

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Site-Specific RNA-protein covalent bioconjugation using tRNA guanine transglycosylase (TGT)

Permalink

<https://escholarship.org/uc/item/62z3x5t7>

ISBN

9798193214830

Author

Guerrero Baker, Karen

Publication Date

2026-09-30

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Site-Specific RNA-protein covalent bioconjugation using tRNA guanine transglycosylase (TGT)

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Chemistry

by

Karen Guerrero Baker

Committee in charge:

Neal K Devaraj, Chair
Pascal Gagneux
Alexis C. Komor
Brian Leigh
Navtej Toor

Copyright

Karen Guerrero Baker, 2026

All rights reserved.

The Dissertation of Karen Guerrero Baker is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2026

DEDICATION

I dedicate this dissertation to my family, especially my parents and husband. Thank you for believing in me, especially during the moments when I struggled to believe in myself. This doctoral program and work have taught me that anything is possible through drive, focus, resilience and determination.

EPIGRAPH

Nothing in life is to be feared,
it is only to be understood.
Now is the time to understand more,
so that we may fear less.

Marie Curie

TABLE OF CONTENTS

DISSERTATION APPROVAL PAGE	iii
DEDICATION	iv
EPIGRAPH	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	ix
LIST OF TABLES	xi
LIST OF ABBREVIATIONS	xii
ACKNOWLEDGEMENTS	xiii
VITA	xvi
ABSTRACT OF THE DISSERTATION	xvii
Chapter 1 Introduction	1
1.1 Motivation	1
1.2 tRNA modifications by tRNA Guanine Transglycosylase (TGT)	3
1.3 Mechanism and Structure of tRNA Guanine Transglycosylase (TGT)	4
Chapter 2 Developing <i>E. coli</i> tRNA guanine transglycosylase (TGT)-RNA covalent complexes	6
2.1 Introduction	6
2.2 <i>In vitro</i> characterization of covalent complex formation between RNA-TAG3 and TGT ...	8
2.3 Stability assessment of the robustness of TGT-RNA conjugates	13
2.4 Test stability of fusing GFP to TGT	16
2.5 Conclusion	18
2.6 Materials and Methods	19
2.6.1 <i>Buffers Used:</i>	19
2.6.2 <i>Wild-type and mutant TGT Plasmid Preparation for In Vitro Studies</i>	20
2.6.3 <i>Expression/Purification of recombinant TGT</i>	20

2.6.4 RNA transcript with one TAG3 plasmids preparation and IVT.....	21
2.6.5 Plasmids Design and Sequences	22
2.6.6 TGT Covalent Conjugate Reaction Set up	31
2.6.7 TGT Covalent Conjugate Reaction Urea PAGE Analysis	31
2.6.8 Addition of PEG 8000.....	32
2.6.9 Excess Guanine, preQ1 and Queuine Treatment.....	32
2.6.10 TGT covalent conjugate reaction with and without RNase Inhibitor.....	32
2.6.11 RNA Refolding	33
2.6.12 SDS Treatment	33
2.6.13 BME Treatment.....	33
2.6.14 GFP-TGTs with TAG3 no denaturing step	34
2.7 Acknowledgements	34
Chapter 3 Cell work of TGT-RNA covalent conjugate.....	35
3.1 Assessment of plasmid GFP expression and toxicity	35
3.2 Fluorescence Microscopy Studies.....	36
3.3 Immunoblot analysis of cellular TGT-RNA conjugation	38
3.4 Conclusion	39
3.5 Materials and Methods.....	40
3.5.1 Plasmid preparation for mammalian studies.....	40
3.5.2 Image acquisition and analysis.....	45
3.5.3 Western Blot.....	46
3.6 Acknowledgements	47
Chapter 4 Conclusion	48
4.1 RNA Biology and its existing research tools	48
4.2 Future Direction of Covalent RNA-Protein Conjugates	49

REFERENCES.....	51
------------------------	-----------

LIST OF FIGURES

Figure 1.1: TGT modification present across all three kingdoms of life.....	4
Figure 1.2: Mechanism of TGT facilitating exchange of guanine (purple) for preQ1 (blue) via formation of a covalent intermediate (red). Adapted from Biela et al.,2013. ²⁸	5
Figure 2.1: (A) RNA-TAG labeling of RNAs with functionalized preQ1 small molecules. (B) RNA-TAG installation of preQ1-benzylguanine for covalent SNAP-tag conjugation to target RNA. (C) Direct site-specific covalent conjugation of a TGT linked functional protein to a target RNA.	7
Figure 2.2: TGT (D264E) mutation increases covalent intermediate stability	8
Figure 2.3: Urea PAGE gel shift analysis of TGT(D264E)-RNA TAG3 complex formation (A) ~200nt (B) ~1000nt.....	9
Figure 2.4: There was no visible TGT-RNA conjugate detected when the reactive guanine in UGU recognition motif is replaced with cytosine.	10
Figure 2.5: (A) The optimal TGT concentration to maximize product formation is 40 μ M. (B) Increasing wild-type TGT concentration to 40 μ M did not enhance TGT-RNA complex.	11
Figure 2.6: The addition of PEG 8000 did not increase TGT-RNA complex yield.	12
Figure 2.7: The removal of TAG3 affected TGT-RNA covalent complex formation.	12
Figure 2.8: (A) The addition of guanine and queuine did not significantly disassociate the covalent conjugate formation, whereas excess preQ1 dissociated ~50% of the complex. (B) The same RNA transcript without TAG3 was assessed in the presence of the small molecules as a negative control and as expected no covalent conjugate was observed.....	13
Figure 2.9: The addition of RNase inhibitor and RNA misfolding are not the cause of the appearance of a secondary species detected.	14
Figure 2.10: (A) Secondary species was suspected to be caused by TGT undergoing structural configuration changes however SDS treatment ruled out this hypothesis. (B) BME treatment indicate the secondary species detected was due to disulfide bonds.	15
Figure 2.11: (A) Covalent conjugation of TAG3 with GFP fused TGTs was confirmed via SDS PAGE gel shift analysis. (B) GFP fluorescence suggest higher molecular weight species correspond to TGT-RNA conjugate.....	17
Figure 2.12: (A) Covalent conjugation of ~200nt and ~1000nt RNA-TAG3 transcripts with GFP fused TGTs was confirmed via urea PAGE gel shift analysis. (B) GFP fluorescence suggest higher molecular weight species correspond to TGT-RNA conjugate.....	18
Figure 2.13: Covalent complex stability of (A) ~200nt and (B) ~1000nt RNA-TAG3 transcripts with GFP fused TGTs assessed with treatment of excess guanine and preQ1. Excess guanine didn't affect the complex stability, whereas preQ1 caused partial dissociation.	18
Figure 3.1: (A) GFP fluorescence persists after 96 hours (B) XTT viability assay suggest GFP fused TGTs are not toxic to HeLa cells.	36

Figure 3.2: Cartoon demonstrating formation of encodable covalent RNA-protein conjugation in cells.	37
Figure 3.3: Fluorescence microscopy of 24-hour transient transfection of GFP-TGT (WT) and mutant fusion. Green signal is from GFP. Blue, fluorescent signal is from Hoechst 33342. Red signal is from mCherry.	37
Figure 3.4: Translocation of GFP-TGT variants from the nucleus to the cytoplasm in the presence of TAG3-containing transcripts is quantified as the cytoplasmic-to-nuclear fluorescence intensity ratio (CNI). The mean cytoplasmic fluorescence is divided by the mean nuclear fluorescence for each cell; error bars represent standard deviation of biological triplicates.....	38
Figure 3.5: Western blot analysis of His-Tagged TGT variants co-expressed with TAG3-containing transcripts in HeLa cells using His Tag antibody (upper) and β -actin antibody (lower).	39

LIST OF TABLES

Table 3.1: Primary antibodies used in this study.	47
--	----

LIST OF ABBREVIATIONS

mRNA	Messenger RNA
RBP	RNA-binding protein
TGT	tRNA guanine transglycosylase
MCP	MS2 bacteriophage-derived coat proteins
<i>Z. mobilis</i>	<i>Zymomonas mobilis</i>
<i>E. coli</i>	<i>Escherichia coli</i>
RNA-TAG	RNA Transglycosylation At Guanosine
WT	Wild type
sfGFP	Superfold green fluorescent protein
IVT	<i>In vitro</i> transcription
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis
BME	β-Mercaptoethanol
DTT	Dithiothreitol
NLS	Nuclear Localization Signal
UTR	Untranslated Regions
ORF	Open Reading Frame
CNI	Cytoplasmic-to-nuclear fluorescence intensity ratio
CLIP	Cross-linking and immunoprecipitation

ACKNOWLEDGEMENTS

As a first-generation college student, navigating an academic system unfamiliar to both myself and my family presented significant challenges in pursuing not only a college degree, but ultimately a PhD. My journey to this program was shaped by unexpected opportunities that led me to this exceptional institution. At a young age, I was convinced that dentistry was the career path I was destined to pursue and was determined on making that dream a reality. However, joining an industrial research group after my undergraduate studies it sparked and deepened my curiosity for science far beyond dentistry. That growing passion drew me towards graduate school. After being rejected from the UCSD chemistry PhD program, I questioned whether graduate school was truly part of my future. After experiencing the waves of emotions that accompany rejection and uncertainty, the opportunity to pursue a master's degree presented itself. Balancing my graduate studies while working full time as a researcher was incredibly demanding, but it was during this experience of intense dedication and perseverance that I realized that pursuing a PhD was truly the path I was meant to follow. Now after many years of hard work, resilience and focus, I am approaching the finish line of this incredible program.

Like many of life's greatest accomplishments, this milestone would not have been possible without the encouragement and guidance of the remarkable mentors, peers, friends and family who stood beside me throughout this adventure. This chapter of my life has been both one of the most beautiful and challenging, but it is one I will always hold close to my heart.

I would like to thank my parents for being my heroes, my role models, my greatest source of motivation. Thank you for teaching me to rise beyond the limits that life placed in my path. Thank you for showing me the importance of resilience, especially during moments of failure and

rejection. Thank you for teaching me to be optimistic, compassionate and loving. Above all, thank you for your unconditional support and for encouraging me to explore my creativity in this lifetime.

I would like to thank my incredible husband, David. I do not have the words to fully express my gratitude for the love, care and encouragement you continuously give me. Thank you for being my greatest supporter and for always cheering me on from the sidelines. This achievement belongs to both of us.

I am deeply grateful to my advisor, Neal Devaraj, for believing in me and for providing me with countless opportunities. Thank you, Neal, for sharing your passion for research and your guidance, unwavering mentorship, encouragement and support throughout my doctoral journey. Your scientific insight and enthusiasm challenged me to grow as both a critical thinker and researcher.

Thank you for allowing me to be part of such a hardworking, curious and extraordinary research group. I would also like to thank my lab mates for being so welcoming and for their immense encouragement and friendship: Ember Tota, Alessandro Fracassi, Satyam Khanal, Alexander Harjung, Stormi Chadwick, Judith Flores, Zulfiqar Mohamedshah, Evan McCormack, Caroline Knittel, Peng Ji, Taeyang An, Hong-Guen Lee, Chih-Chin Chi, Shane Douty, Wei-An Chen, Andrew Ryu, Chuxuan Ling, Sangyoon Kang, Jacob Vance, Suyeong Han, Jiyue Chen, Youngjun Lee, Lalita Tanwar, Kira Podolsky, Christy Cho, and Kayla Busby. In particular, I would like to express my deepest gratitude to Ember Tota. You are one of the most selfless, compassionate, optimistic and driven individuals I have ever met. Your mentorship and guidance throughout this journey are gifts I will always cherish. I genuinely do not have the words to fully express my gratitude for your support and friendship. Thank you for being an exceptional role model and for inspiring me every day.

I would like to thank my undergraduate and master's advisor, Sajith Jayasinghe. You profoundly influenced my decision to pursue a PhD, and your constant encouragement and words of affirmation helped me believe in myself. I have been truly privileged to have your mentorship and guidance, both then and now.

A heartfelt thank you to my committee members: Alexis C. Komor, Navtej Toor, Brian Leigh and Pascal Gagneux. Thank you for offering guidance, support and scientific insight throughout my graduate career. I would particularly like to thank Alexis C. Komor and Navtej Toor for giving me the opportunity to develop valuable skill sets as a rotation student in your lab. It was a wonderful experience getting to know both your research group and you personally. I would also like to thank Brian Leigh for your guidance during my experience as a teacher assistant. You taught me valuable lessons on how to become a good mentor.

Lastly, thank you to all my friends and family for offering their unconditional love and support. To my brothers, Robert and Kenny, you are my rocks and my motivation to become a better person. Thank you for your companionship, love and support especially during stressful moments. I am also very grateful for all the friends who consistently provided encouragement and care. I would not have made it this far in my journey without the love and support of so many people. I will forever be grateful for all the time and energy you all have gifted me.

Chapter 2, in full, is currently being reviewed for submission for publication. Guerrero, Karen; Chadwick, Stormi R.; Tota, Ember M.; Devaraj, Neal K. The dissertation author was a primary researcher and co-author of this paper.

Chapter 3, in full, is currently being reviewed for submission for publication. Guerrero, Karen; Chadwick, Stormi R.; Tota, Ember M.; Devaraj, Neal K. The dissertation author was a primary researcher and co-author of this paper.

VITA

- 2017 Bachelor of Science in Biochemistry, California State University San Marcos
- 2021 Master of Science in Chemistry, California State University San Marcos
- 2026 Doctor of Philosophy in Chemistry, University of California San Diego

PUBLICATIONS

Guerrero, K., Smith, E., Rajkumar, S. S., Cairo, Z., Adame, J., Rawas, R., Rawas, R., Jayasinghe, S. Identification of Amino Acid Conservation in The Curli Accessory Protein CsgF. *MicroPubl Biol.* **2025**

ABSTRACT OF THE DISSERTATION

Site-Specific RNA-protein covalent bioconjugation using tRNA guanine transglycosylase (TGT)

by

Karen Guerrero Baker

Doctor of Philosophy in Chemistry

University of California San Diego, 2026

Neal K. Devaraj, Chair

Nucleic acids have been shown to have roles beyond carrying and delivering genetic information. They have drawn interest and gained popularity as powerful tools, being developed for furthering scientific research and enabling advancements in biotechnology. The use of nucleic acids in the advancement of technologies that enable visualization, tracking, and manipulation of biological macromolecules has significantly broadened their investigation, especially RNA. The

growing interest towards RNA biology due to its wide-ranging functions and impacts in the cell has been a motivator for the development of robust strategies to further study and harness its programmability.

The focus of this work is to develop a strategy to generate fully encodable, robust, site-specific, covalent RNA-protein conjugates using tRNA guanine transglycosylase (TGT) from *E. coli*. A key step in the mechanism of guanine exchange by TGT is the formation of a covalent intermediate between TGT and its RNA substrate. It has been shown that by exchanging aspartate 264 for glutamate in the active site of TGT, a stabilized catalytic intermediate can be formed with its 25 nucleotide RNA substrate (TAG3). In this work, we demonstrate covalent complex formation with longer TAG3-containing RNA substrates *in vitro* and visualize them via denaturing urea-PAGE gel shift analysis. Furthermore, these covalent complexes are stable in the presence of relevant biomolecules, remaining largely intact in the presence of excess guanine and queuine and are only partially disrupted by excess preQ1.

With demonstrated robustness of covalent conjugate formation *in vitro*, we next formed TGT-RNA conjugates in mammalian cells and visualized their localization via fluorescence microscopy. The images revealed trapping of NLS-tagged GFP-TGT in the nucleus and controlled export to the cytoplasm with the inclusion of RNA containing TAG3. Further analysis using cytoplasmic-to-nuclear fluorescence intensity ratio (CNI) revealed greater nuclear export of RNA-TGT(D264E) conjugates relative to that of wild-type TGT conjugates. Furthermore, we showed that cellular covalent TGT-RNA complexes are stable under denaturing conditions by analyzing them via immunoblotting. Together, these results establish a foundation for TGT as a highly site-selective, effective covalent RNA-protein conjugation platform for use in complex biological environments.

Chapter 1 Introduction

1.1 Motivation

In the past decade, the idea that RNA is solely the blueprint for protein production has been discredited.¹ It is now recognized as essential to cell biology with a variety of complex roles and functions, including protein production, gene regulation, cellular localization, and self-regulation.¹⁻³ Beyond the importance of its native biological functions, the value of RNA for chemical biological purposes is becoming increasingly evident, one instance being the use of messenger RNA (mRNA) for the development of highly effective vaccines.⁴ The increased awareness of the significance and usefulness of RNA brings forth the need for additional tools to study and manipulate the biological macromolecule.

RNA visualization facilitates the direct study of RNA localization and dynamics within cells, expanding the range of questions that can be addressed concerning cellular processes. Visualization and manipulation of RNA through site-specific attachment of labels, including affinity tags and detectable reporter groups, are methods widely used in research. Many chemical and enzymatic methods are available to post transcriptionally label RNA, either internally or at the 5' or 3' ends.¹ One method covalently attaches a chemically modified guanosine triphosphate (GTP) to a target adenosine via a 2',5'-phosphodiester bond.⁵ However, the usefulness of this technique is limited as the excess bulk of the GTP could affect the RNA-function. In another method, carbodiimides are employed to label the 5' end with a monophosphate group which is transformed to a primary amine, enabling coupling with NHS esters.⁶ This technique has multiple limitations such as partial degradation of RNA during functionalization, the ability to only label the 5' end, and requires the incorporation of a monophosphate group at the 5' end when designing

the construct. There are a variety of methods actively being used to harness and visualize nucleic acids such as RNA aptamers, fluorescence in-situ hybridization (FISH) and RNA-binding proteins (RBPs).⁷

RNA-binding protein technologies have given researchers the ability to harness and further study RNA behavior in native cellular environments. These systems heavily rely on proteins binding to engineered RNAs containing the protein's recognition motif, facilitating the recruitment of functional conjugates to a target RNA.⁸ An example of a developed RBP system continuously being used in the literature are MS2-MCP systems. These systems utilize an MS2 bacteriophage derived coat protein (MCP) fused to a fluorescent protein to visualize an RNA containing multiple MCP recognition sites (MS2 stem loop motifs) in cellular processes.⁹ There are other RBP systems with analogous recognition mechanisms, like λ N-boxB and PP7-PCP, that facilitate the recruitment of functionalized proteins to an intended RNA.¹⁰⁻¹¹

While RBP technologies are useful for recruiting proteins to RNA via programmed recognition elements, they rely on non-covalent interactions which can become unstable and disrupted in the complex cellular environment. This can be problematic if the RBP-RNA complexes are being purified from cell lysate for characterization, especially if undergoing purifications steps involving high salt concentrations, detergents or denaturing conditions.¹² For instance, involvement of harsh washing or exposure to denaturants such as SDS, heating or Urea would lead to reduction or complete dissociation of RBP-RNA complexes recovered from cells.¹² On the contrary, a covalent bond between RNA and proteins stabilizes the conjugate, overcoming this limitation by providing durable interactions leading to higher retention and detection of the RBP from complex environments for analysis and characterization.¹²⁻¹³ There are existing methods that can covalently link RBP-RNA that withstand dissociation under stringent denaturing

conditions eliminating non-specific, noncovalent interactions. For example, cross-linking and immunoprecipitation (CLIP) is a popular technique utilized for RBP-RNA characterization.¹³ CLIP captures the RBP complexes by covalently linking them via UV irradiation, purifying by immunoprecipitation and analyzing through sequencing.¹²⁻¹³ This highly specific method provides the stability necessary to withstand denaturing conditions and reduce nonspecific interactions.¹² Although it is a beneficial approach for covalent linkage, it is a harsh and technical workflow due to the need of UV irradiation that could jeopardize the integrity of the biomolecules. Therefore, the need to develop a technology that can form covalent interactions within live cells is critical to allow us to assess real-time interactions without the need of fixation or harsh treatments like UV irradiation. Sequence-specific recruitment of functional proteins—such as enzymes that methylate adenosine residues on RNA—can be achieved through nucleic acid-mediated strategies. For instance, a covalent RNA–protein complex could direct N6-adenosine methylation to a specific mRNA, thereby modulating its stability and translation.¹⁴ Overall, RNA–protein covalent complexes hold significant potential for advancing technologies that either target RNA directly or exploit its inherent programmability within biological systems.

1.2 tRNA modifications by tRNA Guanine Transglycosylase (TGT)

tRNA guanine transglycosylase (TGT) exchanges guanine for 7-deazaguanine derivatives in its cognate tRNA substrates to increase translational fidelity.¹⁵⁻¹⁸ This modification occurs in all three kingdoms of life, with slight variations in the associated tRNA and small-molecule substrates (**Figure 1.1**).¹⁹ Eukaryotes and bacterial TGTs recognize and covalently modify the guanine at the wobble position of the anticodon loops of the following four tRNAs: tRNA^{Asn}, tRNA^{Asp}, tRNA^{His} and tRNA^{Tyr}.²⁰⁻²¹ Bacterial TGT inserts preQ1, a queuine precursor, which is then further enzymatically modified to queuine.¹⁷ Alternatively, eukaryotic TGTs directly incorporate queuine salvaged from the environment without the need of an intermediate substrate.²² This post

transcriptional modification is fundamental for the proper translation of mRNAs.¹⁸ Lastly, TGT is also found in archaea but exchanges guanine for preQ0 which is then further modified to archaeosine.

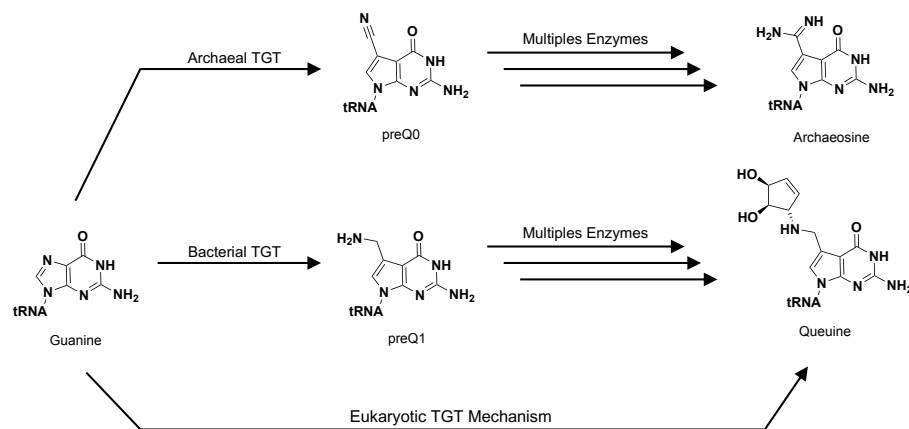


Figure 1.1: TGT modification present across all three kingdoms of life

1.3 Mechanism and Structure of tRNA Guanine Transglycosylase (TGT)

It has been previously shown that TGT does not require an entire tRNA to recognize its RNA substrate.²³ Previous biochemical research has shown that a *E. coli* TGT can recognize and modify a 17-nucleotide hairpin (ECY-A1 or TAG2) derived from the tyrosine tRNA anticodon loop.²³ Furthermore, *E. coli* TGT has proven to be able to more efficiently modify an extended 25-nucleotide hairpin (ECY-MH or TAG3) than ECY-A1.²³ Further studies showed that the minimal recognition unit required for TGT's enzymatic activity is the cognate tRNAs' UGU motif at position 33-35.²⁴ It was observed that the nucleobase at position 32 is required to be a pyrimidine, cytosine or uracil, because the double-ringed purine nucleobases, adenine or guanine reduces enzymatic activity.²⁴

The mechanism of preQ1 insertion by bacterial TGT from *Zymomonas mobilis* (*Z. mobilis*) has been previously described along with the crystal structure.²⁰⁻²¹ While there is no crystal structure of *Escherichia coli* (*E. coli*) TGT, it shares 55% homology to that from *Z. mobilis*, and

the analogous active site residues have been identified (**Figure 1.2**). Briefly, a highly conserved nucleophilic aspartate residue attacks the ribose backbone, breaking the glycosidic bond with guanine, leaving behind a reversible TGT-RNA covalent intermediate.¹⁷ PreQ1 then enters the active site and the nitrogen at the 9 position attacks the anomeric carbon of the ribose, forming a new glycosidic bond between the RNA and preQ1 and releasing TGT.^{17,19,25-27} This irreversible post transcriptional modification occurs under mild 37°C temperatures.

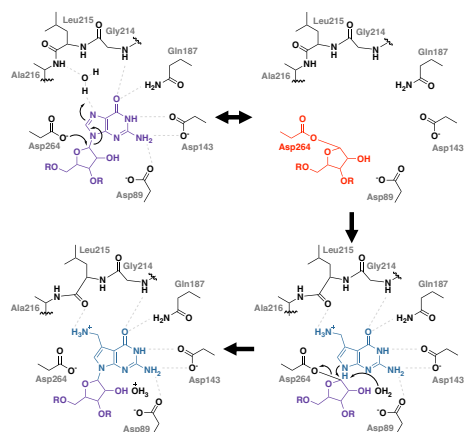


Figure 1.2: Mechanism of TGT facilitating exchange of guanine (purple) for preQ1 (blue) via formation of a covalent intermediate (red). Adapted from Biela et al.,2013.²⁸

Chapter 2 Developing *E. coli* tRNA guanine transglycosylase (TGT)-RNA covalent complexes

2.1 Introduction

Our understanding of cell biology has expanded with the increasing development of technologies that can label, detect, and manipulate various biomolecules. The availability of such tools facilitates the study of DNA and RNA. RNA research has drawn interest because of recent discoveries emphasizing its importance beyond protein production.

While RBP technologies are among the most powerful current tools to study RNA, they depend on non-covalent interactions between the RNA and protein species that can be vulnerable under complex cellular conditions. Therefore, the need for covalent bonds between RNAs and proteins are fundamental when robust interactions are crucial for more complex or variable environments. Some examples of cellular scenarios that would strongly benefit from a RNA-protein covalent complex easing the recruitment of functional proteins via nucleic-acid mediated mechanisms would be directing adenosine methylation to a desired mRNA to increase its stability and translational efficiency.¹⁴ Developing a technology that forms RNA-protein conjugates that are fully nucleic-acid mediated, sequence-specific with covalent bonds would offer stability and specificity in less favorable conditions. Although existing technologies have demonstrated the ability to form sequence-specific covalent RNA-protein conjugates *in vitro*, their performance in mammalian cells is limited by insufficient sequence specificity.²⁹ Consequently, there is still a demand for a modular, genetically encodable system that can form covalent RNA-protein conjugates with high site specificity. In summary, the use of RNA-protein covalent conjugates demonstrates great potential for the development of new technologies that can target desired RNAs or leverage its unique programmability in biological systems.

Previous work in our lab developed an enzymatic RNA modification technology named RNA-TAG (RNA transglycosylation at guanosine) that modifies RNAs containing a TGT recognition sequence (TAG2 or TAG3) with different functional, conjugated preQ1 small molecule handles (**Figure 2.1A**).³⁰⁻³⁸ More recently, RNA-TAG has been adapted to generate covalent RNA-protein conjugates. In this approach, a bifunctional preQ1-benzylguanine (preQ1-BG) probe is used to couple RNA-TAG labeling with the SNAP-tag protein labeling system.³⁹ Specifically, TGT installs preQ1-BG onto an RNA of interest, allowing a SNAP-tag fused protein to bind and form a stable covalent RNA-protein complex (**Figure 2.1B**). Although this use of RNA-TAG is effective, it requires a two-step process to produce the final conjugate. A single-step technology for direct covalent attachment of a functional protein to a target RNA would streamline this process (**Figure 2.1C**). Such a system would eliminate the need to initially label RNA with a specific substrate to facilitate the recruitment step. The catalytic mechanism of TGT and its selectivity towards its RNA substrate suggested the potential for TGT to be developed as a tool for direct covalent modification of RNA.

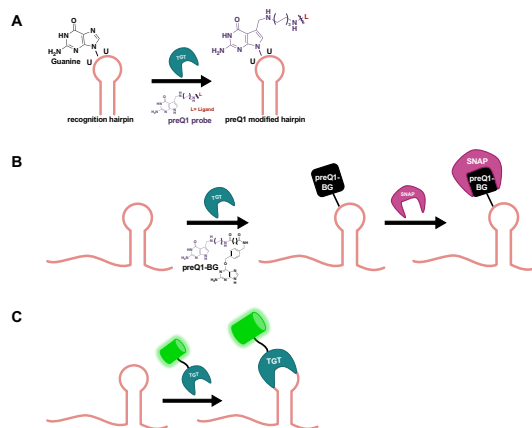


Figure 2.1: (A) RNA-TAG labeling of RNAs with functionalized preQ1 small molecules. (B) RNA-TAG installation of preQ1-benzylguanine for covalent SNAP-tag conjugation to target RNA. (C) Direct site-specific covalent conjugation of a TGT linked functional protein to a target RNA.

TGT's heavy involvement in RNA modification and its high selectivity towards its recognition RNA substrate makes the enzyme a strong candidate with great potential to be used for the development of a covalent RNA-protein technology. Previous mechanistic studies of the TGT-RNA covalent intermediate formed during the endogenous post-transcriptional modification observed that the covalent complex can be stabilized by mutating the nucleophilic residue, aspartate 264 in *E. coli* TGT, to glutamate, even in the presence of excess guanine or preQ1 (**Figure 2.2**).⁴⁰ This robust interaction will impart the stability necessary to study RNA localization and interactions when noncovalent associations are not strong enough. Furthermore, this technology could be expanded to study the stability, function, and regulatory mechanisms of RNAs, as well as facilitate nucleic acid sequence mediated control and degradation of RNA species and associated proteins.

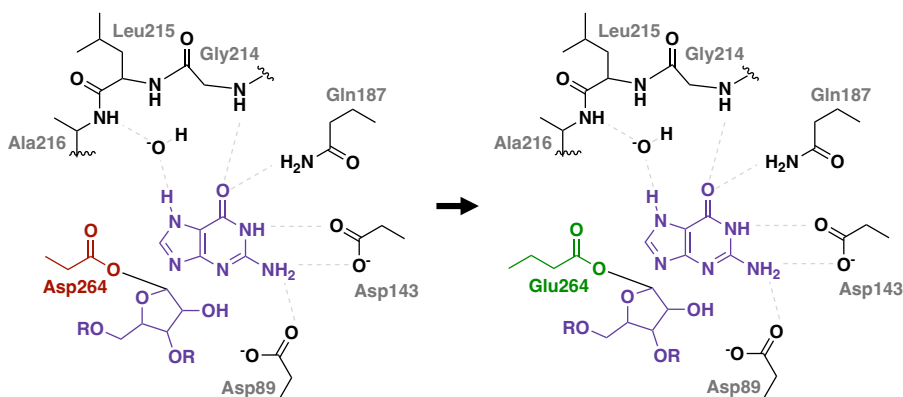


Figure 2.2: TGT (D264E) mutation increases covalent intermediate stability

2.2 *In vitro* characterization of covalent complex formation between RNA-TAG3 and TGT

While it has been previously reported that the D264E mutation increases stability of the *E. coli* TGT-TAG3 covalent conjugate compared to the wild-type, its effectiveness and activity with an RNA transcript have not been tested.⁴⁰ Thus, raising uncertainty whether TGT(D264E) can recognize longer RNA-TAG3 substrates. To address this uncertainty, we began our studies by

investigating whether we could form a covalent complex with RNA transcripts. Due to TGT transglycosylation having such an efficient recognition mechanism, we hypothesized that by incorporating the 25-nucleotide recognition motif (TAG3) to an RNA transcript we would be able to observe the formation of the covalent conjugate. To initially test TGT-RNA formation, we evaluated the stability of a covalent conjugate formed between recombinantly expressed TGT(D264E) and *in vitro* transcribed (IVT) RNA transcripts containing a single TAG3 recognition motif. To ensure transcription was not disrupted the single TAG3 motif was placed in the 3' UTR. We assessed two different RNA transcripts of different lengths, ~200nt and ~1000nt RNA-TAG3. We cloned *E. coli* WT TGT to D264E via site-directed mutagenesis. We expressed, purified and assessed TGT (WT and D264E) activity. Covalent conjugation amongst the TGT and RNA transcript was analyzed using a urea PAGE gel shift analysis. The visualization of an upward shift in the product lane in comparison to the starting material is a clear indicator of an increase of molecular weight suggesting a complex formation. As expected, the two differently sized RNA transcripts demonstrated covalent conjugation with mutant TGT(D264E) but very little visible conjugation with WT TGT (Figure 2.3).

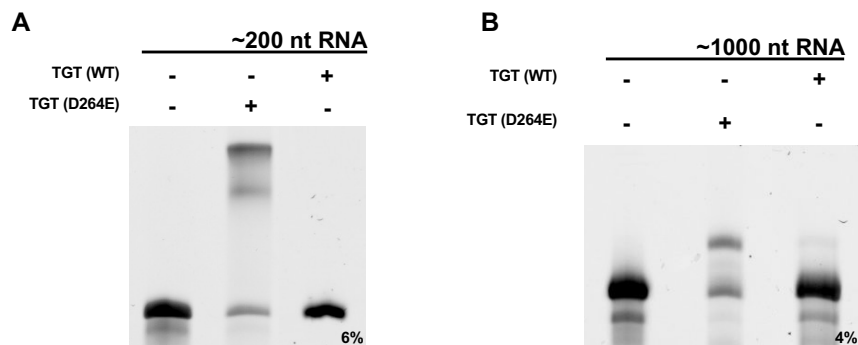


Figure 2.3: Urea PAGE gel shift analysis of TGT(D264E)-RNA TAG3 complex formation (A) ~200nt (B) ~1000nt.

In order to test selectivity towards the cognate guanine at the UGU recognition motif, we incubated both TGT(D264E) and TGT(WT) with the ~200nt and ~1000nt RNA-TAG3 Δ C, in which the guanine is replaced by cytosine. In both RNA-TAG3 Δ C transcripts we did not observe an upward shift. These results confirm that the excised nucleobase in the RNA transcripts is indeed the cognate guanine in the TAG3 recognition motif (**Figure 2.4**).

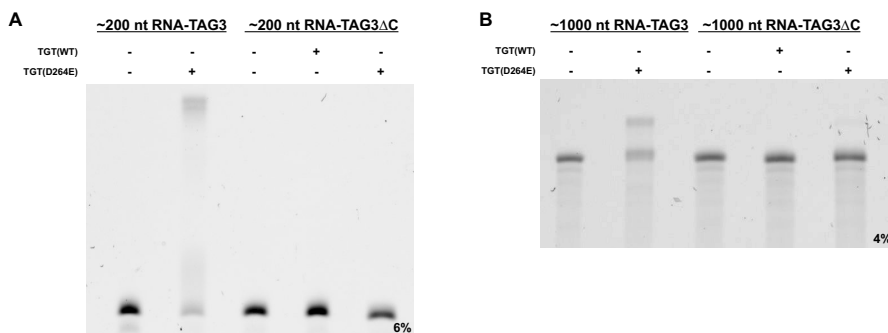


Figure 2.4: There was no visible TGT-RNA conjugate detected when the reactive guanine in UGU recognition motif is replaced with cytosine.

While both RNA transcripts exhibit the ability to promote covalent conjugation, the final product showed to be more robust with the ~200nt RNA transcript than ~1000nt. We hypothesize that this could be due to issues with the RNA transcript itself, possibly from low *in vitro* transcription efficiency leading to truncated RNA lacking complete recognition motifs, as well as RNA degradation caused by RNase contamination. Therefore, follow up *in vitro* studies were conducted with the ~200nt RNA transcript.

To optimize the covalent conjugation yield, we probed the effects of TGT(D264E) concentration in the reaction. After testing 4 different concentrations of enzyme, it was shown that 40 μ M was the optimal enzymatic concentration needed to ensure maximum product formation. Concentrations higher than 40 μ M did not significantly increase product yield (**Figure 2.5A**). Due to the poor detectability of the stable covalent conjugate formed by wild-type TGT, we

hypothesized that it was due to insufficient enzyme concentration added to the reaction. To test this, we increased the concentration of wild-type TGT to 40 μM , however, these new reaction conditions did not increase the yield (**Figure 2.5B**).

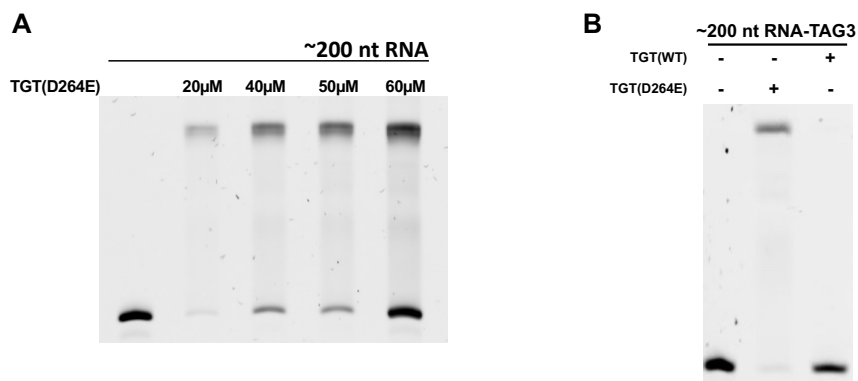


Figure 2.5: (A) The optimal TGT concentration to maximize product formation is 40 μM . (B) Increasing wild-type TGT concentration to 40 μM did not enhance TGT-RNA complex.

In an attempt to obtain complete conjugation of all the starting RNA substrates, we tested whether adding PEG 8000 to the *in vitro* reactions would generate a more crowded environment driving TGT and RNA substrates closer in proximity, resulting in an increase of reaction rates and efficiency. Unfortunately, there was not a significant difference in product conjugation, so the use of PEG 8000 was not added to the reaction conditions (**Figure 2.6**). Further experiments would be needed to optimize the reaction conditions to attain a 100% starting material conversion but for the focus of our studies the attained product satisfies the requirement to continue further developing our technology.

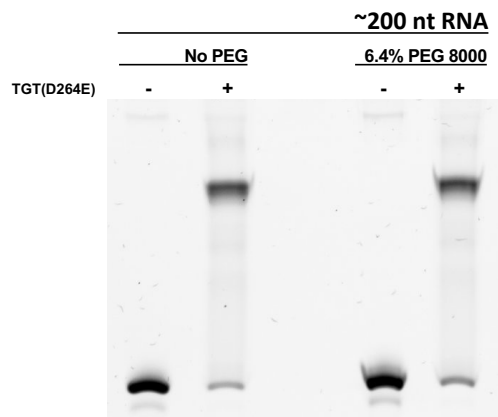


Figure 2.6: The addition of PEG 8000 did not increase TGT-RNA complex yield.

Lastly, it was essential to confirm that the removal of TAG3 from the RNA transcript would eliminate the gel shift in the product lane that was previously observed before proceeding with further studies. Therefore, we removed TAG3 while the remainder of the RNA transcript remained intact and the transcript was treated with 40 μ M of TGT using the same reaction conditions. As expected, no gel shift was observed in the product lane indicating no molecular weight difference with the starting material supporting our previous gel shift analysis data, suggesting the formation of the TGT-RNA complex is due to the interaction between TGT and TAG3 (**Figure 2.7**).

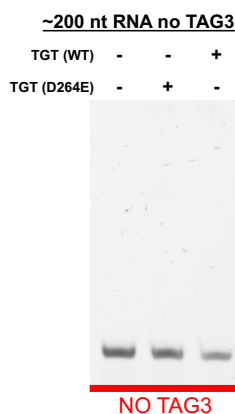


Figure 2.7: The removal of TAG3 affected TGT-RNA covalent complex formation.

2.3 Stability assessment of the robustness of TGT-RNA conjugates

A stability assessment of the TGT-RNA conjugate was conducted to see if the integrity of the complex would be compromised in mammalian cells by endogenous small molecule substrates of TGT. The robustness of the complex was evaluated by adding excess exogenous small molecule competitors (guanine, preQ1 or queuine). The addition of guanine and queuine did not significantly reduce the covalent conjugate formation, whereas the presence of excess preQ1 dissociated ~50% of the complex (**Figure 2.8A**). Additionally, the same RNA transcript without TAG3 was evaluated in the presence of the small molecules as a negative control, and as expected no covalent conjugate was observed in any of the conditions (**Figure 2.8B**). These results suggest that TGT-RNA complexes assembled in mammalian cells are unlikely to be disrupted by guanine and queuine.

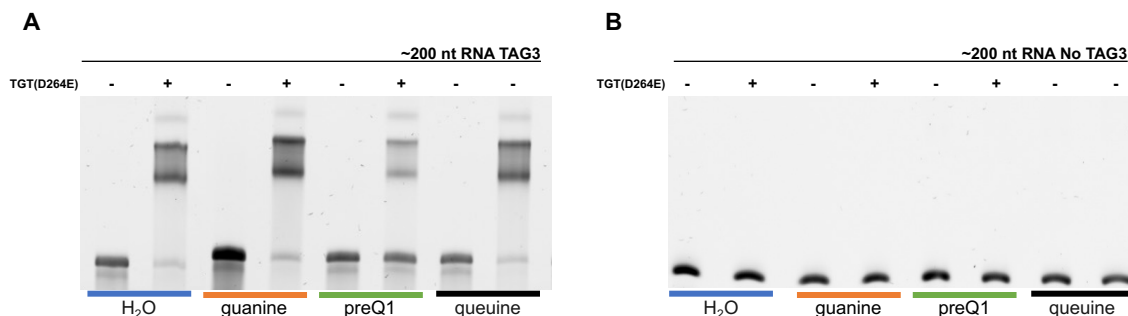


Figure 2.8: (A) The addition of guanine and queuine did not significantly disassociate the covalent conjugate formation, whereas excess preQ1 dissociated ~50% of the complex. (B) The same RNA transcript without TAG3 was assessed in the presence of the small molecules as a negative control and as expected no covalent conjugate was observed

However, upon repeating the stability assessment, we observed inconsistent additional bands in the product lane, preventing confident identification of the TGT-RNA conjugate and raising uncertainty regarding the identity of the other species. To determine the cause, we assessed if the presence of RNase inhibitors affected the molecular weight of the TGT-RNA conjugate. Reactions prepared with and without RNase inhibitors showed the same bands indicating that the

inhibitor does not cause the band heterogeneity observed (**Figure 2.9**). We then hypothesized that the different species detected could be due to RNA misfolding that could occlude the recognition hairpin at the 3'UTR limiting its accessibility to TGT. To test this theory, parallel reactions were conducted with RNA that was either denatured (98°C, 1 min) and refolded (4°C) or untreated. The refolding process did reduce the fraction of the fainter higher molecular weight bands and increased the abundance of the suspected corresponding TGT-RNA band; however, the denaturing process also led to RNA degradation (**Figure 2.9**). These results indicate that the higher molecular weight bands correspond to species that are disrupted by heat suggesting they are non-covalent interactions.

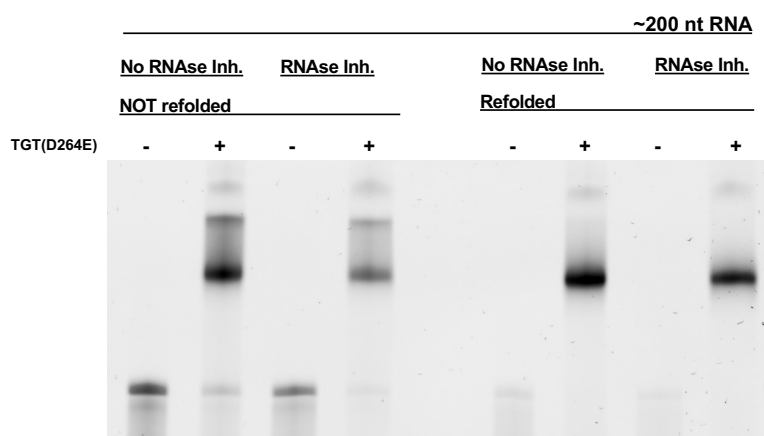


Figure 2.9: The addition of RNase inhibitor and RNA misfolding are not the cause of the appearance of a secondary species detected.

It was then hypothesized that TGT may be undergoing conformational changes before, during or after the reaction affecting how the TGT-RNA conjugate migrates during gel electrophoresis. This alteration in migration behavior can be leading us to believe the multiple bands correspond to different entities but could be a single species with differential migration rates. To decipher if TGT indeed is undergoing structural configuration changes, we added sodium dodecyl sulfate (SDS) to the reaction product to denature the protein and dissociate any non-

covalent interactions present. While treating the sample with SDS did reduce some of the species, there was still a second band observed with a similar molecular weight of the TGT-RNA conjugate (**Figure 2.10A**).

The similar molecular weight bands observed after SDS treatment suggested that they represented a single species but with disulfide bond formation among some TGT-RNA conjugates affecting its migration resulting in the appearance of two different bands. To test this hypothesis, we added more reducing agent, β -Mercaptoethanol (BME), to the reaction because dithiothreitol (DTT) in the reaction buffer may have lost its activity during storage. The addition of BME resulted in a single band and no detectable RNA degradation. These results indicate that the previously observed bands were caused by disulfide bond formation and non-covalent species rather than different molecular products (**Figure 2.10B**).

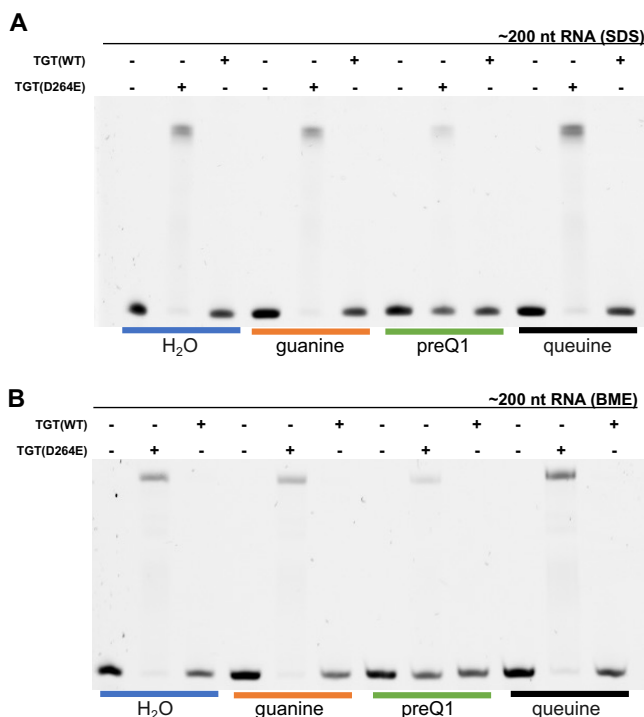


Figure 2.10: (A) Secondary species was suspected to be caused by TGT undergoing structural configuration changes however SDS treatment ruled out this hypothesis. (B) BME treatment indicate the secondary species detected was due to disulfide bonds.

2.4 Test stability of fusing GFP to TGT

In addition to ensuring conjugate stability in the presence of endogenous small molecules, it was imperative to assess the retention of TGT activity toward its RNA hairpin substrate when fused to other functional proteins. As a proof of concept, we fused TGT with superfolder-GFP (sfGFP) to enable visualization of covalent conjugate formation by gel imaging. Four GFP fusion constructs were designed where GFP is positioned at either the N- or C-terminus of both TGT(WT) and TGT(D264E) proteins to determine optimal construct activity. To minimize experimental variability, the fused constructs were designed to closely match the non-fused TGT plasmids. However, the TGT-GFP construct didn't yield sufficient protein for all planned *in vitro* studies, and after troubleshooting the expression protocol, we decided to use the GFP-TGT construct for all subsequent experiments.

The first step was to incubate recombinant GFP-TGTs with TAG3 to determine whether complex formation could be observed. Due to the smaller size of the expected covalent complex, SDS-PAGE was used for the gel shift analysis under reaction conditions previously reported in the literature.⁴¹ As anticipated, an upward shift was observed in the product lane relative to TGT not treated with TAG3 (**Figure 2.11A**). Samples were not heat denatured prior to gel loading to prevent protein unfolding to preserve GFP signal from the fused TGTs, allowing gel imaging to confirm the higher molecular weight band corresponded to GFP-TGT bound to TAG3. Thus, the gel was imaged before staining with Coomassie blue so fluorescent bands could be analyzed in parallel with the stained bands (**Figure 2.11B**). Both imaging methods detected the same bands containing fusion TGT, indicating complete GFP translation and demonstrating that GFP incorporation didn't affect TGT activity.

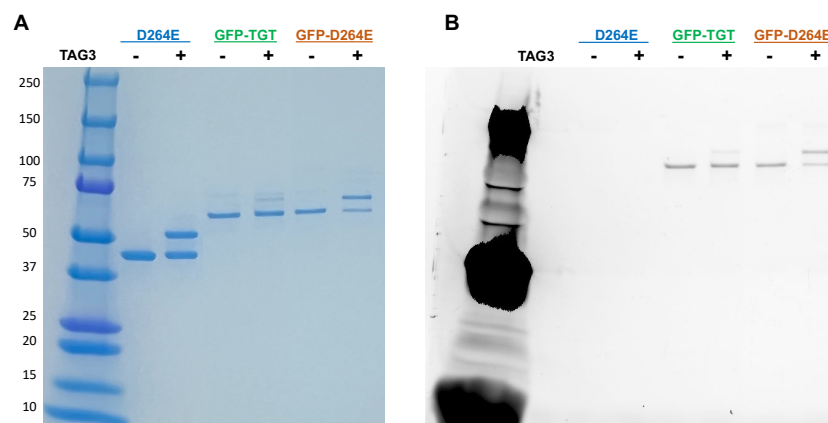


Figure 2.11: (A) Covalent conjugation of TAG3 with GFP fused TGTs was confirmed via SDS PAGE gel shift analysis. (B) GFP fluorescence suggest higher molecular weight species correspond to TGT-RNA conjugate.

The next step was to incubate GFP-TGT with the ~200nt or ~1000nt RNA transcript containing a single TAG3 hairpin to confirm that a fused TGT kept its ability to covalently link to longer RNA transcripts. As anticipated, a higher molecular weight band was detected in the product lane relative to the starting material, suggesting that fusing TGT to a functional protein does not impair covalent conjugation with longer RNAs. It was fundamental to preserve the GFP signal to verify the shifted bands corresponded to the suspected TGT-RNA species, so samples were not heated prior to the gel analysis. The reactions containing the ~200nt RNA transcript were analyzed in a 6% urea gel and those with the ~1000nt RNA were analyzed in a 4% urea gel because they have different migration rates due to their size difference. Both gels were imaged before and after GelRed staining to confirm the presence of GFP-TGT within the higher molecular weight band (**Figure 2.12**). Overlaying of the fluorescent and nucleic acid signal validated the shifted band corresponded to the TGT-RNA complex. In the final analysis, we assessed the stability of the covalent linkage between the fused TGTs and longer RNA transcripts by adding excess exogenous small molecules, guanine and preQ1. Consistent with previous data, only preQ1

disrupted the covalent conjugate, suggesting fusing TGT to functional proteins does not compromise its enzymatic ability (**Figure 2.13**).

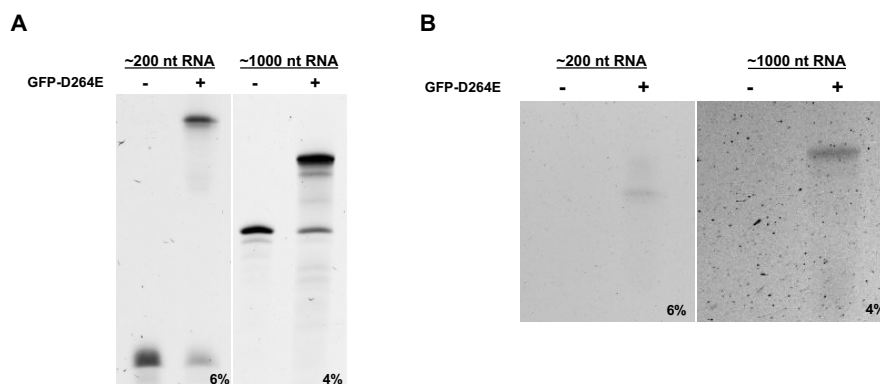


Figure 2.12: (A) Covalent conjugation of ~200nt and ~1000nt RNA-TAG3 transcripts with GFP fused TGTs was confirmed via urea PAGE gel shift analysis. (B) GFP fluorescence suggest higher molecular weight species correspond to TGT-RNA conjugate.

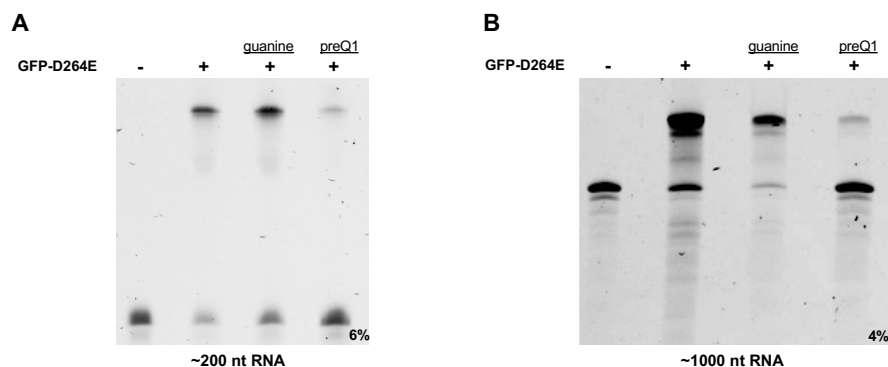


Figure 2.13: Covalent complex stability of (A) ~200nt and (B) ~1000nt RNA-TAG3 transcripts with GFP fused TGTs assessed with treatment of excess guanine and preQ1. Excess guanine didn't affect the complex stability, whereas preQ1 caused partial dissociation.

2.5 Conclusion

The strategy presented herein has demonstrated that *E. coli* TGT can form stable covalent conjugates with RNA transcripts containing a single recognition hairpin. We observed conjugate formation with two different length RNA transcripts that are stable in the presence of denaturants

and excess relevant endogenous small molecules. Additionally, it was fundamental to show that TGT can retain its activity when fused to another protein before proceeding to mammalian studies. Thus, as a proof of concept we fused TGT to superfold green fluorescent protein and observed TGT activity was not disrupted. We believe our TGT-RNA technology will be a valuable tool, allowing the formation of stable covalent conjugates in mammalian cells using fully encodable components in a highly specific, robust, and efficient manner.

2.6 Materials and Methods

2.6.1 Buffers Used:

TGT His Lysis/Wash Buffer (Zymo His-Wash P2003-4): 50 mM sodium phosphate buffer pH 7.7, 300 mM NaCl, 50mM imidazole, 0.03% Triton X-100, 1x Halt and 400 μ M PMSF

TGT His Elution Buffer (Zymo His-Elution P2003-5): 50 mM sodium phosphate buffer pH 7.7, 300 mM NaCl, 250mM imidazole, 1xHalt and 400 μ M PMSF

TGT Strep Wash Buffer (Iba Strep-Wash 2-1003-100): 100 mM Tris/HCl pH 8.0, 150 mM NaCl and 1mM EDTA

TGT Strep Elution Buffer (Iba Strep-Elution 2-1000-025): 100 mM Tris/HCl pH 8.0, 150 mM NaCl, 1mM EDTA and 50mM Biotin

TGT Storage Buffer: 25 mM HEPES, pH 7.3, 2 mM DTT, 1 mM EDTA, and 100 μ M PMSF

TGT Reaction Buffer 1x: 100 mM HEPES, pH 7.3, 5 mM DTT, and 20 mM MgCl₂

TBE Buffer (Corning 46-011-CM): (10x stock) 108 g Tris base, 55 g boric acid, 7.5 g EDTA in 1 L Milli-Q water.

2X RNA Loading Buffer (New England Biolabs B0363S): 95% Formamide, 0.02% SDS, 0.02% bromophenol blue, 0.01% xylene cyanol, 1 mM EDTA

2.6.2 Wild-type and mutant TGT Plasmid Preparation for In Vitro Studies

The TGT WT and mutant were prepared by using plasmid containing the *E. coli* TGT sequence, an N terminal strep tag, and a C terminal 6xHis tag. New England Biolabs Q5 Site-Directed Mutagenesis Kit (NEB M0494) was used for PCR amplification and product ligation. The product was transformed into NEB 5-alpha competent *E. coli* cells and plated on carbenicillin antibiotic Luria Bertani (LB) broth Agar plates to ensure only successfully transformed cells grew. The single colonies were grown for 14-16 hours at 37°C. The following day, single cell colonies were grown in 10mL LB bacterial cultures containing 100 µg/mL carbenicillin overnight at 37°C. After a 14-16 hours incubation period at 37°C, the plasmids were extracted using Biomiga® plasmid miniprep II kit. The purified constructs submitted to Plasmidsaurus for whole plasmid sequencing.

2.6.3 Expression/Purification of recombinant TGT

E. coli TGTs were expressed from the plasmids previously described. The purified plasmids were transformed into *E. coli* BL21(DE3) (Thermo Fisher C606010) competent cells and plated on carbenicillin antibiotic LB Agar plates to ensure only successfully transformed cells grew. The single colonies were grown for 14-16 hours at 37°C. The following day, a single colony of the transformants was grown in a 10mL LB containing 100 µg/mL carbenicillin overnight at 37°C. Approximately 5mL of starting culture was transferred into 250mL of LB broth containing 100 µg/mL carbenicillin. The culture was grown to an optical density of ~0.6 and protein expression was induced by adding IPTG at final concentration of 1mM. Culture was induced at 18°C shaking overnight. Bacterial cells were pelleted at 6000xg for 20 minutes at 4°C. Cell pellets

were either resuspended in a lysis buffer and purified immediately or placed in -20°C for long term storage.

The cell pellets were resuspended with 2mL of lysis buffer and sonicated on ice using the following settings: 4 cycles (30 seconds on/ 2minutes off); output control=4; duty cycle%: 30-40. The lysate was centrifuged at 10,000xg for 30 minutes at 4°C to harvest the supernatant containing the desired protein. During the 30-minute centrifugation, the HisPur Ni-NTA resin (ThermoFisher 88221) needed for His Purification was equilibrated in a spin column with two resin-bed volumes of lysis buffer. After centrifugation, the supernatant was transferred to the capped spin column containing the equilibrated nickel resin and incubated for two hours with end over end mixing at 4°C. After incubation, the supernatant/resin mixture washed multiple times with His wash buffer. The protein will be eluted with His elution buffer. The elutions were immediately transferred into a capped spin column containing equilibrated strep resin (Iba 2-5010-002) and incubated for 30 minutes with end over end mixing at 4°C. The supernatant/resin mixture was washed, and the protein was eluted. The eluted protein was dialyzed against the TGT storage buffer using a 30kDa cut off spin filter (Millipore UFC503024) overnight. Precast protein gels (Bio-Rad 4561096) were run on Bio-Rad Mini-Protean gel system to verify protein purification. The protein's concentration was determined by absorbance on the nanodrop and aliquoted in fractions needed for a single covalent conjugate reaction and stored in -80°C.

2.6.4 RNA transcript with one TAG3 plasmids preparation and IVT

The plasmids are in a pcDNA3.1 backbone containing a single TAG3 in the sequence. These plasmids were prepped using the same protocol mentioned in section 2.6.2. The plasmids were linearized using the restriction endonuclease, XbaI (NEB R0145). The linearized DNA was purified using a DNA cleanup kit (NEB T1130). The purified DNA was *in vitro* transcribed using

HiScribe T7 High Yield RNA Synthesis Kit (NEB E2040). The RNA transcribed was purified using an RNA cleanup kit (NEB T2030).

2.6.5 Plasmids Design and Sequences

Plasmid Design: Gene fragments with repetitive sequences were purchased from Thermo Fisher GeneArt. All DNA Ultramars and oligonucleotides were purchased from IDT.

~200nt RNA transcript with one TAG3 Sequence:

GACCCAAGCTGGCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCTCTAAC
ACAGACTCTCGGTACCATCATTTTCATATCCCCGGGAGCAGACTGTAAATCTGCTCC
CACCACCATCATTTAATGAATTCCATCAGGAATCCCTCACTTCTGCAGACTGGCCGT
CGTTTTACACTCGAGTCTAGA

~1000nt RNA transcript with one TAG3 Sequence:

ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTT
CAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGG
GCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGT
GGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAG
GCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAG
GGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGAC
CCAGGACTCCTCCCTGCAAGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCA
CCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACGATGGGCTGGGAGGCC
TCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAG
GCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGG
CCAAGAAGCCCGTGCAGCTGCCCCGGCGCCTACAACGTGAACATCAAGTTGGACATC

ACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCG
CCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAATATCTAACACAGACTCTCG
GTACCATCATTTTCATATCCCCGGGAGCAGACTGTAAATCTGCTCCCACCACCATCA
TTTAATGAATTCCATCAGGAATCCCTCACTTAAAGCCCCGCCGAAAGGCGGGCTTTTC
TGTGTCTGCAGACTGGCCGTCGTTTTACACTCGAGCATGCATCT

Strep-TGT(WT)-His Coding Sequence:

ATGTGGAGCCATCCGCAGTTTGAAAAAGGCAAATTTGAACTGGACACCACCGACGG
TCGCGCACGCCGTGGCCGCCTGGTCTTTGATCGTGGCGTAGTGGAACGCCTTGTTT
TATGCCTGTTGGCACCTACGGCACCGTAAAAGGGATGACGCCGGAAGAAGTTGAAG
CCACTGGCGCGCAAATTATCCTCGGCAACACCTTCCACCTGTGGCTGCGCCCGGGCC
AGGAAATCATGAACTGCACGGCGATCTGCACGATTTTATGCAGTGGAAGGGGCCG
ATCCTCACCGACTCCGGCGGCTTCCAGGTCTTCAGCCTTGCGGATATTCGTAAAATC
ACCGAACAGGGCGTGCCTTCCGTAACCCGATCAACGGCGATCCGATTTTCCTCGAT
CCTGAAAAATCAATGGAGATTCAGTACGATCTTGGTTCGGATATCGTCATGATCTTT
GATGAGTGTACGCCGTATCCTGCTGACTGGGATTACGCAAAACGCTCCATGGAGATG
TCTCTGCGTTGGGCGAAGCGTAGCCGTGAGCGTTTTGACAGTCTCGGAAACAAAAAT
GCGCTGTTTGGTATCATCCAGGGCAGCGTTTACGAAGATTTACGTGATATTTCTGTTA
AAGGTCTGGTAGATATCGGTTTTGATGGCTACGCTGTCGGCGGTCTGGCTGTGGGTG
AGCCGAAAGCAGATATGCACCGCATTCTGGAGCATGTATGCCCCGCAAATTCCGGCA
GACAAACCGCGTTACCTGATGGGCGTTGGTAAACCAGAAGACCTGGTTGAAGGCGT
ACGTCGTGGTATCGATATGTTTGACTGCGTAATGCCAACCCGCAACGCCCCGAAATGG
TCATTTGTTTCGTGACCGATGGCGTGGTGAAAATCCGCAATGCGAAGTATAAGAGCG
ATACTGGCCCACTCGATCCTGAGTGTGATTGCTACACCTGTCGCAATTATTCACGCG

CTTACTTGCATCATCTTGACCGTTGCAACGAAATATTAGGCGCGCGACTCAACACCA
TTCATAACCTTCGTTACTACCAGCGTTTGATGGCGGGTTTACGCAAGGCTATTGAAG
AGGGTAAATTAGAGAGCTTCGTAACCTGATTTTTACCAGCGTCAGGGGCGAGAAGTA
CCACCTTTGAACGTTGATCACCATCACCACCATCACTAA

Strep-TGT(D264E)-His Coding Sequence:

ATGTGGAGCCATCCGCAGTTTGAAAAAGGCAAATTTGAACTGGACACCACCGACGG
TCGCGCACGCCGTGGCCGCCTGGTCTTTGATCGTGGCGTAGTGGAACGCCTTGTTT
TATGCCTGTTGGCACCTACGGCACCGTAAAAGGGATGACGCCGGAAGAAGTTGAAG
CCACTGGCGCGCAAATTATCCTCGGCAACACCTTCCACCTGTGGCTGCGCCCCGGGCC
AGGAAATCATGAAACTGCACGGCGATCTGCACGATTTTATGCAGTGGAAGGGGCCG
ATCCTCACCGACTCCGGCGGGCTTCCAGGTCTTCAGCCTTGCGGATATTCGTAAAATC
ACCGAACAGGGCGTGCACTTCCGTAACCCGATCAACGGCGATCCGATTTTCCTCGAT
CCTGAAAAATCAATGGAGATTCAGTACGATCTTGGTTCGGATATCGTCATGATCTTT
GATGAGTGTACGCCGTATCCTGCTGACTGGGATTACGCAAAACGCTCCATGGAGATG
TCTCTGCGTTGGGCGAAGCGTAGCCGTGAGCGTTTTGACAGTCTCGGAAACAAAAAT
GCGCTGTTTGGTATCATCCAGGGCAGCGTTTACGAAGATTTACGTGATATTTCTGTTA
AAGGTCTGGTAGATATCGGTTTTGATGGCTACGCTGTCGGCGGTCTGGCTGTGGGTG
AGCCGAAAGCAGATATGCACCGCATTCTGGAGCATGTATGCCCGCAAATTCCGGCA
GACAAACCGCGTTACCTGATGGGCGTTGGTAAACCAGAAGACCTGGTTGAAGGCGT
ACGTCGTGGTATCGATATGTTTGAATGCGTAATGCCAACCCGCAACGCCCGAAATGG
TCATTTGTTTCGTGACCGATGGCGTGGTGAAAATCCGCAATGCGAAGTATAAGAGCG
ATACTGGCCCACTCGATCCTGAGTGTGATTGCTACACCTGTCGCAATTATTCACGCG
CTTACTTGCATCATCTTGACCGTTGCAACGAAATATTAGGCGCGCGACTCAACACCA

TTCATAACCTTCGTTACTACCAGCGTTTGATGGCGGGTTTACGCAAGGCTATTGAAG
AGGGTAAATTAGAGAGCTTCGTAACTGATTTTTACCAGCGTCAGGGGCGAGAAGTA
CCACCTTTGAACGTTGATCACCATCACCACCATCACTAA

Strep-GFP-TGT(WT)-His Coding Sequence:

ATGTGGAGCCATCCGCAGTTTGAAAAACGTAAAGGAGAAGAACTTTTCACTGGTGT
CGTCCCTATTCTGGTGGAACTGGATGGTGATGTCAACGGTCATAAGTTTTCCGTGCG
TGGCGAGGGTGAAGGTGACGCAACTAATGGTAAACTGACGCTGAAGTTCATCTGTA
CTACTGGTAAACTGCCGGTACCTTGGCCGACTCTGGTAACGACGCTGACTTATGGTG
TTCAGTGCTTTGCTCGTTATCCGGACCATATGAAGCAGCATGACTTCTTCAAGTCCGC
CATGCCGGAAGGCTATGTGCAGGAACGCACGATTTCTTTAAGGATGACGGCACGT
ACAAAACGCGTGCGGAAGTGAAATTTGAAGGCGATACCCTGGTAAACCGCATTGAG
CTGAAAGGCATTGACTTTAAAGAAGATGGCAATATCCTGGGCCATAAGCTGGAATA
CAATTTTAACAGCCACAATGTTTACATCACCGCCGATAAACAAAAAAATGGCATTAA
AGCGAATTTTAAAATTCGCCACAACGTGGAGGATGGCAGCGTGCAGCTGGCTGATC
ACTACCAGCAAAACACTCCAATCGGTGATGGTCCTGTTCTGCTGCCAGACAATCACT
ATCTGAGCACGCAAAGCGTTCTGTCTAAAGATCCGAACGAGAAACGCGATCATATG
GTTCTGCTGGAGTTCGTAACCGCAGCGGGCATCACGCATGGTATGGATGAACTGTAC
AAAGGTGGAGGTGGAAGTGGCGGTGGAGGCTCAGGCAAATTTGAACTGGACACCAC
CGACGGTCGCGCACGCCGTGGCCGCCTGGTCTTTGATCGTGGCGTAGTGGAACGCC
TTGTTTTATGCCTGTTGGCACCTACGGCACCGTAAAAGGGATGACGCCGGAAGAAGT
TGAAGCCACTGGCGCGCAAATTATCCTCGGCAACACCTTCCACCTGTGGCTGCGCCC
GGGCCAGGAAATCATGAAACTGCACGGCGATCTGCACGATTTTATGCAGTGGAAGG
GGCCGATCCTCACCGACTCCGGCGGGCTTCCAGGTCTTCAGCCTTGGCGATATTCGTA

AAATCACCGAACAGGGCGTGCACTTCCGTAACCCGATCAACGGCGATCCGATTTTCC
TCGATCCTGAAAAATCAATGGAGATTCAGTACGATCTTGGTTCGGATATCGTCATGA
TCTTTGATGAGTGTACGCCGTATCCTGCTGACTGGGATTACGCAAAACGCTCCATGG
AGATGTCTCTGCGTTGGGCGAAGCGTAGCCGTGAGCGTTTTGACAGTCTCGGAAACA
AAAATGCGCTGTTTGGTATCATCCAGGGCAGCGTTTACGAAGATTTACGTGATATTT
CTGTTAAAGGTCTGGTAGATATCGGTTTTGATGGCTACGCTGTCGGCGGTCTGGCTG
TGGGTGAGCCGAAAGCAGATATGCACCGCATTCTGGAGCATGTATGCCCCGCAAATT
CCGGCAGACAAACCGCGTTACCTGATGGGCGTTGGTAAACCAGAAGACCTGGTTGA
AGGCGTACGTCGTGGTATCGATATGTTTGACTGCGTAATGCCAACCCGCAACGCCCCG
AAATGGTCATTTGTTTCGTGACCGATGGCGTGGTGAAAATCCGCAATGCGAAGTATAA
GAGCGATACTGGCCCACTCGATCCTGAGTGTGATTGCTACACCTGTCGCAATTATTC
ACGCGCTTACTTGCATCATCTTGACCGTTGCAACGAAATATTAGGCGCGCGACTCAA
CACCATTCATAACCTTCGTTACTACCAGCGTTTGATGGCGGGTTTACGCAAGGCTAT
TGAAGAGGGTAAATTAGAGAGCTTCGTAAGTATTTTACCAGCGTCAGGGGCGAG
AAGTACCACCTTTGAACGTTGATCACCATCACCACCATCACTAA

Strep-GFP-TGT(D264E)-His Coding Sequence:

ATGTGGAGCCATCCGCAGTTTGAAAAACGTAAAGGAGAAGAACTTTTCACTGGTGT
CGTCCCTATTCTGGTGGAAGTGGATGGTGATGTCAACGGTCATAAGTTTTCCGTGCG
TGGCGAGGGTGAAGGTGACGCAACTAATGGTAAACTGACGCTGAAGTTCATCTGTA
CTACTGGTAAACTGCCGGTACCTTGGCCGACTCTGGTAACGACGCTGACTTATGGTG
TTCAGTGCTTTGCTCGTTATCCGGACCATATGAAGCAGCATGACTTCTTCAAGTCCGC
CATGCCGGAAGGCTATGTGCAGGAACGCACGATTTCTTTAAGGATGACGGCACGT
ACAAAACGCGTGCGGAAGTGAAATTTGAAGGCGATACCCTGGTAAACCGCATTGAG

CTGAAAGGCATTGACTTTAAAGAAGATGGCAATATCCTGGGCCATAAGCTGGAATA
CAATTTTAACAGCCACAATGTTTACATCACCGCCGATAAACAAAAAATGGCATTAA
AGCGAATTTTAAAATTCGCCACAACGTGGAGGATGGCAGCGTGCAGCTGGCTGATC
ACTACCAGCAAAACACTCCAATCGGTGATGGTCCTGTTCTGCTGCCAGACAATCACT
ATCTGAGCACGCAAAGCGTTCTGTCTAAAGATCCGAACGAGAAACGCGATCATATG
GTTCTGCTGGAGTTCGTAACCGCAGCGGGCATCACGCATGGTATGGATGAACTGTAC
AAAGGTGGAGGTGGAAGTGGCGGTGGAGGCTCAGGCAAATTTGAACTGGACACCAC
CGACGGTCGCGCACGCCGTGGCCGCCTGGTCTTTGATCGTGGCGTAGTGGAACGCC
TTGTTTTATGCCTGTTGGCACCTACGGCACCGTAAAAGGGATGACGCCGGAAGAAGT
TGAAGCCACTGGCGCGCAAATTATCCTCGGCAACACCTTCCACCTGTGGCTGCGCCC
GGGCCAGGAAATCATGAAACTGCACGGCGATCTGCACGATTTTATGCAGTGGAAGG
GGCCGATCCTCACCGACTCCGGCGGCTTCCAGGTCTTCAGCCTTGGCGATATTCGTA
AAATCACCGAACAGGGCGTGCACCTCCGTAACCCGATCAACGGCGATCCGATTTTCC
TCGATCCTGAAAAATCAATGGAGATTCAGTACGATCTTGGTTCGGATATCGTCATGA
TCTTTGATGAGTGTACGCCGTATCCTGCTGACTGGGATTACGCAAAACGCTCCATGG
AGATGTCTCTGCGTTGGGCGAAGCGTAGCCGTGAGCGTTTTGACAGTCTCGGAAACA
AAAATGCGCTGTTTGGTATCATCCAGGGCAGCGTTTACGAAGATTTACGTGATATTT
CTGTAAAGGTCTGGTAGATATCGGTTTTGATGGCTACGCTGTCGGCGGTCTGGCTG
TGGGTGAGCCGAAAGCAGATATGCACCGCATTCTGGAGCATGTATGCCCCGCAAATT
CCGGCAGACAAACCGCGTTACCTGATGGGCGTTGGTAAACCAGAAGACCTGGTTGA
AGGCGTACGTCGTGGTATCGATATGTTTGAATGCGTAATGCCAACCCGCAACGCCCCG
AAATGGTCATTTGTTCTGTGACCGATGGCGTGGTGAAAATCCGCAATGCGAAGTATAA
GAGCGATACTGGCCCACTCGATCCTGAGTGTGATTGCTACACCTGTCGCAATTATTC

ACGCGCTTACTTGCATCATCTTGACCGTTGCAACGAAATATTAGGCGCGCGACTCAA
CACCATTTCATAACCTTCGTTACTACCAGCGTTTGATGGCGGGTTTACGCAAGGCTAT
TGAAGAGGGTAAATTAGAGAGCTTCGTAAGTATTTTACCAGCGTCAGGGGCGAG
AAGTACCACCTTTGAACGTTGATCACCATCACCACCATCACTAA

Strep-TGT(WT)-GFP-His Coding Sequence:

ATGTGGAGCCATCCGCAGTTTGAAAAAGGCAAATTTGAACTGGACACCACCGACGG
TCGCGCACGCCGTGGCCGCCTGGTCTTTGATCGTGGCGTAGTGGAACGCCTTGTTT
TATGCCTGTTGGCACCTACGGCACCGTAAAAGGGATGACGCCGGAAGAAGTTGAAG
CCACTGGCGCGCAAATTATCCTCGGCAACACCTTCCACCTGTGGCTGCGCCCCGGGCC
AGGAAATCATGAAACTGCACGGCGATCTGCACGATTTTATGCAGTGGAAGGGGCCG
ATCCTCACCGACTCCGGCGGCTTCCAGGTCTTCAGCCTTGCGGATATTCGTAAAATC
ACCGAACAGGGCGTGCACCTCCGTAACCCGATCAACGGCGATCCGATTTTCCTCGAT
CCTGAAAAATCAATGGAGATTCAGTACGATCTTGGTTCGGATATCGTCATGATCTTT
GATGAGTGTACGCCGTATCCTGCTGACTGGGATTACGCAAAACGCTCCATGGAGATG
TCTCTGCGTTGGGCGAAGCGTAGCCGTGAGCGTTTTGACAGTCTCGGAAACAAAAAT
GCGCTGTTTGGTATCATCCAGGGCAGCGTTTACGAAGATTTACGTGATATTTCTGTTA
AAGGTCTGGTAGATATCGGTTTTGATGGCTACGCTGTCGGCGGTCTGGCTGTGGGTG
AGCCGAAAGCAGATATGCACCGCATTCTGGAGCATGTATGCCCGCAAATTCCGGCA
GACAAACCGCGTTACCTGATGGGCGTTGGTAAACCAGAAGACCTGGTTGAAGGCGT
ACGTCGTGGTATCGATATGTTTGACTGCGTAATGCCAACCCGCAACGCCCGAAATGG
TCATTTGTTTCGTGACCGATGGCGTGGTGAAAATCCGCAATGCGAAGTATAAGAGCG
ATACTGGCCCACTCGATCCTGAGTGTGATTGCTACACCTGTCGCAATTATTCACGCG
CTTACTTGCATCATCTTGACCGTTGCAACGAAATATTAGGCGCGCGACTCAACACCA

TTCATAACCTTCGTTACTACCAGCGTTTGATGGCGGGTTTACGCAAGGCTATTGAAG
AGGGTAAATTAGAGAGCTTCGTAACTGATTTTTACCAGCGTCAGGGGCGAGAAGTA
CCACCTTTGAACGTTGATGGTGGAGGTGGAAGTGGCGGTGGAGGCTCACGTAAAGG
AGAAGAACTTTTCACTGGTGTCGTCCCTATTCTGGTGGAACTGGATGGTGTGTCAA
CGGTCATAAGTTTTCCGTGCGTGGCGAGGGTGAAGGTGACGCAACTAATGGTAAAC
TGACGCTGAAGTTCATCTGTACTACTGGTAAACTGCCGGTACCTTGGCCGACTCTGG
TAACGACGCTGACTTATGGTGTTCAGTGCTTTGCTCGTTATCCGGACCATATGAAGC
AGCATGACTTCTTCAAGTCCGCCATGCCGGAAGGCTATGTGCAGGAACGCACGATTT
CCTTTAAGGATGACGGCACGTACAAAACGCGTGCGGAAGTGAAATTTGAAGGCGAT
ACCCTGGTAAACCGCATTGAGCTGAAAGGCATTGACTTTAAGAAGATGGCAATAT
CCTGGGCCATAAGCTGGAATACAATTTTAACAGCCACAATGTTTACATCACCGCCGA
TAAACAAAAAATGGCATTAAAGCGAATTTTAAAATTCGCCACAACGTGGAGGATG
GCAGCGTGCAGCTGGCTGATCACTACCAGCAAAACACTCCAATCGGTGATGGTCCT
GTTCTGCTGCCAGACAATCACTATCTGAGCACGCAAAGCGTTCTGTCTAAAGATCCG
AACGAGAAACGCGATCATATGGTTCTGCTGGAGTTCGTAACCGCAGCGGGCATCAC
GCATGGTATGGATGAACTGTACAAACACCATCACCACCATCACTAA

Strep-TGT(D264E)-GFP-His Coding Sequence:

ATGTGGAGCCATCCGCAGTTTGAAAAAGGCAAATTTGAACTGGACACCACCGACGG
TCGCGCACGCCGTGGCCGCCTGGTCTTTGATCGTGGCGTAGTGGAACGCCTTGTTT
TATGCCTGTTGGCACCTACGGCACCGTAAAAGGGATGACGCCGGAAGAAGTTGAAG
CCACTGGCGCGCAAATTATCCTCGGCAACACCTTCCACCTGTGGCTGCGCCCGGGCC
AGGAAATCATGAACTGCACGGCGATCTGCACGATTTTATGCAGTGGAAGGGGCCG
ATCCTCACCGACTCCGGCGGCTTCCAGGTCTTCAGCCTTGGCGATATTCGTAAAATC

ACCGAACAGGGCGTGCACTTCCGTAACCCGATCAACGGCGATCCGATTTTCCTCGAT
CCTGAAAAATCAATGGAGATTCAGTACGATCTTGGTTCGGATATCGTCATGATCTTT
GATGAGTGTACGCCGTATCCTGCTGACTGGGATTACGCAAAACGCTCCATGGAGATG
TCTCTGCGTTGGGCGAAGCGTAGCCGTGAGCGTTTTGACAGTCTCGGAAACAAAAAT
GCGCTGTTTGGTATCATCCAGGGCAGCGTTTACGAAGATTTACGTGATATTTCTGTTA
AAGGTCTGGTAGATATCGGTTTTGATGGCTACGCTGTCGGCGGTCTGGCTGTGGGTG
AGCCGAAAGCAGATATGCACCGCATTCTGGAGCATGTATGCCCCGCAAATTCCGGCA
GACAAACCGCGTTACCTGATGGGCGTTGGTAAACCAGAAGACCTGGTTGAAGGCGT
ACGTCGTGGTATCGATATGTTTGAATGCGTAATGCCAACCCGCAACGCCCCGAAATGG
TCATTTGTTTCGTGACCGATGGCGTGGTGAAAATCCGCAATGCGAAGTATAAGAGCG
ATACTGGCCCCACTCGATCCTGAGTGTGATTGCTACACCTGTCGCAATTATTCACGCG
CTTACTTGCATCATCTTGACCGTTGCAACGAAATATTAGGCGCGCGACTCAACACCA
TTCATAACCTTCGTTACTACCAGCGTTTGATGGCGGGTTTACGCAAGGCTATTGAAG
AGGGTAAATTAGAGAGCTTCGTAACTGATTTTTTACCAGCGTCAGGGGCGAGAAGTA
CCACCTTTGAACGTTGATGGTGGAGGTGGAAGTGGCGGTGGAGGCTCACGTAAAGG
AGAAGA ACTTTTTACTGGTGTCTGTCCTATTCTGGTGGAACTGGATGGTGTATGTCAA
CGGTCATAAGTTTTCCGTGCGTGGCGAGGGTGAAGGTGACGCAACTAATGGTAAAC
TGACGCTGAAGTTCATCTGTACTACTGGTAAACTGCCGGTACCTTGGCCGACTCTGG
TAACGACGCTGACTTATGGTGTTCAGTGCTTTGCTCGTTATCCGGACCATATGAAGC
AGCATGACTTCTTCAAGTCCGCCATGCCGGAAGGCTATGTGCAGGAACGCACGATTT
CCTTTAAGGATGACGGCACGTACAAAACGCGTGCGGAAGTGAAATTTGAAGGCGAT
ACCCTGGTAAACCGCATTGAGCTGAAAGGCATTGACTTTAAAGAAGATGGCAATAT
CCTGGGCCATAAGCTGGAATACAATTTTAAACAGCCACAATGTTTACATCACCGCCGA

TAAACAAAAAATGGCATTAAAGCGAATTTTAAAATTTCGCCACAACGTGGAGGATG
GCAGCGTGCAGCTGGCTGATCACTACCAGCAAAACACTCCAATCGGTGATGGTCCT
GTTCTGCTGCCAGACAATCACTATCTGAGCACGCAAAGCGTTCTGTCTAAAGATCCG
AACGAGAAACGCGATCATATGGTTCTGCTGGAGTTCGTAACCGCAGCGGGCATCAC
GCATGGTATGGATGAACTGTACAAACACCATCACCACCATCACTAA

2.6.6 TGT Covalent Conjugate Reaction Set up

The TGT covalent conjugate conditions used were: 10 μ M RNA substrate, 1X TGT reaction buffer, 1U/ μ L Murine RNasin (NEB M0314) and 40 μ M TGT enzyme. TGT was the last component added to the reaction. Once all components were added they were gently mixed and briefly spun down. The reaction mixture was incubated for 1 hour in a thermocycler set to 37°C.

2.6.7 TGT Covalent Conjugate Reaction Urea PAGE Analysis

Urea PAGE gel shift analysis was the primary method of analysis for TGT labeling reactions. Following established protocols, a 6% denaturing Urea-TBE-PAGE gel was prepared using the SequaGel® UreaGel 19:1 Acrylamide: Bisacrylamide Denaturing Gel System. Gels were separated by the Bio-Rad Mini-Protean system. Approximately 100ng of control or reaction product was denatured with 2X RNA loading dye from New England Biosciences at 98°C for 5 mins. Loaded samples and ran at 100V for 15 minutes to allow the material to enter the gel followed by 200V for 60 minutes (analyzing 200nt RNA transcript) or 90 minutes (analyzing 1000nt RNA transcript) to separate starting material from labeled product. The gels were stained using gel red nucleic acid stain and imaged on a Bio-Rad ChemiDoc-MP gel imager. The gel shift was observed visually and quantified via densitometry using ImageJ software from NIH, an open-source image analysis program. An upward gel shift in the product lane signified an increase in molecular weight therefore indicating that the RNA substrate was successfully conjugated to TGT.

2.6.8 Addition of PEG 8000

Dissolved PEG 8000 (Promega V3011) in water to make a 20% stock solution and vortexed to make a homogenous solution. TGT covalent conjugation reaction used are as followed: 10 μ M RNA substrate, 1X TGT reaction buffer, 1U/ μ L Murine RNAsin, 6.4% PEG 8000 and 40 μ M TGT enzyme (TGT added last). The reaction was gently mixed and briefly spun down. The reaction mixture was incubated for 1 hour in a thermocycler set to 37°C. Reactions were analyzed using the same method mentioned in section 2.6.7.

2.6.9 Excess Guanine, preQ1 and Queuine Treatment

Made homogenous aqueous solutions of 500 μ M guanine, preQ1 and queuine. TGT covalent conjugate reaction conditions used: 10 μ M RNA substrate, 1X TGT reaction buffer, 1U/ μ L Murine RNAsin and 40 μ M TGT enzyme (TGT added last). Reaction was gently mixed and briefly spun down. The reaction mixture was incubated for 1 hour in a thermocycler set to 37°C. After reaction had been completed it was divided into four 2 μ L aliquots and to each aliquot 2 μ L of either 500 μ M guanine, 500 μ M preQ1, 500 μ M queuine or ddH₂O. They were incubated at 37°C for 30 minutes. Reactions were analyzed using the same method mentioned in section 2.6.7.

2.6.10 TGT covalent conjugate reaction with and without RNase Inhibitor

There were two identical sets of reactions set up. Both had the following reaction conditions: 10 μ M RNA substrate, 1X TGT reaction buffer and 40 μ M TGT enzyme (TGT added last). One set had 1U/ μ L Murine RNAsin added before the addition of TGT. Both sets of reactions were gently mixed and incubated for 1 hour at 37°C in thermocycler. Reactions were analyzed using the same method mentioned in section 2.6.7.

2.6.11 RNA Refolding

There were two identical sets of reactions set up using the same conditions mentioned in section 2.6.6. One set underwent RNA refolding before initiating the covalent conjugate reaction of 1hr at 37°C. The RNA refolding protocol required to only set up the RNA transcript in reaction buffer and water. The RNA transcript was heated to 98°C for 1 minute on the thermocycler and cooled at a rate of 1°C/sec until the reactions are reached 4°C. Once the RNA refolding thermocycler protocol was completed RNase inhibitor and TGT (last) were added. The second set of reactions was set up using the same protocol mentioned in section 2.6.6. Both set of reactions were then placed in the thermocycler set to 37°C for 1hr. Reactions were analyzed using the same method mentioned in section 2.6.7.

2.6.12 SDS Treatment

TGT covalent conjugate reaction was set up using the same conditions and protocol mentioned in section 2.6.6. During the TGT covalent conjugate reaction incubation time, an aqueous homogenous solution of 10% SDS was prepared. After reaction was completed, 10% SDS was added to the reaction to have a final concentration of 1% SDS. Reactions were analyzed using the same method mentioned in section 2.6.7.

2.6.13 BME Treatment

TGT covalent conjugate reaction was set up using the same conditions and protocol mentioned in section 2.6.6. After reaction had completed, 100% β -Mercaptoethanol (BME) was added to the reaction to have a final concentration of 5% BME. Reactions were analyzed using the same method mentioned in section 2.6.7.

2.6.14 GFP-TGTs with TAG3 no denaturing step

TGT covalent conjugate conditions used were: 20 μ M RNA substrate, 1X TGT reaction buffer, 1U/ μ L Murine RNAsin (NEB M0314) and 10 μ M TGT enzyme. Reactions were incubated at 37°C for 1 hour in a thermocycler. Reactions were analyzed using the same method mentioned in section 2.6.7 without heating the sample to 98°C.

2.7 Acknowledgements

Chapter 2, in full, is currently being reviewed for submission for publication. Guerrero, Karen; Chadwick, Stormi R.; Tota, Ember M.; Devaraj, Neal K. The dissertation author was a primary researcher and co-author of this paper.

Chapter 3 Cell work of TGT-RNA covalent conjugate

3.1 Assessment of plasmid GFP expression and toxicity

With the *in vitro* data obtained supporting the ability of TGT(D264E) to form stable covalent conjugates with RNA transcripts bearing the TGT recognition element, the next step was to decipher if the covalent complex could be formed in mammalian cells. The preliminary experiment before diving deep into cell experimental design was to assess the toxicity of exogenous *E. coli* TGT. We designed and prepared 2 different mammalian codon optimized GFP fused TGT plasmids: GFP-TGT(WT) and GFP-TGT(D264E) for transient transfection of HeLa cells. Additionally, a GFP only positive control was designed using the same plasmid backbone as the TGT fusions. Post transfection fluorescent signal and cell viability were monitored over 96 hours to observe changes in GFP expression and cell health. Fluorescence was observed amongst all constructs, including the positive control, at 96 hours post transfection (**Figure 3.1A**). In the *in vitro* studies, recombinantly expressed GFP-TGT fusion constructs retained TGT activity, and the observed fluorescence in the HeLa cells suggests that the fusion design did not disrupt GFP translation or folding. Additionally, the effects of the constructs on cell health were monitored using an XTT viability assay at 24, 48, 72 and 96 hours (**Figure 3.1B**). Vehicle transfection and Staurosporine-induced cell death conditions were included as controls. The results suggest the constructs are not toxic to HeLa cells for at least up to 96 hours. With this data in hand, the next step is to test the formation of TGT-RNA conjugates in mammalian cells.

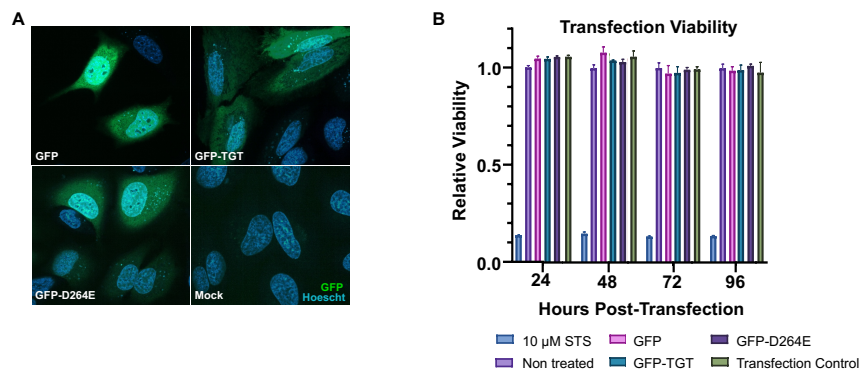


Figure 3.1: (A) GFP fluorescence persists after 96 hours (B) XTT viability assay suggest GFP fused TGTs are not toxic to HeLa cells.

3.2 Fluorescence Microscopy Studies

As an initial proof-of-concept, we adapted a previously established strategy used to optimize RNA-labeling systems such as MS2, in which multiple RNA-binding protein (RBP) recognition sites are inserted into an RNA of interest to recruit fluorescently fused RBPs to the transcript. A nuclear localization sequence (NLS) was fused to the RBP to retain unbound protein in the nucleus to reduce background fluorescent signal, thereby confining cytoplasmic fluorescence to RBPs specifically bound to the target RNA.^{9,11,42} Inspired by this method, we utilized fluorescence microscopy to visualize GFP-labeled TGT enzymes bound to multiple TAG3 recognition sites encoded at the 3' end of an RNA of interest. An NLS was fused to N terminus of GFP-TGT, and ten TAG3 hairpins were incorporated into the 3' untranslated region (UTR) of an mCherry transcript to minimize disruption of RNA function. We expected to see unbound GFP-TGT fusions remain confined in the nucleus while TGT-RNA covalent conjugates get exported via RNA transport mechanisms (**Figure 3.2**).⁴³ Since our *in vitro* data indicates that the TGT(D264E) mutant forms stronger covalent linkage with an RNA transcript than TGT(WT), we expected cytoplasmic translocation solely to covalent conjugates formed with the mutant fusion

protein. Interestingly, both GFP-TGT(WT) and GFP-TGT(D264E) fusions translocated from the nucleus to the cytoplasm in the presence of the RNA transcript bearing TAG3 hairpins (**Figure 3.3**). Quantification of cytosolic vs nucleus GFP fluorescence signal from the microscopy images via cytoplasmic-to-nuclear fluorescence intensity ratio (CNI) reveal greater nuclear export of TGT(D264E)-RNA conjugates relative to wild-type TGTs (**Figure 3.4**). These results suggest that stability is fundamental for complex integrity and thereby leading to greater translocation from the nucleus due to its high-affinity and multivalent TGT-substrate interactions.

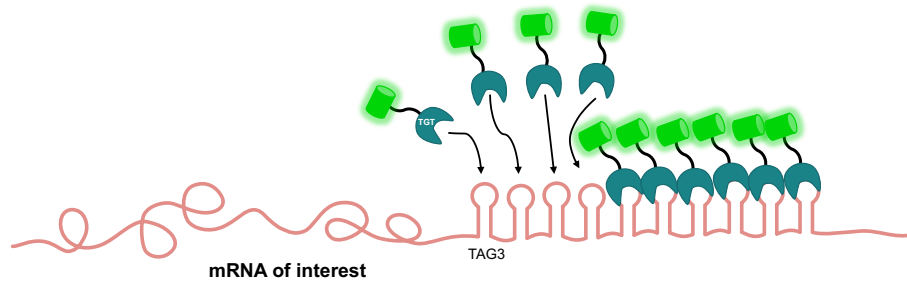


Figure 3.2: Cartoon demonstrating formation of encodable covalent RNA-protein conjugation in cells.

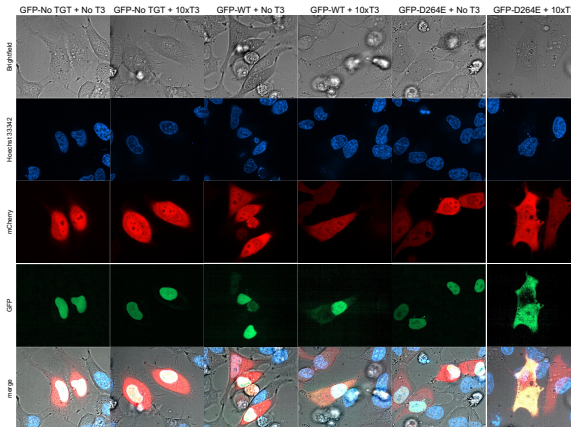


Figure 3.3: Fluorescence microscopy of 24-hour transient transfection of GFP-TGT (WT) and mutant fusion. Green signal is from GFP. Blue, fluorescent signal is from Hoechst 33342. Red signal is from mCherry.

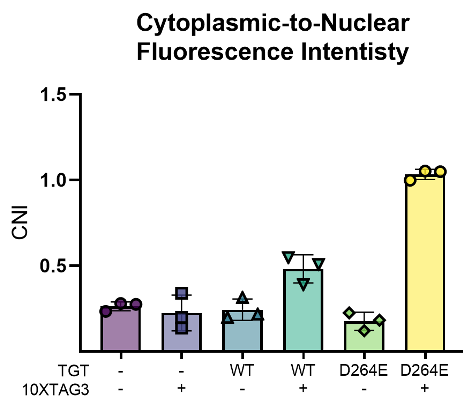


Figure 3.4: Translocation of GFP-TGT variants from the nucleus to the cytoplasm in the presence of TAG3-containing transcripts is quantified as the cytoplasmic-to-nuclear fluorescence intensity ratio (CNI). The mean cytoplasmic fluorescence is divided by the mean nuclear fluorescence for each cell; error bars represent standard deviation of biological triplicates.

3.3 Immunoblot analysis of cellular TGT-RNA conjugation

Given that both TGT(WT) and TGT(D264E) conjugates exported from the nucleus into the cytoplasm, it was important to verify whether the complexes formed in the mammalian cellular environment are covalent. Therefore, to assess their covalency HeLa cells were transiently co-transfected with plasmids encoding (1) an mCherry open reading frame (ORF) bearing either one or ten TAG3 recognition hairpins in the 3' UTR and (2) an NLS-tagged GFP-TGT fusion (WT or D264E). After 48 hours, cell lysate underwent denaturing treatment, was resolved on a gel, and analyzed via immunoblotting to quantify a gel shift. An upward gel shift was observed in the sample co-transfected with the plasmids encoding for the RNA transcript bearing 10 TAG3 recognition elements and GFP-TGT(D264E), indicating the presence of a higher molecular weight species consistent with anticipated covalent complex, TGT(D264E)-RNA (**Figure 3.5**). On the contrary, no apparent gel shift was detected in the TGT(WT)-RNA sample lane indicating the wild-type conjugate is less stable than the mutant complex under denaturing conditions. There was no obvious higher molecular weight species observed in the samples treated with the RNA

transcript bearing a single TAG3 hairpin for either WT or D264E mutant fusions, suggesting a single TAG3 hairpin is not sufficient to form an ample amount of complex to be detected by immunoblotting. In summary, these findings are consistent with the known tendency of TGT to dimerize during binding to its RNA substrate, reiterating the importance of including multiple TAG3 hairpins to enhance conjugate formation and stability, and increase detection sensitivity in a complex cellular environment.²⁰

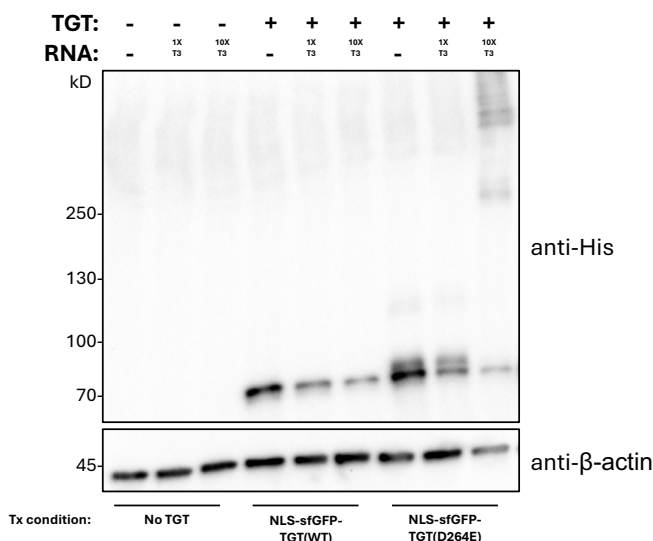


Figure 3.5: Western blot analysis of His-Tagged TGT variants co-expressed with TAG3-containing transcripts in HeLa cells using His Tag antibody (upper) and β-actin antibody (lower).

3.4 Conclusion

Due to the rising interest in RNA research, there is a need for the advancement of an innovative strategy to assemble covalently linked RNA-protein conjugates in mammalian cells. Existing methodologies to form RNA-protein complexes are dependent on noncovalent interactions that are unable to maintain association under denaturing conditions. Furthermore, technologies that covalently link these complexes capture the interactions after cell lysis rather than in real-time within living cells and rely heavily on UV irradiation, which can cause structural

alterations to the species. In this work, we demonstrate that our TGT-RNA technology can overcome these limitations, demonstrating covalent conjugation stability via fluorescence microscopy and immunoblotting. With the results outlined above, we believe our technology could be utilized for various applications and expand the scientific understanding of RNA function and utility.

3.5 Materials and Methods

3.5.1 Plasmid preparation for mammalian studies

All plasmids were prepared as previously mentioned in section 2.6.2.

mCherry RNA with 10 TAG3 plasmid: A previously reported plasmid (Addgene 40651) was modified by replacing the existing transcript with one encoding for mCherry and 10 TAG3 recognition sites.⁴²

NLS-GFP-TGT fusion plasmids: A previously reported plasmid (Addgene 40649) was modified by removing the MS2 bacteriophage coat protein (MCP) and adding TGT.⁴² An N terminal nuclear localization sequence and C terminal 6xHis purification tag was added via Gibson assembly (New England Biolabs).

mCherry RNA transcript with one TAG3 Sequence:

ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTT
CAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGG
GCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGT
GGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAG
GCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAG
GGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGAC

CCAGGACTCCTCCCTGCAAGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCA
CCAAC TTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACGATGGGCTGGGAGGCC
TCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAG
GCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGG
CCAAGAAGCCCGTGCAGCTGCCCCGGCGCCTACAACGTGAACATCAAGTTGGACATC
ACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCG
CCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAATATCTAACACAGACTCTCG
GTACCATCATTTTCATATCCCCGGGAGCAGACTGTAAATCTGCTCCCACCACCATCA
TTTAATGAATTCCATCAGGAATCCCTCACTTAAAGCCCGCCGAAAGGCGGGCTTTTC
TGTGTCTGCAGACTGGCCGTCGTTTTACACTCGAGCATGCATCT

mCherry RNA transcript with 10 TAG3 Sequence:

ATGGTGAGCAAGGGCGAGGAGGACAACATGGCCATCATCAAGGAGTTCATGCGCTT
CAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGG
GCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGG
CGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAA
GGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGA
GGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGA
CCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGC
ACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGC
CTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGA
GGCTGAAGCTGAAGGACGGCGGCCACTACGACGCCGAGGTCAAGACCACCTACAAG
GCCAAGAAGCCCGTGCAGCTGCCCCGGCGCCTACAACGTCAACATCAAGCTGGACAT
CACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAGCGCGCCGAGGGCC

GCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGGATCCTAAGG
TACCTAATTGCCTAGAAAGGGAGCAGACTGTAAATCTGCTCCCCTGCAGGTCGACTC
TAGAAAGGGAGCAGACTGTAAATCTGCTCCCCTGCAGTATTCCCGGGTTCATTAGAT
CCTAAGGTACCTAATTGCCTAGAAAGGGAGCAGACTGTAAATCTGCTCCCCTGCAGG
TCGACTCTAGAAAGGGAGCAGACTGTAAATCTGCTCCCCTGCAGTATTCCCGGGTTC
ATTAGATCCTAAGGTACCTAATTGCCTAGAAAGGGAGCAGACTGTAAATCTGCTCCC
CTGCAGGTCGACTCTAGAAAGGGAGCAGACTGTAAATCTGCTCCCCTGCAGTATTCC
CGGGTTCATTAGATCCTAAGGTACCTAATTGCCTAGAAAGGGAGCAGACTGTAAATC
TGCTCCCCTGCAGGTCGACTCTAGAAAGGGAGCAGACTGTAAATCTGCTCCCCTGCA
GTATTCCCGGGTTCATTAGATCCTAAGGTACCTAATTGCCTAGAAAGGGAGCAGACT
GTAAATCTGCTCCCCTGCAGGTCGACTCTAGAAAGGGAGCAGACTGTAAATCTGCTC
CCCTGCAGTATTCCCGGGTTCATTAGATCCTAAGGTACCTAATTGCCTAGAAAACAT
GAGGATCACCCATGT

NLS-GFP-TGT(WT) sequence:

ATGGGCCCCAAAAAGAAAAGAAAAGTTGGCTACCCCTACGACGTGCCCGACTACGC
CATCGAAGGCCGCCATATGAGAAAGGGCGAAGAGCTGTTACCGGTGTGGTTCCAA
TCCTTGTGGAGCTCGACGGCGACGTCAACGGACACAAGTTTTTCAGTGCGCGGTGAG
GGAGAGGGCGACGCTACTAACGGAAAACCTACCCTGAAGTTCATCTGTACTACCGG
AAAACCTGCCCCGTGCCCTGGCCAACTCTGGTGACTACCCTGACCTATGGAGTGCAATG
TTTTGCTAGATACCCGGATCACATGAAGCAGCATGATTTTTTTTAAGAGCGCCATGCC
AGAGGGCTACGTGCAGGAACGCACAATAAGCTTCAAAGATGATGGGACGTACAAAA
CAAGAGCAGAGGTGAAATTCGAGGGAGATACCCTTGTTAATCGGATCGAGCTTAAG
GGGATTGACTTCAAGGAGGACGGGAACATACTGGGGCACAAGCTGGAATATAACTT

CAATTCACATAACGTGTATATAACCGCCGACAAGCAGAAGAACGGCATTAAGGCTA
ATTTCAAGATTAGACACAACGTTGAGGACGGCTCAGTACAGTTAGCCGATCACTACC
AACAAAATACACCAATCGGGCGACGGGCCAGTCTTACTGCCGGACAATCACTACCTG
AGTACCCAGTCTGTTTTAAGTAAAGACCCTAACGAAAAAAGGGATCATATGGTTCTC
CTTGAATTTGTAAGTGTGCTGCCGGTATTACGCACGGGATGGACGAGCTCTACAAGGGT
GGAGGTGGAAGTGGCGGTGGAGGCTCAAAGTTTGAGCTGGATACTACTGATGGGCG
GGCTAGAAGAGGCAGACTGGTGTTTGATAGAGGCGTCGTCGAGACCCCCTGTTTTAT
GCCAGTGGGCACCTACGGCACAGTGAAGGGAATGACCCCAGAGGAGGTGGAGGCT
ACAGGCGCTCAGATCATCCTGGGCAACACCTTCCACCTGTGGCTGCGGCCTGGCCAG
GAGATCATGAAGCTGCACGGCGACCTGCACGACTTCATGCAGTGAAGGGCCCAAT
CCTGACCGACTCCGGCGGCTTCCAGGTGTTTTCTCTGGGCGACATCCGCAAGATCAC
AGAGCAGGGCGTGCACCTCCGGAACCCCATCAATGGCGACCCTATCTTCCTGGACCC
CGAGAAGAGCATGGAGATCCAGTATGACCTGGGCTCCGACATCGTGATGATCTTCG
ACGAGTGCACACCCTACCCTGCCGACTGGGATTATGCTAAGCGCAGCATGGAGATG
TCCCTGCGGTGGGCCAAGAGATCTCGCGAGCGGTTTCGATAGCCTGGGCAACAAGAA
TGCTCTGTTTGGCATCATCCAGGGCAGCGTGTACGAGGACCTGAGGGACATCAGCGT
GAAGGGCCTGGTGGACATCGGCTTTGATGGCTATGCTGTGGGAGGACTGGCTGTGG
GAGAGCCAAAGGCCGACATGCACAGAATCCTGGAGCACGTGTGCCACAGATCCCC
GCTGACAAGCCCAGATACCTGATGGGCGTGGGCAAGCCTGAGGATCTGGTGGAGGG
CGTGCGGAGGGGAATCGACATGTTTCGATTGCGTGATGCCTACCAGGAACGCCAGAA
ATGGCCACCTGTTTGTGACAGACGGCGTGGTGAAGATCCGCAATGCCAAGTACAAG
TCCGATACCGGCCCTCTGGACCCAGAGTGCGATTGTTATACATGCCGCAACTACTCT
CGGGCCTATCTGCACCACCTGGACAGGTGTAATGAGATCCTGGGCGCTAGACTGAA

CACCATCCACAATCTGAGGTACTATCAGAGACTGATGGCTGGACTGAGGAAGGCTA
TCGAGGAGGGCAAGCTGGAGAGCTTCGTGACAGACTTCTACCAGCGGCAGGGCAGA
GAGGTCCCACCACTGAATGTCGATCATCACCACCATCACCCTAA

NLS-GFP-TGT(D264E) sequence:

ATGGGCCCCAAAAAGAAAAGAAAAGTTGGCTACCCCTACGACGTGCCCGACTACGC
CATCGAAGGCCGCCATATGAGAAAGGGCGAAGAGCTGTTACCGGTGTGGTTCCAA
TCCTTGTGGAGCTCGACGGCGACGTCAACGGACACAAGTTTTTCAGTGCGCGGTGAG
GGAGAGGGCGACGCTACTAACGGAAAACCTACCCTGAAGTTCATCTGTACTACCGG
AAAACCTGCCCGTGCCCTGGCCAACTCTGGTGACTACCCTGACCTATGGAGTGCAATG
TTTTGCTAGATACCCGGATCACATGAAGCAGCATGATTTTTTTAAGAGCGCCATGCC
AGAGGGCTACGTGCAGGAACGCACAATAAGCTTCAAAGATGATGGGACGTACAAAA
CAAGAGCAGAGGTGAAATTCGAGGGAGATACCCTTGTTAATCGGATCGAGCTTAAG
GGGATTGACTTCAAGGAGGACGGGAACATACTGGGGCACAAGCTGGAATATAACTT
CAATTCACATAACGTGTATATAACCGCCGACAAGCAGAAGAACGGCATTAAGGCTA
ATTTCAAGATTAGACACAACGTTGAGGACGGCTCAGTACAGTTAGCCGATCACTACC
AACAAAATACACCAATCGGCGACGGGCCAGTCTTACTGCCGGACAATCACTACCTG
AGTACCCAGTCTGTTTTAAGTAAAGACCCTAACGAAAAAAGGGATCATATGGTTCTC
CTTGAATTTGTAACCTGCTGCCGGTATTACGCACGGGATGGACGAGCTCTACAAGGGT
GGAGGTGGAAGTGGCGGTGGAGGCTCAAAGTTTGAGCTGGATACTACTGATGGGCG
GGCTAGAAGAGGCAGACTGGTGTTTGATAGAGGCGTCGTCGAGACCCCCTGTTTTAT
GCCAGTGGGCACCTACGGCACAGTGAAGGGAATGACCCCAGAGGAGGTGGAGGCT
ACAGGCGCTCAGATCATCCTGGGCAACACCTTCCACCTGTGGCTGCGGCCTGGCCAG
GAGATCATGAAGCTGCACGGCGACCTGCACGACTTCATGCAGTGGAAGGGCCCAAT

CCTGACCGACTCCGGCGGCTTCCAGGTGTTTTCTCTGGGCGACATCCGCAAGATCAC
AGAGCAGGGCGTGCACTTCCGGAACCCCATCAATGGCGACCCTATCTTCCTGGACCC
CGAGAAGAGCATGGAGATCCAGTATGACCTGGGCTCCGACATCGTGATGATCTTCG
ACGAGTGCACACCCTACCCTGCCGACTGGGATTATGCTAAGCGCAGCATGGAGATG
TCCCTGCGGTGGGCCAAGAGATCTCGCGAGCGGTTCGATAGCCTGGGCAACAAGAA
TGCTCTGTTTGGCATCATCCAGGGCAGCGTGTACGAGGACCTGAGGGACATCAGCGT
GAAGGGCCTGGTGGACATCGGCTTTGATGGCTATGCTGTGGGAGGACTGGCTGTGG
GAGAGCCAAAGGCCGACATGCACAGAATCCTGGAGCACGTGTGCCACAGATCCCC
GCTGACAAGCCCAGATACCTGATGGGCGTGGGCAAGCCTGAGGATCTGGTGGAGGG
CGTGCGGAGGGGAATCGACATGTTCGAATGCGTGATGCCTACCAGGAACGCCAGAA
ATGGCCACCTGTTTGTGACAGACGGCGTGGTGAAGATCCGCAATGCCAAGTACAAG
TCCGATACCGGCCCTCTGGACCCAGAGTGCGATTGTTATACATGCCGCAACTACTCT
CGGGCCTATCTGCACCACCTGGACAGGTGTAATGAGATCCTGGGCGCTAGACTGAA
CACCATCCACAATCTGAGGTACTATCAGAGACTGATGGCTGGACTGAGGAAGGCTA
TCGAGGAGGGCAAGCTGGAGAGCTTCGTGACAGACTTCTACCAGCGGCAGGGCAGA
GAGGTCCCACCACTGAATGTCGATCATCACCACCATCACCCTAA

3.5.2 Image acquisition and analysis

For confocal analysis of GFP-TGT translocation, 100 ng each of TGT- and TAG3-containing plasmids were transiently transfected using Fugene HD transfection reagent (Promega E2311) into HeLa cells grown in 8-well chamber slides (Ibidi 80826). Transgenes were allowed to express for 24 hours before cells were stained with NucBlue Live reagent (Invitrogen R37605) for 20 minutes at 37°C and 5% CO₂. Live cell images were acquired on a Zeiss Z1 spinning-disk microscope and analyzed in ImageJ (<https://imagej.net/ij/>). Background was subtracted using the

rolling ball algorithm with sliding paraboloid. Nuclear ROIs were defined from the DAPI channel and whole-cell ROIs from the mCherry channel using automatic thresholding for each individual image. Cytoplasmic ROIs were generated by excluding the nuclear ROI from the whole-cell ROI using the XOR function. Mean GFP fluorescence intensity was measured in nuclear and cytoplasmic ROIs for individual cells, and cytoplasmic enrichment was expressed as the cytoplasmic-to-nuclear GFP intensity ratio (CNI).

3.5.3 Western Blot

For Western blot analyses of covalently modified RNA transcripts, 2 µg each of TGT- and TAG3-containing plasmids were transiently transfected using Fugene HD transfection reagent (Promega E2311) into HeLa cells grown in six-well dishes (ThermoFisher 140675). Transgenes were allowed to express for 48 hours before cells were collected by trypsinization with TrypLE (Gibco 12604013) followed by neutralization in complete growth media. The cells were pelleted at 200xg for 5 minutes, washed once with HBSS (Gibco 14-025-092), and re-pelleted. The pellets were resuspended in 150 µL RIPA lysis buffer (ThermoFisher 89900) containing 1x Halt Protease Inhibitor Cocktail (ThermoFisher 87786), 0.1 mM PMSF (ThermoFisher 215740010), and 0.02 U/µL SUPERase•In RNase Inhibitor (ThermoFisher AM2694). After resuspension in lysis buffer, the cells were incubated on ice for 15 minutes followed by sonication with a tip sonicator at 50% amplitude for 10 seconds. Lysed cells were incubated for another 5 minutes on ice before being spun at 14,000xg for 15 minutes at 4°C to separate the insoluble and soluble fractions. Protein concentrations of the lysate supernatants were determined via absorbance on a nanodrop instrument (ThermoFisher UX-83061-00). Lysate supernatants were mixed with one quarter volume of 4x Laemmli sample buffer (Bio-Rad 1610747) containing 10% BME (Gibco 21985023). 20 µg of protein material was separated per well by SDS-PAGE in 15-well, Mini-

PROTEAN TGX Precast Protein Gels (Bio-Rad 4561096) in 1x tris/glycine/SDS buffer (Bio-Rad 1610732). Samples were transferred to PVDF membranes (Bio-Rad) using the Trans-Blot Turbo Transfer System (Bio-Rad) and blocked for 2 h at room temperature in 5% BSA (Sigma A4737-1G) in TBS (Cell Signaling Technology). Membranes were incubated with primary antibodies in 5% BSA in TBST (Cell Signaling Technology 12498) overnight at 4°C with shaking as indicated in **Table 3.1**. Membranes were washed six times with TBST over the course of 30 min and then exposed to goat anti-mouse or goat anti-rabbit HRP-conjugated secondary antibody (Abcam; 1:2,000) in TBST with 5% BSA for 1 h at room temperature. Membranes were washed six times over the course of 30 min and exposed to HRP substrate (ECL Western Blotting Substrate, Pierce). Chemiluminescent signals were recorded using a ChemiDoc (Bio-Rad).

Table 3.1: Primary antibodies used in this study.

Antigen	Supplier	Catalog Number	Host Species	Dilution
β -actin	Cell Signaling Technology	3700S	mouse	1:1000
6XHIS	Cell Signaling Technology	12698S	rabbit	1:1000

3.6 Acknowledgements

Chapter 3, in full, is currently being reviewed for submission for publication. Guerrero, Karen; Chadwick, Stormi R.; Tota, Ember M.; Devaraj, Neal K. The dissertation author was a primary researcher and co-author of this paper.

Chapter 4 Conclusion

4.1 RNA Biology and its existing research tools

When the central dogma of molecular biology was proposed RNA was considered to be a key player in protein production, which led to the belief that coding for protein was RNA's main function.⁴⁴ Since the development of the central dogma, RNA has further shown to have diverse functions fundamental in cell biology such as gene expression regulation,⁴⁵ catalysis,⁴⁶ genome stability and maintenance,⁴⁷ DNA damage repair,⁴⁸ and the list goes on. Due to the significance RNA has demonstrated to have in cellular machinery, the interest to further study RNA has increased and as such the need for tools to observe it in its natural environment is crucial. Additionally, novel strategies for RNA visualization and manipulation have become increasingly important due to a drastic expansion in the demand for RNA research and tool development.

In addition to its endogenous biological functions, the use of RNA in chemical biology applications is gaining recognition and momentum due to its unmatched versatility and programmability. By using its unparalleled coding ability to channel an intended function it has influenced the innovation of transformative technologies such as mRNA vaccines,⁴ fluorescent RNA aptamers,⁴⁹ CRISPR-Cas systems,^{50,51} antisense oligonucleotides like siRNA,⁵² and RNA fluorescence in situ hybridization.⁵³

An avenue to enhance the development of more groundbreaking technologies is to improve visualization and manipulation tools to facilitate the study of RNA. There are various labeling strategies commonly used in RNA research for visualization but are often found to have limitations such as the need of chemical modification, lack of stability or low specificity *in vitro* or in mammalian cells.^{12, 29} Consequently, there is still a demand for an RNA labeling system that surpasses these limitations. Driven by the need of more robust technologies, our lab found the

motivation to develop a robust RNA transglycosylation at guanine (RNA-TAG) system that enzymatically, site-specifically label RNAs with functional handles.³⁰⁻³⁸ One developed application of RNA-TAG facilitates the formation of a covalent RNA-protein conjugate via preQ1-benzylguanine insertion followed by treatment with a SNAP-tagged protein of interest.³⁹ However, this is a dual-step strategy requiring prior *in vitro* insertion of the benzylguanine handle by TGT. The technology presented here eliminates the necessity of the prior small molecule insertion, facilitating an encodable, single step, direct covalent attachment of a functional TGT fusion to an RNA of interest.

4.2 Future Direction of Covalent RNA-Protein Conjugates

The development of technologies that facilitate the assembly of complex biological macromolecular systems has been a productive area of study due to the unique, specific, and indispensable roles they naturally play in biological processes. These technologies have enabled innovation in research, diagnostic technologies, and therapeutic developments. Among biomolecules, enzymes are widely employed for research and technology development due to their uniquely specific activities.⁵⁴ Additionally, the programmability of nucleic acids remains unmatched. The coupling of enzymatic function with this programmability has paved the way for revolutionary tools in modern molecular biology including nucleic acid sequencing, gene editing, cloning, and nucleic acid labeling.⁵⁵ In light of these advancements, there is a strong motivation for the development of covalent nucleic acid-protein conjugates to improve existing technologies.

Currently there are fully encodable engineered RNA binding protein systems that can form RNA-protein associations in a single-step. While these RBP technologies are valuable for protein recruitment, they are dependent on non-covalent interactions, which could be problematic if the interaction or structure is disrupted. The focus of this work was to develop a platform that couples

the benefits of a fully encodable modification system, similar to existing RBP approaches, with the robustness of covalent conjugation. Due to the simplicity of the covalent complex formation mechanism between a TGT and its recognition site, we hypothesized that by increasing the stability of TGT-RNA covalent complex we could surpass existing limitations of nucleic acid-protein conjugates. Taken together, these results establish that the TGT(D264E)-RNA covalent complex can be formed not only *in vitro* but also in mammalian cells. We demonstrate that the assembled complex is sufficiently robust to withstand cellular small-molecule competition and is stable under denaturing conditions. More broadly, this system combines the advantages of a fully genetically encodable RNA conjugation platform with the added stability of covalent attachment needed to survive a complex cellular environment, while facilitating the recruitment of functional proteins to an RNA of interest.

The ability of TGT(D264E)-RNA covalent complexes to be formed not only *in vitro* but also in mammalian cells make our technology a powerful tool for diverse applications in RNA biology. The gel-shift, microscopy, and immunoblotting data herein support the feasibility of our approach for modifying RNA and suggest broader utility in applications requiring durable recruitment of functional proteins to engineered RNAs. Furthermore, this technology could be expanded to study the stability, function, and regulatory mechanisms of RNAs, as well as RNA imaging, interaction analysis, facilitate sequence-mediated control, control translation, and degradation of RNA species and associated proteins, among others.

REFERENCES

1. Hanspach, G., Trucks, S. & Hengesback, M. Strategic labelling approaches for RNA single-molecule spectroscopy. *RNA Biol.* **9**, 1119-1132 (2019).
2. Singh, G., Pratt, G., Yeo, G. W. & Moore, M. J. The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annu Rev Biochem.* **84**, 325-54 (2015).
3. Lehmann, R. & Nüsslein-Volhard, C. Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of oskar, a maternal gene in *Drosophila*. *Cell.* **47**, 141–52 (1986).
4. Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J. L., Pérez Marc, G., Moreira, E. D., Zerbini, C., Bailey, R., Swanson, K. A., Roychoudhury, S., Koury, K., Li, P., Kalina, W. V., Cooper, D., Frenck, R. W. Jr., Hammitt, L. L., Türeci, Ö., Nell, H., Schaefer, A., Ünal, S., Tresnan, D. B., Mather, S., Dormitzer, P. R., Şahin, U., Jansen, K. U. & Gruber, W. C. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med.* **383**, 2603-2615 (2020).
5. Büttner, L., Javadi-Zarnaghi, F. & Höbartner, C. Site-Specific Labeling of RNA at Internal Ribose Hydroxyl Groups: Terbium-Assisted Deoxyribozymes at Work. *J Am Chem Soc.* **136**, 8131-8137 (2014).
6. Nakajima, N. & Ikada, Y. Mechanism of amide formation by carbodiimide for bioconjugation in aqueous media. *Bioconjug Chem.* **6**, 123-30 (1995).
7. Mannack, L. V., Eising, S. & Rentmeister, A. Current techniques for visualizing RNA in cells. *Fluorescence.* **5**, 775 (2016).
8. Julio, A. R. & Backus, K. M. New approaches to target RNA binding proteins. *Curr Opin Chem Biol.* **62**, 13-23 (2021).
9. Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H. & Long, R. M. Localization of ASH1 mRNA Particles in Living Yeast. *Mol Cell.* **2**, 437-445 (1998).
10. Baron-Benhamou, J., Gehring, N. H., Kulozik, A.E. & Hentze, M. W. Using the λ N Peptide to Tether Proteins to RNAs. *Methods Mol Biol.* **257**, 135-153 (2004).
11. Wu, B., Chen, J. & Singer, R.H. Background free imaging of single mRNAs in live cells using split fluorescent proteins. *Sci Rep.* **4**, 3615 (2014).
12. McHugh, C. A., Russell, P. & Guttman, M. Methods for comprehensive experimental identification of RNA-protein interactions. *Genome Biol.* **15**, 203 (2014).

13. Sugimoto, Y., König, J., Hussain, S., Zupan, B., Curk, T., Frye, M. & Ule, J. Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. *Genome Biol.* **13**, R67 (2012).
14. Wilinski, D. & Dus, M. N6-adenosine methylation controls the translation of insulin mRNA. *Nat Struct Mol Biol.* **30**, 1260–1264 (2023).
15. Harada, F. & Nishimura, S. Possible anticodon sequences of tRNA^{His}, tRNA^{Asn}, and tRNA^{Asp} from *Escherichia coli*. Universal presence of nucleoside O in the first position of the anticodons of these transfer ribonucleic acid. *Biochemistry* **11**, 301-308 (1972).
16. Johannsson, S., Neumann, P. & Ficner, R. Crystal Structure of the Human tRNA Guanine Transglycosylase Catalytic Subunit QTRT1. *Biomolecules* **8**, 81 (2018).
17. Goodenough-Lashua, D. M. & Garcia, G. A. tRNA-guanine transglycosylase from *E. coli*: a ping-pong kinetic mechanism is consistent with nucleophilic catalysis. *Bioorg Chem.* **31**, 331-44 (2003).
18. Müller, M., Legrand, C., Tuorto, F., Kelly, V. P., Atlasi, Y., Lyko, F. & Ehrenhofer-Murray, A. E. Queuine links translational control in eukaryotes to a micronutrient from bacteria. *Nucleic Acids Res.* **47**, 3711–3727 (2019).
19. Stengl, B., Reuter, K. & Klebe, G. Mechanism and substrate specificity of tRNA guanine transglycosylases (TGTs): tRNA-modifying enzymes from the three different kingdoms of life share a common catalytic mechanism. *ChemBioChem* **6**, 1926–1939 (2005).
20. Romier, C., Reuter, K., Suck, D. & Ficner, R. Crystal structure of tRNA-guanine transglycosylase: RNA modification by base exchange. *EMBO J.* **15**, 2850-7 (1996).
21. Xie, W., Liu, X. & Huang, R. Chemical trapping and crystal structure of a catalytic tRNA guanine transglycosylase covalent intermediate. *Nat Struct Mol Biol.* **10**, 781–788 (2003).
22. Fergus, C., Barnes, D., Alqasem, M. A. & Kelly, V. P. The queuine micronutrient: charting a course from microbe to man. *Nutrients* **7**, 2897-929 (2015).
23. Curnow, A. W., Kung, F. L., Koch, K. A. & Garcia, G. A. tRNA-Guanine Transglycosylase from *Escherichia coli*: Gross tRNA Structural Requirements for Recognition. *Biochemistry* **32**, 5239–5246 (1993).
24. Nakanishi, S., Ueda, T., Hori, H., Yamazaki, N., Watanabe, N. & Okada, K. A UGU sequence in the anticodon loop is a minimum requirement for recognition by *Escherichia coli* tRNA-guanine transglycosylase. *J Biol Chem.* **269**, 32221–32225 (1994).
25. Romier, C., Meyer, J. E. W. & Suck, D. Slight sequence variations of a common fold explain the substrate specificities of tRNA-guanine transglycosylases from the three kingdoms. *FEBS Lett.* **416**, 93–98 (1997).

26. Farkas, W. R., Jacobson, K. B. & Katze, J. R. Substrate and inhibitor specificity of tRNA-guanine ribosyltransferase. *BBA - Gene Struct Expr.* **781**, 64–75 (1984).
27. Abold Todorov, K. & Garcia, G. A. Role of aspartate 143 in *Escherichia coli* tRNA-guanine transglycosylase: alteration of heterocyclic substrate specificity. *Biochemistry* **45**, 617–625 (2006).
28. Biela, I., Tidten-Luksch, N., Immekus, F., Glinca, S., Nguyen, T. X. P., Gerber, H. D., Heine, A., Klebe, G. & Reuter, K. Investigation of specificity determinants in bacterial tRNA-guanine transglycosylase reveals queuine, the substrate of its eucaryotic counterpart, as inhibitor. *PLoS One.* **8**, e64240 (2013).
29. Huang, R. & Ting, A. Y. Directed evolution of a sequence-specific covalent protein tag for RNA labeling. *Proc Natl Acad Sci. U.S.A.* **122**, e2422085122 (2025).
30. Alexander, S. C., Busby, K. N., Cole, C. M., Zhou, C. Y. & Devaraj, N. K. Site-Specific Covalent Labeling of RNA by Enzymatic Transglycosylation. *J Am Chem Soc.* **137**, 12756-12759 (2015).
31. Busby, K. N. & Devaraj, N. K. Enzymatic Covalent Labeling of RNA with RNA Transglycosylation at Guanosine (RNA-TAG). *Methods Enzymol.* **641**, 373-99 (2020).
32. Busby, K. N., Fulzele, A., Zhang, D., Bennett, E. J. & Devaraj, N. K. Enzymatic RNA Biotinylation for Affinity Purification and Identification of RNA–Protein Interactions. *ACS Chem Biol.* **15**, 2247-2258 (2020).
33. Ehret, F., Zhou, C. Y., Alexander, S. C., Zhang D. & Devaraj, N. K. Site-Specific Covalent Conjugation of Modified RNA by tRNA Guanine Transglycosylase. *Mol Pharm.* **15**, 737-742 (2017).
34. Zhou, C. Y., Alexander, S. C. & Devaraj, N. K. Fluorescent Turn-on Probes for Wash-Free RNA Imaging via Covalent Site-Specific Enzymatic Labeling. *Chem Sci.* **8**, 7169-7173 (2017).
35. Zhang, D., Zhou, C. Y., Busby, K. N., Alexander, S. C. & Devaraj N. K. Light-activated control of translation by enzymatic covalent RNA labeling. *Angew Chem Int Ed.* **57**, 2822-2826 (2018).
36. Zhang D., Jin, S., Piao, X. & Devaraj, N. K. Multiplexed Photoactivation of RNA with Single Cell Resolution. *ACS Chem Biol.* **15**, 1773-1779 (2020).
37. Busby, K. N., Fulzele, A., Zhang, D., Bennett, E. J. & Devaraj N. K. Enzymatic RNA Biotinylation for Affinity Purification and Identification of RNA-protein Interactions. *ACS Chem Biol.* **15**, 2247-2258 (2020).

38. Zhang, D., Liu, L., Jin, S., Tota, E., Li, Z., Piao, X., Zhang, X., Fu, X. & Devaraj N. K. Site-Specific and Enzymatic Cross-Linking of sgRNA Enables Wavelength-Selectable Photoactivated Control of CRISPR Gene Editing. *J Am Chem Soc.* **144**, 4487-4495 (2022).
39. Tota, E. M. & Devaraj, N. K. RNA-TAG Mediated Protein-RNA Conjugation. *ChemBioChem* **24**, e202300454 (2023).
40. Kittendorf, J. D., Sgraja, T., Reuter, K., Klebe, G. & Garcia, G. A. An essential role for aspartate 264 in catalysis by tRNA-guanine transglycosylase from *Escherichia coli*. *J Biol Chem.* **278**, 42369-76 (2003).
41. Xie, W., Liu, X. & Huang, R. H. Chemical trapping and crystal structure of a catalytic tRNA guanine transglycosylase covalent intermediate. *Nat Struct Biol.* **10**, 781-8 (2003).
42. Wu, B., Chao, J. A. & Singer, R. H. Fluorescence Fluctuation Spectroscopy Enables Quantitative Imaging of Single mRNAs in Living Cells. *Biophys J.* **102**, 2936-2944 (2012).
43. Wickramasinghe, V. & Laskey, R. Control of mammalian gene expression by selective mRNA export. *Nat Rev Mol Cell Biol.* **16**, 431–442 (2015).
44. Cobb, M. 60 years ago, Francis Crick changed the logic of biology. *PLoS Biol.* **15**, e2003243 (2017).
45. Morris, K. V. The emerging role of RNA in the regulation of gene transcription in human cells. *Semin Cell Dev Biol.* **22**, 351-8 (2011).
46. Strobel, S. A., Cochrane, J. C. RNA catalysis: ribozymes, ribosomes, and riboswitches. *Curr Opin Chem Biol.* **11**, 636-43 (2007).
47. Wickramasinghe, V. O. & Venkitaraman, A. R. RNA Processing and Genome Stability: Cause and Consequence. *Mol Cell.* **61**, 496-505 (2016).
48. Tsao, N., Ashour, M. E. & Mosammaparast, N. How RNA impacts DNA repair. *DNA Repair (Amst).* **131**, 103564 (2023).
49. Svensen, N. & Jaffrey, S. R. Fluorescent RNA Aptamers as a Tool to Study RNA-Modifying Enzymes. *Cell Chem Biol.* **23**, 415-25 (2016).
50. Burmistrz, M., Krakowski, K. & Krawczyk-Balska, A. RNA-Targeting CRISPR-Cas Systems and Their Applications. *Int J Mol Sci.* **21**, 1122 (2020).
51. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. & Zhang, F. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-23 (2013).

52. Chery, J. RNA therapeutics: RNAi and antisense mechanisms and clinical applications. *Postdoc J.* **4**, 35-50 (2016).
53. Young, A. P., Jackson, D. J. & Wyeth, R. C. A technical review and guide to RNA fluorescence in situ hybridization. *PeerJ.* **8**, e8806 (2020).
54. Robinson, P. K. Enzymes: principles and biotechnological applications. *Essays Biochem.* **59**, 1-41 (2015).
55. Cozzolino, F., Iacobucci, I., Monaco, V. & Monti, M. Protein-DNA/RNA Interactions: An Overview of Investigation Methods in the -Omics Era. *J Proteome Res.* **20**, 3018-3030 (2021).