex, we set out to determine where the TS1 R33 RNA binding event was occurring on the TS protein. This was accomplished using mass spectrometry (MS) to analyze peptides present in the elution profiles of trypsin digested TS protein, and compare it with the crosslinked TS protein-RNA complex. The rationale behind the experiment is that the RNA should remain

intact through the trypsin digestion, and the presence of a large 10 kD negatively charged RNA should cause any crosslinked protein fragments to disappear from the MS detection compared to the control of the unmodified protein. Since the upper detection limit in this particular matrix-assisted laser desorption/ionization (MALDI) Time of Flight (TOF) mass spectrometry is roughly 4kD, fragments containing the crosslinked 10kD RNA would no longer be detectable.

The band corresponding to the TS protein dimer crosslinked to the 33mer RNA (Figure 3.5) was excised and used as the crosslinked sample. Unreacted "TS" protein from minute 0 was used as the negative control. The elution profiles of the two samples had many similarities, but the crosslinked sample, had several peaks that had disappeared (Figure 3.7). Peptide fragment analysis of the two samples indicated that while they both had over 85% coverage of the entire thymidylate synthase sequence, there were several key peptides in the crosslinked sample that were either significantly diminished in intensity, or missing entirely (Figure 3.6). When examined more closely, it appears that there are three major "hot spots" of RNA binding.

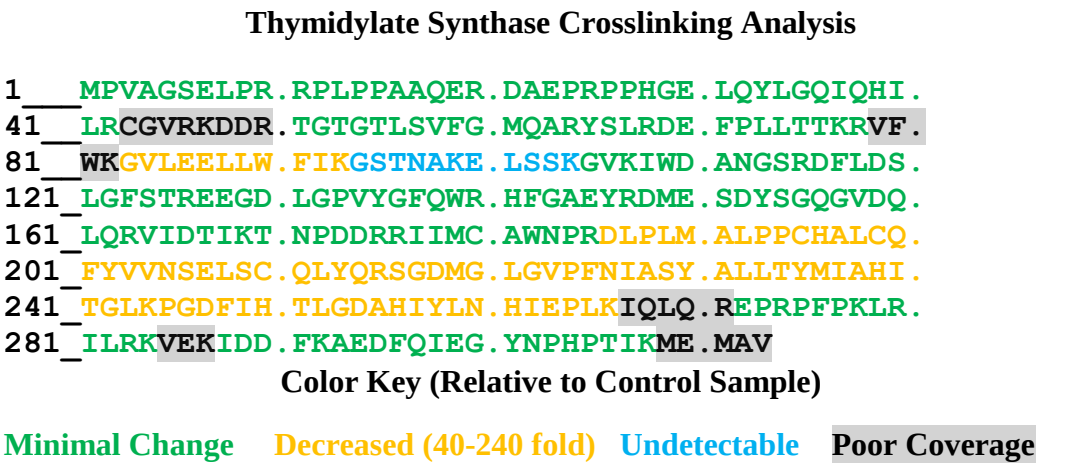


Figure 3.6 Color mapping of the mass spectrometry sequence analysis of respective ratios of TS peptides present in the control sample, compared to the RNA crosslinked sample.

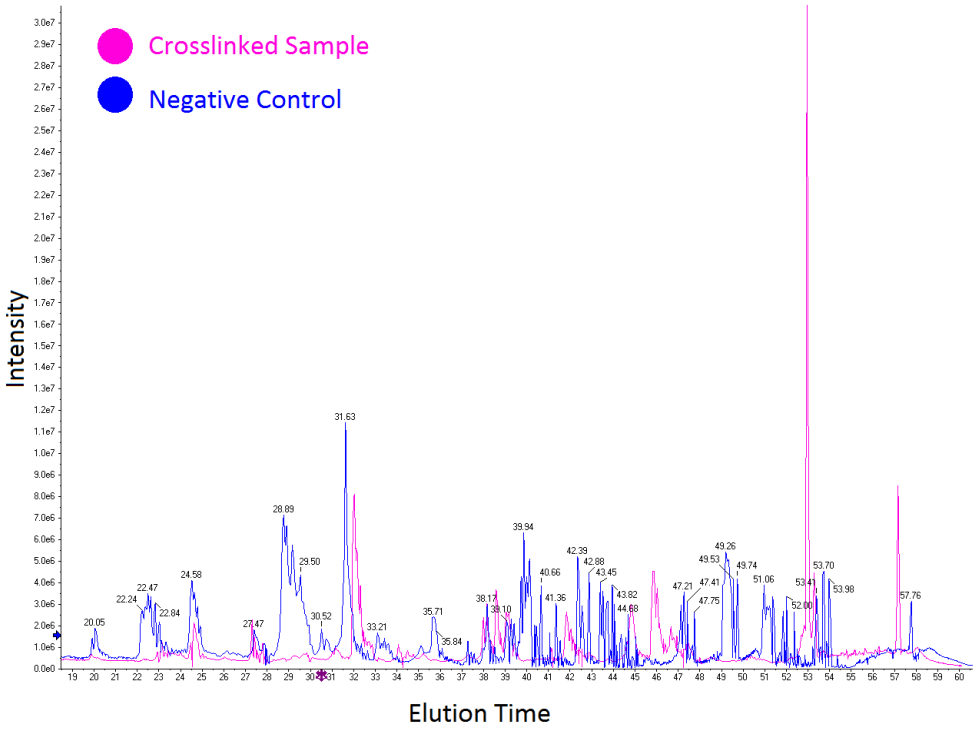


Figure 3.7 Elution profile of trypsin digested thymidylate synthase protein (blue) overlaid with the R33 RNA crosslinked sample (magenta).

It should not be surprising that there are multiple crosslinking sites, as most RNA-protein interactions are dynamic in nature. The multiple crosslinking sites are likely an artifact of the design of the crosslinking experiment, i.e. that the RNA could be prematurely crosslinked prior to reaching its optimal binding site.

Structural Examination of Protein-RNA Binding Regions

Our crosslinking analysis allows us to map the TS protein-RNA interactions in three-dimensional space. It was previously unknown where the TS mRNA binds to the TS protein, but it was hypothesized that this occurs in the catalytic cavity which is often occupied with either a phosphate, sulfate, or substrate group in crystal structures (15). Previous studies have shown that mutating the catalytic cysteine residue does not affect the

RNA binding capacity of TS protein (60), suggesting that the majority of the interactions are in the peripheral regions around the catalytic active site. When peptides with significantly decreased abundance in the complex were mapped on either of the previously determined hTS crystal structures, a helix-loop-helix motif (G84-K104) adjacent to the active site cleft emerged as the likely crosslinking domain (Figure 3.9). In the crosslinked complex, a peptide involved in a solvent-exposed helix (G94-K104) was undetectable, and the abundance of a preceding peptide involved in a second helix (G84-K93) was reduced by about 150-fold. Additionally, peptides in adjacent helices and a β-sheet within the core of the protein were reduced, however, this was likely through indirect effects of RNA crosslinking at the neighboring helix-loop-helix motif. The G84-K104 helix-loop-helix domain is located opposite the active site, though within the same cleft. The G94-K104 helix protrudes from the protein surface, then continues outward. The helix involving residues G84-K93 has one solvent exposed face, while the other side provides one of the walls of the active site cleft. The accessibility of the helix-loop-helix motif on the protein exterior and crosslinking data are consistent with a proposed function as the hTS RNA binding site. However, the sequence of the G84-K104 helix-loop-helix domain bears no homology to known RNA-binding motifs, suggesting that hTS recognizes the TS site 1 RNA via a unique structural element. Deeper insight into details of the protein-RNA interaction will require a crystal structure of the complex.

Examination of the structure of the thymidylate synthase solved in “low salt” conditions, shows the two homodimeric subunits sharing two amino acid residues (Arg 175-Arg 176) in the other’s active sites (Figure 3.8). And when we map the peptide crosslinking hotspots (Figure 3.6) onto the thymidylate synthase structure (Figure 3.9 and Figure 3.10) we can see that the crosslinking appears to be occurring within the active site cleft, opposite

of the active site (Figure 3.9, 3.10). In the absence of a co-crystal structure, a 20 nucleotide hairpin from a nucleolin-RNA complex (PDB ID 1RKJ), which is of similar size to the TS site 1 RNA hairpin, was extracted and placed into a position contacting the three major RNA "hot spots" in order to visualize this interaction (Figure 3.11).

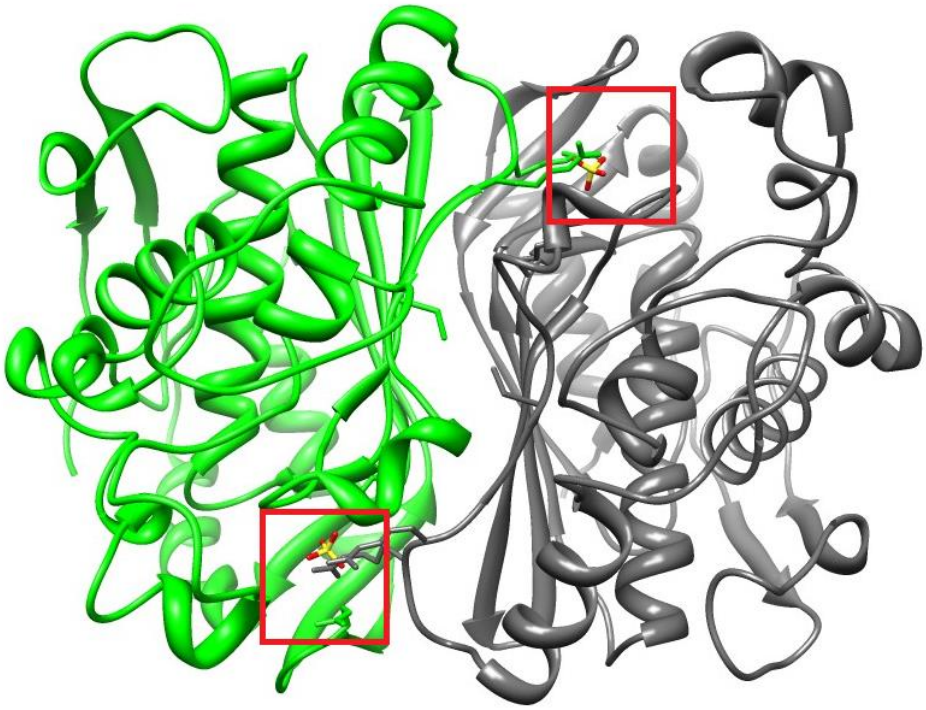


Figure 3.8 Structure of the human thymidylate synthase solved at low salt (active) conditions. Two subunits colored green and grey respectively, and red boxes indicate respective active site locations. PDB ID 4H1I.

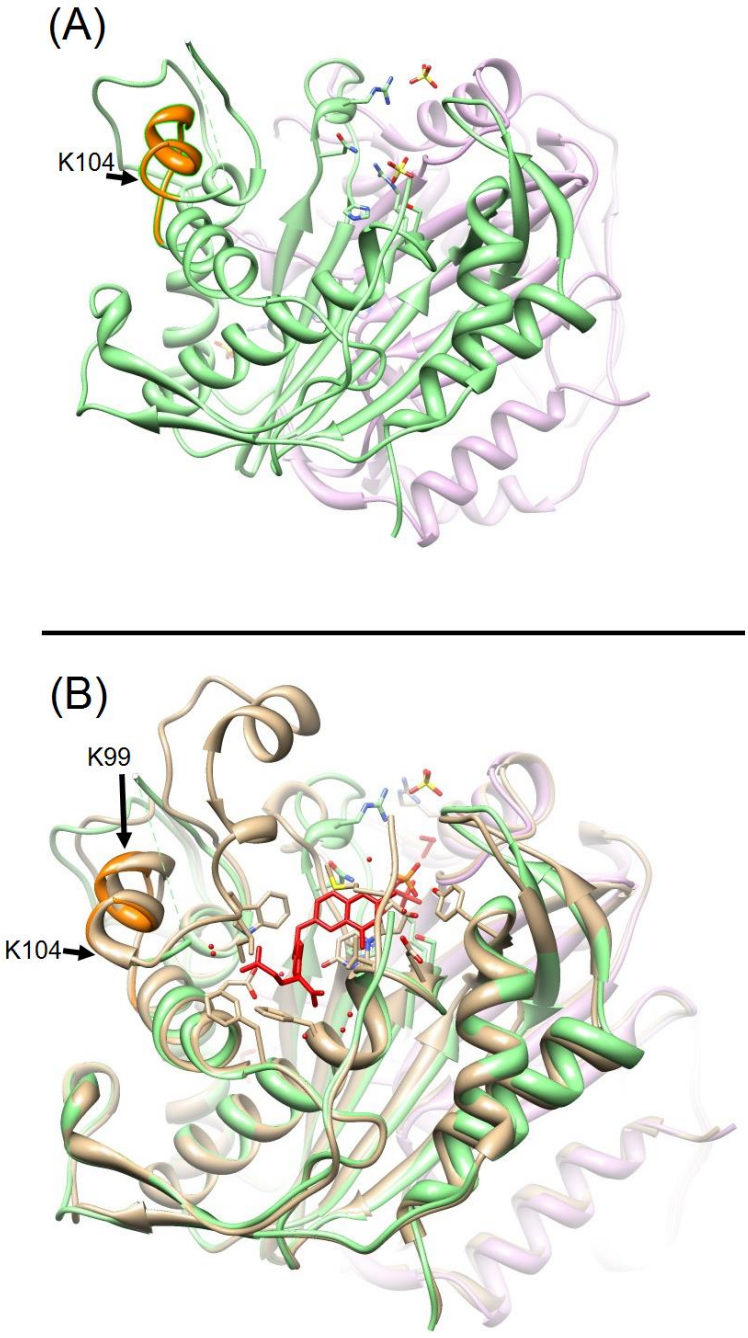
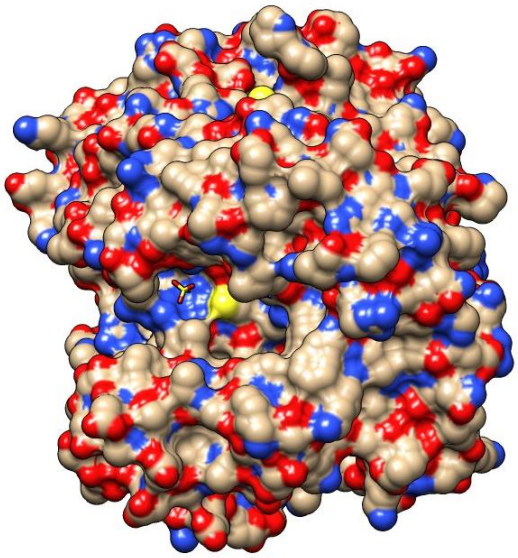


Figure 3.9 (A) Results from TS R33 RNA crosslinking mapped on the TS crystal structure determined at high salt (green & magenta, PDB ID 4GHY). Mass-spectrometric analysis indicates that the RNA crosslinks to the region between G94-K104 (orange).In (B) the structure of hTS in its active conformation has been superimposed (tan, PDB code 1HVY) which has been crystalized in complex with a TS inhibitor raltitrexed (red). The region from K107-E128 is disordered in the high (inactive) salt structure, but ordered in the active structure.

(A)



(B)

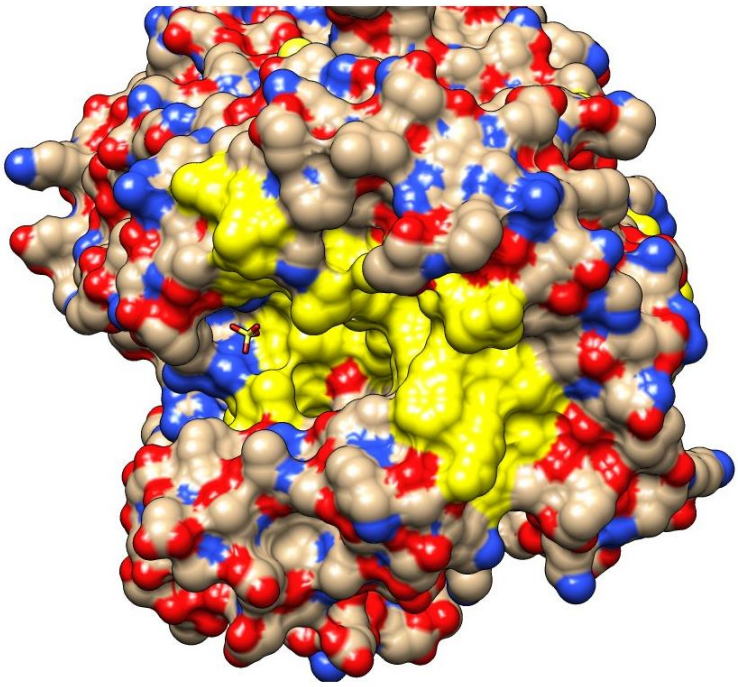


Figure 3.10 Thymidylate synthase structure with electrostatic potential surface mapping (A). Blue signifies acidic residues, red signifies basic. (B) Peptide mapping from figure 3.6 overlaid onto figure (A). Yellow indicated >40 fold decrease in abundance. PDB ID 4H1I.

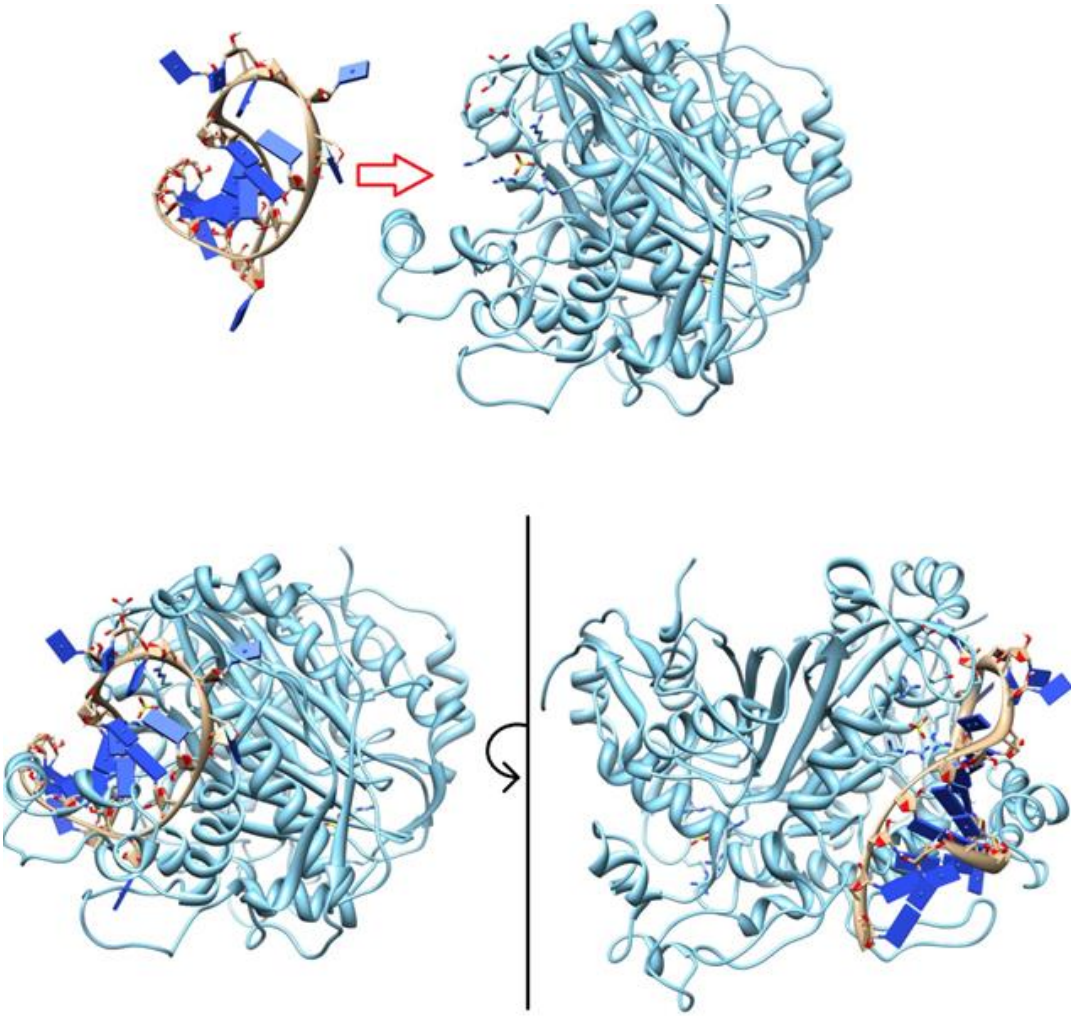


Figure 3.11 Model obtained by positioning a 20 nucleotide hairpin RNA at the hTS crystal structure obtained at low salt conditions. RNA location predicted by peptide crosslinking map. Alpha helices modeled to fit into the RNA groove, beta sheet loops modeled to interact with phosphate backbone.

In the absence of a three dimensional TS RNA structure, we can observe that the nucleolin RNA terminal loop can fit into the TS crystal structure, and can do so in a way that not only conforms to the crosslinking results, but also in a way that explains several observations of the crosslinking experiments. We observe that only a single 33mer RNA can

crosslink to the TS dimer. Based on the likely binding region of the RNA, this may be due to a conformational shift where, upon RNA binding within the active site cleft, the two neighboring arginines of the opposite TS homodimer are shifted into a conformation that disrupts the TS RNA binding interaction. This feedback from one enzyme to another through its active site may explain why the TS protein will bind and crosslink to only one 33mer RNA. When the RNA length is shortened to a 14mer we no longer see this feedback, and observe multiple crosslinking interactions (Figure 3.4). If the residues from glycine 83 through lysine 104 were binding near the RNA terminal loop region, than an RNA with less than approximately 20 nucleotides would not likely be long enough to impact the active site. (Figure 3.12)

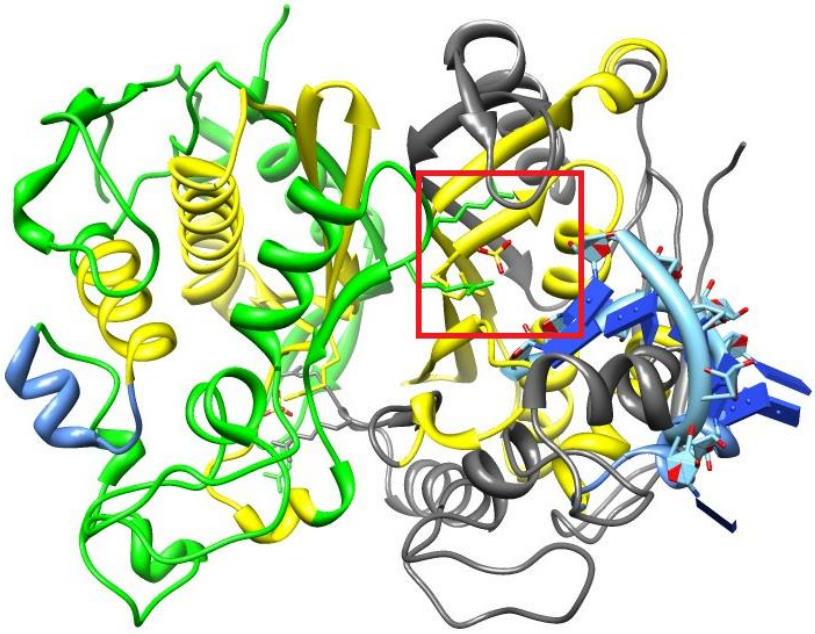


Figure 3.12 TS subunits in the inactive conformation (green/grey, PDB 4GHY) with the mass spectrometric analysis of the crosslinking activity mapped on to the structure. Blue indicated the G94-K104 crosslinking region, and the yellow indicates regions effected by crosslinking. The RNA is truncated to a 14mer terminal loop RNA. Red box indicates active site.

RNA Binding Activity of Peptides

Short peptide sequences of the highly active "hot spot" (residues 82-104) were tested for their *in vitro* binding activity against a 23mer 2AP modified TS1 RNA. Previously tested peptides, and the TS protein were both used as controls for positive binding interaction. (Figure 3.13). To further characterize the peptide-RNA interaction, as well as identify any nonspecific precipitation, a gel shift assay was conducted on the peptides (Figure 3.14).

The results of the 2AP assay suggest that there are specific regions of the TS protein, that when utilized in the context of a peptide, are capable of binding to the TS mRNA. The hTS region 82-89 interacts with the TS RNA in a dose dependent fashion. Interestingly, shifting just two amino acids upstream ablates this binding activity, as we see no RNA interaction with peptide TS84-92. This would suggest that the regions of the TS protein that crosslinks with the RNA may not be the regions of the protein responsible for the binding interaction. Rather, the proximal regions of the protein around the "hot spots" are likely forming the majority of the protein-RNA interaction.

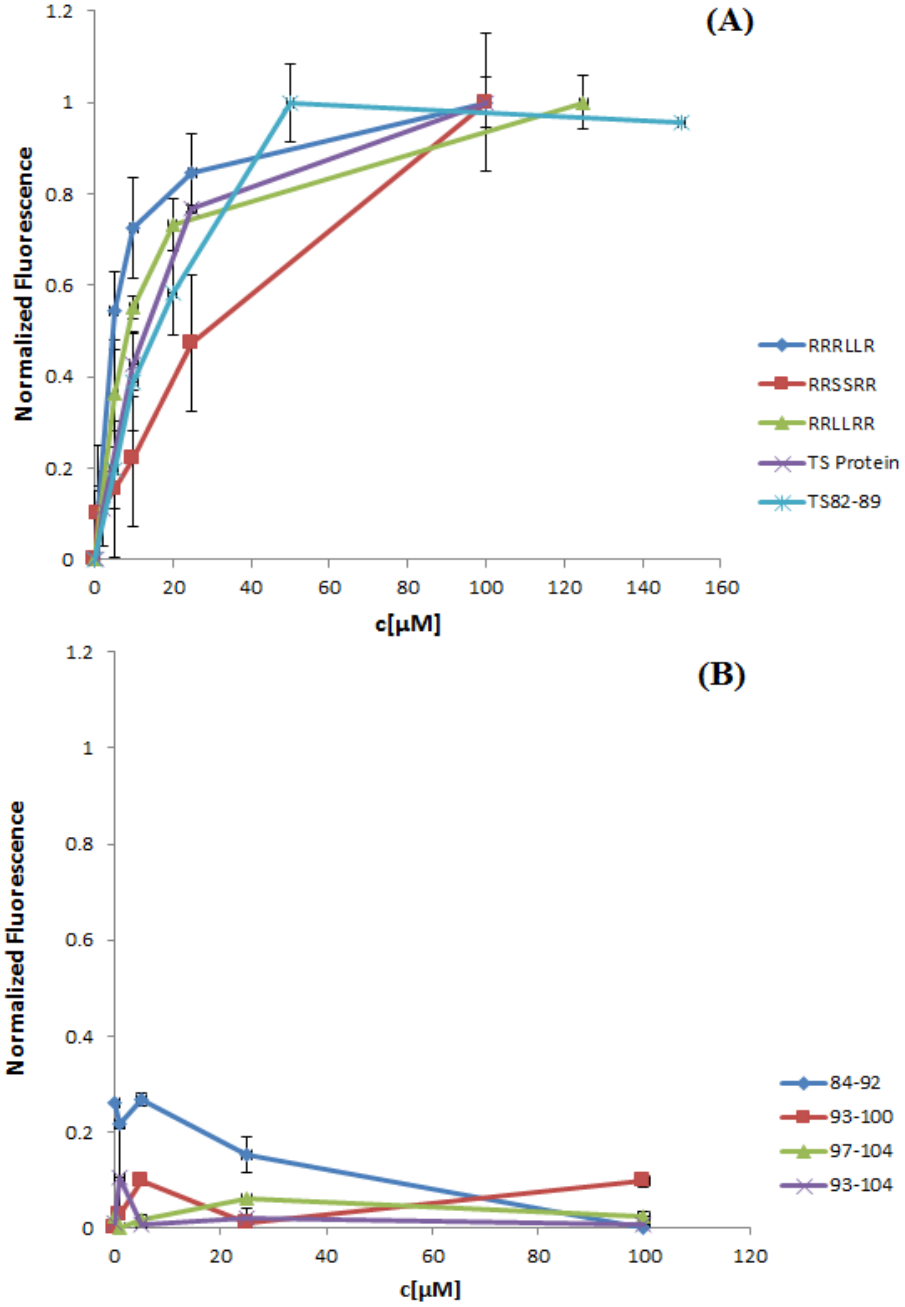


Figure 3.13 2AP assay examining various TS1 RNA ligands, "hits" from the crosslinking analysis are peptides TS82-89, TS84-92, TS93-100, TS93-104, TS97-104. Active compounds (A) and inactive (B). Peptide 84-92 shows interactions consistent with RNA aggregation. This observation is confirmed in the gel shift analysis (Figure 3.14) Values are normalized to a compound free control. Results are averages of triplicate experiments. Error bars correlate to 1σ.

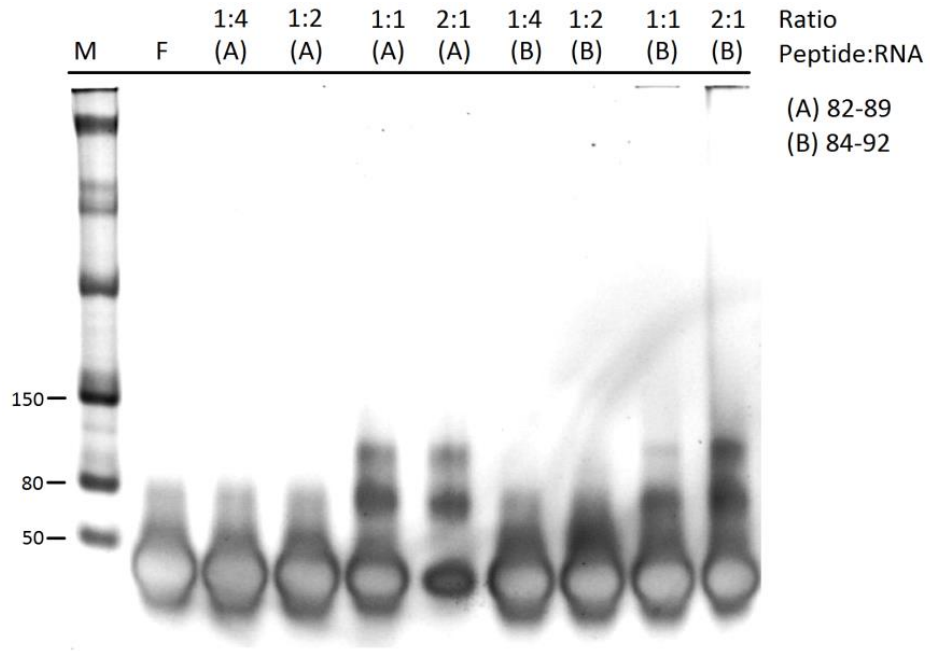


Figure 3.14 Gel shift of peptides TS82-89 and TS84-92. M indicates the single stranded marker, F is the free 33mer RNA.

Applications of 2AP Screen

Based on the results of the 2AP TS peptide analysis, we set out to probe whether the 2AP assay extends to more than just peptides, and whether the results of the 2AP assay were indicative of the *in vitro* activity of a compound or peptide in the more biologically complex IVT assay. Several known RNA binding aminoglycoside antibiotics were tested, but exhibited no change in the relative 2AP fluorescence levels, indicating that non-specific RNA binding was insufficient to yield an increase in fluorescence. However upon screening several polymyxin antibiotics, it was discovered that polymyxin B did in fact lead to an increase in relative fluorescence (Figure 3.15). Interestingly other structurally similar polymyxins were also tested in the 2AP assay, but did not affect the fluorescence levels (Figure 3.15).

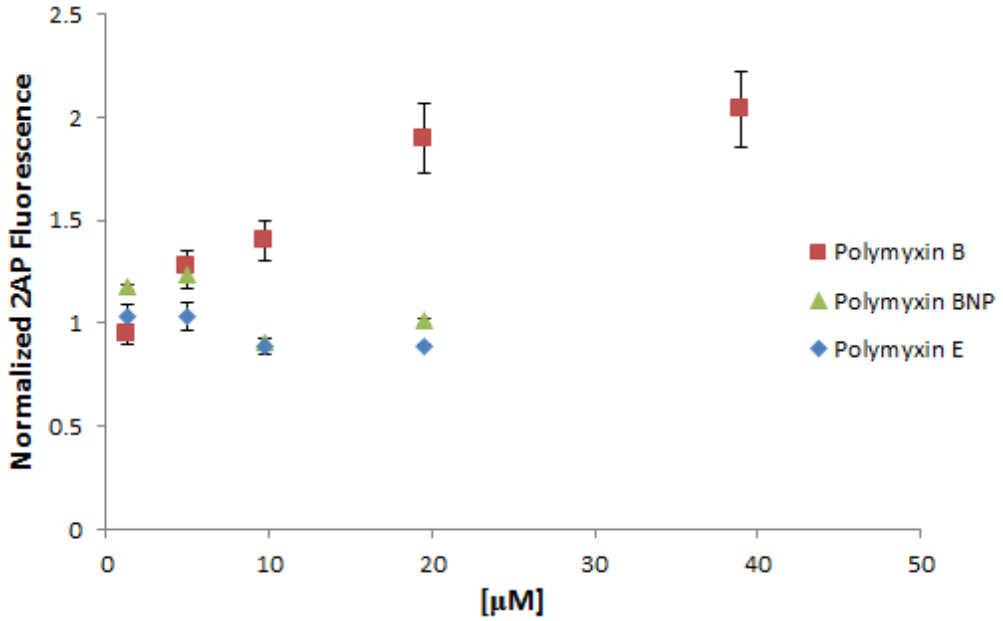


Figure 3.15 2AP assay of polymyxin B, E, and B-nonapeptide (BNP). Values are normalized to a negative control. Results are averages of triplicate experiments. Error bars correlate to 1σ.

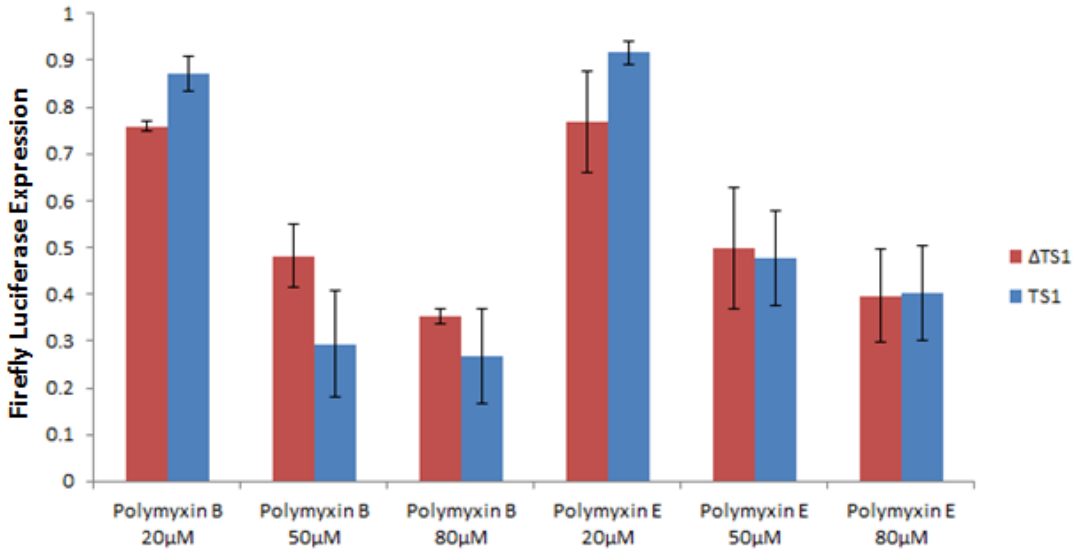


Figure 3.16 IVT assay dose response of polymyxin B and E in monocistronic construct with TS1 sequestered start codon (TS1) or unmodified monocistronic firefly reporter (ΔTS1). Values are normalized to a compound free control. Results are averages of triplicate experiments. Error bars correlate to 1σ.

To determine whether the selectivity of the polymyxins extended to the biologically more representative IVT assay, polymyxin B and E were tested in a dose response analysis in either a monocistronic reporter construct bearing the TS1 AUG sequestering secondary structure (TS1) or the unmodified monocistronic firefly reporter (ΔTS1) (Figure 3.16).While it is apparent that the polymyxins are capable of inhibiting translation in a non-specific fashion at higher dosages, it appears that polymyxin B has a modest window of effectiveness for preferentially inhibiting translation of the TS1 regulated firefly reporter. We then probed whether the 2AP active TS peptide 82-89 exhibited similar activity and selectivity compared to the TS 84-92 peptide in the IVT assay. The peptides were tested at two concentrations, and their respective activities were plotted (Figure 3.17). Peptides 93-100, 93-104, and 97-104 were also examined at 200μM in the IVT assay, but no change in translation was observed for either TS1 or the ΔTS1 constructs.

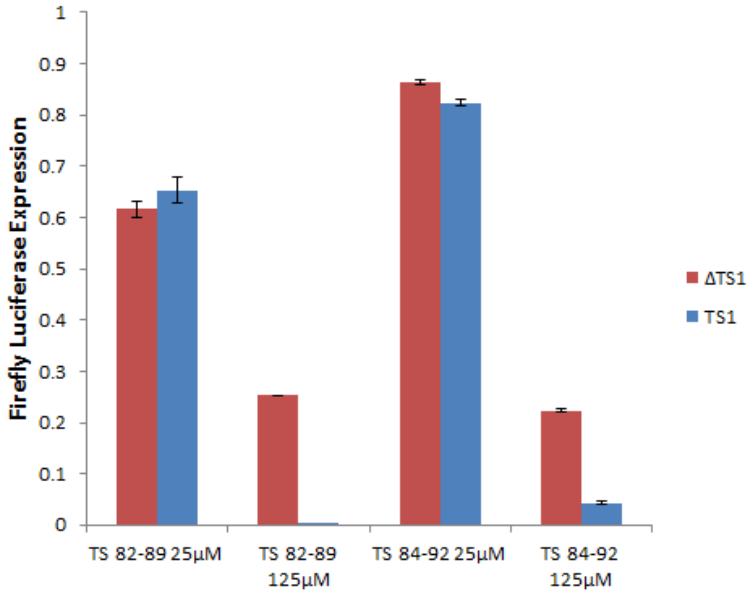


Figure 3.17 IVT assay of two TS peptides (82-89 and 84-92) implicated in the binding interactions of the TS1 RNA. Peptides were run at 25 μM and 125μM in both the TS1 and the ΔTS1 monocistronic constructs. Values are normalized to a compound free control. Results are averages of triplicate experiments. Error bars correlate to 1σ.

Based on their respective expression levels, it appears that the TS 82-89 peptide is more potent, and moderately more selective for the TS1 containing construct than the TS 84-92. However, contrary to the 2AP assay, the TS 84-92 peptide does exhibit selectivity towards the TS1 construct as well. The explanation is likely that the addition of the tryptophan and phenylalanine to the TS 84-92 peptide resulted in its propensity to precipitate RNA, previously observed in the peptides studied in Chapter 2 (Figure 2.14). This effect is likely amplified in the 2AP assay where there is only buffer and salt present (and we observe a dose dependent decrease in fluorescence), but more mild in the IVT assay with the cell lysate present to mitigate some of its precipitation potential. Based on the gel shift assay, it is clear that the peptide recognizes, and binds to the TS1 RNA, but that there is aggregation occurring simultaneously.

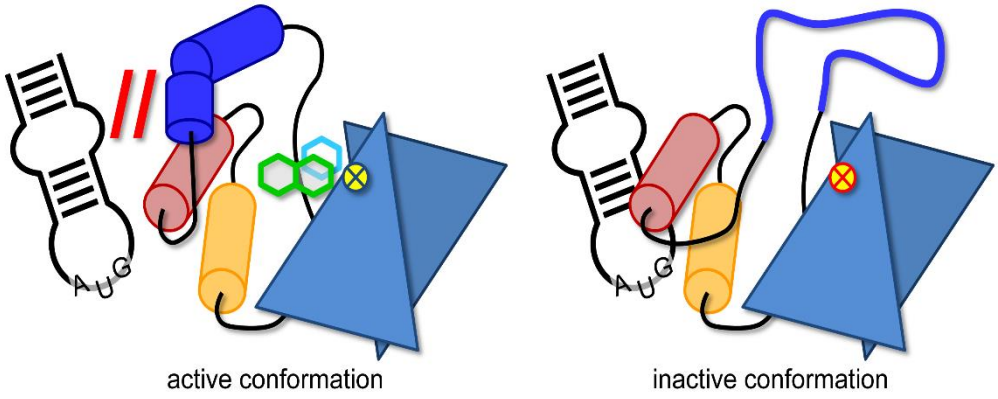


Figure 3.18 Proposed model for the hTS protein binding to the TS site 1 RNA. In the "Active" conformation, the RNA binding capacity of the G94-K104 helix (orange) is limited due to the structured region from K107-E128. Whereas in the "inactive" conformation, the K107-E128 adopts a disordered and dynamic fold, allowing recognition and binding of the G84-K104 helix to the terminal hairpin loop of the TS site 1 RNA.

Conclusions

A combination of X-ray crystallography, crosslinking assay, gel shifts, and IVT data, suggests that the RNA binding activity of the TS protein likely occurs in regions near the crosslinking "hot-spots" specifically around the G84-K104 helix. It is important to keep in mind that interactions between nucleic acids and proteins often occur in the RNA minor and major grooves, often employing a protein helix-turn-helix motif (29), and at apical loops where, in case of the TS site 1 RNA, the thiouridine was located. Regions upstream and downstream of the G84-K104 helix are likely participating in the RNA recognition and binding. This is observed in gel shifts and IVT assay with the peptides from these regions, whereby G94-K104 does not bind to the RNA (Figure 3.13) but the peptides 82-89 and even 84-92 bind quite efficiently. Based on the truncation mutants, a simple apical loop with a semi-stable secondary structure is sufficient for hTS protein recognition and binding. The affinity of the various peptides for the TS1 RNA was found to be sequence specific and not simply dependent on the number of acidic residues on the peptide, as peptides 82-89, 84-92, and 93-104 both have three lysine residues, and a net charge of +1 yet behave strikingly different. Previous peptide studies have attempted to determine RNA binding regions of the thymidylate synthase protein but only within arginine rich regions. Of those previously shown to bind to the TS1 RNA are peptides from residues 56-72, 131-147, 175-192, and 200-217, with peptide 175-192 binding with the most efficiently (10). However the binding affinity for those peptides appears to be correlated with the number of arginine residues located within the peptide. Our *in vitro* crosslinking data, coupled with the results of the 2AP assays and gel shifts, suggest that G84-K104, an arginine free region, appear to be critical for the protein-RNA interaction at the RNA apical loop. Based on these observations, we

hypothesize that the TS protein and its cognate RNA behave in a manner depicted in Figure 3.18, where, in the presence of its enzymatic substrates, the helices in region K107-E128 adopt a well ordered conformation, which precludes RNA recognition and binding by the TS protein. Conversely, in its enzymatically inactive conformation, the K107-E128 region adopt a flexible, disordered conformation, thereby allowing TS protein to recognize, and access the RNA, with the RNA apical loop in close proximity to the G84-K104 helices.

Table 3.1 Mass spectrometry analysis of trypsin digestion of TS protein (Control) and TS-crosslinked protein (Crosslinked). Red indicates peptide fragments whose respective levels are significantly compromised, likely due to crosslinking.

Peptide Sequence	Control Area	Crosslinked Area	Control/Crosslinked
AEDFQIEGYNPHPTIK	3779314.085	5194576.599	0.727549977
GVLEELLWFIK	44242140.21	310232.9881	142.6093997
DEFPLLTTK	9675430.68	12985314.98	0.745105582
YSLRDEFPLLTTK	8666479.765	8519434.519	1.017259977
IWDANGSR	3508683.7	14003466.66	0.250558221
DFLDSLGFSTR	16031000.4	21201090.1	0.756140384
EEGDLGPVYGFQWR	11907240.9	11455258.5	1.039456325
HFGAEYR	4683714	16349054.6	0.286482253
DMESDYSGQGVDQLQR	5277623.08	7664003.8	0.688624799
IIMCAWNPR	7596998.77	4175914.776	1.819241813
DLPLMALPPCHALCQFYVVNSELSCQLYQR	759375.2675	17230.06	44.07270012
SGDMGLGVPFNIASYALLTYMIAHITGLKPGD	633434.8901	2899.23	218.4838354
EPRPFPK	1656057.7	6637252.9	0.249509507
TGTGTLSVFGMQAR	10977068.89	11627560.23	0.944056076
DAEPRPPHGELQYLGQIQHILR	8813577.036	5985859.2	1.472399658
RPLPPAAQER	4055451.85	17048520.6	0.237877054

Acknowledgements

Portions of the text and Figures 3.1-3.18 are part of a manuscript submitted for publication entitled, “Molecular recognition of mRNA by human thymidylate synthase” The dissertation author is the primary author, co-authors include Sergey Dibrov, and Melody B. Kao. Thomas Hermann was the corresponding author of this article. The UCSD Bimolecular/Proteomics Mass Spectrometry Facility and Majid Ghassemian was responsible for the mass spectrometry analysis.

Chapter 4. Targeting Translation in the HCV IRES

Introduction

In this work, we investigate small molecule and peptide inhibitors of the Hepatitis C Virus internal ribosome entry site (HCV IRES) and explored their effects on translation as determined by an *in vitro* transcription-translation (IVT) assay.

The hepatitis C virus is a positive-sense single strand RNA virus from the flaviviridae family. The virus consists of a lipid enveloped coat protein containing its genetic material, making it difficult to detect by the immune system (11) HCV infection affects over 200 million people worldwide, and is the primary cause of liver failure and hepatocellular carcinoma (12). Prior to the recently approved HCV protease (NS3) inhibitors boceprevir and telaprevir, therapy consisted of extensive rounds of pegylated interferon-α and ribavirin (13). However, due to the limited efficacy of these drugs, as well as the prevalence of drug resistant mutations in HCV quasispecies, there remains an urgent need for additional antiviral agents with unique HCV targets.

One target that presents itself is the HCV IRES, located in the 5' untranslated region of the HCV RNA genome. The HCV IRES is a highly conserved sequence (14), in contrast to the remainder of its genome. This is likely due to the extensive structural and folding requirements necessary for the HCV IRES to undergo cap independent translation (15). While cap dependent translation involves multiple steps, and numerous cellular co-factors, IRES driven translation has only minimal co-factor requirements (11). In the case of the HCV IRES, the RNA structure bypasses the need for these co-factors by directly recruiting the 40S ribosomal subunit to the start codon (16). However, the strengths of these cap independent translational pathways, their simplicity and structurally driven functions, is also

their biggest vulnerability. A small molecule or peptide that disrupts the structure of the IRES RNA could significantly impact, or abolish IRES driven translation.

The functions of the IRES are achieved through several independently folding RNA domains (17, 18). HCV IRES domains II, III, and IV have all been shown to be critical in ribosome recruitment (Figure 4.1), elF3 binding, and initiation codon placement

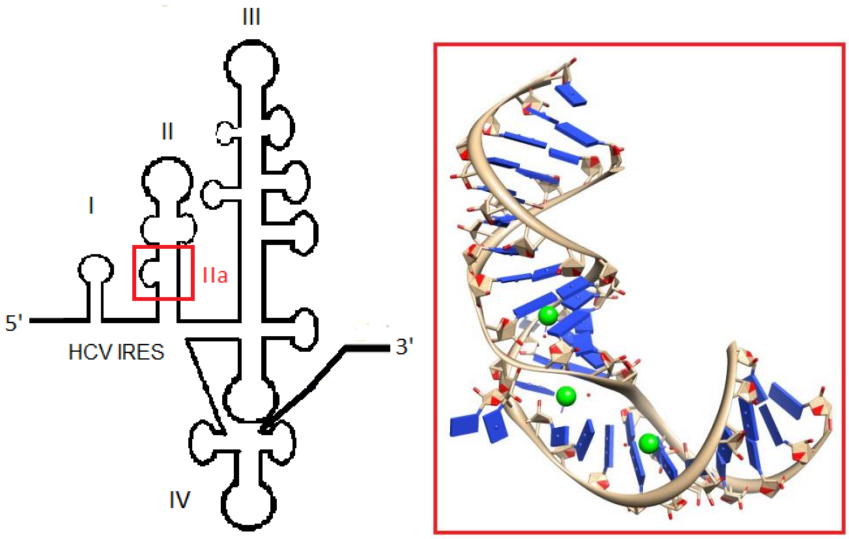


Figure 4.1 Secondary structure of the HCV IRES, and the crystal structure of subdomain IIa. PDB-ID 2PN4.

(19, 20). This work focuses on measuring HCV IRES translation as a whole, quantified through an *in vitro* transcription translation assay (IVT), previously discussed in Chapter 1.

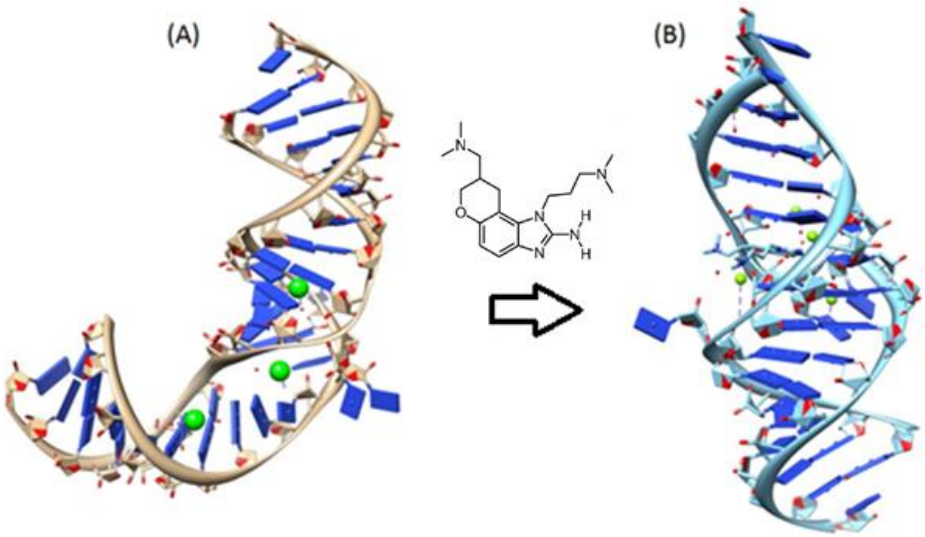


Figure 4.2 HCV IRES Subdomain IIa crystallized in either the absence (A) or presence (B) of a benzimidazole translation inhibitor. PDB-ID 2PN4 (left) 3TZR (right).

Compound Testing

The small molecules tested in the IVT assay are biased towards a benzimidazole scaffold previously demonstrated to have low micromolar binding affinity for the subdomain IIa (21). Additionally, we have demonstrated that a benzimidazole translation inhibitor captures, and stabilizes the IRES domain IIa in a translation-compromised conformation (Figure 4.2).

In order to obtain a positive control for a known HCV inhibitor in the context of the IVT assay, translational inhibition by the previously characterized benzimidazole was tested (Figure 4.3). The DNA construct used for the IVT assay was the ΔTS1 bicistronic reporter, previously used in Chapter 2 (Figure 1.5) with a cap driven firefly luciferase now acting as the positive control, and an internal HCV IRES driven renilla luciferase acting as the experimental reporter. An exploratory set of approximately 100 RNA friendly synthetic

small molecules and peptides were tested in the IVT assay in 96-well format (Figure 4.4) at both 10μM and 100μM.

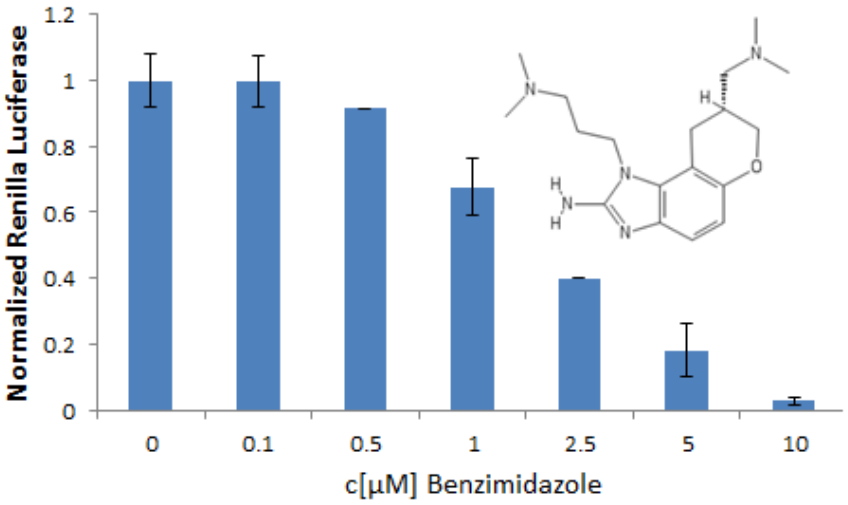


Figure 4.3 Dose response IVT assay on benzimidazole IRES inhibitor. Error bars correspond to 1σ. Results are the average of samples tested in triplicate.

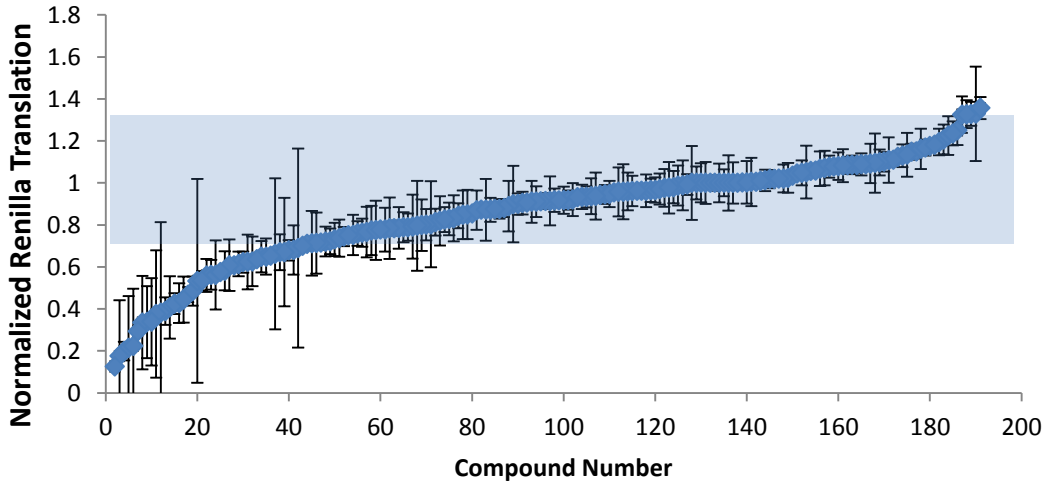


Figure 4.4 Compound screen. Tested compounds are arranged by inhibition potency. Values are triplicate averages of normalized renilla luciferase translation. The shaded rectangle indicated the +-3σ error bar for the respective controls.

Approximately 60 of the compounds decreased the renilla luciferase enough to warrant further analysis. 51 compounds were eliminated from the results based on non-specific translation inhibition, as determined in the IVT assay using the monocistronic cap driven firefly luciferase reporter construct. This is likely due to nonspecific binding to ribosomes. Two other compounds inhibited the renilla luciferase enzyme. The remaining compound scaffolds (Figure 4.5) exhibited significant reduction in HCV IRES driven translation specifically, with minimal obstruction to cap driven translation as determined by a monocistronic cap driven firefly luciferase reporter.

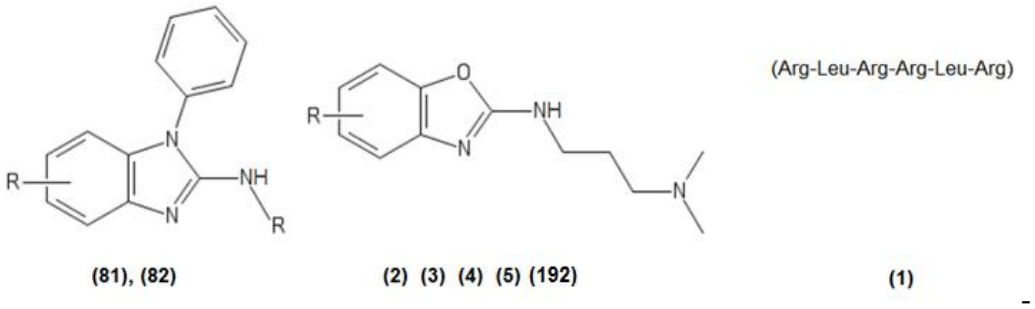


Figure 4.5 Result of IVT assay and subsequent control screen of synthetic compounds (81, 82, 2, 3, 4, and 5) and peptide (1) demonstrate that the above compound scaffolds and peptides are active hits against HCV IRES driven translation, as well as presenting favorable non-specific inhibition profiles.

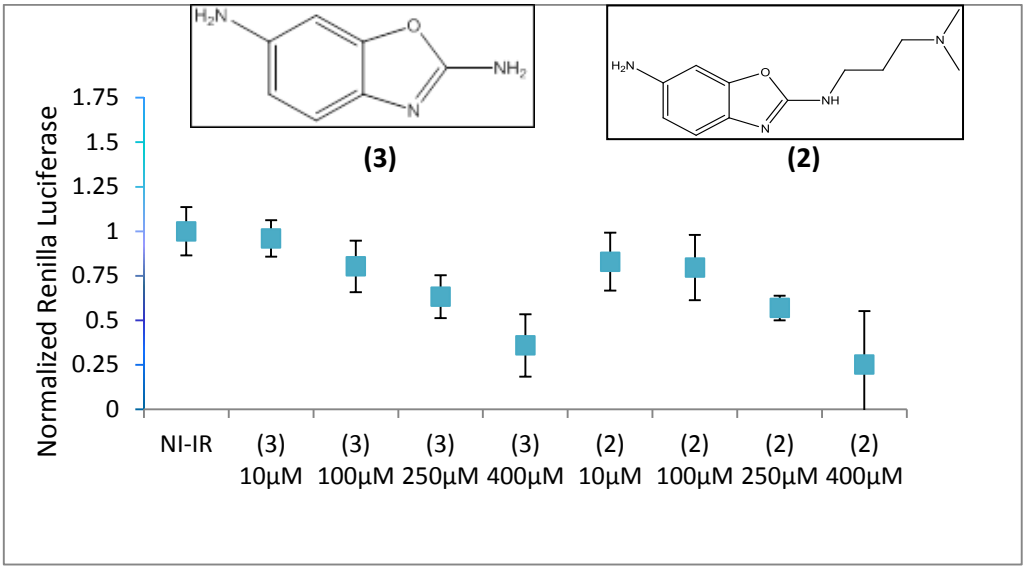


Figure 4.6 Dose response curves for compounds (2) and (3) in in the IVT assay. NI-IR indicates the negative control. Values are triplicate averages of normalized renilla luciferase translation. The error bars correlate to 1σ.

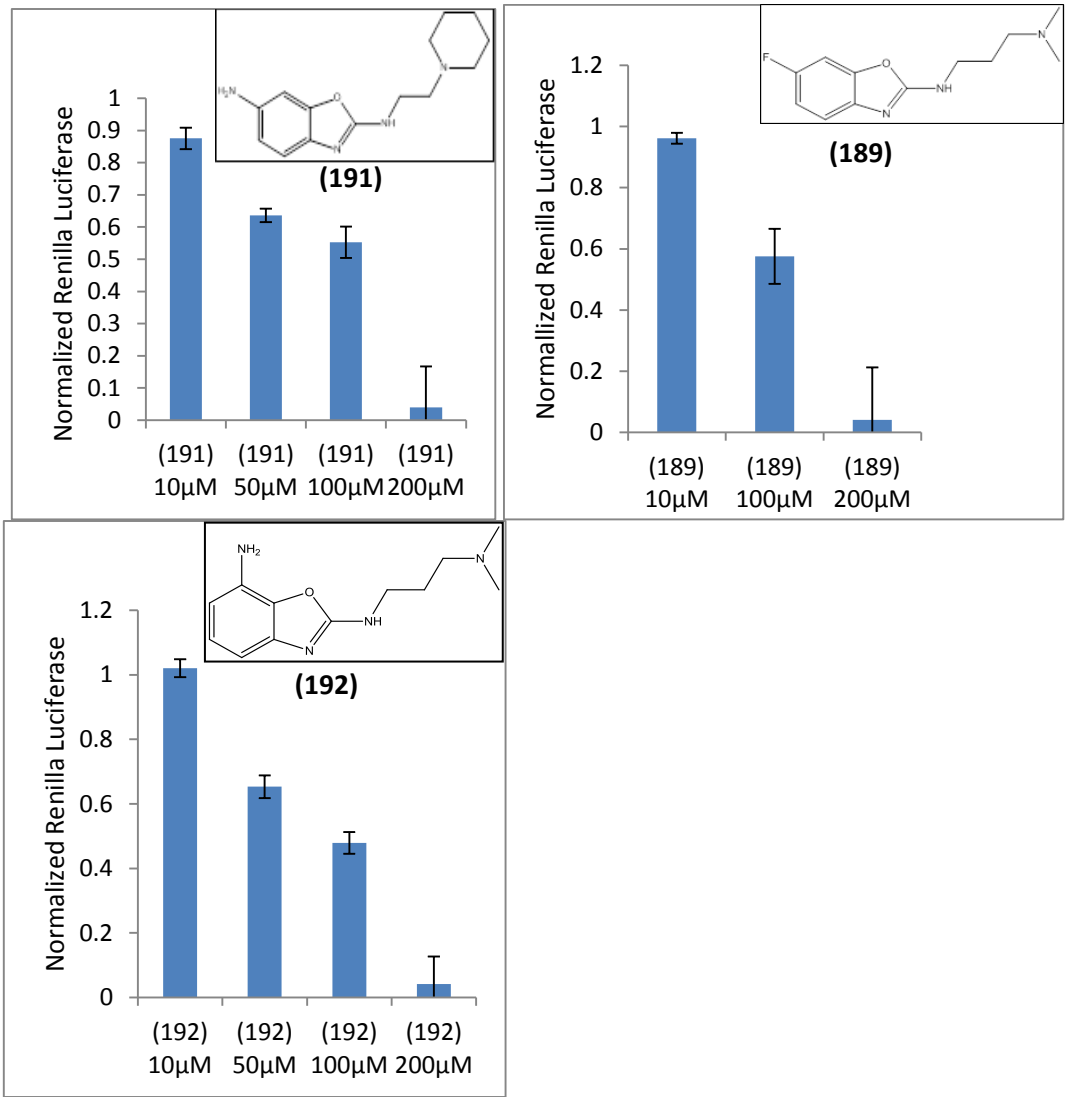
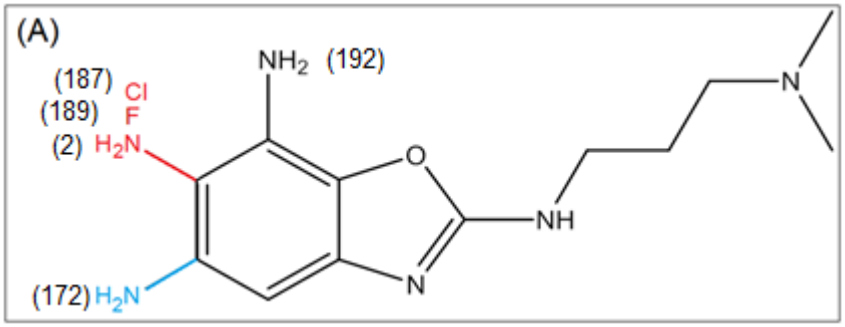


Figure 4.7 Dose response curves of compounds (189) (A), (191) (B) and (192) (C) from the scaffold class of the previous IRES hits (2), (3), (4), and (5) in IVT assay. Values are triplicate averages of normalized renilla luciferase translation. The error bars correlate to 1σ.



(B) Chemical Group	IVT Assay EC_{50}
Ortho-NH2	98 µM
Meta-NH2	270 µM
Para-NH2	90 µM
Meta-F	106 µM
Meta-Cl	NA

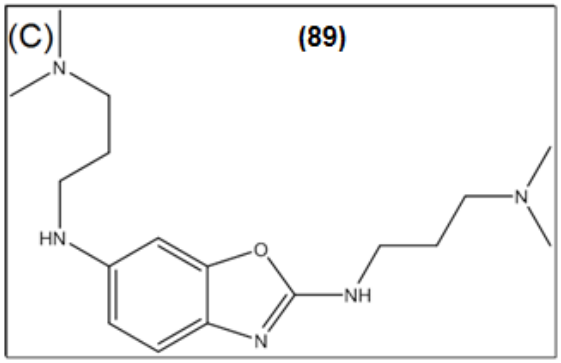


Figure 4.8 Structure activity relationship analysis of a small subset of benzoxazoles. The relative structures of the benzoxazoles and the key primary amine are shown (A) with their respective substituents color coordinated to their position. Black (ortho-substituted), red (meta-substituted) and blue (para-substituted). The respective activities for these compounds is shown in table (B). Compound 89 is shown (C) with both sides of the molecule substituted with dimethylpropyl-amine groups.

Structure Activity Relationship Analysis

In order to optimize the affinity for the benzimidazole and benzoxazole ligands, we performed a closer examination of the effects that particular chemical groups exerted on the HCV IRES, as monitored through the IVT assay. A small subset of similar benzoxazole compounds were analyzed in their respective IRES inhibition potential (Figure 4.8). For this examination, the benzoxazole with the right side substituted with the dimethylpropane-1,3-diamine had its single primary amine location moved from either the ortho, meta, or para position (Figure 4.8 A). In addition, the primary amine was then replaced with either a fluorine, or a chlorine atom.

The results of this examination highlight the activity of several important functional groups. While the ortho-amine and para-amine compounds behaved quite similarly, the meta-amine saw its EC_{50} decrease 3 fold. Due to the rotational freedom of the benzoxazole core around the secondary amine, this could result in the ortho and para-amine benzoxazoles behaving very similarly in the binding site, and therefore, having similar EC_{50} values. Whereas the meta-amine may not be optimally positioned to interact with the IRES. Curiously, we discovered that when fluorine was substituted for the meta-amine, this resulted in a rescue of the previously diminished binding activity. When a chlorine was substituted at this position, it was unable to reproduce this same effect, possibly indicating the chemical group size limitations on the left side of the benzoxazole. This is supported by the EC_{50} value of compound 89 (Figure 4.8 C) which has the meta position substituted with a second dimethylpropane-1,3-diamine group, and which displays minimal IRES inhibition.

Insights into IRES driven translation

Using the previously mentioned benzimidazole as our "optimal ligand" we set out to probe the mechanism of domain IIa mediated IRES inhibition. In the initial studies of IRES inhibiting compounds, it was apparent that any "active" compound not only decreased IRES driven translation, but cap driven translation as well (Table 2). This was also the case for the benzimidazole inhibitor (Figure 4.9), which proved curious because previous studies had revealed only minimal cellular toxicity of this compound (22). Based on our observation that cap driven translation decreased simultaneously with IRES driven renilla luciferase, we hypothesize that ligand binding of the subdomain IIa by the benzimidazole, stabilizes the IRES domain II, thereby trapping it in a translation-stalled complex on the ribosome. Any upstream translation in a bicistronic construct would then be ablated due to the "ribosome

sink" created by the presence of ribosome binding-capable, but translation-incapable HCV IRES. To test this hypothesis, we designed a chimeric bicistronic construct which lacked the domain II on the IRES driven renilla, but contained a fully functional cap driven firefly luciferase (Figure 4.10).

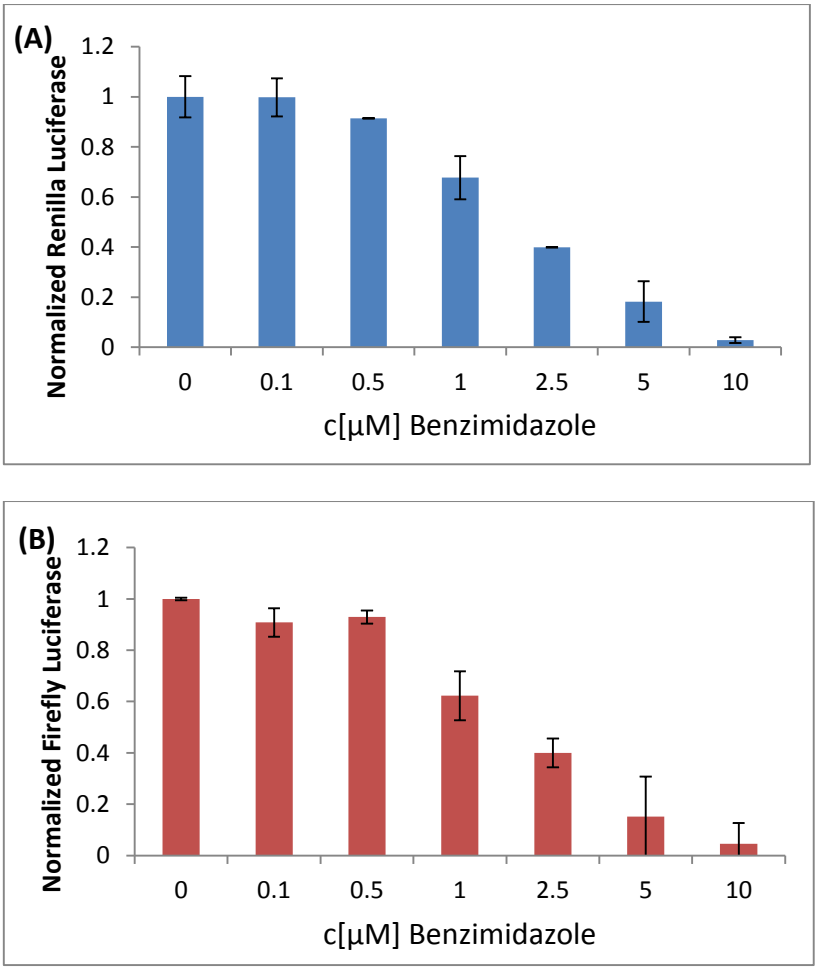


Figure 4.9 Dose response of benzimidazole in bicistronic reporter construct and respective levels of translation as measured by either renilla luciferase (A) or firefly luciferase (B). Values are triplicate averages of normalized renilla luciferase translation. The error bars correlate to 1σ.

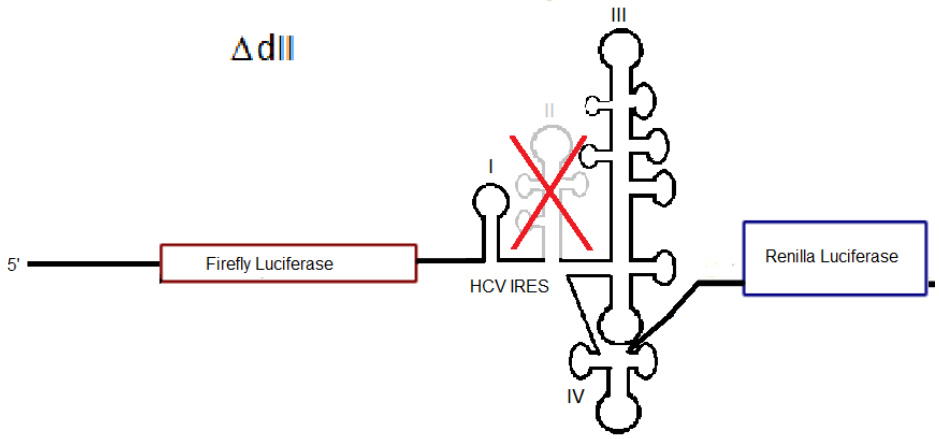


Figure 4.10 Bicistronic construct ΔdII containing a cap driven firefly luciferase, and an internal HCV ΔdII IRES driven renilla luciferase.

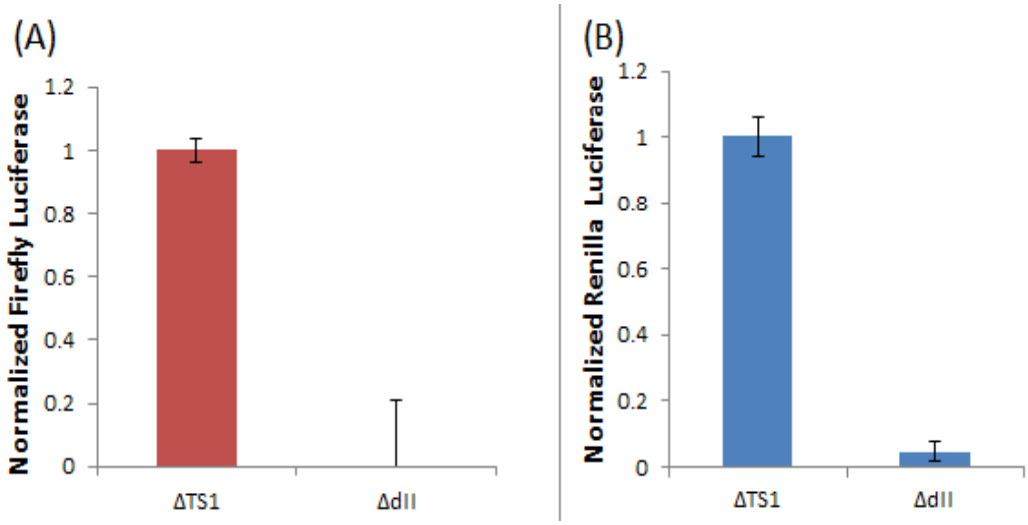


Figure 4.11 Normalized translation levels of firefly luciferase (A) and renilla luciferase (B) in bicistronic constructs with either wild type domain II (ΔTS1) or domain II removed mutant (ΔdII). Values are triplicate averages of normalized renilla luciferase translation. The error bars correlate to 1σ.

Due to the elimination of the HCV IRES domain II, the translation of the HCV IRES driven renilla luciferase was essentially eliminated. Interestingly, the upstream cap driven translation was also diminished (Figure 4.11). These findings support our hypothesis that without a functional domain II, the HCV IRES is unable to release from the ribosome. This likely results in stalled IRES complexes sequestering ribosomes thereby ablating cap driven translation through ribosome depletion from the assay mixture.

.

Conclusions

The *in vitro* transcription translation assay provides us a very useful tool in the analysis of IRES driven translation, and ways to inhibit its proper function. Based on our mutational analysis, the presence of a potent inhibitor targeting the domain IIa creates a stalled IRES complex that is unable to release from the ribosomes. This leads us to believe that the benzimidazole inhibitor acts by blocking the pathway of HCV IRES releasing from the ribosome to continue protein synthesis. Based on results from screening small molecule candidates, the substituted benzimidazole and the benzoxazole scaffolds offer promise, and further potential for further IRES domain II inhibitor development.

Chapter 5. Crystallography

Introduction

To truly understand the interactions of the aforementioned targets in the thymidylate synthase study, namely the TS1 RNA and TS protein, attempts were made to crystallize the protein bound to the RNA, as well as the RNA alone. This presents a challenge, as the crystallization conditions for the protein alone are already quite well understood. But based on the crystal packing observed in the crystal structures (1, 3), co-crystallizing a large RNA hairpin in those conditions was unlikely to be successful. Current TS protein structures demonstrate that the protein is dynamic in solution. It populates two distinct conformational states, switching between a catalytically "active" and "inactive" state by undergoing a rearrangement of the- loop consisting of residues 181-197 (1, 3, 23). The amino terminus of the protein is proline rich and disordered in all crystal structures. It has been shown to play an important role in substrate binding, likely indicating that it also plays a role in RNA binding and recognition (23). In the presence of enzymatic substrates, the TS protein prefers to adopt a "catalytically active" conformation (3). In addition, mutational analysis has shown that if the "inactive" state is destabilized, an increase in enzymatic activity of ~33% is observed, suggesting that at equilibrium, TS protein exists as one-third "inactive" state and two-thirds "active" state enzyme (24). To our knowledge, there were no previous crystal structures of full length, ligand free TS in the active conformation.

Protein-RNA Complex

To overcome the challenges in obtaining a crystal structure of a TS protein-RNA complex, several approaches were taken. The TS protein would have to be purified to be

RNase free, the length of the RNA would be varied, and the “inactive” stabilized mutant M190K (23) would also be utilized in the crystal screens. Additionally, the RNA was either set in the drops, or soaked over the course of an hour, into previously grown TS crystals. When set in the drop with the protein, the RNA was added in 4 fold excess.

The diffraction quality of the crystals was checked on our in-house x-ray source, and if they were of sufficient quality, a data set was collected. Data were processed using HKL-2000, and a structure solution was obtained by molecular replacement using previously solved TS structures (Table 1). Once the structure was sufficiently refined, the electron density between the respective asymmetric units was examined for potential density corresponding to an RNA fragment. Our studies were unable to locate any ordered electron density outside the TS protein, which were indicative of bound RNA, and any additional density was likely either due to noise, or the disordered amino terminus of the TS protein.

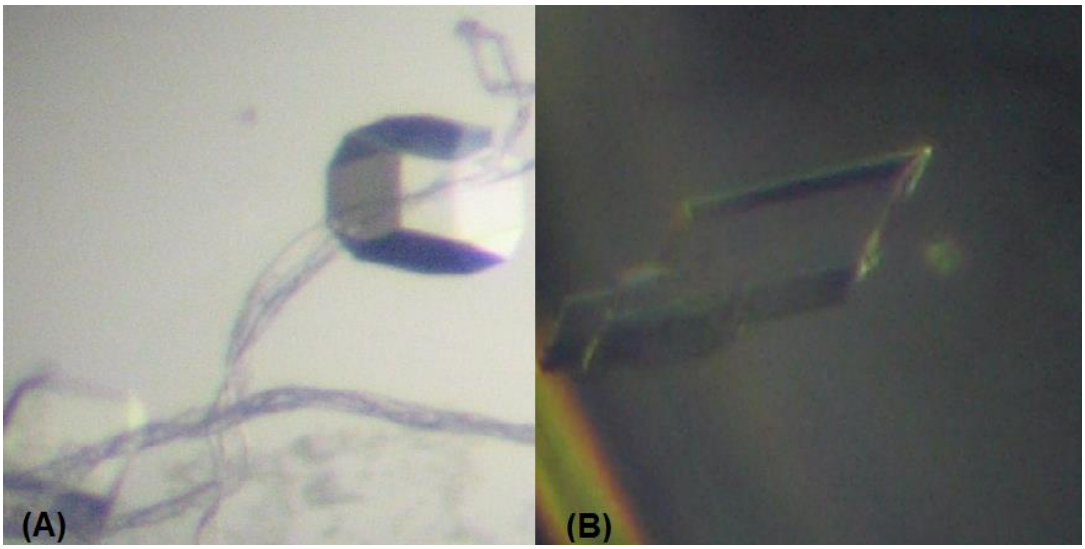


Figure 5.1 TS protein crystallized in the presence of 4 fold excess 33mer TS1 RNA in high salt (A) and low salt (B) conditions.

As illustrated in Figure 5.1, the thymidylate synthase enzyme crystallizes into a cubic morphology in high salt conditions, and distorted hexagonal prism in low salt conditions. In both salt conditions, the crystals are sensitive to additives, including RNA. In the absence of RNA, the crystal morphology of the low salt conditions changes into triangular plates, whereas the high salt conditions they adopt a triangular pyramid morphology (Figure 5.2).

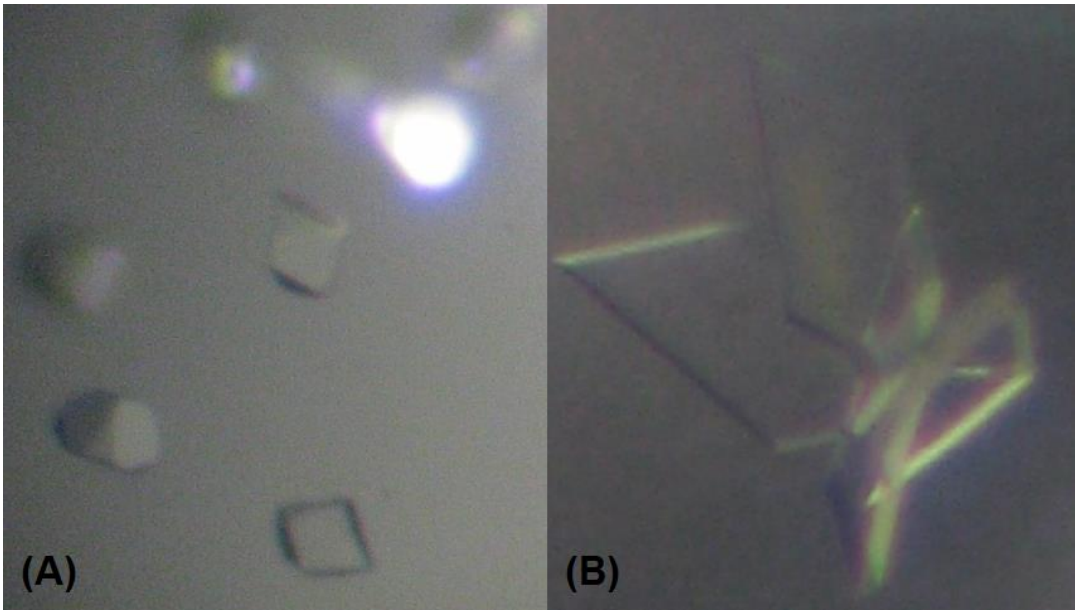


Figure 5.2 TS protein crystalized in the absence of additives or RNA in high salt (A) and low salt (B) conditions.

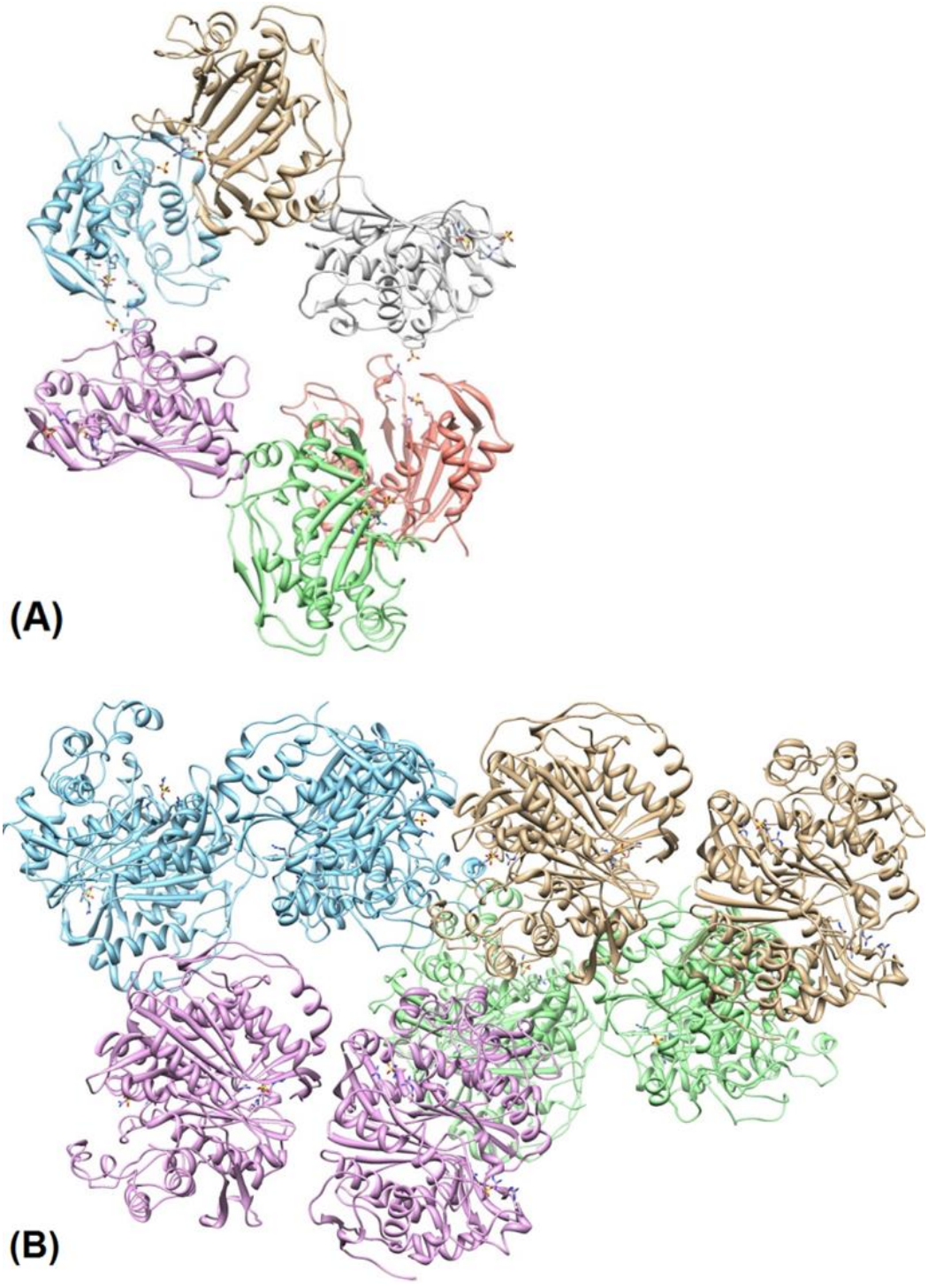


Figure 5.3 Unit cells of high salt (A) and low salt (B) crystal structures exemplifying the crystal packing of the high salt (monomers, inactive) and the low salt conditions (dimer of dimers, active).

After collecting and solving data sets for the respective crystals in Figure 5.1, we can observe that the TS protein crystalizes in the same conformation in the presence of RNA, and that the RNA is not detected in either of the crystal structures (Figure 5.3). After interpreting the data in Chapter 4, this should not be too surprising. Based on the observation that the TS site 1 RNA prefers to bind around the active site, and that this likely precludes substrate binding, than it suggest this interaction was occurring near the sulfate ions present in the crystal structure. If a large 18-20Å nucleic acid were present in that region, this would likely disrupt the crystal packing observed in both the high salt structure, as well as the low salt structure. The crystal packing of the high salt (inactive) structure seems to offer the most space to accommodate a RNA while maintaining its original packing, in addition to being in the conformation most likely to bind to the RNA. That being said, it is unlikely that this approach will succeed for two reasons. First, based on the size-specificity of the RNA-protein interaction, if a short RNA were used, it is likely to populate several positions on the protein, and thereby disrupt the crystal packing. Whereas if a longer and more specific RNA were used, its size would likely be sufficient to disrupt the crystal packing. Secondly, as we can observe from the two separate crystal structures (Figure 5.4), the protein is quite dynamic in solution, and it is unknown if the RNA bound protein exists in a transitory conformation that has yet to be observed.

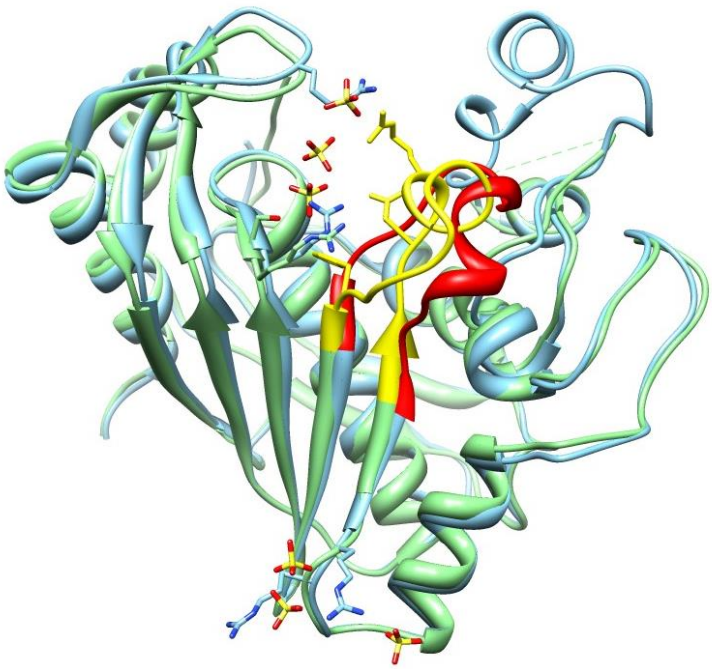


Figure 5.4 Overlay of the TS protein in high salt conditions (green) and low salt conditions (blue). The most notable change in the structures comes from the loop 180-197 (active conformation in red, inactive in yellow) rearranging in order to move the active site catalytic cysteine residue out of the active site. In addition, the loop from residues 106-129 appear disordered in the high-salt (inactive) structure, but ordered in the low-salt (active) structure.

RNA Crystallography

In order to better understand the physiology of the TS site 1 RNA and its interaction with the TS protein, attempts were made to crystalize the terminal loop. Difficulties arise when working with GC rich RNA hairpin loops, as is demonstrated by the fact that a 20mer TS RNA crystallizes as a dimer, due to the sequence similarities of the 5' and 3'ends (25). To circumvent the problem of dimerization, the constructs used for crystallography were designed to be longer than 30 nucleotides, so as to provide sufficient sequence dissimilarities. Obtaining milligram quantities of RNA becomes cost prohibitive if constructed my chemical synthesis, therefore T7 RNA polymerase runoff transcription was utilized for all the RNA constructs.

Initial attempts at RNA crystallography used RNA transcription reactions with RNA cleaving ribozymes located at the 3' terminus of the RNA of interest (Figure 5.5). The

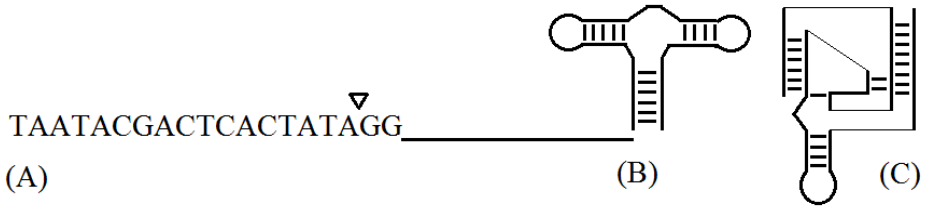


Figure 5.5 T7 RNA transcription scheme, the T7 promoter sequence (A) is followed by the sequence of interest, and flanked by either the hammerhead ribozyme (B) or the HDV ribozyme (C).

hepatitis D virus (HDV) ribozyme or the hammerhead ribozyme were used based on sequence constraints. The use of ribozymes enhances the 3' homogeneity of the RNA, due to the nature of the T7 RNA polymerase to add additional nucleotides at the end of the 3' sequence (26, 27). The drawback of using this approach is that the yields of RNA suffer, as additional purification steps are required in order to isolate the RNA species of interest from the unincorporated nucleotides, template, cleaved ribozyme, and un-cleaved product. This would generally lead to an approximate 10 fold reduction in yield.

Upon the discovery that this particular RNA was not amenable to crystallography, -a slightly different approach was undertaken, wherein the ribozyme was removed from the DNA template and runoff transcription was used (Figure 5.6). The RNA would have heterogonous 3' ends, but would not require the additional purification steps, and could be simply isolated from the transcription reaction. In addition to the removal of the ribozyme, the binding site for the RNA binding protein U1A was added at the 3' end. This particular RNA binding protein has been shown to increase the resolution of previously diffracting crystals, as well as cause previously un-crystalizing RNAs to crystalize (28).

After generating sufficient quantity of a particular RNA, it was subjected to a series of sparse matrix crystallization screens. Hits or promising conditions would be examined further, and if a particular RNA appeared to be amenable to crystallization, then it was re-transcribed with homogenous 3' ends in order to increase the odds of orderly crystal packing.

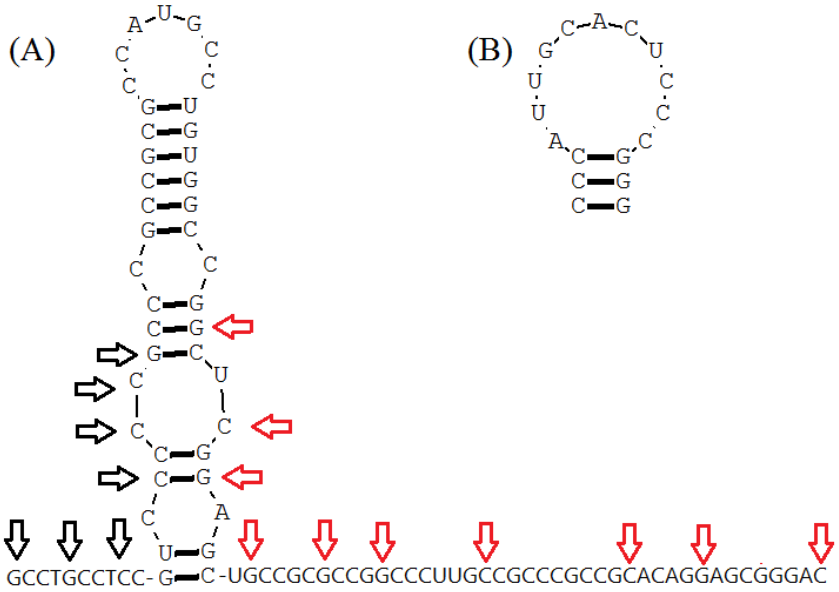


Figure 5.6 Secondary structure of the TS site 1 RNA (A) and the U1A RNA hairpin (B) with the respective 5' ends (black arrows) and 3' ends (red arrows) used in the design of crystallization constructs.

Unfortunately neither of these methods have yet yielded useful crystals. However this approach might still be promising for future crystallization trials, in addition to efforts using constructs carrying the well characterized GAAA tetraloop to promote intermolecular RNA interactions..

Conclusions

We have characterized a systematic approach to RNA crystallography of the TS site 1 RNA, as well as offered insight into RNA-protein crystallography with the thymidylate

synthase that may help develop future crystallization screens of these two targets. While crystals occasionally formed in our screens, they have been either too small, or non-diffracting, suggesting that some combination of approaches may yield usable RNA crystals. Additionally, we successfully harnessed the TS1 RNA to act as an additive in our protein-crystallography screens, allowing us to determine a previously disordered region of the TS protein.

Chapter 6. Materials, Methods, and Supplemental Figures

Materials and methods

TS protein expression and purification – A DNA construct coding for a codon optimized C-terminal hexahistidine tagged human TS (huTS) was purchased from GenScript (Piscataway, NJ) and inserted into the pET22b+ plasmid (Invitrogen, Carlsbad, CA) which was transformed into Neb5α competent *e*-coli cells. The T7 Express E.coli (Neb) bacteria expressing the hexahistidine tagged huTS protein was grown in Luria Broth (LB) at 37 °C until it reached an OD_{600} of 0.6. At this point, Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5mM. The induction period lasted 8 hours at 37 °C, at which point the bacteria were isolated by centrifugation at 10,000*g for 15 minutes at 4 °C. After resuspension in 1X phosphate buffered saline (PBS) pH 7.5 plus 2mM imidazole and 1mM DTT (1ml buffer per 10ml bacteria), the cells were homogenized using sonic disruption (Fisher Model 505). The enzyme was then loaded on a 1ml bed volume of cOmplete His-Tag Purification Resin (Roche). The crude extract was subjected to two 30 column-volume washes of PBS plus 1mM DTT, 2mM imidazole and 5mM imidazole respectively. An additional wash step utilizing 10 column volumes of PBS plus 1% 3-[(3-Cholamidopropyl)-dimethylammonio] 1-propane sulfonate (CHAPS) buffer was performed prior to elution to remove any bacterial TS from the solution. The huTS was eluted on a linear gradient of 5mM-50mM imidazole. Fractions were pooled, concentrated, and dialyzed into 20mM tris(hydroxymethyl)aminomethane (TRIS) pH 8, 1mM DTT.

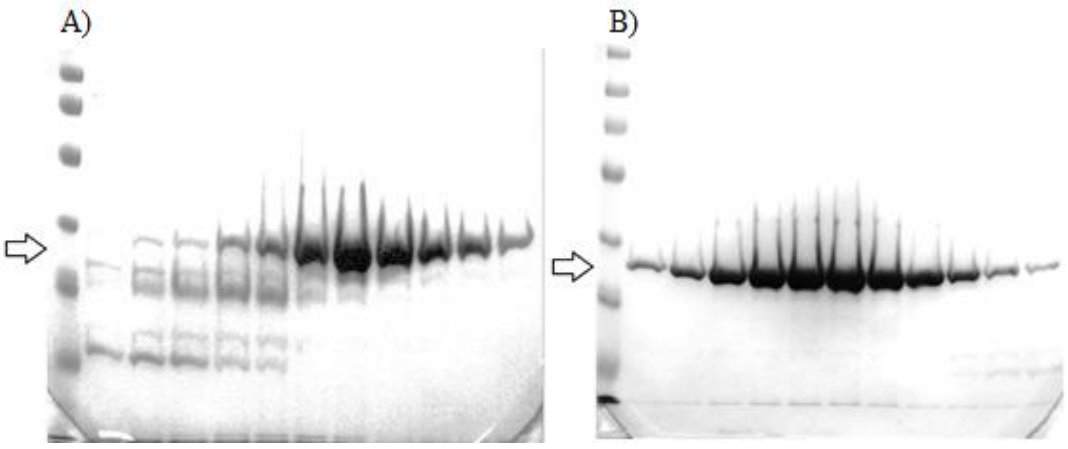


Figure 6.1 SDS-PAGE gel of Ni^{2+} affinity purification of e-coli expressed hTS protein (A) eluted on a linear gradient (10mM-200mM) imidazole. Pooled and concentrated fractions were collected and run on a 100ml column containing sephacryl s-100. Fractions containing protein were analyzed on a SDS-PAGE gel (B).

Construction of the bicistronic reporter constructs – Cloning of all bicistronic reporter constructs (Fig. 2A) followed identical procedures with the exception of ΔTS1. Oligonucleotides for the cloning are outlined in Table 6.2. All TS1 containing constructs had oligonucleotides (sense/antisense) ligated into the pSP-luc+NF plasmid (Promega, Madison, WI). The resulting TS site 1 hairpin+firefly luciferase sequence was then digested in its entirety and ligated together with equimolar amounts of HCV IRES sequence (see below), and pRL-null vector. This three-component ligation mixture was transformed into Neb5α cells, plated on LB-Amp medium and incubated at 37º C overnight. Individual clones were sequenced to confirm correct insertion.

Oligos encoding TS1 and TS-mt 1-4 were phosphorylated, annealed at 65°C, then ligated into BstEII and HinDIII digested pSP-luc+NF plasmid. Once sequenced, the constructs were digested with NdeI and EcoRI to extract the TS1+Firefly Luciferase, gel purified, then ligated together with the EcoIR-BstBI digested HCV IRES for 30 minutes. Then 250ng of NdeI-BstBI digested pRL-null was added to the ligation mixture which was

incubated for another 30 minutes. The resulting product is (5'-3') TS Hairpin+Firefly Luciferase→44bp→HCV IRES→Renilla luciferase.

The ΔTS1 construct had the NdeI-EcoRI restriction product (firefly luciferase) ligated with the IRES/pRL-null mixture without undergoing any insertions in the 5' end. The resulting product is (5'-3') Firefly luciferase→44bp→HCV IRES→ Renilla luciferase

Construction of the HCV IRES component – Oligonucleotides for the cloning of the HCV IRES are outlined in Table 6.3. Synthetic oligonucleotides were phosphorylated, annealed pairwise (IR-nT with IR-nB), and then ligated together. To increase the yield of full length IRES construct, a polymerase chain reaction (PCR) was conducted on the resulting ligation mix. Primers used for the PCR reaction were designed to amplify the IRES region, while retaining the original EcoRI and BstBI restriction sites (Table 6.2). The PCR product was purified with a QIAquick PCR Purification Kit, digested with EcoRI and BstBI. After further purification the IRES was saved for ligation into the TS constructs.

In vitro translation assay and luciferase detection – The in vitro transcription-translation assay (IVT) was conducted according to manufacturer recommendations (Promega). The reaction volume was 7.5µl. Briefly, each 7.5µl reaction contained 1.5µl of 100ng/ul plasmid DNA, 1.5µl of either H2O or compound solution, and 5µl of reaction buffer containing reticulocyte lysate, SP6 polymerase, and RNase inhibitor. Incubation was at 30º C.

Detection of firefly and Renilla luciferase levels was done using the Dual-Glo Luciferase Assay System (Promega) according to manufactures recommendations. Briefly, 5µl of the reaction mixture was mixed with 25µl of Dual-Glo luciferase substrate on a 96 well plate. After incubation for 60 seconds at RT the firefly luminescence was measured on

a SpectraMax Gemini XS by Molecular Devices (Sunnyvale, CA). Then, 25µl of the Dual-Glo Stop & Glo substrate was added to each well, incubated for 60 seconds, and the luminescence of the Renilla luciferase was quantified.

IVT time course – Time course experiments were conducted as per the protocol above, with the variation of having 9 triplicate samples. Samples were removed from 30º C and placed immediately on ice at each time point (0, 5, 10, 20, 30, 45, 60, 90, 120 minutes) for luciferase quantification. Results of the time course experiment are shown in Figure 7.2. A time point of 45 minutes was chosen for the following experiments to ensure that translation activity was recorded before reaching saturation through substrate depletion.

IVT TS enzyme titration – Previously frozen TS enzyme was diluted in buffer to a 5-fold concentrated sample (i.e. 5, 25, and 125µM), and 1.5µl was added to the transcription-translation reaction in order to achieve a final concentration of 1, 5, or 25µM. No other conditions were modified.

IVT Compound screening – 1.5ul aqueous compound stock was added to the assay mixture to a final concentration of 10µM or 100µM, respectively unless otherwise indicated. Controls were H2O or 1% DMSO.

IVT Construction of truncation mutants – The TS1 DNA bicistronic construct was used as the template for a series of QuickChange Mutagenesis reactions, whereby the roughly 42 nucleotide TS1 terminal loop was systematically truncated utilizing the QuickChange protocol according to manufacture protocols (Agilent). Briefly, oligonucleotide primers containing the desired TS hairpin were synthesized (IVT) and contained 12 complementary

nucleotides upstream and downstream of the mutation site. The non-complementary sequences inserted into the respective constructs are as follows;

TS1 23

(Sense) 5′-3′ GCCGCGCCATGCCTGTGGCCGTC

(Antisense) 5′-3′ GACGGCCACAGGCATGGCGCGGC

TS1 20

(Sense) 5′-3′ CGCGCCATGCCTGUGGCCGTC

(Antisense) 5′-3′ GACGGCCACAGGCATGGCGCG

TS1 15

(Sense) 5′-3′ CGCCATGCCUGTGGTC

(Antisense) 5′-3′ GACCACAGGCATGGCG

These QuickChange primers were then utilized per the manufactures protocol, and clones were sequenced to determine correct insertion.

IVT Truncation mutant screening – Previously frozen TS enzyme was thawed at room temperature, and diluted in 1x PBS buffer to a 25μM stock solution. Then 1.5μl of the enzyme was added to a 7.5 μl transcription-translation reaction, utilizing the various TS1 truncation mutant DNA constructs. The negative controls were simply the TS1 truncation mutant DNA constructs with 1.5μl of dilution buffer added, rather than the TS enzyme. No other conditions were modified.

TS immunodepletion – To deplete the TNT extract of cellular TS, mouse antibody (MAb-TS(106) at 200μg/ml; Thermo Fisher, Waltham, MA) was mixed with the extract at a 1/250

dilution. This mixture was gently agitated at 4º C for 45 minutes, at which point Red agarose conjugated goat-anti-mouse antibody was mixed 1/250 with the antibody/reticulocyte lysate mixture and allowed to gently agitate for 30 minutes at 4ºC. The agarose conjugated complex was removed from the mixture by centrifugation at 1000x g for 1 minute.

TS enzyme assay – The TS enzyme activity assay was conducted as described in the literature (58). Briefly, the enzymatic conversion of 5,10-methylenetetrahydrofolate (5,10-mTHF) to dihydrofolate results in an increased absorbance at 340nm. The reaction conditions were 50mM TRIS pH 7.0, 125mM KCl, 5mM dUMP, and 5mM 5,10-mTHF. TS protein was added to a final concentration of 5µM. Enzymatic reactions were carried out at 37º C for 15 minutes and quenched by the addition of 0.36N HCl. Absorbance was measured on a Shimadzu UV-2401PC spectrophotometer.

RNA free energy calculations – Free energies of secondary structures for the site 1 motif RNA and the stabilized mutants mt1-mt4 for preparation of the graph in Fig. 4 were calculated using RNAfold in the ViennaRNA package (59).

Peptide RNA binding gel shift assay – RNA representing the TS site 1 sequence (nt C80 – G112, Figure 1.5) was obtained as HPLC-purified oligonucleotide from IDT (Coralville, IA). Peptides were purchased from Genscript and dissolved in water to a final concentration of 1-5mM. Peptide concentrations were determined by the BCA Protein Assay Kit from Thermo Scientific (Waltham, MA). Individual reaction mixtures consisted of 5µl of 2x reaction buffer (12% glycerol, 25mM TRIS pH 7.0, 5mM DTT, 2mM MgCl2), 50µM of RNA, an appropriate aliquot of the desired peptide, and brought to a final volume of 10µl

with dH2O. Mixtures were allowed to incubate at room temperature for 5 minutes, then placed on ice for another 5 minutes.

Native polyacrylamide gels (6%) were prepared using 1.5ml of 40% acrylamide bis-acrylamide solution, (19:1), 1µl of 10x 3-(N-morpholino)propanesulfonic acid (MOPS) buffer (200mM MOPS pH 7.0, 80mM sodium acetate) diluted with 7.3ml dH2O, and polymerized by addition of 20µl TEMED and 0.2ml 10% ammonium persulfate. The gel was pre-run in 1x MOPS running buffer (20mM MOPS pH 7.0, 8mM sodium acetate) at 4 ^{0}C for 30 minutes at 100V, then loaded with the reaction mixtures and allowed to run for another 40 minutes. Visualization of RNA was done using SYBR Gold Nucleic Acid Gel Stain (Invitrogen) according to manufacturer protocol.

RNA Transcription- RNA transcription was conducted according to the New England BioLabs protocol for *HiScribe T7 In Vitro Transcription Kit*. Briefly, 1µg per µl of PCR amplified RNA template bearing the T7 polymerase initiation site was added to the *HiScribe* reaction mixture. This was then incubated at 42 ℃ for 4 hours, at which point the RNA product was purified by phenol/chloroform extraction and isopropanol precipitation.

Protein Crosslinking- The crosslinking of the TS protein to the R33 and R14 RNAs was conducted at room temperature. The concentrations of the TS protein and the TS RNA were 200µM each in the initial crosslinking experiments, and 25µM TS protein with 10µM RNA in the time course crosslinking experiment (Figure 3.4 A and B respectively). The final buffer conditions were 10mM HEPES pH 7.0, 100mM NaCl, 2mM DTT, and 10% glycerol. The RNA-Protein complex was incubated on ice for 5 minutes prior to exposure to UV radiation of 360nm. After UV exposure of 10 minutes (Figure 3.4 A) or the designated time

points (Fig. 3.4 B), the samples were placed on ice until their subsequent loading on to either a native, or denaturing polyacrylamide gel. After electrophoretic separation of the protein-RNA complex for a desired duration (30-60 minutes), the protein was visualized using SimplyBlue SafeStain (Invitrogen) according to manufacture protocols.

Kinetic Analysis of hTS and RNA Crosslinking- The kinetic analysis of the crosslinking reaction between the TS protein to the R33 was conducted at room temperature. The concentrations of the TS protein and the TS RNA were 25μM TS protein with 10μM RNA (Figure 3.4 A and B respectively). The final buffer conditions were 10mM HEPES pH 7.0, 100mM NaCl, 2mM DTT, and 10% glycerol as in the previous crosslinking experiments. The RNA-Protein complex was incubated on ice for 5 minutes prior to exposure to UV radiation of 360nm (at room temperature). After UV exposure of 0.5, 1, 3, 10, 25, and 50 minutes, samples were placed on ice until their subsequent loading on to a native polyacrylamide gel. After electrophoretic separation of the protein-RNA complex for a desired duration (30-60 minutes), the protein was visualized using SimplyBlue SafeStain (Invitrogen) according to manufacture protocols.

Protease digestion of crosslinked hTS-RNA complexes- Mass-spectrometric analysis of peptide fragments was performed following established procedures (64). Slices of polyacrylamide gel from electrophoretic separation containing crosslinked complex, or free hTS as a control, were cut to 1 mm by 1 mm cubes and destained 3 times by first washing with 100 μl of 100 mM ammonium bicarbonate for 15 minutes, followed by addition of the same volume of acetonitrile (ACN) for 15 minutes. The supernatant was removed and

samples were dried in a speedvac. Samples were then transferred to reducing conditions by mixing with 200 µl of 100 mM ammonium bicarbonate and 10 mM DTT, and incubated at 56°C for 30 minutes. After removal of the liquid, 200 µl of 100 mM ammonium bicarbonate and 55 mM iodoacetamide was added to gel pieces and incubated at room temperature in the dark for 20 minutes. After the removal of the supernatant and one wash with 100 mM ammonium bicarbonate for 15 minutes, an equivalent volume of ACN was added to dehydrate the gel pieces. The solution was then removed and samples were dried in a speedvac. For protease digestion, enough solution of ice-cold trypsin (0.01 µg/µl) in 50 mM ammonium bicarbonate was added to cover the gel pieces and kept on ice for 30 min. After complete rehydration, the excess trypsin solution was removed, replaced with fresh 50 mM ammonium bicarbonate, and left overnight at 37°C. Peptides were extracted twice by the addition of 50 µl of 0.2% formic acid and 5 % ACN followed by vortexing at room temperature for 30 min. The supernatant was removed and saved. A total of 50 µl of 50% ACN-0.2% formic acid was added to the sample, which was vortexed again at room temperature for 30 min. The supernatant was removed and combined with the supernatant from the first extraction. The combined extraction solutions were analyzed directly by high pressure liquid chromatography (LC) coupled with tandem mass spectrometry (MS/MS) using electrospray ionization.

Mass-spectrometric analysis of crosslinked hTS-RNA complexes- Electrospray ionization of peptide fractions was performed using a TripleTof 5600 hybrid mass spectrometer (ABSCIEX) interfaced with nano-scale reversed-phase HPLC (Tempo) using a 10 cm-100 micron ID glass capillary packed with 5-µm C18 Zorbax™ beads (Agilent, CA). Peptides were eluted from the C18 column into the mass spectrometer using a linear gradient (5–60%) of ACN at a flow rate of 250 µl/min for 1h. The buffers used to create the ACN

gradient were: Buffer A, 98% H_2O, 2% ACN, 0.2% formic acid, and 0.005% trifluoroacetic acid (TFA); Buffer B, 100% ACN, 0.2% formic acid, and 0.005% TFA. MS/MS data were acquired in a data-dependent manner in which the MS1 data was acquired for 250 ms at m/z of 400 to 1250 Da and the MS/MS data was acquired from m/z of 50 to 2,000 Da. The independent data acquisition (IDA) parameters were MS1-TOF 250 milliseconds, followed by 50 MS2 events of 25 milliseconds each. The IDA criteria were, over 200 counts threshold, charge state +2-4 with 4 seconds exclusion. The collected data were analyzed using MASCOT® (Matrix Sciences) and Protein Pilot 4.5 (ABSCIEX) for peptide identification. Peakview (ABSCIEX) was used for peptide quantification analysis. For this analysis the m/z corresponding to each peptide was extracted from the MS1 chromatogram and the area under the peak was determined. Peak areas for the peptides that showed no change in their relative area were used to normalize different chromatograms analyzed in these experiments.

Protein Images- Molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311) (61).

Protein Crystallography- Crystallization was performed by hanging drop vapor diffusion with hTS protein and TS1-33 RNA both at 150 μM in the presence of either 1.75 M ammonium sulfate, 0.1 M TRIS pH 8.25, 20 mM β-mercaptoethanol ("high salt") or 30% PEG 1500 0.1M TRIS pH 8, 20mM β-mercaptoethanol and 3% 1,5-diaminopentane dihydrochloride ("low salt"). The final drop volume was 2-4 μL. Crystals appeared within 24

hours, and grew to full size within 3-5 days. Crystals were flash-frozen in liquid nitrogen before data collection. Diffraction data were collected in a nitrogen stream at 110K on a Rigaku rotating anode X-ray generator (λ = 1.54 Å) equipped with a MAR345 imaging plate detector system. Datasets collected were processed, integrated, and scaled with the HKL2000 package. Structures were solved by molecular replacement with the program Phaser using search models derived from published hTS coordinates and refined by the program Refmac both within the CCP4 package. Subsequent iterative rounds of manual building and refinement, alternating between Refmac and manual rebuilding in Coot, were based on the obtained $2F_o$-F_c and F_o-F_c maps. Positions of sulfate ions were initially assigned based on electron density as well as geometry of coordinating ligands. Final refinement was carried out in PHENIX with individual isotropic atomic displacement parameters and water picking. Coordinates and structure factors for the have been deposited in the RCSB Protein Data Bank under accession codes 4H1I (low salt structure) and 4GHY (high salt structure).

X-ray diffraction data measurement and processing- Crystallographic diffraction experiments were carried out at the Biomolecule Crystallography Facility at the University of California, San Diego. Data was collected on a Rigaku MicroMax-007HF rotating anode X-ray generator with mar345 image plate detectors from a single crystal. Data was indexed and processed with HKL2000 (62). Data collection and processing statistics are described in Table 6.1.

2AP Assay- The 2-aminopurine fluorescence assay was conducted as previously described (8). Briefly, the RNA was synthesized by Integrated DNA Technologies (Coralville, Iowa) with the 5′→3′ sequence: GCCGCGCC2ApUGCCUGUGGCCGGC.

The RNA concentration used in the assay was 5μM, compounds were titrated in their respective concentrations to yield a final reaction volume of 50μl. The final concentration of the reaction buffer was 10mM HEPES pH 7.5 and 2mM $MgCl_2$. Before the addition of the experimental compounds, each sample was analyzed for the base level of fluorescence of RNA alone. This was set as the "base level", or blank, and the respective changes in fluorescence were monitored upon the addition of the experimental compound. The plate reader used was the Spectra Max Gemini XS by Molecular Devices.

Table 6.1 Crystallographic data and refinement statistics for thymidylate synthase at high and low salt conditions.

	Low Salt	High Salt
X-ray source	Rotating Anode (Rigaku)	Rotating Anode (Rigaku)
Detector	MAR 345mm Plate	MAR 345mm Plate
Wavelength (Å)	1.5418	1.5418
Temperature (K)	173	173
No. of frames	360	360
Space group	C121	$P3_121$
Unit-cell parameters		
a (Å)	157.85	96.19
b (Å)	94.93	96.19
c (Å)	131.44	83.28
Resolution range (Å)	19.9-3.1	19.6-3
Observed Criterion Sigma (F)	1.34	1.35
Observed Criterion Sigma(I)	1.34	1.35
Number Reflections(All)	29851	10060
Number Reflections(Observed)	29732	8628
Percent Possible(Observed)	99.6	94.4
Redundancy	6.2	7.9
Refinement Statistics		
Structure Solution Method	MOLECULAR REPLACEMENT	MOLECULAR REPLACEMENT
reflnsShellList	3.095	3.005
Resolution(Low)	19.888	19.632
Cut-off Sigma(F)	1.34	1.34
Number of Reflections(all)	29851	10060
Number of Reflections(Observed)	29732	8628
Number of Reflections(R-Free)	1507	869
Percent Reflections(Observed)	99.6	94.39
R-Factor(Observed)	0.2344	0.2216
R-Work	0.2317	0.215
R-Free	0.2849	0.2802
R.m.s.d. bond lengths (Å)	0.77	1.212
R.m.s.d bond angles (°)	0.003	0.009
R-Free Selection Details	RANDOM	RANDOM
Ligands	2 sulfate ions	2 sulfate ions
Structure solution	Phaser	Phaser
Structure refinement	PHENIX	PHENIX

Table 6.2. Compound name, compound number, and relative renilla and firefly luciferase expression.

Compound Number	1	2	3	4	5	6	7	8	9	10	11	12
Compound Name	KR 203 10uM	KR 203 100uM	KR101 10uM	KR101 100uM	KR100	KR100 100uM	KR202	KR202 100uM	RRSSRR	Z1142498670	Z1142498670 100uM	Z1096981391 10uM
Normalized Cap Driven Translation	1.16260839	1.075044747	0.952369956	1.171027563	1.19	1.23609177	1.173	0.994860573	0.9364	1.179085824	1.180962446	1.250955647
Normalized IRES Driven Translation	1.32542426	0.711950271	0.98824227	0.669053614	0.91	1.079515356	0.829	1.051442106	1.0122	0.709307215	0.914025626	0.788210923
Compound Number	13	14	15	16	17	18	19	20	21	22	23	24
Compound Name	Z1096981391	Z1096977298	Z1096977298	Z1097085056 1	Z10970	Z1134133815 1	Z11341	Z1117862893	Z111786	Z1024508848	Z1024508848 100uM	Z1147614813 10uM
Normalized Cap Driven Translation	0.74053269	0.568314625	0.952153224	0.96578257	0.941	0.605865631	1.114	0.972674063	1.1333	0.798183359	0.981290878	0.9205268
Normalized IRES Driven Translation	0.90429498	0.94531687	0.789361093	1.067610662	1.023	1.034666261	1.13	1.001855876	0.8495	1.044953328	0.872227629	0.843334271
Compound Number	25	26	27	28	29	31	32	33	34	35	36	37
Compound Name	Z1147614813	Z1147634441	Z1147634441	Z1147632017 1	Z11476	Z1147641139 1	Z11476	Z1155788729	Z115578	Z1155792478	Z1155792478 100uM	Z1096982797 10uM
Normalized Cap Driven Translation	1.09278951	0.958796618	0.818933578	0.609903908	0.924	1.009442181	0.641	0.862664013	0.966	1.052778121	1.004018851	1.059990177
Normalized IRES Driven Translation	1.00443104	1.07716488	0.762781646	0.941956272	0.759	0.933095044	0.855	0.678851459	0.7092	0.899386942	0.796646547	0.937840255
Compound Number	38	39	40	41	42	43	44	0	46	47	48	49
Compound Name	Z1096982797	Z1082771278	Z1082771278	Z1095381795 1	Z10953	Z1123715484 1	Z11237	KD-II-49a 10u	KD-II-49	KD-II-85a 10u	KD-II-85a 100uM	KD-99a 10uM
Normalized Cap Driven Translation	0.94164971	0.745659605	0.279982079	0.807408322	0.812	0.875421603	0.937	0.509438654	0.8887	0.202139985	0.931055213	0.384864284
Normalized IRES Driven Translation	0.9412505	0.937275701	0.766644181	0.759612375	0.647	0.647988732	0.607	1.03312	0.8182	1.01039	0.76352	0.49478
Compound Number	50	51	52	53	54	55	56	57	58	59	60	61
Compound Name	KD-99a 100u	KD-II-122a 10	KD-II-122a 10	KD-II-123a 10	KD-II-1	KD-II-90a 10uI	KD-II-9	KD-II-101b 10	KD-II-10	KD-99a 10uM	KD-99a 100uM	KD-II-101b 100uM
Normalized Cap Driven Translation	0.96372557	0.967364552	0.568735933	0.897663812	1.328	1.002256643	0.646	1.363730842	1.3841	1.276623242	0.945027873	0.729167996
Normalized IRES Driven Translation	0.35656667	0.84893	1.404693333	1.405826667	1.189	1.15181	1.276	0.79594	0.6582	0.962068153	0.962326378	0.766000115
Compound Number	62	63	64	65	66	67	68	69	70	71	72	73
Compound Name	KD-II-113a 1	KD-II-113a 10	KD-118a 10uM	KD-118a 100u	KD-II-1	KD-II-145a 10	KD-I-8	KD-I-88a	KD-II-12	KD-II-123a	KD-125a	KD-125a
Normalized Cap Driven Translation	0.92868219	1.205477885	0.816581444	1.016638412	1.122	1.397847974	1.214	0.876610305	1.1136	0.879395092	0.961427956	0.931736862
Normalized IRES Driven Translation	1.10834564	0.785522505	1.056181234	0.81921101	0.581	0.484834208	1.326	0.703166141	0.9985	0.420352154	0.600264114	0.760563752
Compound Number	74	75	76	77	78	79	80	81	82	83	84	85
Compound Name	KD-128a	KD-128a	KD-II-126a	KD-II-126a	KD_II1	KD_II132a	KD-26a	KD-26a	KD-II-20	KD-II-20a	KD-II-25a	KD-II-25a
Normalized Cap Driven Translation	1.14632201	0.986405499	1.169831329	0.60445458	0.638	0.49501015	1.018	0.396189644	0.3969	1.052190158	0.701451531	0.709832648
Normalized IRES Driven Translation	0.9269585	0.382749664	0.594583773	0.195227285	0.529	0.625889166	0.794	0.711640833	0.2128	-0.088235	0.792884166	0.663191666

Table 6.2 Continued.

Compound Number	86	87	88	89	90	91	92	93	94	95	96	97
Compound Name	KD-23b	KD-23b	ES-275	ES-275	ES-260	ES-260	ES-261	ES-261	ES-263	ES-263	ES_ME	ES_ME
Normalized Cap Driven Translation	0.57735927	0.550780611	0.592847222	0.494717654	0.477	0.563274508	1.135	0.739805909	1.0548	0.574037771	0.972710279	0.736868982
Normalized IRES Driven Translation	0.67551	0.647245833	1.101472699	1.32542426	1.059	1.32542426	1.079	1.32542426	0.8319	1.32542426	1.177202899	1.32542426
Compound Number	98	99	100	101	102	103	104	105	106	107	108	109
Compound Name	ES-211	ES-211	ES-259	ES-259	ES-271	ES-271	JM-91	JM-91	ES-280	ES-280	ES-290S_10uM	ES-290S_100uM
Normalized Cap Driven Translation	1.08307929	0.515607509	0.680343326	0.450765074	0.693	1.088920054	0.594	0.093052308	0.8116	0.350723228	0.832231324	0.708091786
Normalized IRES Driven Translation	0.97962234	1.108946292	0.931965048	1.008450129	1.002	1.066179882	0.756	1.066971868	1.1026	0.896820973	0.933664806	0.899053137
Compound Number	110	111	112	113	114	115	116	117	118	119	120	121
Compound Name	ES-290SW_1	ES-290SW_10	ES-291_10uM	ES-291_100uM	1m_10	1m_100uM	1n_10	1n_100uM	2a_10uM	2a_100uM	2b_10uM	2b_100uM
Normalized Cap Driven Translation	0.99147244	1.014680143	0.930039942	1.029205707	0.918	0.8497344	1.1	0.977474728	0.9869	0.876491753	1.107241696	1.161002877
Normalized IRES Driven Translation	1.09267689	0.936290768	0.825651656	0.979260531	0.876	0.871846705	1.25	0.777343047	1.0182	0.970722402	1.225405752	1.235664388
Compound Number	122	123	124	125	126	127	128	129	130	131	132	133
Compound Name	2f_10uM	2f_100uM	1h_10uM	1h_100uM	1e_10	1e_100uM	1o_10	1o_100uM	2g_10uM	2g_100uM	2c_10uM	2c_100uM
Normalized Cap Driven Translation	1.02223535	1.158202623	1.01356	0.919950614	0.929	0.876332353	0.889	0.905868964	0.9549	0.856452775	0.739323437	0.7193735
Normalized IRES Driven Translation	1.13408335	0.407421898	0.62634598	0.335083498	1.163	0.381099718	0.375	0.916996462	0.8995	0.913923761	1.168349629	0.918061061
Compound Number	134	135	136	137	138	139	140	141	142	143	144	145
Compound Name	2d_10uM	2d_100uM	2e_10uM	2e_100uM	2i_10u	2i_100uM	2l_10u	2l_100uM	2h_10uM	2h_100uM	2j_10uM	2j_100uM
Normalized Cap Driven Translation	0.99927933	1.051635817	0.972661771	0.817944997	1.112	1.164218406	1.064	1.343405135	1.1193	1.269908362	1.115802042	1.212534647
Normalized IRES Driven Translation	1.08379302	0.937350685	0.662221547	0.175314081	0.562	0.534048217	0.992	1.182379258	1.0012	0.424708192	0.9171279	0.22390834
Compound Number	146	147	148	149	150	151	152	153	154	155	156	157
Compound Name	2k_10uM	2k_100uM	2m_10uM	2m_100uM	1c_10u	1c_100uM	1d_10u	1d_100uM	1a_10uM	1a_100uM	1b_10uM	1b_100uM
Normalized Cap Driven Translation	1.00330394	1.041155259	0.995504667	1.076306808	1.095	1.063594755	0.766	1.011437558	0.7784	1.092911855	0.93942076	1.081892315
Normalized IRES Driven Translation	1.00243513	0.338515873	0.912805571	0.889423099	0.562	0.791597321	0.977	0.961401225	0.4446	0.795890168	0.957695214	0.880021373
Compound Number	158	159	160	161	162	163	164	165	166	167	168	169
Compound Name	1g_10uM	1g_100uM	1f_10uM	1f_100uM	1k_10u	1k_100uM	1l_10u	1l_100uM	1i_10uM	1i_100uM	1j_10uM	1j_100uM
Normalized Cap Driven Translation	0.87563815	0.978604339	1.111263387	1.11156401	0.686	1.05947448	0.947	0.827950596	0.8847	1.10221	0.978674244	1.268328872
Normalized IRES Driven Translation	0.86976362	0.623648124	0.737272587	0.871847153	1.082	0.910894252	0.957	1.094319807	1.0888	1.048253075	1.327081058	0.670531664
Compound Number	172	173	174	175	180	181	182	183	184	185		
Compound Name	TS-IR 125uM	TS-IR 125uM F	TS-IR 125uM I	TS-IR 125uM R	TS-IR 2	TS-IR 202F 10(	TS-IR 2	TS-IR 203249 :	TS-IR 41	TS-IR 418 100uM		
Normalized Cap Driven Translation	0.95672578	1.508671497	0.871678359	0.921220316	0.807	0.871123731	0.922	0.70155122	0.9812	0.879927951		
Normalized IRES Driven Translation	0.87992795	0.734095428	0.054065157	0.619013412	1.055	0.950497192	0.277	0.847724881	1.0809	0.608468708		

Table 6.3 Oligonucleotide sequences for cloning the TS site 1 motif and mutants mt1-mt4 in the bicistronic reporter constructs (Figure. 1.2).

TS1 sense:
AGCTTCGTCCCCCGCCCGCCGCGCCATGCCTGTGGCCGGCTCGGAGCTG

TS1 antisense:
GTGACCAGCTCCGAGCCGGCCACAGGCATGGCGCGGCGGGCGGGGGACGA

TS1 mt1 sense:
AGCTTCGTCCCCCGCCGGCCGCGCATGCCTGTGGCCGGCTCGGAGCTG

TS1 mt1 antisense:
GTGACCAGCTCCGAGCCGGCCACAGGCATGCGCGGCCGGCGGGGGACGA

TS1 mt2 sense:
AGCTTCGTCCCGAGCCGGCCGCGCCATGCCTGTGGCCGGCTCGGAGCTG

TS1 mt2 antisense:
TGACCAGCTCCGAGCCGGCCACAGGCATGGCGCGGCCGGCTCGGGACGA

TS1 mt3 sense:
AGCTTCGTCCCGCGCCGGCCGCGCATGCCTGTGGCCGGCTCGGAGCTG

TS1 mt3 antisense:
GTGACCAGCTCCGAGCCGGCCACAGGCATGCGCGGCCGGCGCGGGACGA

TS1 mt4 sense:
AGCTTCGTCCCGAGCCGGCCGCGCATGCCTGTGGCCGGCTCGGAGCTG

TS1 mt4 antisense:
GTGACCAGCTCCGAGCCGGCCACAGGCATGCGCGGCCGGCACGGGACGA

Table 6.4 Oligonucleotide sequences for cloning the HCV IRES in the bicistronic reporter constructs (Figure. 1.3).

IR-1T	AATTCCCAGCCCCCGA
IR-1B	CAATCGGGGGCTGGG
IR-2T	TTGGGGGCGACACTCCACCATAGATC
IR-2B	GAGTGATCTATGGTGGAGTGTCGCCCC
IR-3T	ACTCCCCTGTGAGGAACTACTGTCTTCACGCAGAAAGCGT
IR-3B	GCTAGACGCTTTCTGCGTGAAGACAGTAGTTCCTCACAGGG
IR-4T	CTAGCCATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGACCCC
IR-4B	AGGGGGGGTCCTGGAGGCTGCACGACACTCATACTAACGCCATG
IR-5T	CCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGT
IR-5B	ACTCACCGGTTCCGCAGACCACTATGGCTCTCCCGGG
IR-6T	GAGTACACCGGAATTGCCAGGACGACCGGGTCCTTTCTTGG
IR-6B	TGATCCAAGAAAGGACCCGGTCGTCCTGGCAATTCCGGTGT
IR-7T	ATCAACCCGCTCAATGCCTGGAGATTTGGGCGTGCCCCC
IR-7B	TCGCGGGGGCACGCCCAAATCTCCAGGCATTGAGCGGGT
IR-8T	GCGAGACTGCTAGCCGAGTAGTGTTGGGTCGCGAAA
IR-8B	AGGCCTTTCGCGACCCAACACTACTCGGCTAGCAGTC
IR-9T	GGCCTTGTGGTACTGCCTGATAGGGTGCTTGCGAGTGCCC
IR-9B	CCCGGGGCACTCGCAAGCACCCTATCAGGCAGTACCACA
IR-10T	CGGGAGGTCTCGTAGACCGTGCACCATGAGCACGAATCCTT
IR-10B	CGAAGGATTCGTGCTCATGGTGCACGGTCTACGAGACCT
IRES Forward	AGAGAATTCCCAGCCCCCGATTGG
IRES Reverse	ACTTTCGAAGGATTCGTGCTCATGG

Table 6.5 IVT assay results of the TS mRNA truncation mutants in the bicistronic constructs. Normalized values are indicated at the bottom row in bold.

RNA Truncation Mutants				Firefly Luciferase			
TS 23	TS23+ 5uM huTS	TS20	TS 20 Redo	TS 20+5uM huTS	TS15	TS 15 Redo	TS 15+5uM huTS
538	271.945	937	992	303.668	1403	1203	920.891
468	298.992	968	966	416.457	1466	1487	873.758
503	**285.4685**	**953**	**979**	**360.0625**	**1435**	**1345**	**897.3245**
49.5	19.125117	22.2	18.4	79.753867	44.92	201	33.32806392
0.1	0.0669955	0.02		0.2215001	0.031	0.15	0.037141596
0.91	**0.5164729**	**1.72**		**0.3678927**	**2.596**	**2.68**	**0.667184037**

Table 6.6 IVT assay results of the TS protein titration in either the TS1 or the ΔTS1 construct.

TS protein titration TS1 Bicistronic				
	TS Hairpin No huTS	Wt+1uM huTS	Wt+5uM huTS	Wt+25uM huTS
Firefly	652.2395	511	312	-5.7255
	453.2145	509	352	13.9525
Average	552.727	510	332	4.1135
stdev	140.73193	1.49	28.4	13.91444724
	0.1273069	0	0.04	0.049436001
Normalized	**1**	**0.92**	**0.6**	**0.007442191**
TS Protein titration ΔTS1 Bicistronic				
Firefly	4462.419	4278	5340	4382.639
	4147.02	5000	4281	5264.627
	4304.7195	4639	4811	4823.633
	1	**1.08**	**1.12**	**1.120545253**
stdev	223.02077	510	749	623.6596957
error (0.5)	0.0259042	0.06	0.08	0.064646263

Table 6.7 IVT assay results of the bicistronic reporter containing the stabilizing mutations.

Firefly	1	2	3	Average RLU	Normalized	Stdev	Norm std
Blank	-1.842	3.089	-1.246	0.000	0.000	2.691	0.050
ΔTS1	2999.180	3059.119	2708.483	2922.260	1.639	187.547	0.032
TS1	1658.802	1687.869	2001.702	1782.791	1.000	190.139	0.053
Mt1	907.294	890.908	911.952	903.385	0.507	11.053	0.006
Mt2	164.511	179.150	145.775	163.145	0.092	16.729	0.051
Mt3	1323.054	1276.271	1282.172	1293.832	0.726	25.478	0.010
Mt4	1018.446	987.172	970.291	991.969	0.556	24.433	0.012

Renilla	1	2	3	Average Renilla RLU	Normalized		
Blank	-2.105	-1.830	3.934	0.000	0.000	3.410	0.050
ΔTS1	432.351	381.079	302.964	372.132	1.000	65.156	0.088
TS1	341.000	329.824	434.596	368.474	0.990	57.536	0.078
Mt1	376.359	306.506	339.022	340.629	1.179	34.954	0.051
Mt2	398.906	412.222	387.270	399.466	1.073	12.485	0.016
Mt3	487.506	488.132	475.925	483.855	1.300	6.874	0.007
Mt4	437.742	406.664	472.214	438.874	1.179	32.790	0.037

Table 6.8 IVT assay results of the stabilizing mutations within the TS depleted extract.

Firefly	TS1	ΔTS1	Mt1	Mt2	Mt3	Mt4
	1252.07	1265.56	886.13	258.21	1070.82	877.49
	975.40	1040.33	763.37	257.84	1184.27	848.52
	1102.79	1213.67	779.46	281.47	1143.29	936.48
Average	1110.09	1173.19	809.65	265.84	1132.79	887.49
Stdev	138.48	117.94	66.72	13.53	57.45	44.83
Norm.	1	1.056841	0.729361	0.239477	1.0205	0.79948
Renilla	TS1	ΔTS1	Mt1	Mt2	Mt3	Mt4
	101.77	81.95	70.57	99.44	103.67	98.52
	60.65	69.62	*28.21*	90.98	74.42	64.06
	75.60	*127.19*	65.53	97.64	93.00	95.25
Average	79.34	75.78	68.05	96.02	90.36	85.94
Stdev	20.81428	8.713677	23.13847	4.4588284	14.802	19.02039
Norm Renilla	1	0.955186	0.857738	1.2102337	1.1389	1.08323

Table 6.9 2AP assay: peptides and the fluorescence dose response.

Peptide	0uM	1uM	5uM	10uM	25uM	125uM
RRRLLR						
Norm						
Fluorescence	1	1.083	1.43	1.5768	1.672	1.79526
norm to max	0.557	0.603	0.8	0.8783	0.931	1
	0	0.046	0.24	0.3213	0.374	0.44298
	0	0.104	0.55	0.7253	0.845	1
Net Fluorescence	1	1.149	1.23	1.3275	1.698	2.47738
RRSSRR	0	1	5	10	25	125
	0.404	0.464	0.5	0.5358	0.685	1
	0	0.06	0.09	0.1322	0.282	0.59635
Norm to max	0	0.101	0.15	0.2217	0.473	1
RRLLRR		0uM	5uM	10uM	20uM	125uM
		0	0.7	1	2.097	
	Norm to Max	0.476	0.67	0.7655	0.859	1
		0	0.19	0.2896	0.384	0.52413
Norm to max		0	0.36	0.5526	0.732	1

TS Protein	1	2uM	10uM	25uM	125uM
	377.1	394	405.1	466.5	542.569
	89.65	99	100.7	136.3	199.437
	265.4	267	268.91	265.2	265.155
	287.4	295	304.39	330.1	343.132
	1.083	1.11	1.1729	1.245	1.29408
	0.837	0.86	0.9064	0.962	1
	0	0.02	0.0695	0.125	0.1631
Norm to max	0	0.11	0.426	0.768	1

TS82-89	0uM	0.5uM	1uM	2uM	5uM	25uM
	1	1.09	1.19	1.3	1.4	1.618
	0.673865	0.738	0.8	0.86	1	0.986
Norm to max	0	0.196	0.39	0.58	1	0.956

Table 6.10 2AP assay: peptides that displayed no fluorescence altering characteristics.

Peptide	1uM TS97	2uM TS97	5uM TS97	25uM TS97	1uM TS93+	2uM TS93+	5uM TS93+	25uM TS93+	1uM TS93	2uM TS93	5uM TS93	25uM TS93
Normalized Fluor.	0.0325271	0	0.1369107	0.26167553	0.01471522	0.02677755	0.0038721	0	0	0.018016	0.0204136	0.0088245

Table 6.11 IVT assay results of the lead benzimidazole compound in the cap driven firefly luciferase (A) and the HCV IRES driven renilla luciferase (B). The bottom row indicates the relative luciferase values, while the second to last row indicates the absolute luciferase values in the assay.

A) Benzimidazole Dose Response Firefly Luciferase								
WT 0uM	0.1uM	0.5uM	1uM	2.5	5uM	10uM	20uM	100uM
2540.726	2321.464	2501.0042	1267	910	350.9	129.3	63.398	83.728
2549.974	2478.012	2266.6472	1868	988	516.6	122.5	36.491	84.515
2500.839	2118.046	2436.3612	1590	1134	284.5	94.28	45.067	13.854
2530.513	2305.84	2401.3375	1575	1011	384	115.4	48.318	60.699
1	0.911215	0.9489528	0.622	0.4	0.152	0.046	0.0191	0.024
B) Renilla Luciferase								
WT 0uM	0.1uM	0.5uM	1uM	2.5	5uM	10uM	20uM	100uM
411.3019	467.2317	457.25467	252.6	186	75.65	20.39	-11.33	-13.74
493.7427	514.4039	421.29067	346.1	196	100.5	5.92	-44.52	-52.03
519.6579	414.5629	498.96467	347.1	219	77.81	13.71	-23.11	-13.69
474.9008	465.3995	459.17	315.3	200	84.65	13.34	-26.32	-26.49
1	0.979993	0.9668756	0.664	0.42	0.178	0.028	-0.055	-0.056

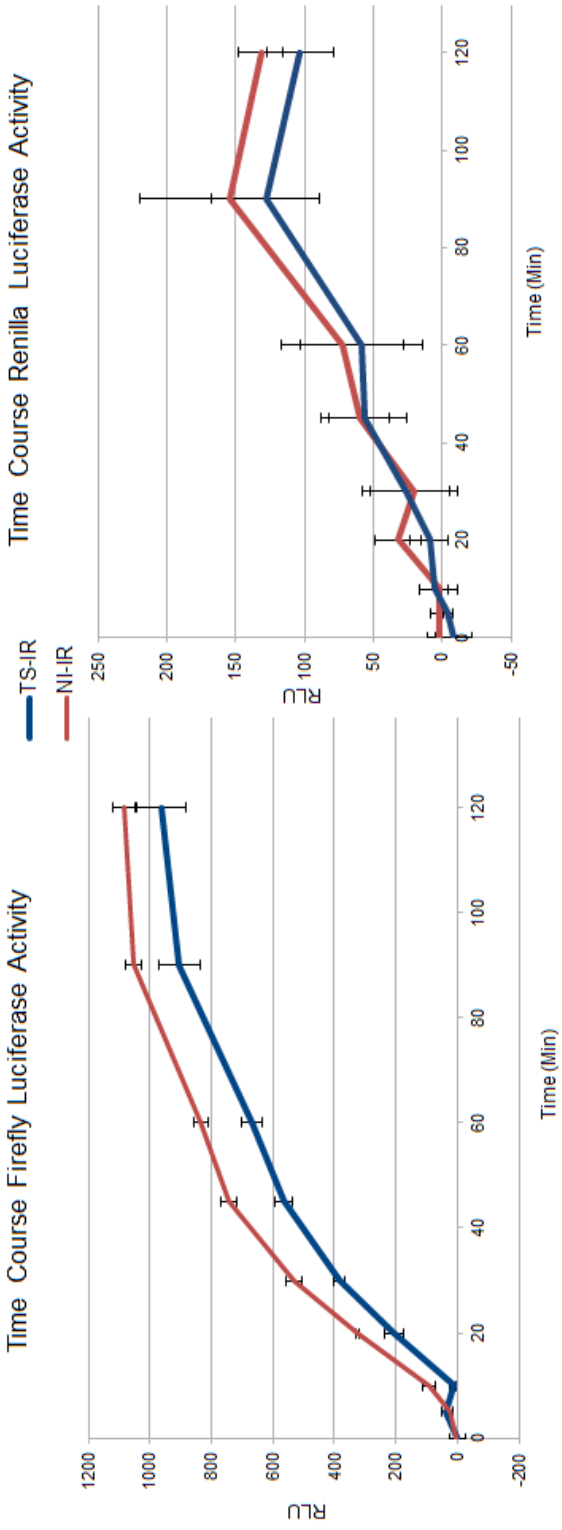


Figure 6.2 Time course of the bicistronic reporter constructs in the IVT assay.

REFERENCES

(1) Schiffer, Celia A., et al. "Crystal structure of human thymidylate synthase: a structural mechanism for guiding substrates into the active site." *Biochemistry* 34.50 (1995): 16279-16287.

(2) Berger, Sondra H., Franklin G. Berger, and Lukasz Lebioda. "Effects of ligand binding and conformational switching on intracellular stability of human thymidylate synthase." *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1696.1 (2004): 15-22.

(3) Phan, Jason, et al. "Structure of human thymidylate synthase suggests advantages of chemotherapy with noncompetitive inhibitors." Journal of Biological Chemistry 276.17 (2001): 14170-14177.

(4) Genovese, Filippo, et al. "Dimer–monomer equilibrium of human thymidylate synthase monitored by fluorescence resonance energy transfer." *Protein Science* 19.5 (2010): 1023-1030.

(5) Miller, Michael R., et al. "TU-tagging: cell type–specific RNA isolation from intact complex tissues." *Nature methods* 6.6 (2009): 439-441.

(6) Luo, BeiBei, et al. "Human thymidylate synthase with loop 181–197 stabilized in an inactive conformation: Ligand interactions, phosphorylation, and inhibition profiles." *Protein Science* 20.1 (2011): 87-94.

(7) Dibrov, S.M., Johnston-Cox, H., Weng, Y.H. and Hermann, T. (2007) Functional architecture of HCV IRES domain II stabilized by divalent metal ions in the crystal and in solution. *Angew Chem Int Ed Engl*, **46**, 226-229.

(8) Barbieri, C.M., Kaul, M. and Pilch, D.S. (2007) Use of 2-aminopurine as a fluorescent tool for characterizing antibiotic recognition of the bacterial rRNA A-site. *Tetrahedron*, **63**, 3567-6574.

(9) Sinkeldam, R.W., Greco, N.J. and Tor, Y. (2010) Fluorescent analogs of biomolecular building blocks: design, properties, and applications. *Chem Rev*, **110**, 2579-2619.

10) Voeller, Donna M., et al. "The identification of thymidylate synthase peptide domains located in the interface region that bind thymidylate synthase mRNA." *Biochemical and biophysical research communications* 297.1 (2002): 24-31.

11) Penin, F., Dubuisson, J., Rey, F.A., Moradpour, D. and Pawlotsky, J.M. (2004) Structural biology of hepatitis C virus. *Hepatology*, **39**, 5-19.

12) Chung, Raymond T., et al. "Peginterferon Alfa-2a plus ribavirin versus interferon alfa-2a plus ribavirin for chronic hepatitis C in HIV-coinfected persons." *New England Journal of Medicine* 351.5 (2004): 451-459.

13) Enserink, Martin. "First Specific Drugs Raise Hopes for Hepatitis C." *Science* 332.6026 (2011): 159-160.

14) Gallego, José, and Gabriele Varani. "Targeting RNA with small-molecule drugs: therapeutic promise and chemical challenges." *Accounts of chemical research* 34.10 (2001): 836-843.

15) Ji, Hong, et al. "Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA." *Proceedings of the National Academy of Sciences of the United States of America* 101.49 (2004): 16990-16995.

16) Fraser, Christopher S., and Jennifer A. Doudna. "Structural and mechanistic insights into hepatitis C viral translation initiation." *Nature Reviews Microbiology* 5.1 (2006): 29-38.

17) Kieft, Jeffrey S., et al. "The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold." *Journal of molecular biology* 292.3 (1999): 513-529.

18) Lukavsky, Peter J. "Structure and function of HCV IRES domains." *Virus research* 139.2 (2009): 166-171.

19) (J.S. Kieft), K. Zhou, R. Jubin, J.A. Doudna Mechanism of ribosome recruitment by hepatitis C IRES RNA. RNA, 7 (2001), pp. 194–206

20) Berry, Katherine E., et al. "Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning." *Structure* 19.10 (2011): 1456-1466.

21) Seth, Punit P., et al. "SAR by MS: discovery of a new class of RNA-binding small molecules for the hepatitis C virus: internal ribosome entry site IIA subdomain." *Journal of medicinal chemistry* 48.23 (2005): 7099-7102.

22) Liu, Shuanghu, et al. "Measuring antiviral activity of benzimidazole molecules that alter IRES RNA structure with an infectious hepatitis C virus chimera expressing Renilla luciferase." *Antiviral research* 89.1 (2011): 54-63.

23) Lovelace, Leslie L., et al. "Variants of human thymidylate synthase with loop 181–197 stabilized in the inactive conformation." *Protein Science* 18.8 (2009): 1628-1636.

24) Gibson, Lydia M., Leslie L. Lovelace, and Lukasz Lebioda. "The R163K Mutant of Human Thymidylate Synthase Is Stabilized in an Active Conformation: Structural Asymmetry and Reactivity of Cysteine 195†‡." *Biochemistry* 47.16 (2008): 4636-4643.

25) Dibrov SM, McLean JE & Hermann T. Structure of an RNA dimer of a regulatory element from human thymidylate synthase mRNA (2011). Acta Cryst. D 67, 97-104

26) Ferré-D'Amaré, Adrian R., and Jennifer A. Doudna. "Use of cis-and trans-ribozymes to remove 5′ and 3′ heterogeneities from milligrams of in vitro transcribed RNA." *Nucleic acids research* 24.5 (1996): 977-978.

27) Pleiss, Jeffrey A., Maria L. Derrick, and Olke C. Uhlenbeck. "T7 RNA polymerase produces 5' end heterogeneity during in vitro transcription from certain templates." *Rna* 4.10 (1998): 1313-1317.

28) Ferré-D'Amaré, Adrian R., and Jennifer A. Doudna. "Crystallization and structure determination of a hepatitis delta virus ribozyme: use of the RNA-binding protein U1A as a crystallization module." *Journal of molecular biology* 295.3 (2000): 541-556.

29) Stefl, Richard, Lenka Skrisovska, and Frédéric H-T. Allain. "RNA sequence-and shape-dependent recognition by proteins in the ribonucleoprotein particle." *EMBO reports* 6.1 (2005): 33-38

30) Garg, D., Henrich, S., Salo-Ahen, O.M., Myllykallio, H., Costi, M.P. and Wade, R.C. (2010) Novel approaches for targeting thymidylate synthase to overcome the resistance and toxicity of anticancer drugs. *J Med Chem*, **53**, 6539-6549.

31) E. Chu, M. A. Callender, M. P. Farrell, J. C. Schmitz, *Cancer Chemother. Pharmacol.* 2003, 52, S80– S89.

32) C. Heidelberger, Cancer Treat. Rep. 1981, 65, S3–S9; b) J. M. Davies, R. M. Goldberg, Semin. *Oncol*. 2011, 38, 552 –560.

33) M. Scartozzi, E. Maccaroni, R. Giampieri, M. Pistelli, A. Bittoni, M. Del Prete, R. Berardi, S. Cascinu, *Pharmacogenomics* 2011, 12, 251 –265.

34) Chu, Edward, et al. "Identification of an RNA binding site for human thymidylate synthase." *Proceedings of the National Academy of Sciences* 90.2 (1993): 517-521.

35) Lin, Xiukun, et al. "Characterization of a cis-acting regulatory element in the protein coding region of thymidylate synthase mRNA." *Nucleic acids research* 28.6 (2000): 1381-1389.

36) Tai, Ningwen, et al. "Translational autoregulation of thymidylate synthase and dihydrofolate reductase." *Front Biosci* 9 (2004): 2521-2526.

37) Carreras, Christopher W., and Daniel V. Santi. "The catalytic mechanism and structure of thymidylate synthase." *Annual review of biochemistry* 64.1 (1995): 721-762.

38) Chu, Edward, et al. "Autoregulation of human thymidylate synthase messenger RNA translation by thymidylate synthase." *Proceedings of the National Academy of Sciences* 88.20 (1991): 8977-8981.

39) Tan, Ruoying, and Alan D. Frankel. "Structural variety of arginine-rich RNA-binding peptides." *Proceedings of the National Academy of Sciences* 92.12 (1995): 5282-5286.

40) Pestova, Tatyana V., Jon R. Lorsch, and Christopher UT Hellen. "4 The Mechanism of Translation Initiation in Eukaryotes." *Cold Spring Harbor Monograph Archive* 48 (2007): 87-128.

41) Unbehaun, Anett, et al. "Release of initiation factors from 48S complexes during ribosomal subunit joining and the link between establishment of codon-anticodon base-pairing and hydrolysis of eIF2-bound GTP." *Genes & development* 18.24 (2004): 3078-3093.

42) Pestova, Tatyana V., Jon R. Lorsch, and Christopher UT Hellen. "4 The Mechanism of Translation Initiation in Eukaryotes." *Cold Spring Harbor Monograph Archive* 48 (2007): 87-128.

43) Pestova, Tatyana V., and Victoria G. Kolupaeva. "The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection." *Genes & development* 16.22 (2002): 2906-2922.

44) Jackson, Richard J. "The ATP requirement for initiation of eukaryotic translation varies according to the mRNA species." *European Journal of Biochemistry* 200.2 (1991): 285-294.

45) Reynolds, J. E., et al. "Unique features of internal initiation of hepatitis C virus RNA translation." *The EMBO journal* 14.23 (1995): 6010.

46) Tsukiyama-Kohara, Kyoko, et al. "Internal ribosome entry site within hepatitis C virus RNA." *Journal of virology* 66.3 (1992): 1476-1483.

47) Kieft, Jeffrey S., et al. "Mechanism of ribosome recruitment by hepatitis C IRES RNA." *Rna* 7.02 (2001): 194-206.

48) Kieft, Jeffrey S., et al. "Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation." *Nature Structural & Molecular Biology* 9.5 (2002): 370-374.

49) Komar, Anton A., and Maria Hatzoglou. "Internal ribosome entry sites in cellular mRNAs: mystery of their existence." *Journal of Biological Chemistry* 280.25 (2005): 23425-23428.

50) Vattem, Krishna M., and Ronald C. Wek. "Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells." *Proceedings of the National Academy of Sciences of the United States of America* 101.31 (2004): 11269-11274.

51) Gingras, Anne-Claude, Brian Raught, and Nahum Sonenberg. "eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation." *Annual review of biochemistry* 68.1 (1999): 913-963.

52) Paraskeva, Efrosyni, et al. "Ribosomal pausing and scanning arrest as mechanisms of translational regulation from cap-distal iron-responsive elements." *Molecular and cellular biology* 19.1 (1999): 807-816.
53) Muckenthaler, Martina, Nicola K. Gray, and Matthias W. Hentze. "IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F." *Molecular cell* 2.3 (1998): 383-388.

54) Gebauer, Fátima, and Matthias W. Hentze. "Molecular mechanisms of translational control." *Nature reviews Molecular cell biology* 5.10 (2004): 827-835.

55) Cazenave, C., et al. "Comparative inhibition of rabbit globin mRNA translation by modified antisense oligodeoxynucleotides." *Nucleic acids research* 17.11 (1989): 4255-4273.

56) Pelham, Hugh RB, and Richard J. Jackson. "An efficient mRNA-dependent translation system from reticulocyte lysates." *European Journal of Biochemistry* 67.1 (1976): 247-256.

57) Buratowski, Stephen, and Lewis A. Chodosh. "Mobility Shift DNA-Binding Assay Using Gel Electrophoresis." *Current protocols in molecular biology* (2001): 12-2.

58) Wahba, A. J., and Friedkin, M. (1961) *The Journal of biological chemistry* 236, PC11-12

59) Lorenz, R., Bernhart, S. H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., and Hofacker, I. L. (2011) *Algorithms for molecular biology : AMB* 6, 26

60) Lin, Xiukun, et al. "Role of cysteine amino acid residues on the RNA binding activity of human thymidylate synthase." *Nucleic acids research* 31.16 (2003): 4882-4887.

61) UCSF Chimera--a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. J Comput Chem. 2004 Oct;25(13):1605-12.

62) Z. Otwinowski and W. Minor, " Processing of X-ray Diffraction Data Collected in Oscillation Mode ", Methods in Enzymology, Volume 276: Macromolecular Crystallography, part A, p.307-326, 1997,C.W. Carter, Jr. & R. M. Sweet, Eds., Academic Press (New York).

63) Brunn, Nicholas D., et al. "Targeting a Regulatory Element in Human Thymidylate Synthase mRNA." *ChemBioChem* 13.18 (2012): 2738-2744.

64) Shevchenko, Andrej, et al. "Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels." *Analytical chemistry* 68.5 (1996): 850-858.