Engineering of TEV Protease for Manipulation of Biosystems

by

Xi Chen

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Institute of Biomaterials and Biomedical Engineering
University of Toronto

© Copyright by Xi Chen 2013
Engineering of TEV protease for Manipulation of Biosystems

Xi Chen

Doctor of Philosophy

Institute of Biomaterials and Biomedical Engineering
University of Toronto

2013

Abstract

Synthetic biology is a nascent discipline that aims to design and construct new biological systems beyond those found in nature, ultimately using them to probe, control, or even replace existing biological systems. The success of synthetic biology depends on the assembly of a set of well-defined and modular tools. These tools should ideally be mutually compatible, reusable in different contexts, and have minimum crosstalk with endogenous proteins of the subject. The tobacco etch virus protease (TEV protease, TEVp) is a suitable candidate for such a tool due to its unique substrate specificity and high efficiency. Importantly, TEVp is capable of imitating proteolysis, a ubiquitous mechanism in nature for post-translational modifications and signal propagation. Here, TEVp is employed as a self-contained proteolytic device capable of executing biological tasks that are otherwise governed by endogenous proteins and processes. Consequently, the goal of using TEVp for synthetic manipulation of biosystems is achieved.

First, a single-vector multiple gene expression strategy utilizing TEVp self-cleavage was created. This approach was used for the robust expression of up to three genes in both bacterial and mammalian cells with consistent stoichiometry. The products can then be individually purified or targeted to distinct subcellular compartments respectively.
Second, a temperature-inducible TEVp was created by incremental truncation of TEVp. The 18th truncation of TEVp (tsTEVp) resulted in negligible activity at 37 °C, but retained sufficient activity at 30 °C for rapid processing of its substrates in several mammalian cell cultures.

Finally, tsTEVp was applied in the context of other synthetic modules to generate a variety of biological responses. Its versatility was demonstrated as cellular processes including protein localization, cellular blebbing, protein degradation, and cell death were rewired to respond to the physical stimulus of temperature.
Acknowledgments

My supervisor, Dr. Kevin Truong,

for your enthusiastic ideas, constant encouragement, and patient guidance

My lab mates,

for your friendship, technical assistance, and informative discussions

My thesis committee,

for your fresh insights, helping me focus, and challenging me

My friends and family,

for your company, support, and helping me find balance in life
Table of Contents

Acknowledgments ........................................................................................................................ iv
Table of Contents ........................................................................................................................ v
List of Publications ...................................................................................................................... viii
List of Abbreviations .................................................................................................................... ix
List of Figures ............................................................................................................................... x
List of Movies ............................................................................................................................... xii

1 Introduction .................................................................................................................................. 1
   1.1 Motivation ............................................................................................................................... 1
   1.2 Specific objectives .................................................................................................................. 2
   1.3 Organization .......................................................................................................................... 2

2 Background .................................................................................................................................... 4
   2.1 Protein synthetic biology ....................................................................................................... 4
       2.1.1 Overview of synthetic biology ....................................................................................... 4
       2.1.2 Simultaneous expression of multiple genes .................................................................... 6
       2.1.3 Protein switches .............................................................................................................. 7
   2.2 Temperature sensitive proteins ............................................................................................. 8
       2.2.1 Generation of TS proteins by sequence modifications ...................................................... 9
       2.2.2 Generation of TS proteins by indirect procedures .......................................................... 10
   2.3 TEV protease ......................................................................................................................... 11
       2.3.1 Fundamentals of TEV protease ..................................................................................... 11
       2.3.2 The role of TEV protease in synthetic biology ............................................................... 12
   2.4 Model proteins and systems .................................................................................................. 15
       2.4.1 N-degrons ...................................................................................................................... 15
       2.4.2 CaRQ ............................................................................................................................ 16
       2.4.3 Caspase-7 ..................................................................................................................... 17

3 Experimental Procedures .......................................................................................................... 18
   3.1 Plasmid construction .............................................................................................................. 18
       3.1.1 Materials and reagents for subcloning ......................................................................... 18
3.1.2 Generating cassette from PCR amplified gene ........................................ 22
3.1.3 Constructing fusion proteins by combining cassettes .............................. 24
3.1.4 Removing fluorescent protein tag ......................................................... 25
3.1.5 Generating cassettes for short peptides .................................................. 25
3.1.6 Performing point mutations .................................................................... 26

3.2 In vitro protein analysis .............................................................................. 27
3.2.1 Protein production and extraction from E. coli ........................................ 27
3.2.2 Affinity tag purifications ......................................................................... 28
3.2.3 SDS-PAGE ............................................................................................. 28
3.2.4 Fluorescence spectroscopy for FRET measurements ............................... 29

3.3 Cell culture and live-cell fluorescence imaging .......................................... 29
3.3.1 Materials and reagents for cell culture .................................................... 30
3.3.2 Cell culture protocols ............................................................................ 30
3.3.3 Protein extraction from mammalian cells for SDS-PAGE ....................... 32
3.3.4 Imaging equipment ................................................................................... 32
3.3.5 Imaging procedures ................................................................................ 33

3.4 Data analysis ............................................................................................... 33
3.4.1 Statistical comparisons ............................................................................ 34
3.4.2 Absolute quantification of fluorescent proteins ....................................... 34
3.4.3 Ratiometric quantification of fluorescent proteins ................................... 34

4 TEVp-facilitated stoichiometric delivery of multiple genes .............................. 36
4.1 Introduction and Aims ................................................................................ 36
4.2 Results ......................................................................................................... 38
4.2.1 Characterization of the TEV protease .................................................... 38
4.2.2 TEVp self-cleavage in E. coli ................................................................. 40
4.2.3 TEVp self-cleavage in mammalian cells ................................................ 43
4.2.4 Stoichiometry of TEVp self-cleavage products ..................................... 49

4.3 Discussion .................................................................................................. 51
4.4 Conclusion .................................................................................................. 53

5 Engineering temperature sensitive TEVp ..................................................... 54
5.1 Introduction and Aims ................................................................................ 54
5.2 Results ................................................................................................................. 55
  5.2.1 TEVpSensor for real-time monitoring of TEVp activity ................................. 55
  5.2.2 Creation of temperature sensitive TEVp by incremental N-terminal shortening... 56
  5.2.3 Real-time quantification of tsTEVp activity using TEVpSensor ....................... 60
5.3 Discussion ............................................................................................................ 64
5.4 Conclusion ........................................................................................................... 66

6 Synthetic manipulation of biosystems using tsTEVp ........................................... 68
  6.1 Introduction and Aims ....................................................................................... 68
  6.2 Results .............................................................................................................. 69
    6.2.1 Protein translocation from nucleus to membrane mediated by tsTEVp and DAGR. 69
    6.2.2 Cellular blebbing mediated by tsTEVp-controlled localization of CaRQ .......... 71
    6.2.3 tsTEVp mediated protein degradation by N-degron activation ..................... 74
    6.2.4 Synthetic activation of caspase-7 by tsTEVp .............................................. 78
  6.3 Discussion ......................................................................................................... 85
  6.4 Conclusions ....................................................................................................... 90

7 Summary ............................................................................................................... 92

8 Future work .......................................................................................................... 94
  8.1 Further development of tsTEVp and associated applications ......................... 94
  8.2 Further development of inducible proteolytic devices ..................................... 96
  8.3 Using tsTEVp as a tool for understanding biology .......................................... 99

9 References ............................................................................................................ 102

Appendices ............................................................................................................. 113
  Appendix A: DNA and protein sequences ............................................................. 114
  Appendix B: List of oligonucleotide primers ......................................................... 120
  Appendix C: Supplemental movie legends ......................................................... 121
  Appendix D: Additional experiments on N-degrons ........................................... 124
List of Publications


## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>Adenosine-5'-triphosphate</td>
</tr>
<tr>
<td>CaM</td>
<td>Calmodulin</td>
</tr>
<tr>
<td>CaRQ</td>
<td>Ca2+-activated Rho protein with embedded IQp</td>
</tr>
<tr>
<td>Ceru</td>
<td>Cerulean, equivalent to CFP</td>
</tr>
<tr>
<td>CFP</td>
<td>Cyan fluorescent protein, equivalent to Ceru</td>
</tr>
<tr>
<td>CID</td>
<td>Chemical inducers of dimerization</td>
</tr>
<tr>
<td>Crk2</td>
<td>Cdc2-related kinase 2</td>
</tr>
<tr>
<td>DAG</td>
<td>Diacylglycerol</td>
</tr>
<tr>
<td>DAGR</td>
<td>DAG receptor</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s modified eagle medium</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal Bovine Serum</td>
</tr>
<tr>
<td>FKBP12</td>
<td>12-kDa FK506 and rapamycin-binding protein</td>
</tr>
<tr>
<td>FRB</td>
<td>FKBP-rapamycin binding domain</td>
</tr>
<tr>
<td>FRET</td>
<td>Fluorescence resonance energy transfer</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>His</td>
<td>Polyhistidine-tag</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal ribosomal entry site</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Broth</td>
</tr>
<tr>
<td>LOV</td>
<td>Light-oxygen-voltage domain, LOV2 domain</td>
</tr>
<tr>
<td>LOVS1K</td>
<td>LOV2 domain fused with a Stim1 C-terminal fragment</td>
</tr>
<tr>
<td>MCS</td>
<td>Multiple Cloning Site</td>
</tr>
<tr>
<td>MLCKp</td>
<td>Myosin light chain kinase peptide</td>
</tr>
<tr>
<td>mRFP</td>
<td>Monomeric red fluorescent protein, equivalent to RFP</td>
</tr>
<tr>
<td>NES</td>
<td>Nuclear export signal</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear localization signal</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDBu</td>
<td>Phorbol 12,13-dibutyrate</td>
</tr>
<tr>
<td>PKC</td>
<td>Protein kinase C</td>
</tr>
<tr>
<td>pLyn</td>
<td>Palmitoylation sequence of Lyn kinase</td>
</tr>
<tr>
<td>RFP</td>
<td>Red fluorescent protein, equivalent to mRFP</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SH3</td>
<td>SRC Homology 3 Domain</td>
</tr>
<tr>
<td>TEV</td>
<td>Tobacco etch virus</td>
</tr>
<tr>
<td>TEVp</td>
<td>Tobacco etch virus protease</td>
</tr>
<tr>
<td>TS</td>
<td>Temperature-sensitive</td>
</tr>
<tr>
<td>tsTEVp</td>
<td>Temperature-sensitive tobacco etch virus protease</td>
</tr>
<tr>
<td>TVMVp</td>
<td>Tobacco vein mottling virus protease</td>
</tr>
<tr>
<td>Ven</td>
<td>Venus, equivalent to YFP</td>
</tr>
<tr>
<td>YFP</td>
<td>Yellow fluorescent protein, equivalent to Ven</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2.1</td>
<td>Overview of cellular processes</td>
</tr>
<tr>
<td>2.2</td>
<td>Design of protein synthetic systems</td>
</tr>
<tr>
<td>2.3</td>
<td>Summary of current methods for inducing TEVp activity</td>
</tr>
<tr>
<td>3.1</td>
<td>Overview of the pCfvtx vector</td>
</tr>
<tr>
<td>3.2</td>
<td>Flowchart for creating cassettes</td>
</tr>
<tr>
<td>3.3</td>
<td>Schematic of using MCSs to create fusion cassettes</td>
</tr>
<tr>
<td>3.4</td>
<td>Primer design for creating peptides</td>
</tr>
<tr>
<td>4.1</td>
<td>Overview of fusion proteins employed for chapter 4</td>
</tr>
<tr>
<td>4.2</td>
<td><em>In vitro</em> characterization of TEVp activity</td>
</tr>
<tr>
<td>4.3</td>
<td>TEVp self-cleavage in <em>E. coli</em></td>
</tr>
<tr>
<td>4.4</td>
<td>TEVp self-cleavage and subsequent purification of a binding complex</td>
</tr>
<tr>
<td>4.5</td>
<td>Differential localization of TEVp self-cleavage product in mammalian cells</td>
</tr>
<tr>
<td>4.6</td>
<td>Expression of multiple genes in COS-7 cells</td>
</tr>
<tr>
<td>4.7</td>
<td>Transfection efficiency of TEVpCNDVR compared with its individual modules</td>
</tr>
<tr>
<td>4.8</td>
<td>Expression of multiple genes in HEK-293 and HeLa cells</td>
</tr>
<tr>
<td>4.9</td>
<td>Stoichiometry of TEVp self-cleavage products</td>
</tr>
<tr>
<td>5.1</td>
<td>Overview of fusion proteins employed for chapter 5</td>
</tr>
<tr>
<td>5.2</td>
<td>Characterization of TEVpSensor</td>
</tr>
<tr>
<td>5.3</td>
<td>Protein sequence of truncated TEVp variants</td>
</tr>
</tbody>
</table>
Figure 5.4: Temperature sensitivity of truncated TEVp variants. ................................. 58
Figure 5.5: tsTEVp activity in HEK-293, HeLa and CHO cell lines. ................................. 60
Figure 5.6: Quantitative analysis of tsTEVp activity using TEVpSensor. .............................. 62
Figure 5.7: Quantitative analysis of tsTEVp activity at various temperatures. ..................... 63
Figure 6.1: Overview of fusion proteins employed for Chapter 6. ........................................ 69
Figure 6.2: tsTEVp and DAGR mediated protein translocation. .......................................... 70
Figure 6.3: tsTEVp and CaRQ mediated control over cell blebbing. ................................. 72
Figure 6.4: tsTEVp and CaRQ mediated control over cell blebbing. ................................. 73
Figure 6.5: Protein sequence of RFP-tevX-Deg-YFP .............................................................. 75
Figure 6.6: Characterization of tevX-Deg variants. ............................................................... 76
Figure 6.7: Synthetic activation of protein degradation by tsTEVp. ...................................... 77
Figure 6.8: Protein sequence of Casp7tevS .......................................................................... 79
Figure 6.9: Synthetic activation of a Caspase-7 chimera ...................................................... 80
Figure 6.10: Propidium iodide staining. ............................................................................... 81
Figure 6.11: Characterization of CaspSensor. ...................................................................... 82
Figure 6.12: Casp7tevS activity indicated by CaspSensor cleavage. ...................................... 84
List of Movies

Movie 4.1  PDBu stimulated migration of DAGRVen in COS-7 cells expressing TEVpCNDVR

Movie 5.1  COS-7 cell expressing tsTEVp and TEVpSensor at 30 °C

Movie 5.2  COS-7 cell expressing tsTEVp and TEVpSensor at 37 °C

Movie 6.1  HEK293 cells expressing tsTEVp/pLyn-tevS-CaRQ at 37 °C stimulated by ATP

Movie 6.2  HEK293 cells expressing tsTEVp/pLyn-tevS-CaRQ/DualCMVLOVS1K/Orai1 at 30 °C stimulated by light

Movie 6.3  COS-7 cells expressing tsTEVp/pLyn-Casp7tevS/CaspSensor at 30 °C

Movie 6.4  COS-7 cells expressing tsTEVp/pLyn-Casp7tevS/CaspSensor at 37 °C
1 Introduction

This chapter introduced the general motivation for this work, states the specific research goals, and provides an overview of the organization of this thesis.

1.1 Motivation

Living cells use complex molecular signaling circuits to monitor external and internal signals and translate them into appropriate physiological responses. These circuits are made up of a dynamic network of interacting proteins which execute specific tasks such as translocation, activation/inhibition and degradation. In other words, regulation of protein activity is fundamental to cellular function. The growing field of synthetic biology aims to exert artificial control over protein activity and create protein interactions beyond those found in nature (1). The ultimate objective is to use these new creations to probe, control, or even replace existing biological systems.

To achieve this aim, it is necessary to assemble a kit of well-characterized “tools” for the execution of specific tasks. These tools should ideally be mutually compatible, reusable in different contexts, and have minimum crosstalk with endogenous proteins of the subject. The tobacco etch virus protease (TEV protease or TEVp) is a suitable candidate for such a tool. One can envision many ways in which TEVp can be used to manipulate biology as various physiological processes are linked to the cleavage of peptide bonds, such as cell proliferation, development, apoptosis, immunity, blood clotting, and wound healing (2, 3).

Due to its specificity for the rare recognition motif, TEVp is a highly specific protease, and usually non-toxic in E. coli (4), S. cerevisiae (5), D. melanogaster (6), and mammalian cell cultures (7). Although TEVp is most often used as a recombinant protein in vitro for affinity tag purification, it can potentially be used as an orthogonal protease for a variety of other purposes in biological research, especially in the creation of synthetic networks. This reasoning is well
supported by recent works that employed various variants of TEVp to influence or probe cellular events, such as apoptosis (8), cell-cycle (6), and inflammation (9).

1.2 Specific objectives

The overall nature of this project is to engineer TEVp as a tool for studying and manipulating biological systems. This thesis achieves its goals in an objective-oriented manner. There are three specific research objectives:

1. Develop a reliable method for the delivery of multiple genes based on TEVp self-cleavage.

2. Develop an inducible TEVp that can be conditionally activated in cells by an artificial stimulus.

3. Demonstrate the versatility of the inducible TEVp by using it for conditional manipulation of synthetic networks and biological events.

1.3 Organization

This thesis consists of nine chapters organized to demonstrate the ideas and design strategies that were utilized to engineer TEVp for manipulating biosystems. The current chapter provides the overall motivation of this work in the context of synthetic biology and set forth specific research objectives. The second chapter provides background information in three general categories: synthetic biology, the current and potential role of TEVp in synthetic biology, and several other proteins, systems and concepts employed in this work. The third chapter describes the experimental techniques used, and is comprised of four sections: plasmid construction, *in vitro* protein analysis, live-cell fluorescence imaging, and data analysis.
Chapters four, five and six describe the progressive development of TEVp and associated applications. In chapter four, a versatile method was developed for the expression of multiple genes with the help of TEVp. This method was shown to be effective in cleaving a polyprotein into separate components in bacterial and mammalian cells with consistent stoichiometry. In chapter five, a temperature-inducible TEVp was created by incremental truncation of TEVp. The 18th truncation of TEVp resulted in negligible activity at 37 °C, but retained sufficient activity at 30 °C for rapid processing of its substrates. In chapter six, the temperature-inducible TEVp was applied in a variety of contexts to demonstrate its ability to generate temperature-induced biological responses, namely protein translocation, cell blebbing, protein degradation and cell death. This concludes our vision of using TEVp as a versatile “plug-and-play” tool for the synthetic manipulation of biosystems.

The final chapters provide a general summary of the work and suggest reasonable future directions. Chapter seven recapitulates the contributions described in chapters four to six. Chapter eight discusses future directions that could stem from this work. Chapter nine is a list of references. Additional information and experiments can be found in the appendices.
2 Background

2.1 Protein synthetic biology

2.1.1 Overview of synthetic biology

Synthetic biology is a nascent discipline that aims to design and construct new biological systems beyond those found in nature, ultimately using them to probe, control, or even replace existing biological systems (1, 10, 11). An essential feature of synthetic biology is the conversion of signals from the exterior of the cell into specific cellular events, such as protein production, translocation, degradation, and modification (Figure 2.1).

![Figure 2.1: Overview of cellular processes.](image)

Cellular processes can be regulated on two levels: transcriptional and post-translational. Transcriptional regulation results in concentration change of a gene of interest. Post-translational regulations such as protein localization, degradation and modification can be applied at faster time scales and is spatially resolved.

Initial strategies to build synthetic circuits were mainly focused on regulation of DNA transcription, and included famous examples such as the “Repressilator” (12) and the “Toggle
Switch” (13). These strategies control protein activity through regulation of its concentration by varying the rate of transcription. However, post-translational processes occur at much smaller time scales (from milliseconds to minutes) (1), and are functionally dependent on subcellular localization (14). Post-translational, or protein level control of processes allows the construction of more sophisticated signaling circuits, which can be fast, robust and spatially resolved (15). The subfield of synthetic biology for post-translational process will henceforth be called protein synthetic biology.

![Diagram of protein synthetic systems](image)

**Figure 2.2: Design of protein synthetic systems.**

(Left) Activation of a natural circuit involving multiple interacting proteins by a natural stimulus leads to an observable phenotype below (e.g. a distinctive morphology). (Middle) A protein switch can be introduced to rewire the natural circuit to an engineered input stimulus. (Right) Multiple protein switches can replace the natural circuit to create streamlined regulation. Taken together, the expression of multiple proteins and the engineering of protein switches are two key steps to creating protein synthetic systems. Black arrows indicate natural stimulus, red arrows indicate engineered stimulus. Squares are engineered proteins, circles are natural proteins.
The general idea of protein synthetic biology is summarized in Figure 2.2. The two key tools required for protein synthetic biology are: simultaneous expression of multiple genes, and proteins switches. The ability to express multiple genes is crucial for the development of protein-based synthetic systems because artificial protein networks involve multiple protein modules. Co-delivery of biosensors is also often required to monitor their activities and validate synthetic systems. Furthermore, consistent stoichiometry of the simultaneously expressed genes is important because many protein-protein interactions involve stoichiometric binding partners, and a consistent stoichiometry minimizes cell-to-cell variability when interpreting results. A protein switch is a protein modified to respond to engineered/unnatural input stimulus (11). These inputs could either be physical (light or temperature), chemical (ions or small molecules), or biological (protein interactions). Protein switches allow rewiring of native interactions to exogenous factors that is tunable and orthogonal to other endogenous processes. As such, specific pathways can be isolated for study, and ultimately, advanced pathways with streamlined functionalities can be designed to bypass native circuits (Figure 2.2).

2.1.2 Simultaneous expression of multiple genes

Conventionally, multiple genes can be expressed in a cell through simultaneous or sequential transfections of multiple vectors. However, this method becomes tedious and inefficient as the number of vectors increases. The expression of multiple genes from a single vector can be achieved by incorporating expression cassettes each with its own independent promoter (16). Unfortunately, this method leads to the competition for a limited pool of transcriptional factors and is susceptible to gene suppression from promoter interference (17). The internal ribosomal entry site (IRES) mediated polycistronic vector is an improvement over the aforementioned multiple expression vector as it expresses all its genes from a single promoter through a translation re-initiation mechanism (18). However, the translations are uncoupled in the sense that they are invoked from different translational initiation events. Consequently, the proteins are not produced in stoichiometric proportions, with a heavier bias on upstream expression (19).
The 2A sequences from picornaviruses have recently become a popular tool to express multiple genes from a single vector. Through a proposed translational skipping mechanism, the 2A sequence is able to disrupt peptide bond formation during protein synthesis between the Gly and Pro residues at its C-terminus (20). 2A sequences have functioned in all eukaryotic systems to date (mammal, insect, yeast, fungi and plant) (21). Notably, 2A sequences have found diverse applications by co-expressing proteins necessary for adoptive cell therapies (22, 23), genetic engineering of human stem cells (24), and the induction of pluripotent stem cells (25, 26). A significant advantage of 2A sequences over the previously discussed methods is that protein products are more likely to retain the stoichiometry of their genes on the expression vector.

However, the 2A sequences have three major limitations. First, the 2A-associated translational skipping mechanism is exclusive to eukaryotic ribosomes (27). Therefore, 2A sequences cannot be used for protein production and purification where bacterial cells are involved. Second, of the 20-30 residues in a 2A sequence, all but one of them will remain on the C-terminal of its upstream protein after separation. This sequence can potentially interfere with protein function, especially if the protein’s innate C-terminus is involved in binding interaction with other proteins. This problem was manifested in one case where an enzyme’s activity was severely diminished when it was expressed upstream to a 2A sequence (28). This may also explain why the ordering of genes around 2A sequences can lead to different physiological responses (29). Third, the separation efficiency of the 2A sequences is largely dependent on its length. Short 2A sequences (14-18 residues) often result in low (50-70%) separation efficiencies (30). The separation efficiency can be improved by increasing the length of the 2A sequences to 20-30 residues (in one case as long as 56 residues) (31). As a result, an even longer trailing peptide sequence is left on upstream protein, which may be undesirable as previously discussed.

### 2.1.3 Protein switches

Two examples of engineered protein switches will be discussed below. Further examples of protein switches will be discussed in the TEV protease section (section 2.3).
**Allosteric control of proteins through modular recombination:** In a classic example of protein switches by Dueber et al., the autoinhibitory domains for native N-WASP, CDC42-binding GBD and PIP2-binding B domain, were replaced by combinations of SH3 and PDZ domains (32, 33). The unnatural peptides for SH3 and PDZ were used to control the new N-WASP variant. By duplicating and rearranging the SH3 and PDZ domains, input and output behaviors of the switch were analyzed and revealed a broad range of signaling modes, including linear response, constitutively active, constitutively repressed, and highly cooperative.

**Allosteric control of proteins by light-induced conformational change of LOV domain:** By fusing the light oxygen voltage (LOV) domain from *A. sativa* Phototropin 1, several proteins were put under control by the physical input of blue-light. The prime example of this strategy was the engineering of photo-activable Rac1 (PA-Rac1) (34). Irradiation of PA-Rac1 produced instantaneous membrane protrusions and ruffling. Localized irradiation of PA-Rac1 near the cell edge caused direction migration. This demonstrates that this mode of protein control is both spatially and temporally precise. The versatility of LOV fusions gave rise to many other light-activated switches to control apoptosis-inducing caspase-7 (35) and stromal interaction molecule 1 (STIM1) (36).

### 2.2 Temperature sensitive proteins

The use of temperature sensitive (TS) mutants is a well-established method for probing gene function (37). It is applicable to organisms in which temperature changes can be tolerated, including viruses (38), *E. coli* (39), *S. cerevisiae* (40), mammalian cells (41), *C. elegans* (42), *D. melanogaster* (37), plants (43), and even fish (41). Typically, a TS mutant is one in which there is a marked drop in activity when expressed above a certain temperature (restrictive temperature). At a lower temperature (permissive temperature), the activity of the mutant is very similar to that of the wild-type (44-46). Strategies for generating TS proteins can be grouped into direct and indirect methods. Direct methods involve amino acid modifications on the target protein itself to evolve temperature-sensitive characteristics. Indirect methods make use of third
party temperature-sensitive systems to modify the target protein resulting in either loss or gain of function.

2.2.1 Generation of TS proteins by sequence modifications

TS proteins are typically generated by random mutagenesis followed by screening for the desired TS phenotype in a large number of progeny (46). This method is usually laborious and limited to systems that allow large-scale screening. The number of genes that can be mutated to temperature sensitivity is limited and many TS mutants remain leaky at the restrictive temperature even after extensive mutagenesis (37, 39). In direct mutagenesis, the need for screening thousands of mutants is avoided by predicting a small number of substitutions that are more likely to produce a TS phenotype (47). One of these techniques focuses on predicting and modifying buried amino acids in globular proteins based on the fact that buried hydrophobic amino acids are key to thermo-stability of proteins (44, 48). This method is more efficient compared to random mutagenesis (44), but it is not practical to design specific TS mutations for each individual protein. Both random and direct mutagenesis yield a modified protein that usually have appreciably lower activity than the wild-type even at the permissive temperature (46).

A less common strategy for creating TS proteins involves truncating the N or C terminal ends of the target protein. The functional and structural importance of these terminal regions of proteins varies from protein to protein. Some proteins readily tolerate the removal of terminal residues without impairment of native function, which is coherent with the observation that terminal regions are frequently poorly ordered in crystal structures (49). However, some proteins are particularly sensitive to terminal truncations. In most of these cases, step-wise reduction in activity was observed in conjunction with incremental truncation of terminal residues. This has been demonstrated in a number of proteins including rhodanese (50), Stoffel fragment of Taq polymerase (51), heat labile enterotoxin (52), staphylococcal nuclease R (53), β-lactamases (54), chloramphenicol acetyltransferase (55), and ribonuclease HI (56). In general, there exist a
threshold truncation at which the function of the protein is significantly affected. Further truncations beyond this threshold leads to complete loss of function. However, at this threshold, some of these truncated proteins can still regain their function under certain physical conditions, usually at a lower temperature (50, 51, 54).

2.2.2 Generation of TS proteins by indirect procedures

The use of a TS degron provides a more general approach for creating TS proteins primarily in yeast (40, 57-60). The TS degron was constructed using the ubiquitin fusion technique (61) along with a TS mutant of dihydrofolate reductase (DHFR) bearing an N-terminal destabilizing Arg residue (Ub-Arg-DHFRts). This technique utilizes the N-end rule for protein degradation (reviewed in section 2.4.1). At low temperatures (~ 23 °C), Ub-Arg-DHFRts cannot be rapidly degraded through the ubiquitin pathway. Degradation can be activated at 37 °C, possibly because the unfolding of Ub-Arg-DHFRts reveals Lys residues for ubiquitin attachment (58, 60). The target protein is expressed as a fusion protein to the Ub-Arg-DHFRts module. In essence, Ub-Arg-DHFRts triggers the irreversible inactivation of any coupled target proteins by temperature-induced degradation. One limitation of this technique is that it cannot be applied to proteins whose functions will be altered by the N-terminal Ub-Arg-DHFRts fusion. Also, only nuclear and cytoplasmic proteins are currently known to be accessible by the N-end rule degradation pathway, therefore proteins present in other compartments, such as the secretory pathway, cannot be inactivated using this approach. Finally, the intrinsic stability of the target protein may cause variability in degradation rates and may affect applications where temporal precision is desired (59, 62).

Another indirect approach involves TS splicing variants of inteins. These conditional inteins splice themselves out at the permissive temperature to regenerate a wild-type host protein while remaining in the host at the restrictive temperature (37, 46). Therefore, strategic placement of the TS inteins within the host could disrupt normal function at the restrictive temperature but recover wild-type function through temperature-induced splicing. Unlike TS degron, this method results
in irreversible protein activation after temperature shift. Initially, TS splicing variants of GyrA and recA inteins were identified (63). The more versatile Sce VMA inteins were later successfully implemented in proof of principle studies in plants, E. coli, S. cerevisiae, and D. melanogaster (37, 39, 46, 64). However, the TS intein splicing approach is not universally transferable between proteins (46, 65). There are many factors that determine the successful application of splicing inteins in regulating protein activity. First, the intein must be located at a site that allows adequate protein expression and stability. Second, the intein must be able to sufficiently block protein activity when unspliced. Third, the local sequence around the intein must be able to allow it to be efficiently spliced out, and leave a native-like protein with wild-type activity (65). Each target protein folds in a unique manner, the insertion site for inteins needs to be specifically designed for each subject in order to ensure proper splicing (46). As a result, the generation of TS intein proteins has only met with limited success in the past decade.

2.3 TEV protease

This section focuses on the key protein involved in this thesis, the TEV protease (TEVp). First, the fundamentals of the protease will be introduced, including its substrate specificity, temperature sensitivity, and application in biotechnology. Second, the current role of TEVp in synthetic biology will be discussed, including a review of several prominent studies in using various forms of inducible TEVp to manipulate cellular processes.

2.3.1 Fundamentals of TEV protease

Tobacco etch viruses (TEVs) express all of their proteins in the form of polyproteins before they are processed by the TEV nuclear inclusion α protease, also known as the TEV protease (TEVp) (66). Natural substrates of TEVp exist as a seven-residue long sequence in the form of E-X-X-Y-X-Q-(G/S) (67). Cleavage occurs after Q, resulting in G or S becoming the new N-terminal residue of the downstream fragment. Although not existent in nature, other residues can also be
 tolerated at the P1’ position in place of G or S, albeit with a small impact on the cleavage efficiency (68, 69). In most protein engineering applications, substrates ENLYFQS and ENLYFQG are used due to their high efficiency (70, 71). A salient feature of TEVp is its high specificity and good activity towards its substrates, which makes it a popular tool for cleaving genetically engineered fusion proteins, especially in the removal of affinity tags in tandem affinity purification (72). The high specificity of TEVp also allows it to be expressed in a number of model systems such as *E. coli* (4), *S. cerevisiae* (5), *D. melanogaster* (6), and mammalian cells cultures (7) without adverse effects on development and viability.

TEVp is active over a wide range salt concentrations and temperatures from 4 °C to 37 °C, with optimum activity at around 30 °C, while retaining appreciable activity at 4 °C and 37 °C (73, 74). Point mutation studies have been conducted to improve the catalytic efficiency and solubility of TEVp (71, 75, 76). The most widely used mutated variant of TEVp is the TEVp(S219V) mutation which prevents auto-inactivation and increase the stability of wild-type TEVp by ~100-fold (71). The TEVp(S219V) mutant is used throughout this thesis.

### 2.3.2 The role of TEV protease in synthetic biology

The irreversible modification of proteins by proteolytic cleavage represents a simple and ubiquitous mechanism for signal propagation in living cells, and governs important biological processes such as cell proliferation, development, apoptosis, immunity, blood clotting, and wound healing (2, 3). Considering the functional relevance of proteases in biology, proteases and their targets can be of great interest in the engineering of synthetic biological systems. The development and utilization of TEVp in synthetic biosystems is still in its infancy, but has enormous potentials. One of the main strengths of TEVp is its high specificity and lack of endogenous cleavage sites in most biological systems of interest (6, 7). Furthermore, its short recognition sequence could allow it to be embedded within target proteins with minimum effect on their function (77). Several recent studies have successfully demonstrated the application of TEVp in synthetic biology. Within these studies, three different mechanisms for conditional
control of TEVp activity were implemented. A graphical summary of the three mechanisms is shown in Figure 2.3.

![Figure 2.3](image)

**Figure 2.3: Summary of current methods for inducing TEVp activity.**

(A) TEVp can be activated by induced expression from a conditional promoter. In this case, the inducer of activity (yellow circle) can be a small molecule such as galactose, or temperature in the case of heat shock promoters. (B) TEVp can be activated by reconstitution of split-TEVp fragments, inducible by rapamycin (Rap) mediated dimerization of FKBPs/FRBs. (C) TEVp can be activated by induced-proximity. In the example shown, closer proximity of TEVp to its cut site is caused by recruitment of a tethered arrestin to GPCR following ligand (red square) binding. Due to high efficiency of TEVp, a weakened cut site has to be used.

**Induced TEVp expression to disrupt protein function in vivo:** TEVp can be activated by induced expression from a conditional promoter (Figure 2.3A). For example, TEVp was expressed in a cell-type-specific and/or temporally controlled manner in *D. melanogaster* to cleave and deactivate an engineered Rad21 protein bearing a TEVp cleavage site (6). The induced expression of TEVp was achieved using a heat-shock promoter or the UAS/GAL4 system. It was the first study where TEVp was stably expressed in a complex metazoan organism. As a
component of cohesin, Rad21 plays an important role in the separation of sister chromatids during mitosis. Phenotypic comparisons were made between cells expressing wild-type Rad21 and TEVp-cleavable Rad21 and revealed essential roles of cohesin in mitotic and even non-mitotic cells. A distinctive feature of this study was that direct deactivation of a protein by TEVp cleavage can occur on a much faster time scale than gene deletion or RNA interference studies. This allowed deactivation of Rad21 at various stages of the cell cycle with temporal precision.

Reconstitution of split-TEVp fragments by rapamycin-induced protein dimerization: TEVp activity can be temporally controlled by complementation of split-TEVp fragments (Figure 2.3B). This technique was originally developed to study protein-protein interactions in cell cultures (7, 78). In this technique, TEVp was split into two non-functioning fragments and attached to two interacting protein partners, such as receptor tyrosine kinases. The heterodimerization of this partner brings the split-TEVp fragments together, resulting in the assembly of a functional TEVp. A useful extension to this technique was to use rapamycin-induced FRB-FKBP12 interaction to bring together the split-TEVp fragments (7, 79). In this way, TEVp activity can be temporally controlled by the simple addition of a small molecule. The rate-limiting step of this technique is the reconstitution of the split-TEVp fragments. The low inducible activity could be compensated by higher split-TEVp expression, but at the cost of higher background activity from spontaneous reconstitution (79-82). A major application of this technique was demonstrated in the activation of several executioner caspases and subsequent apoptosis (8). In this study, rapamycin-induced split-TEVp activation was used to selectively activate caspases 3, 6, and 7 by cleaving off autoinhibitory elements. The selective activation of these otherwise non-orthogonal caspases revealed their non-redundant roles in the orchestration of apoptosis.

Regulated localization of TEVp to measure receptor activation (Tango assay): TEVp can be activated by induced-proximity (Figure 2.3C). In a study by Barnea et al., TEVp was recruited to the plasma membrane by tethering to a receptor signaling protein following receptor activation (9). The induced-proximity of TEVp allowed it to cleave and release a transcription factor into the nucleus and activate the transcription of a reporter gene. This technique was tested on three
classes of receptors: G protein-coupled receptors, receptor tyrosine kinases, and steroid hormone receptors. Consequently, some insights about the role of GPR1 in inflammation were revealed. The activation of TEVp was determined by its regulated localization within the cell, it therefore functions at an even faster time scale than the previously mentioned induced-expression method. However, because the catalytic efficiency of TEVp towards its natural cleavage sequence is very high, a weaker substrate had to be engineered in order to minimize background activity. Even so, the signal to noise ratio of TEVp activation was high which likely explains why this method has yet to be reproduced for other applications.

2.4 Model proteins and systems

The ideal characteristics of a synthetic device are portability and expandability. For a synthetic device to be considered robust, it has to demonstrate the ability to be applied in a variety of contexts without complete redesign. This allows the assembly of a catalogue of well-characterized tools for synthetic biology. In this section, several model proteins and systems used to demonstrate the versatility of TEVp will be discussed.

2.4.1 N-degrons

Targeting selective proteins for degradation is a basic cellular mechanism that governs important processes such as cell cycle progression, cell proliferation, and apoptosis (83). This mechanism is triggered by degradation signals called degrons. Degrons confer metabolic instability of proteins by reducing their \textit{in vivo} half-lives and target them to the 26S proteasome for degradation. The N-degron is a class of degrons where the N-terminal residue of a protein plays an important role in determining their stability (83, 84). The N-end rule relates the \textit{in vivo} half-life of a protein to the identity of its N-terminal residue and is present in all organisms examined, including bacteria, yeasts and mammalian cells (85). In eukaryotes, the N-degron is made up of both a specific destabilizing N-terminal residue, and an N-terminal sequence which functions as
a substrate for the E3/E2 ubiquitin ligase complex (85-87). The primary N-terminal destabilizing residues can be any one of Cys, Arg, His, Lys, Leu, Phe, Trp, Tyr and Ile in eukaryotes, while four other residues Asp, Gln, Glu and Asn can be converted into a primary destabilizing residue by enzymatic reactions (84). For the N-terminal sequence to function as a substrate for the E3/E2 complex, it has to also contain a strategically positioned internal Lys residue that can be captured by the E3/E2 complex (85, 88). The other residues in this N-terminal sequence can vary and a number of these sequences have been discovered in both yeast (89) and E. coli (90). It is logical to deduce that the composition of the N-terminal sequence needs to fulfill two requirements: 1) the location of the internal Lys residue in the substrate has to be within a reasonable distance from the N-terminal residue so that the E3/E2 complex can bind to both the N-terminal residue and the internal Lys simultaneously, and 2) the sequence needs to be unfolded so that the internal Lys is exposed, unless the Lys residue is already exposed in a folded structure.

2.4.2 CaRQ

CaRQ is a synthetic Ca\textsuperscript{2+}-sensitive RhoA which uses Ca\textsuperscript{2+} signaling to control cell morphology and migration (35). RhoA is a prominent member of the Rho family of small GTPases known to regulate the actin cytoskeleton (91). RhoA activation leads to formation of blebs and actinomyosin-based migration (92). The constitutively active mutant, RhoA(Q63L), is capable of inducing dynamic non-apoptotic blebbing in several epithelial-like cell lines such as HEK293, HeLa, and CHO. CaRQ was created by genetically embedding calmodulin (CaM)-binding peptide IQp within the secondary structure of RhoA(Q63L), and placing CaM along with another CaM-binding peptide MLCKp on the N-terminus of RhoA(Q63L), resulting the final fusion protein CaM-MLCKp-RhoA(Q63L)/IQp. At resting Ca\textsuperscript{2+} concentration, CaM will bind to IQp and inhibit RhoA activity, while a strong Ca\textsuperscript{2+} signal will force CaM to bind with a higher affinity to MLCKp and release RhoA inhibition, resulting in associated morphological changes and cell migration. CaRQ can be activated by a variety of Ca\textsuperscript{2+} generating modules such as ionophores, purinergic receptors, α4-acetylcholine receptors, and store-operated calcium entry. An important feature of CaRQ is that it responds differentially to different spatial patterns of
Ca\textsuperscript{2+} signaling. In other words, the cellular localization of CaRQ defines its ability to be activated by a local Ca\textsuperscript{2+} transient.

2.4.3 Caspase-7

Apoptosis or programmed cell death is activated by three major pathways: (1) death receptor binding, (2) Granzyme B incorporation and (3) stimulation by cell stress or damage. Each of these pathways leads to the activation of the caspase signaling cascade, which is made up of a series of cysteine proteases called caspases (93). Caspase-7, in particular, carries out much of the proteolysis during the demolition phase of apoptosis (93). Characteristic substrates of caspase-7 include ICAD (DNA fragmentation) (94), ROCK1 (95) (contraction of actin cytoskeleton and membrane blebbing) and PARP (inactivation of DNA repair mechanism) (96). Caspase-7 is translated as a 303-residue inactive zymogen which can be activated by proteolytic removal of its prodomain at Asp24 and separation of its p20 and p11 subunits at Asp198 (97), a process typically performed by initiator caspases in response to extrinsic signals or intrinsic cellular stress (93). Because caspase-7 is directly responsible for cleaving most of the substrates involved in apoptosis, the activation of caspase-7 itself is sufficient to induce cell death (8, 35). A truncated version of caspase-7 without 56 N-terminal residues, known as 57casp7, has been previously created (98). 57casp7 is capable of cleaving its substrate DEVD and induce apoptotic morphologies when overexpressed in several mammalian cell lines (35, 98).
3 Experimental Procedures

This chapter will describe the experimental procedures used in this thesis in four sections. First, gene construction methods will be described including the complete procedures involved in creating a cassette-based gene library and making point mutations. Second, procedures used for \textit{in vitro} protein analysis will be given including protein purification, SDS-PAGE, and fluorescent spectroscopy. Third, methods and protocols for cell culture and live-cell fluorescence microscopy will be provided. Last, several data analysis techniques used throughout the thesis will be discussed including methods for statistical comparisons, fluorescent quantification and ratiometric calculations.

3.1 Plasmid construction

Gene construction in this work was based on an efficient cassette-based approach that allows the rapid generation of a library of constructs (99).

3.1.1 Materials and reagents for subcloning

\textbf{Standard expression vector for subcloning:} The expression vector, named pCfvtx (Cassette Fused with Venus), is based on the pTriEx1.1-Hygro vector (EMD Chemicals, Mississauga, ON) that allows for expression in both prokaryotic and eukaryotic cells (99). The structure of the pCfvtx vector is show in Figure 3.1. The pCfvtx vector contains both ampicillin and hygromycin resistance genes, as well as a fluorescent protein, allowing for three levels of screening. The T7lac promoter enables the vector to be expressed in \textit{E. coli}, a tool for rapid clonal expansion of both the vector itself and the protein it codes for. The CMV promoter enables expression in mammalian cells. Central to the pCfvtx vector are the two multiple cloning sites, MCS1 and MCS2 downstream of the transcriptional start site. Both sites contain recognition sequences for
several restriction enzymes. In between these cloning sites is a stop codon that prevents the expression of the fluorescent protein unless the gene of interest has been inserted. This provides an easy way to detect and select colonies containing the gene of interest. The multiple cloning sites MCS1 and MCS2 contains eight restriction enzyme recognition sequences, of which six of them form compatible ends: SpeI/NheI (sticky CTAG ends), BamHI/BglII (sticky GATC ends) and StuI/SmaI (blunt ends). The Pmel sites flanking the fluorescent protein allows for the excision of the fluorescent protein, while XhoI facilitates the recombination of genes together with the fluorescent protein.

![Diagram of pCfvtx vector](image)

**Figure 3.1: Overview of the pCfvtx vector.**

This vector contains antibiotic resistance gene Amp. Multiple cloning sites (MCSs) are found downstream of CMV and T7lac promoters. Gene insertion is made between restriction enzyme recognition sequences located in MCS1 and MCS2. SpeI/NheI and BamHI/BglII have compatible cohesive ends. Fluorescent protein used for screening can either be Venus (YFP) or Cerulean (CFP).

**PCR:** PCR reaction was performed using Eppendorf MasterCycler Personal. The 50 µL reaction mixture contained 5 µL of 5 mM dNTP, 5 µL 10x PCR buffer with MgSO₄, 5 µL of each forward and reverse primers diluted to 25 pM, 1 µL of cDNA or plasmid, 0.5 µL Pfu polymerase
(all above from Fermentas, Burlington, ON) and 28.5 µL RNase/DNase-free water (Invitrogen, Oakville, ON).

**Enzyme digestion and ligation:** Enzyme digestion and ligation were performed using Eppendorf MasterCycler Personal. For enzyme digestion, the 30 µL reaction mixture contained 3 µL 10x enzyme buffer (supplied by manufacturer), 1 µL 30x BSA, 1 µL of each restriction enzyme (all above from NEB), 5-10 µL of plasmid DNA or 24 µL of PCR fragment, and filled up to 30 µL with RNase/DNase-free water (Invitrogen). For ligation, the 20 µL reaction mixture contained 2 µL 10x ligase buffer (supplied by manufacturer), 1 µL T4 ligase (all above from NEB) and a mixture of 17 µL DNA containing insert and host in a 4:1 molar ratio.

**Gel electrophoresis and DNA purification:** Agarose gels were prepared with a 0.5x TAE solution with 1% agarose (all above from Invitrogen). Prior to loading, DNA was mixed with 6x Orange Loading Dye (10 µL dye with 50 µL PCR product or 6 µL die with 30 µL enzyme digest product). The O’Gene DNA Ladder (100bp to 10kbp range) was used for mass comparisons (both dye and ladder from Fermentas). DNA bands were viewed on UV transilluminator and were excised for purification if needed. Purification of DNA fragments was done using the PureLink Quick Gel Extraction kit (Invitrogen).

**Bacterial transformation, culturing, plating, screening, and plasmid purification.** Transformation was done using DH5α Subcloning Efficiency chemically competent E. coli cells (Invitrogen). Transformed cells were cultured in Luria Broth (LB) with 100 µg/mL ampicillin or 80 µg/mL kanamycin (for several purchased plasmids). Bacterial colonies were grown on LB agar plates with 200 µg/mL ampicillin. LB and LB agar were prepared from their respective powders (Sigma, St. Louis, MO) in distilled, autoclaved water. Colonies were observed using the Illumatool Tunable Lighting System (Light Tools Research, Encinitas, CA) using excitation filters (440 nm, 488 nm and 540 nm) and longpass emission filters (480 nm, 525 nm and 560 nm, for CFP (Cerulean), YFP (Venus) and RFP (mRFP) respectively). Purification of plasmids after bacterial culture was done using the PureLink Quick Plasmid Miniprep kit (Invitrogen).
Figure 3.2: Flowchart for creating cassettes.

The three main procedures for creating cassettes (inserting a PCR product, creating a fusion cassette, and removing the fluorescent tag) are shown, separated by dashed lines.

3.1.2 Generating cassette from PCR amplified gene

The procedure for inserting a gene amplified by PCR into the standard expression vector is presented here. This procedure along with the procedure for combining cassettes (section 3.1.3) and removing fluorescent tags (section 3.1.4) is summarized in a flowchart (Figure 3.2).

1. Primers for amplifying the gene of interested were designed with overhangs containing a restriction site compatible with a restriction site from each MCS of the vector.

2. The 50 µL PCR mixture was prepared as described above. PCR began with 10 minutes at 95 °C for initial denaturation. Each PCR cycle consisted of 3 minutes of denaturation at 95 °C, 2 minutes of annealing at 60 °C, and 2 minutes for every 1 kbp extension at 72 °C. A total of 32-35 cycles were run. The final extension was 15-20 minutes at 72 °C. Slight variations were made to optimize performance.

3. Each PCR product was mixed with loading buffer (as described previously) and run with gel electrophoresis at 100V for 20 minutes in cold 0.5x TAE buffer. After electrophoresis, the gel was placed in ethidium bromide solution for 20 minutes. They were then observed under UV transilluminator and bands of correct size (compared to DNA ladder) and sufficient quality were excised. For poor or absent bands, the procedure was repeated from step 2 with adjusted parameters. In case of repeated failure, the procedure was repeated from step 1 with redesigned primers.

4. After band excision, they were purified with the PureLink Quick Gel Extraction kit (Invitrogen) using manufacturer’s protocols.

5. The purified PCR fragment, along with the host vector, were enzyme digested using appropriate restriction enzymes, as previously described. Enzyme digestions were
performed at 37 °C for 3 hours. The digested fragments and vectors run with gel electrophoresis and purified.

6. Ligation mixture between the PCR fragment and host vector was prepared as previously described. Ligations were typically performed at 16 °C for 3 hours.

7. Ligation products were transformed into 20 μL aliquots of DH5α competent cells. In this process, 1 μL of ligation product was mixed with the competent cells, heat-shocked at 42 °C for 45 seconds and cooled at 0 °C for 5 minutes. The transformed cells were cultured at 37 °C in 1 mL LB supplemented with 100 μg/mL ampicillin overnight in a shaking incubator.

8. The next day, the culture was diluted 1:1000 in LB, 40 μL of it was spread onto LB agar plates with ampicillin, and incubated at 37 °C overnight.

9. The next day, colonies were screened for fluorescence using the Illumatool system and appropriate filters (as described previously). Successful (fluorescent) colony was picked off the dish by a clean pipette tip and cultured in 2 mL LB supplemented with 100 μg/mL ampicillin overnight in a shaking 37°C incubator. If no successful colonies were spotted, the procedure will be repeated from step 6. Repeated failures may call for the procedure to be repeated from steps 5, 2, or 1.

10. Plasmid DNA was extracted and purified using PureLink Quick Plasmid Miniprep kit (Invitrogen) using manufacturer’s protocol. The plasmid was then analyzed by enzyme digestion and gel electrophoresis for correct size, and sequenced using custom primers (TCAG, Toronto, ON).
Figure 3.3: Schematic of using MCSs to create fusion cassettes.

(A) Two genes of interest (A and B) in their respective cassettes. (B) C-terminal fusion of B (insert) to A (host). (C) N-terminal fusion of A (insert) to B (host). The A-B fusion cassette contains exactly the same MCSs as A and B, thus allowing the same procedure from (A) to (C) to be repeated.

3.1.3 Constructing fusion proteins by combining cassettes

Figure 3.3 demonstrates how fusion proteins can be constructed by making use of two pairs of restriction enzyme recognition sequences, of which one is compatible. In this example, SpeI/NheI makes up the compatible cut sites. The plasmid containing a fluorescent protein was generally selected as the insert. Most plasmids contain the Venus gene between NheI and XhoI. There are two pathways to create the fusion cassette. The first pathway involves inserting using SpeI/XhoI into NheI/XhoI and the second pathway involves inserting using NcoI/NheI into NcoI/SpeI. Note that the end product of the recombination between any two sequences results in a cassette with the same flanking cut sites. This enables sequential recombinations using one single protocol. The procedure for creating the fusion cassette is similar to that described in section 3.1.2, except that the first four steps were skipped and the insert was treated as the PCR fragment.
3.1.4 Removing fluorescent protein tag

Removing of Pmel-flanked fluorescent tag may be needed if it is undesired for the experiment or complicates the screening process. The procedure is largely similar to that described in section 3.1.2. The differences are summarized below.

- The first 4 steps were skipped. For step 5, only one digestion was performed, for the host vector bearing the undesirable fluorescent tag. Only one restriction enzyme was present, Pmel, so 1 extra µL of water was added.

- For step 6, self-ligation was performed between 2 Pmel sites on one plasmid. 17 µL of it was used instead of a combination of an insert and vector.

- For step 9, the colonies were screened for lack of fluorescence.

3.1.5 Generating cassettes for short peptides

Short peptides (< 100 nucleotides) used in this thesis cannot be created using the conventional method described in section 3.1.2 because they are too short to be resolved on a gel. Instead, the peptides were amplified as an extension to the Venus gene. Primers for Venus were designed with the 3’ primer bearing an overhang containing the peptide of interest (Figure 3.4). The cassette was constructed by inserting the peptide bearing Venus gene into the standard cloning vector following the procedure described in section 3.1.2. The primers were also designed such that Venus was flanked by SpeI sites so that it could be removed using the procedure described in section 3.1.4. Excision of Venus and self-ligation at SpeI yielded the peptide of interest in the standard cassette format which was used to create fusion cassettes described in section 3.1.3.
3.1.6 Performing point mutations

Point mutations were introduced using overlap extension PCR with self-hybridizing primers. The procedure is described below.

1. Four primers were needed. Primers A1 and A2 will be the forward and reverse primers for inserting the to-be-mutated gene into the cloning vector. Primer B1 was designed such that it annealed approximately 10-12 base pairs to the 5’ and 3’ sides of the mutation site. The primer should contain the desired mutant sequence near the middle of the primer and have a melting temperature of approximately 72°C. Primer B2 was designed as the reverse compliment of primer 2a.

2. Two separate PCR reactions were performed in parallel (procedure as described in section 3.1.2, step 2). In the first, primers A1 and B2 were the forward and reverse primers, respectively, amplifying fragment F1. In the second, primers B1 and A2 were the forward and reverse primers, respectively, amplifying fragment F2. The PCR products were checked by gel electrophoresis and gel purified.

3. A 40 µL PCR mixture was prepared with 20 µL F1, 20 µL F2, 5 µL dNTP’s, 5 µL buffer, 0.5 µL. This PCR was run for 35 cycles.

4. This gene was inserted into cassette (section 3.1.2, starting from step 3).
3.2 *In vitro* protein analysis

Expression of some proteins from *E. coli* cells were analyzed *in vitro*. Techniques such as affinity-tag purification, SDS-PAGE and fluorescent spectroscopy will be discussed.

3.2.1 Protein production and extraction from *E. coli*

Proteins were expressed from *E. coli* using the T7lac promoter of the standard subcloning vector and extracted by mechanical sonication.

1. *E. coli* transformed with the plasmid expressing the protein of interest were cultured overnight in 6-12 mL (depending on desired concentration) of LB supplemented with 100 µg/mL ampicillin in 37 °C shaking incubator.

2. Cells were harvested by centrifugation at 13200 rpm. The pellet was washed three times in Tris buffer (50 mM Tris and 100 mM NaCl) and re-suspended in 200-300 µL Tris buffer.

3. The cell suspension was sonicated for 1 minute using a Branson Sonicator Model 250 (Thermo Scientific, Ottawa, ON) with output intensity setting 2.5 and a 25% duty cycle.

4. After sonication, the suspension was centrifuged at 13,200 rpm for 3 minutes to separate the supernatant (which should contain the protein of interest if it was soluble) and the cell debris. The fluorescent intensity of the supernatant was assessed and compared to that of the cell debris to determine if sufficient protein was extracted, unless a non-fluorescent protein was being expressed.

5. The supernatant was transferred to a clean 1.5 mL Eppendorf tube and centrifuged at 13,200 rpm for 3 more minutes to weed out any remaining cell debris. The final supernatant was extracted into a clean 1.5 mL Eppendorf tube and used for any downstream applications.
3.2.2 Affinity tag purifications

Affinity tag purification was used to purify several His-tagged and GST-tagged proteins along with their binding partners. For His-tag purifications, the PopCulture His-Mag purification kit (Novagen) was used according to manufacturer’s protocol. For GST-tag purification, immobilized GST sepharose beads were used (Novagen). The procedure for GST-tag purification is detailed below.

1. 100 µL of bead slurry was removed from stock (approximately 50% of the slurry volume was beads).

2. The beads were allowed to settle in a 1.5 mL tube. The storage buffer was removed and 500 µL of Tris buffer was added to wash the beads. This process was repeated twice.

3. The final wash buffer was removed and an excess of the GST-tagged protein lysate was added, mixed and allowed to settle.

4. Excess protein lysate was removed from the beads and 500 µL of Tris buffer was added to wash the beads twice.

5. The final wash buffer was removed and 100 µL of Tris buffer was added to create a final suspension of bead slurry bound by GST-tagged proteins.

6. Unlike His-tag purification, proteins were not eluted from GST sepharose beads because the commonly used glutathione elution buffer questions Venus fluorescence. The beads suspension was loaded directly onto the polyacrylamide gel for SDS-PAGE.

3.2.3 SDS-PAGE

SDS-PAGE was used to analyze existence and size of proteins with fluorescent tags. Equal amounts of fluorescence-normalized lysates or purified protein solutions were mixed with 1x
NuPAGE LDS sample buffer (Invitrogen) to a total volume of 30 µL. The resulting mixture was heated at 70 °C for 10 minutes before they were loaded on NuPAGE Novex 4-12% Bis-Tris pre-cast polyacrylamide gels with 1x SDS-MOPS buffer (both from Invitrogen). The gels were run on Owl EC-105 Vertical Electrophoresis System (Thermo Scientific) at 120 V for 60-90 minutes until desired separation. PageRuler protein ladder (Fermentas) was run simultaneously and used for size reference. Gels were stained overnight in PageBlue Comassie blue-based stain (Fermentas) to visualize protein ladder. Fluorescent bands were visualized using Illumatool system, with appropriate excitation and emission filters as described in section 3.1.1. Photographs of the fluorescent bands are taken by a Canon A350 Powershot camera (chapter 4) and Sony Nex5N camera (chapter 5).

3.2.4 Fluorescence spectroscopy for FRET measurements

Fluorescence spectra were collected with an UV/Vis/NIR fluorescence spectrometer model LS55 running the FL WinLab software (PerkinElmer, Woodbridge, ON). Protein lysate samples (as described above) were loaded into UV/Vis cuvettes. Spectra for CFP-YFP based FRET were collected by exciting the sample at 440 nm and recording the emission spectra from 460 nm – 560 nm.

3.3 Cell culture and live-cell fluorescence imaging

Live-cell fluorescence imaging was used as the primary technique to characterize proteins and associated cellular events. Cell culture reagents, equipment, and procedures as well as imaging equipment and procedures will be described.
3.3.1 Materials and reagents for cell culture

The cell lines used in this work were COS-7, HEK293, HeLa, and CHO.

**Cell culture and transfection:** Cells were cultured in DMEM supplemented with 25 mM D-glucose, 1 mM sodium pyruvate, 4 mM L-glutamine and 10% FBS (Invitrogen). Culture medium was also supplemented with 100 U/mL penicillin and 100 µg/mL streptomycin (Sigma). Cells were passaged at 95% confluency with 0.05% trypsin-EDTA (Sigma). Lipofectamine 2000 was used as the transfection reagent (Invitrogen). Pre-mixed cell freezing medium was used containing DMSO and methylcellulose (Sigma).

**Cell culture materials and equipment:** Cell cultures were maintained in screw-top T-25 flasks (Sarstedt, Montreal, QC) in a HeraCell incubator maintained at 37 °C and 5% CO₂ (Thermo Scientific). All cell manipulations were performed in a Class IIA biosafety cabinet (Thermo Scientific). In preparation for imaging experiments, cells were passaged into 35 mm dishes with 14 mm glass bottom wells (MatTek, Ashland, MA). In preparation for protein extraction experiments, cells were passaged into T-25 flasks.

**Stimulatory reagents:** Stimulatory reagents were added directly by diluting 1:10 in imaging medium. The following stimulatory reagents were used at the indicated concentration: PDBu (10 µM, Sigma), ATP (10 µM, Fermentas), ionomycin (1 µM, Sigma), CaCl₂ (1 mM, Sigma), staurosporine (2 µM, Sigma). All reagents with the exception of CaCl₂ and ATP were prepared in DMSO stocks. After dilution, the final concentration of DMSO in the cell culture was not more than 0.1%.

3.3.2 Cell culture protocols

Cell passaging procedures, especially seeding ratios, were modified depending on the particular cell line in use, as described below.
1. When cells reached 95% confluency, medium was removed from cells by pipette. Cells were washed in 5 mL PBS (Invitrogen) and rocked gently.

2. 1 mL trypsin was added to cells for 5 minutes at 37 °C. Cells were loosened by tapping the flask and transferring the suspension to a clean 1.5 mL tube.

3. The tube was centrifuged at 2,000 rpm for 3 minutes to pellet the cells. The supernatant was removed and cells were re-suspended in 100 µL warm cell culture medium.

4. For propagation into a new T-25 flask, the flask was filled with 5 mL warm complete growth medium and the cell suspension dilution factor was 1:10. This allowed for approximately 3-4 days of culture before the next passage. For seeding into glass bottom wells, the well was filled with 2 mL medium and the dilution factor was 1:15 for HEK293, 1:12 for COS7, CHO and HeLa.

For cell freezing, Styrofoam trays were used to slow the freezing process and prevent ice crystals formation from damaging cell membranes.

1. Cell freezing medium was added drop-wise to the cell suspension by slowly dripping 0.5 mL of medium around the edge of the cryovial. Once added, the mixture was gently pipetted up and down to evenly distribute the cells.

2. After labeling, cryovials were placed in Styrofoam trays and put in a -20 °C freezer overnight.

3. Cells in the tray were then transferred to -80 °C freezer overnight. The next day cells were removed from the Styrofoam tray and put in cardboard freezer boxes for long-term storage.

Transfections were performed using Lipofectamine 2000 according to the manufacturer’s directions, which are summarized below.

1. One 1.5 mL tube was filled with 100 µL DMEM (without serum) and 1 µg plasmid. Another 1.5 mL tube was filled with 100 µL DMEM and 5 µL Lipofectamine 2000.
2. After 5-minute incubation, the contents were mixed together, and incubated for 25 minutes.

3. After the second incubation, the medium was removed from the cells in the glass-bottom dish that are the target of the transfection procedure and replaced with 0.8 mL warm DMEM.

4. The liposome-DNA complex mixture was added to the cell dish and incubated at 37 °C for 4 hours. Then, the medium was removed, 2 mL warm complete growth medium was added and the cells were tested for transgene expression after 18 to 24 hours.

3.3.3 Protein extraction from mammalian cells for SDS-PAGE

Protein expressions from COS-7 cells using SDS-PAGE were studied in chapter 5. The procedure for extracting proteins from COS-7 cells was largely similar to that of E. coli (section 3.2.1). The differences will be discussed below.

- For step 1, 6-8 T25 flasks of COS-7 cells were transfected with the each construct and cultured overnight (the transfection procedure detailed in section 3.3.2 was scaled up by 16 times for transfecting each flask). According to experimental demand, the flask cultures may be subject to further incubation at a different temperature.

- For step 2, cells were harvested by trypsinization as described in section 3.3.2. PBS was used to wash the cells instead of Tris buffer. The cells were centrifuged at 2,000 rpm instead of 13,200 rpm.

3.3.4 Imaging equipment

Imaging was performed using an inverted IX-81 microscope (Olympus, Markham ON) with a Lambda DG4 xenon lamp and QuantEM 512SC CCD camera. There were 10x (air) and 20x,
40x, 60x and 100x (oil immersion) objectives. Bandpass filters (Semrock, Rochester, NY) for CFP, YFP and RFP excitation were 438/24 nm, 500/24 nm and 580/20 nm, respectively, and for emission were 482/32 nm, 542/27 nm and 630/60 nm, respectively. The light intensity at the microscope stage, for CFP excitation light, was 25 mW/cm². On-stage temperature was controlled by a desktop incubation chamber (Solent Scientific, Segensworth, UK). Image acquisition and analysis were done using QEDInVivo (chapter 4) and MetaMorph Advanced (chapters 5 and 6).

3.3.5 Imaging procedures

Imaging experiments were conducted in cell culture medium, except for ATP and ionomycin experiments where PBS supplemented with 1 mM Ca²⁺ was used. For experiments where stimulants were added, 200 µL of 10x concentration were added dropwise around wells to improve even diffusion throughout the dish. For LOVS1K illumination experiments, CFP excitation was pulsed for 1 s every 30 s unless otherwise stated. All temperature induction steps were done at the indicated temperature either in an incubator (with no CO₂), or on the microscope stage. The cells were maintained in culture medium unless otherwise stated.

3.4 Data analysis

This section briefly discusses several techniques that were used to analyze data in this work. Statistical comparisons were made throughout. Fluorescent quantification and ratiometric calculations were used to quantify protein stoichiometry and biosensor cleavage.
3.4.1 Statistical comparisons

All experiments were performed at least in triplicates and graphed data is presented as mean ± standard deviation, unless otherwise stated. For comparing means, significance was calculated using unpaired Student’s t test assuming equal variance. For comparing medians, Kruskal-Wallis test was used. P < 0.05 was considered significant.

3.4.2 Absolute quantification of fluorescent proteins

For stoichiometry calculations in chapter 4, the amount of Cerulean and Venus expression was estimated. Using MetaMorph Advanced software, the relative amount of fluorescent proteins in a region of interest \((X_{ROI})\) was estimated by the following formula with the assumption that the cells were flat (2-dimensional):

\[
X_{ROI} = I_{ROI} \times N_{ROI} \times \varepsilon \times \Phi \times T_{filter}
\]

The region of interest (ROI) for calculating Cerulean and Venus fluorescence were the nucleus and the cytoplasm respectively. \(I_{ROI}\) is the background corrected average pixel intensity within the ROI and \(N_{ROI}\) is the total number of pixels within the ROI. The relative brightness of each fluorescent protein is represented by the product of its extinction coefficient (\(\varepsilon\)) and quantum yield (\(\Phi\)). \(T_{filter}\) is the transmission efficiency of the respective filter set in the microscope for each fluorescent protein. The \(\varepsilon\), \(\Phi\) and \(T_{filter}\) values for Venus are 92,200 M\(^{-1}\)cm\(^{-1}\), 0.57 and 0.77 respectively. The \(\varepsilon\), \(\Phi\) and \(T_{filter}\) values for Cerulean are 43,000 M\(^{-1}\)cm\(^{-1}\), 0.62 and 0.49 respectively.

3.4.3 Ratiometric quantification of fluorescent proteins

Nuc/Mem ratios for the TEVpSensor was calculated by the following formula:
\[
\text{Nuc/Mem} = \frac{I_{\text{CFP(nucleus)}}}{I_{\text{RFP(nucleus)}}} / \frac{I_{\text{CFP(membrane)}}}{I_{\text{RFP(membrane)}}}
\]

\(I_{\text{CFP}}\) and \(I_{\text{RFP}}\) are background-corrected average pixel intensities of CFP and RFP respectively within the indicated region of interest (nucleus or membrane). For N-degron reporters, unnormalized YFP and RFP values were background-corrected average pixel intensities of the cytoplasm. YFP/RFP values were normalized to the average YFP/RFP value of RFP-tevX-Deg-YFP when expressed on its own.
4 TEVp-facilitated stoichiometric delivery of multiple genes

The content of this chapter has been modified from the following peer-reviewed journal article (100).


4.1 Introduction and Aims

The development of proteins that can be used as biosensors or protein switches to monitor or control cellular processes is a key step in the investigation of a broad range of biological questions. One of the technical challenges faced is how one can reproducibly deliver multiple protein components to the cell. This chapter addresses these fundamental challenges by developing a method that can deliver multiple genes to cell in a way that is efficient, inducible and stoichiometric by using TEV protease (TEVp).

The specific aims for the work in this chapter are:

1. Design fusion proteins to demonstrate TEVp self-cleavage in *E. coli*
2. Demonstrate that TEVp self-cleavage products can be co-purified as a complex
3. Design fusion proteins to demonstrate TEVp self-cleavage in mammalian cells
4. Validate that the cleavage products are consistent in stoichiometry

Several fusion proteins were created during the characterization of TEVp self-cleavage (Figure 4.1).
Figure 4.1: Overview of fusion proteins employed for chapter 4.
Schematic of the fusion proteins are shown along with their names used in the text.
4.2 Results

4.2.1 Characterization of the TEV protease

The TEVp used was a variant developed by Kapust et al., which contains a S291V mutation (71). This mutation has been shown to reduce TEVp autoproteolysis and in turn enhance its proteolytic activity on canonical substrates (71, 72). To confirm the activity of this TEVp, a fluorescence resonance energy transfer (FRET) sensor (Ven-tevS-Ceru), consisting of the TEVp substrate tevS (ENLYFQS) flanked by the acceptor Venus (YFP variant (101)) and the donor Cerulean (enhanced CFP (102)), was constructed. A control sensor (Ven-linker-Ceru), with the TEV substrate tevS replaced by an arbitrary linker (DAPVRSLNCT), was also constructed. In vitro activity assay was conducted by mixing Venus-tagged TEVp (TEVpVen) and the sensors in a 1:1 ratio. The cleavage of the Ven-tevS-Ceru sensor was immediately observed by a gradual decline in FRET efficiency (Figure 4.2). In contrast, Ven-linker-Ceru remains unchanged in the experimental time frame of 40 minutes.

TEVp expression has been previously shown to be non-toxic in E. coli (4) and mammalian cells (7). To confirm this, TEVp was expressed in COS-7, HeLa and HEK-293 cells. Cell morphologies were unaffected, and the expression of TEVpVen was robust with distribution in the cytoplasm, nucleus and nucleolus of the cell. This indicates the unlikelihood for the existence of endogenous TEVp substrates in these cell types and establishes the crucial prerequisite for using TEVp in this study.
Figure 4.2: *In vitro* characterization of TEVp activity

(A) Emission spectra of the Ven-TeV-Ceru FRET sensor measured before TEVp addition (solid line) and 40 minutes after TEVp addition (dotted line). The FRET sensor was excited at 440 nm. The Cerulean and Venus emission peaks were identified at 475 nm and 528 nm respectively. (B) Time-lapse behaviour of the Ven-TeV-Ceru FRET sensor (diamond) and the non-cleavable control Ven-linker-Ceru (square) over 40 minutes after the addition of TEVp. A decrease in FRET ratio indicates the sensor being cleaved. Error bars represent s.d. (n = 3).
4.2.2 TEVp self-cleavage in E. coli

The TEVp self-cleavage of fusion proteins in E. coli effectively expressed multiple proteins which could then be purified individually or in protein complexes. TEVp self-cleavage refers to the cleavage of TEVp substrates residing on the same molecule as TEVp itself. The mode of interaction between TEVp and its substrate in self-cleavage events can be both intermolecular and intramolecular. Intramolecular cleavages are more efficient due to the closer proximity between the TEVp active site and its substrate, albeit only possible in cases where the substrate can be accessed by the active site through molecular bending and rotation. A natural example of TEVp self-cleavage is the autoproteolysis of wild-type TEVp between residues 218 and 219 (71, 72).

As an initial test for TEVp self-cleavage, TEVp was fused to the N-terminal of both Ven-tevS-Ceru and the control sensor Ven-linker-Ceru, creating TEVp-tevS-Ven-tevS-Ceru and TEVp-tevS-Ven-linker-Ceru respectively. FRET measurements were taken after one day of expression in E. coli cells. The FRET ratio (Venus/Cerulean) in cells expressing TEVp-tevS-Ven-linker-Ceru was 2.08 ± 0.13 (n = 3), comparable to the FRET ratio of 2.28 ± 0.12 in cells expressing Ven-linker-Ceru. On the other hand, cells expressing TEVp-tevS-Ven-tevS-Ceru recorded a much lower FRET ratio than cells expressing Ven-tevS-Ceru (2.68 ± 0.07 to 1.56 ± 0.04), indicating a separation of the two fluorescent proteins facilitated by TEVp self-cleavage.

To further investigate TEVp self-cleavage, a fusion protein was constructed with three separable modules: RFPHisTEVp-tevS-Ceru-tevS-VenGST (RHTeVpCVG) (Figure 4.3A). The three reporter genes used (mRFP, Cerulean and Venus) have distinctive spectral properties, allowing each subunit to be identified visually. TEVp itself was tethered to mRFP in order to determine if it could excise itself through an adjacent tevS site. The polyhistidine tag (His) and glutathione S-transferase tag (GST) allowed purification of the modules and served as size fillers so that each module can be spatially distinguished on an electrophoretic gel. Lysate extracted from E. coli cells expressing this fusion protein was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The major products were matched up against the controls and identified as RFPHisTEVp at ~40 kDa, Ceru at ~25 kDa and VenGST at ~50 kDa (Figure 4.3B).
His-purification and GST-purification were conducted to show that individual modules can be isolated from RHTEVpCVG. The His-purified and GST-purified samples produced the same bands as RFPHisTEVp and VenGST respectively (Figure 4.3C). While some uncleaved (all three modules) and intermediate cleavage products (two modules) were expected to be seen in both the lysate and purified samples, they did not appear on the gel, suggesting a complete or near complete cleavage of all TEV substrates between the modules.

Figure 4.3: TEVp self-cleavage in E. coli.

(A) Schematic of construct used in (B) and (C). Cleavage occurs at the sites containing tevS (ENLYFQS). (B) SDS-PAGE analysis for post-translational products of the construct shown in a compared against constructs encoding VenGST, Ceru and RFPHisTEVp individually. The three panels, from top to bottom, are pictures of the gel taken using Venus, mRFP and Cerulean excitation/emission filters respectively. Venus bands were detected under the mRFP filter due to cross excitation of Venus by mRFP excitation wavelength. The relative intensities of the bands
are not indicative of the expression levels for their corresponding fluorescent proteins. (C) SDS-PAGE analysis of His-purified and GST-purified samples of the construct shown in (A). The three panels follow the same pattern as described in (B).

Many cellular processes involve enzymatic or regulatory multi-protein complexes. It is often necessary to purify these complexes for biochemical studies. The conventional approach relies on in vitro reconstitution of proteins after they have been individually expressed and purified. This is a tedious process and may not work in cases where chaperone proteins are required for proper complex formation. Here, a utility of the TEVp self-cleavage system was demonstrated by expressing and purifying a two-protein complex from one TEVp self-cleavable fusion construct. In this case, the complex is reconstituted in vivo from the co-expressed proteins and only needs to be purified once. The construct used, HisSH3Ceru-tevS-VenCB1-tevS-TEVp (HSCVCB1TEVp) (Figure 4.4A), contained an N-terminal Src homology 3 domain of human Crk2 (SH3) and its binding peptide CB1. SH3 and CB1 were tagged with Cerulean and Venus respectively for visualization purposes. The post-cleavage products of HisSH3Ceru-tevS-VenCB1-tevS-TEVp were identified as HisSH3Ceru at ~30 kDa and VenCB1 at ~25 kDa by SDS-PAGE (Figure 4.4B). The same bands were produced after His-purification, suggesting that VenCB1 was co-purified together with HisSH3Ceru. On the other hand, the control construct HisSH3Ceru-tevS-Ven-tevS-TEVp (HSCVTEVp) in which CB1 was not present (Figure 4.4A), only yielded the HisSH3Ceru band after His-purification. This experiment demonstrated the feasibility of purifying a protein complex from a TEVp self-cleavable fusion protein.
**Figure 4.4: TEVp self-cleavage and subsequent purification of a binding complex.**

(A) Schematic of constructs used in (B). An additional tevS was added at the cleavage points compared to previous constructs, because having a double tevS site may improve cleavage efficiency. (B) SDS-PAGE analysis of post-translational products of the constructs shown in a together with their His-purified samples, showing successful purification of the SH3-CB1 protein complex expressed from the TEVp self-cleavage system. The two panels, from top to bottom, are pictures of the gel taken using Venus and Cerulean excitation/emission filters respectively.

### 4.2.3 TEVp self-cleavage in mammalian cells

The TEVp self-cleavage of fusion proteins in mammalian cells expresses multiple proteins that were localized into different subcellular compartments. As with the work with *E. coli* cells, the distinctive fluorescent properties of Venus, Cerulean and mRFP were utilized to noninvasively and simultaneously visualize the expression of up to three proteins in mammalian cells. When plasmids encoding Venus, Cerulean and mRFP were transfected into COS-7 cells, a uniform cytoplasmic and nucleoplasmic distribution was observed (data not shown). The size of these fluorescent proteins (~26 kDa) allowed them to cross nuclear pores.
To confine each fluorescent protein in a specific cellular compartment, two localization motifs were introduced: 1) NLS: the nuclear localization sequence (RIRKKLR) derived from the p54 protein (103), and 2) DAGR: the Cys1 domain of Protein Kinase C β-isoform, which translocates from the cytosol/nucleus to either the plasma or nuclear membrane upon phorbol 12,13-dibutyrate (PDBu) stimulation (104). A nuclear localizing module was built by attaching two NLSs in tandem to the C-terminus of Cerulean (CeruNLS), and a plasma membrane (PM) localizing module by attaching DAGR to the N-terminus of Venus (DAGRVen). To demonstrate TEVp self-cleavage and the subsequent subcellular targeting of the cleavage products, a four-module fusion protein TEVp-tevS-CeruNLS-tevS-DAGRVen-tevS-RFP (TEVpCNDVR) (Figure 4.5A) was constructed, containing mRFP as well as the aforementioned CeruNLS and DAGRVen modules. The principle of differential localization is illustrated in Figure 4.5B, showing the predicted targeting patterns of each module after cleavage.

After transfection of TEVpCNDVR in COS-7 cells, the subcellular localizations of the processed modules led to a distinctive fluorescence pattern as predicted. CeruNLS was observed only in the nucleus and nucleolus, whilst both DAGRVen and mRFP were evenly distributed in the entire cell (Figure 4.6). The addition of 10 µM of PDBu led to the migration of only DAGRVen to both plasma and nuclear membranes within 2 mins (Figure 4.6, Movie 4.1). Cytoplasmic and nuclear boundaries were confirmed by bright field microscopy. The same localization patterns were observed when plasmids encoding CeruNLS, DAGRVen and mRFP were individually transfected (data not shown).
Figure 4.5: Differential localization of TEVp self-cleavage product in mammalian cells.

(A) Schematics of the self-cleavage construct (TEVpCNDVR) and its non-cleavable control (CNDVR). (B) Conceptual illustration of what is expected to happen when the self-cleavage construct from (A) is expressed in eukaryotes. The CeruNLS module will enter the nucleus, while DAGRVen will bind to membrane-bound diglyceride (DAG) molecules upon PDBu stimulation.

The control construct CeruNLS-tevS-DAGRVen-tevS-RFP (CNDVR), in which TEVp was not present, produced intense Cerulean, Venus and mRFP fluorescence in the nucleus and was only able to localize to the nuclear membrane in response to 10 µM of PDBu (Figure 4.6). The display
of all three fluorescent signals and the ability for DAGR to correctly respond to its stimulus confirm that all modules were properly folded and their functionality preserved. The apparent fluorescent intensities of both fusion constructs were comparable to that of the single gene constructs, suggesting that multiple genes can be transfected and expressed as a fusion without a significant change in expression levels. The localizations of the three fluorescent modules also appeared to be mutually exclusive, indicating that the cleavage was complete within the detection capability of the microscope.

**Figure 4.6: Expression of multiple genes in COS-7 cells.**

Images of live COS-7 cells expressing fusion constructs introduced in Fig. 4A. The top 2 rows show the fluorescent patterns of the TEVp self-cleaving construct TEVp-tevS-CeruNLS-tevS-DAGRVen-tevS-RFP. The separate modules were distinctively localized, with only the DAGRVen module able to respond to PDBu by migrating to the membranes. The bottom two
rows show the fluorescent patterns of the non-cleavable control CeruNLS-tevS-DAGRVen-tevS-RFP. It is clearly locked in the nucleus by the strong NLS signal. PDBu induced migration could be barely observed as intensely fluorescent pockets (of all three colors) were formed on the inner circumference of the nuclear membrane. Scale bar is 30 µm and applies to all panels.

The transfection efficiency of the TEVpCNDVR fusion construct was approximately 14% following our standard transfection procedure (Figure 4.7). This low value is expected from constructs of its size, which may diminish its use in applications requiring high transfection efficiencies. However, it is possible to improve the transfection efficiency by optimizing DNA to liposome ratios or eliminate non-transfected cells by stable selection.

![Transfection Efficiency Graph](image)

**Figure 4.7: Transfection efficiency of TEVpCNDVR compared with its individual modules.**

COS-7 cells (at ~90% confluency) were transfected with each construct at a DNA to lipofectamine mass ratio of 1.6:4. Transfection efficiencies were calculated from the percentage of fluorescent cells under 20X magnification. Error bars represent s.d. (n = 4).

Differential localization experiments were repeated in HeLa and HEK-293 cells to show that TEVp self-cleavage is not cell-type specific. The expression of TEVpCNDVR was robust in both cell lines with identical localization patterns as in COS-7 cells (Figure 4.8). The expression of
the non-cleavable control construct was also identical. This highlights the advantage of having a self-contained cleavage system rather than relying on endogenous mechanisms for protein separation.

**Figure 4.8: Expression of multiple genes in HEK-293 and HeLa cells.**

Images of live HEK-293 and HeLa cells expressing TEVp-tevS-CeruNLS-tevS-DAGRVen-tevS-RFP, showing identical localization patterns as COS-7 cells (Figure 4.6). Scale bar is 30 µm and applies to all panels.
4.2.4 Stoichiometry of TEVp self-cleavage products

The single transfections of vectors encoding TEVp cleavage protein products yielded a consistent stoichiometry in contrast to a standard co-transfection of the separate protein products. To assess the stoichiometry of the cleavage products, three self-cleaving TEV fusions were constructed, each containing a combination of the Venus gene and the nuclear localizing CeruNLS gene in different ratios: TEVp-tevS-CeruNLS-tevS-Ven (TEVpCNV), TEVp-tevS-CeruNLS-tevS-Ven-tevS-Ven (TEVpCNVV), and TEVp-tevS-CeruNLS-tevS-CeruNLS-tevS-Ven (TEVpCNCNV) (Figure 4.9A). It is anticipated that the molar ratios of Venus and CeruNLS protein products will be dictated by the frequency of representation of their genes on the vector constructs. For example, if the construct contains one CeruNLS gene to one Venus gene, the protein products should be expressed in equimolar ratio.

The transfection of the aforementioned constructs in COS-7 cells each resulted in an evenly distributed Venus signal and an intense nuclear localized Cerulean signal (Figure 4.9B). To estimate the expression levels of Ven and CeruNLS, the fluorescent intensities of cytosolic Venus and nucleoplasmic Cerulean were measured respectively. For the TEVpCNV construct, the ratio between CeruNLS and Venus (CeruNLS/Venus) was 1.03 ± 0.13 (Figure 4.9C). With an additional Venus gene compared to TEVpCNV, the TEVpCNVV construct doubled its Venus expression as shown by its CeruNLS/Venus ratio of 0.51 ± 0.04 (Figure 4.9C). Similarly, TEVpCNCNV doubled its CeruNLS expression with a CeruNLS/Venus ratio of 2.03 ± 0.35 (Figure 4.9C). To compare against the TEV fusion constructs, cells co-transfected with Venus and CeruNLS were examined. Unsurprisingly, Venus/CeruNLS expression ratio in the co-transfected cells was 0.77 ± 0.57, which was non-stoichiometric and extremely inconsistent across the population (Figure 4.9C).
Figure 4.9: Stoichiometry of TEVp self-cleavage products.

(A) Schematics of constructs used for b and c. (B) Images of live COS-7 cells expressing each of the constructs shown in a. (C) CeruNLS/Venus expression ratios as represented by the ratio between nucleoplasmic Cerulean and cytoplasmic Venus fluorescence (n = 40). Error bars represent s.d. Scale bar is 30 µm and applies to all panels.
4.3 Discussion

This work describes a versatile approach for the stoichiometric delivery of multiple genes using a single expression vector by taking advantage of TEVp self-cleavage. The ability to express multiple genes is crucial for the development of protein-based synthetic systems. Artificial protein networks involve multiple protein modules, and often require co-delivery of biosensors to monitor their activities. The TEVp self-cleavage expression method mimics the protein expression strategy of RNA viruses, where multiple proteins are first expressed as a polypeptide before being cleaved by a protease into functional components (66). This strategy for expressing multiple proteins from a single fusion vector in eukaryotes was first conducted using endogenous proteases such as furin (105). However, there are practical difficulties in using an endogenous protease. First, endogenous proteases often have localized activities (e.g. furin in the Golgi) therefore the genes expressed must reside or at least pass through the same compartment as the protease. Second, expression of endogenous proteases is not consistent across cell types. Third, overexpression of such a protease is not feasible because it will result in undesirable cleavage of endogenous substrates.

TEVp is able to cleave its substrate ENLYFQS between QS with high specificity, leaving a 6-residue on the upstream protein and a single Ser residue on the downstream protein (68). TEVp is widely used in purification strategies to separate the passenger proteins from affinity tags due to its high specificity (71, 106). The high specificity of TEVp also allows it to be expressed in a number of model systems such as *E. coli* (4), *S. cerevisiae* (5), *D. melanogaster* (6), and mammalian cells cultures (7) without adverse effects on development and viability. As such, several studies utilized TEVp in eukaryotes to cleave engineered fusion proteins and study subsequent biological effects (6, 8, 9). However, TEVp has mostly been expressed separately from the fusion protein that it cleaves. To achieve genuine single vector expression, TEVp needs be delivered with the genes of interest as a fusion, requiring TEVp to “self-cleave” within the same molecule. Self-cleavage characteristics of TEVp has been previously explored in *E. coli* (107) and plants (108, 109), but was not used for the coordinated expression of multiple genes from a single vector. In this study, self-cleavage of TEVp was demonstrated in both *E. coli* and
mammalian cells (COS-7, HeLa, and HEK-293) for robust separation of up to three genes and implicated for the novel application of reconstitution and purification of protein complexes.

A salient feature of TEVp-assisted gene expression presented in this study is that the products are consistent in stoichiometry. Consistent stoichiometry of the simultaneously expressed genes is important because many protein-protein interactions involve stoichiometric binding partners (110, 111), and a consistent stoichiometry minimizes cell-to-cell variability when interpreting results. The most reliable stoichiometric multiple expression system to date is the 2A technology, which relies on a conserved translational skipping mechanisms in eukaryotes (20, 22, 23, 25). The 2A technology, however, has not been shown to work in bacterial cells (27), widely vary in their efficiency of peptide bond disruption (21, 30), and leave long residual trailing sequences (20-30 residues) to proteins that may perturb their function. Furthermore, in the case of some proteins targeted to the exocytic pathway, the cellular localization of proteins downstream of the 2A sequences were dictated by their upstream partners (29). In essence, 2A technology is not a self-contained system and requires specialized endogenous mechanisms for function. On the other hand, the TEVp self-cleavage system can theoretically be applied to any organism because TEVp is encapsulated into the same fusion protein along with the multiple genes of interest. However, this system also carries a disadvantage. The co-expression of TEVp results in its accumulation in the system. Although TEVp is well tolerated in many organisms (4-7), it is still unclear whether long-term expression will result in undesirable parasitic effects. Nevertheless, technologies capable of limiting protein half-life can be adapted to TEVp in order facilitate its clearance after cleavage (112).

This work will find compelling applications where the expression stoichiometry is particularly important. For example, using this approach in structural biology, the reconstitution of protein complexes can happen within the bacterial host cell at the specific stoichiometry of its components. In synthetic signaling pathways, the response of synthetic system can be optimized on the relative concentration of multiple signaling proteins. Last, in the induction of pluripotent stem cells (25) or neuron cells (113) it might be possible to improve the currently low induction rate of cells by a better control over the stoichiometry of the transcriptional factors.
4.4 Conclusion

In this chapter, a single-vector multiple gene expression strategy utilizing TEVp self-cleavage was created. The feasibility of this approach for robust expression of multiple genes was demonstrated in both bacterial (*E. coli*) and mammalian cells (COS-7, HeLa, and HEK-293).

1. In *E. coli* cells, up to three fluorescent proteins were expressed and could be individually purified using respective purification tags.

2. SH3 and its binding peptide CB1 were expressed simultaneously in *E. coli* cells and purified as a complex using only one purification tag.

3. In mammalian cells, up to three fluorescent-tagged genes were simultaneously expressed. Each gene was differentially targeted to subcellular compartments either constitutively by a localization motif (NLS), or conditionally by small-molecular induced interactions (DAGR recruitment to membrane-bound DAG).

4. The stoichiometry of protein products was consistent with the frequency of appearance of their genes on the expression vector.

Research objective 1 outlined in chapter 1 was addressed. A reliable and versatile method based on TEVp self-cleavage was developed for the simultaneous expression of multiple genes. The demonstration of this method lays the grounds for further developments of TEVp and associated applications in succeeding chapters.
5 Engineering temperature sensitive TEVp

5.1 Introduction and Aims

One can envision many ways in which TEVp can be used to manipulate biology as various physiological processes are linked to the cleavage of peptide bonds. The unique substrate ENLYFQ(G/S) of TEVp ensures targeting specificity and allows TEVp to be utilized without adverse effects on native biology. Temporal control of TEVp activity provides a potentially powerful tool to manipulate biological processes by specifically cleaving proteins of interest which could in turn direct protein localization or alter protein function. This reasoning is well supported by recent works that employed various inducible forms of TEVp to influence cellular events, such as apoptosis (8), cell-cycle (6), and inflammation (9). There are currently two existing methods for inducing TEVp activity on a post-translational level, the split-TEVp (7, 79-82) and the Tango assay (9). Both methods have considerably low signal-to-background ratios. In particular, the Tango assay is designed to only target plasma membrane substrates. Here, a novel modification to TEVp in the form of N-terminal truncations was presented which allows its activity to be temperature-induced in several mammalian cell cultures.

The specific aims for the work in this chapter are:

1. Design a TEVp biosensor for monitoring TEVp activity

2. Screen for a temperature-sensitive TEVp variant from step-wise terminal truncations

3. Demonstrate that temperature-sensitive TEVp can be used for efficient induced proteolysis in mammalian cells

Several proteins were constructed in this chapter to address the specific aims of creating an inducible TEVp (Figure 5.1).
Figure 5.1: Overview of fusion proteins employed for chapter 5.
Schematic of the fusion proteins are shown along with their names used in the text. Non-fluorescent versions of the TEVp variants were used for SDS-PAGE analysis and were omitted from this figure for simplicity.

5.2 Results

5.2.1 TEVpSensor for real-time monitoring of TEVp activity

A translocating fluorescent sensor, pLyn-RFP-tevS-CFP-NLS (TEVpSensor), was created. This sensor consists of RFP and CFP separated by the canonical TEVp recognition site (ENLYFQS, named tevS). RFP is N-terminally tagged with pLyn, the myristoylation and palmitoylation sequence from Lyn kinase, and CFP is C-terminally tagged with a nuclear localization sequence (NLS) (Figure 5.2A). In the absence of TEVp activity, both RFP and CFP will be tethered to the plasma membrane (PM). In the presence of TEVp activity, CFP will be released from the plasma membrane and translocate into the nucleus, creating a distinct color pattern. Differential localization of CFP and RFP was observed when TEVpSensor was co-expressed with TEVp (Figure 5.2B), but not with the inactive mutant TEVp(C151A) (Figure 5.2C).
Figure 5.2: Characterization of TEVpSensor.

(A) Schematic of TEVpSensor. (B) COS-7 cells co-expressing TEVpSensor and TEVp showing differential localization of RFP (membrane) and CFP (nucleus) due to cleavage at tevS. (C) COS-7 cells expressing TEVpSensor with inactive TEVp(C151A) mutant, showing co-localization of RFP and CFP on the membrane. All scale bars represent 30 µm. Cell outlines are marked by dashed lines.

5.2.2 Creation of temperature sensitive TEVp by incremental N-terminal shortening

The first secondary structure of TEVp is a short α-helix between Asn12 and Ser16. Immediately following the α-helix is a series of anti-parallel β-sheets which forms the characteristic β-barrel structure of TEVp (72, 114). Using the S219V mutant of TEVp as the template (71), incremental truncations beginning at Asn12 were conducted with the assumption that the unstructured peptide from residues 1-11 do not contribute to the stable conformation and activity of TEVp. A total of 10 truncated TEVp variants were created, named N12TEVp (Δ1-11) to L21TEVp (Δ1-20) (Figure 5.3). Activity from the truncated TEVp variants expressed and extracted from E. coli
could not be detected using a cell-free *in vitro* assay. It is likely because the truncations have worsened its solubility, which is already low without the truncations (71, 75, 76). Instead, their activity was screened in live COS-7 cells using TEVpSensor.

<table>
<thead>
<tr>
<th>TEVp:</th>
<th>MGLTSGSESFLKGRFYNDISSSTICHLTNEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>N12TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
<tr>
<td>P13TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
<tr>
<td>I14TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
<tr>
<td>S15TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
<tr>
<td>S16TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
<tr>
<td>T17TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
<tr>
<td>I18TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
<tr>
<td>C19TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
<tr>
<td>H20TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
<tr>
<td>L21TEVp:</td>
<td>MGLTSGSEWISSTICHLTNEA</td>
</tr>
</tbody>
</table>

α:  

<table>
<thead>
<tr>
<th>31</th>
<th>41</th>
<th>51</th>
<th>61</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHTTSLYGIFGPFITNKHLFRRNGLTVQLHGVFKVNTT</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>71</th>
<th>81</th>
<th>91</th>
<th>101</th>
<th>111</th>
<th>121</th>
<th>131</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLQQLIDGRDMIIRMPKDFPPPFFPOKLKRFEPQOREERCLVTNFO TKSMVSDTSCTFSSDGIFW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>141</th>
<th>151</th>
<th>161</th>
<th>171</th>
<th>181</th>
<th>191</th>
<th>201</th>
</tr>
</thead>
<tbody>
<tr>
<td>KSNIQTGDQCGPLVSTRGFTIGIHSANFTNTNYFTSVPKNXELLTNQEAQQWGSWLNASVL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RED: Residues from cloning sites  
Underline: β-strand  
Wave underline: α-helix  
*: C151A

**Figure 5.3: Protein sequence of truncated TEVp variants.**

The exact protein sequence of each truncated TEVp variant along with the location of C151A mutation, represented using one letter amino acid code.

Each truncated TEVp variant was co-expressed with TEVpSensor in COS-7 cells. Following a one-day post transfection period at 37 °C, the proteolytic activity of each variant was assessed by the proportion of transfected cells displaying a cleaved sensor (judged by a clear CFP-stained nucleus) (Figure 5.4A).
Figure 5.4: Temperature sensitivity of truncated TEVp variants.

(A) Percent of COS-7 cells displaying cleaved TEVpSensor when co-expressed with indicated truncated TEVp variants. TEVp truncated beyond I14 suffered reduced catalytic activity at 37 °C (data are mean ± std, n = 3 experiments/50 cells each). (B) Comparison of the number of cells displaying cleaved TEVpSensor after a 4-hour incubation at the indicated temperature. Results show improved catalytic activity at 30 °C for S15, I18 and C19 variants (data are mean ± std, n = 3/50, *p < 0.05, #p < 0.01). (C) Number of cells displaying cleaved TEVpSensor over a period of 16 hours under indicated expression and temperature conditions (n = 3/50). (D) SDS-PAGE analysis of TEVpSensor cleavage in COS-7 cells after the indicated duration under the indicated transfection and temperature conditions. Image was taken using RFP excitation/emission filters.

The deletion of the first 13 residues did not have a noticeable impact on the catalytic activity of TEVp. From S15TEVp onwards, the catalytic activity was significantly reduced, while H20TEVp and H21TEVp did not register any activity within the experimental time frame and temperature conditions. The greatest drop in catalytic activity occurred at I18TEVp (Δ1-17), which was also where the first β-strand residue was removed (Thr17).

Each variant was then subjected to a further 4-hour incubation at both 30 °C and 37 °C. Consequently, the catalytic efficiency for the five consecutive truncations from S15TEVp to
C19TEVp displayed temperature dependence (Figure 5.4B), with all of them registering higher activity at 30 °C than at 37 °C. The background activity of S15TEVp, S16TEVp and T17TEVp were large with over 50% of the cells displaying a cleaved sensor at 37 °C, which takes away their practical value unless expressed at much lower concentrations. The activity of T19TEVp was well suppressed at 37 °C but could only be slightly compensated by lowering the temperature to 30 °C. I18TEVp displayed the greatest contrast in activity between the two temperatures and the background activity of I18TEVp remained suppressed at 37 °C over a course of 16-hours (Figure 5.4C), and was therefore selected as the representative temperature-sensitive TEVp (henceforth referred to as tsTEVp) for the scope of this study.

SDS-PAGE analysis of lysates extracted from COS-7 cells expressing tsTEVp and TEVpSensor presented further evidence of temperature dependent activity of tsTEVp (I18TEVp) (Figure 5.4D). The two bands shown are for the full length TEVpSensor at ~50 kDa and the RFP component of the cleaved TEVpSensor (pLyn-RFP) at ~25 kDa. The CFP component of the cleaved TEVpSensor was omitted for simplicity. The SDS-PAGE analysis showed corresponding evidence as the cell imaging data shown in Figure 5.4C, with the exception that it provided an averaged indication of cleavage across all the cells within the population regardless of transfection success. Background cleavage at 37 °C was evident, but well suppressed within 16 hours after induction, as indicated by a much fainter band at pLyn-RFP level than at TEVpSensor level. Cleavage proceeded towards completion after 16 hours at 30 °C as indicated by disappearance of the band at TEVpSensor size.
Figure 5.5: tsTEVp activity in HEK-293, HeLa and CHO cell lines.

(tsTEVp and TEVpSensor were co-expressed in each cell type. 30 °C-induction for four hours resulted in greater number of cells displaying cleaved TEVpSensor for all cell types (*p < 0.01, n = 20/3).

Temperature-dependent activity of tsTEVp was also observed in HEK-293, HeLa and CHO cells which indicates that tsTEVp is not cell-type specific and can be applied in several other major cell lines without reduction in efficiency (Figure 5.5).

5.2.3 Real-time quantification of tsTEVp activity using TEVpSensor

A schematic of how TEVpSensor reports tsTEVp activity is shown in Figure 5.6A. Time-lapse fluorescent imaging were performed to monitor the progression of tsTEVp activity in COS-7 cells co-transfected with tsTEVp and TEVpSensor (Figure 5.6B and 5.6C). The cells were observed after a one-day post-transfection period at 37 °C. When the cells were placed at the permissive temperature of 30 °C, the activation of tsTEVp was visualized by the gradual increase in nuclear CFP intensity (Figure 5.6B, Movie 5.1). Cells maintained at 37 °C did not display differential localization of TEVpSensor during the same time frame (Movie 5.2). The extent of sensor cleavage was quantitatively represented by a ratiometric value Nuc/Mem (see Methods for formula) which estimates the differential localization of RFP and CFP of the sensor between
the nucleus and the membrane (Figure 5.6C). If CFP and RFP are completely co-localized (when TEVpSensor remains uncleaved), Nuc/Mem should have a value of 1.

Cells without initial TEVpSensor cleavage but with sufficient tsTEVp expression were selected to perform time-lapse activation experiment. The initial Nuc/Mem value of TEVpSensor when co-expressed with tsTEVp is $1.11 \pm 0.12$ which suggests a small background activity accumulated over the one-day post-transfection period before the experiment was conducted. However, this value is not significantly different from $1.03 \pm 0.07$ of the inactive mutant tsTEVp(C151A) ($p > 0.05, n = 3$). Nuc/Mem values of TEVpSensor co-expressed with tsTEVp increased from $1.11 \pm 0.12$ to $3.97 \pm 0.47$ over the span of four hours at $30 ^\circ C$ ($n = 3$), while remaining relatively stagnant at $37 ^\circ C$ or when co-expressed with the catalytically inactive tsTEVp(C151A).
Figure 5.6: Quantitative analysis of tsTEVp activity using TEVpSensor.

(A) Cartoon representation of how tsTEVp activity can be detected by translocation of TEVpSensor. (B) Progression of tsTEVp activity as indicated by TEVpSensor. Arrow indicates
visible nuclear staining by CFP after two hours. Scale bars represent 30 µm. Cell outlines are marked by dashed lines in selected panels (C) Time-lapse quantification of TEVpSensor cleavage under indicated expression and temperature conditions (data are mean ± std, n = 3). (D) Scatter plot showing Nuc/Mem values of individual cells (hollow circles) transfected with indicated construct along with TEVpSensor. The cells were incubated for four hours at 37 °C unless indicated otherwise. Solid line represents the median (n = 20 cells, *p < 0.001).

The population distribution of Nuc/Mem values of TEVpSensor under different expression and temperature conditions is shown in Figure 5.6D and 5.7 (n = 20). The median Nuc/Mem values of tsTEVp after four hours at 30 °C and 37 °C was not significantly different to that of full length TEVp and inactive TEVp(C151A) mutant respectively (p > 0.5), suggesting that tsTEVp is capable of rapid activation at 30 °C, while remaining sufficiently inactive at 37 °C. However, based on the distribution of Nuc/Mem values, isolated cases can be identified where there were high tsTEVp activity even at 30 °C. The variability was most likely a result of expression level. Higher expression of tsTEVp will inevitably result in faster accumulation of cleaved TEVpSensor from the background activity alone. An optimum level of TEVp expression (one that leads to sufficient inducible activity with minimum background activity) is desired and often existed within a population of cells but the variability in expression levels between cells was responsible for the broad distribution of results.

Figure 5.7: Quantitative analysis of tsTEVp activity at various temperatures.
Scatter plot showing Nuc/Mem values of individual cells (hollow circles) transfected with tsTEVp and TEVpSensor and incubated for four hours at indicated temperature. Solid line represents the median (n = 20 cells, *p < 0.001).
5.3 Discussion

Considering the functional relevance of proteases in biology (2), proteases and their targets can be of great interest in the engineering of synthetic biosystems. An inducible protease and substrate system, or a proteolytic device, can be engineered with the following design principles in mind: 1) it must be a self-contained system and has no off-target effects, 2) it must be minimally toxic to cells and organisms where it will be expressed, 3) the recognition sequence of this protease must be well tolerated within the context of the target protein, and 4) it must have minimum background activity in the absence of induction and sufficient inducible activity for proteolysis within the timescale of the experiment. TEVp has excellent specificity towards its cleavage sequence (71) and its expression is well tolerated in model systems such as E. coli (4), S. cerevisiae (5), D. melanogaster (6) and several mammalian cell cultures (7). Therefore, TEVp provides a solid foundation from which a proteolytic device can be engineered.

There are two steps to the process of engineering a proteolytic device: 1) the engineering of the protease itself to allow its activity to be temporally controlled by exogenous factors, and 2) the engineering of targets such that their functions and properties can be discriminately controlled by the protease. The prospects of engineering a conditional TEVp have been previously explored. There are currently two methods for inducing TEVp activity on a post-translational level, the split-TEVp (7, 79-82) and the Tango assay (9). In this work, novel modifications to TEVp in the form of N-terminal truncations was conducted, which yielded a variant capable of executing temperature-induced cleavages in a number of mammalian cell lines. While temperature sensitive mutants were typically created by random mutagenesis (44), the more systematic method of N-terminal truncations was used in this study on the basis of two previous findings: 1) Incremental truncations can result in step-wise reduction of activity (50-56), and 2) TEVp has an innate temperature sensitivity with peak enzymatic activity at 30 °C with a linear decrease in reactivity above and below this temperature (73, 74). Consequently, TEVp was weakened by incremental N-terminal truncations to the extent that its activity became negligible at 37 °C, while still retaining sufficient activity at 30 °C for rapid cleavage of its substrates. This new
variant, named tsTEVp, achieved efficient proteolysis on a time scale of four hours after 30 °C induction, while remaining sufficiently inactive at 37 °C.

The primary challenge in engineering an inducible TEVp device lies in striking a balance between low background activity and high inducible activity in order to achieve an optimum signal-to-noise ratio. This challenge was well manifested in both the split-TEVp and Tango method (7, 9). In the Tango method, a low-affinity TEVp cleavage site was employed to improve signal-to-noise ratio because a co-expressed cytoplasmic full-length TEVp cleaved substrates independent of co-localization (9). However, the catalytic efficiency of TEVp towards the low-affinity substrate became a severe rate-limiting step resulting in low inducible activity (81). The background activity of split-TEVp was also substantial, especially at high expression levels, possibly due to spontaneous reconstitution of the split fragments (8, 79, 81, 82). However, higher expression is often desired for efficient proteolysis because split-TEVp has intrinsically low catalytic efficiency (78, 79). The same situation holds true for tsTEVp. The variable nature of transfection led to differences in tsTEVp expression from cell to cell and was responsible the high background activity in some cells. However, an optimum expression level often existed within a population of cells therefore single cells can be selected for further downstream experiments. Also, such variability may pose less of a problem in population studies or studies involving cell lysates, where observations and measurements are averaged. Nevertheless, a balancing point between high cleavage efficiency and low background needs to be adjusted depending on the nature of the application, as some applications will not be able to tolerate background activity, while other applications will require faster cleavage kinetics.

Temperature induction has several unique advantages over other systems, including fast temporal response, good penetrability, and applicability to any tissue type or development stage of the organism. However, this method is only limited to systems where the transition between permissive and non-permissive temperatures is feasible. The operational range of tsTEVp is suitable for mammalian cells because mammals constrain their body temperature at 37 °C and can recover from a cold shock at as low as 24 °C without any long-term consequences (115-117). In this study, prolonged incubation of over 16 hours at the lower temperature of 30 °C did not
have any observable effects on cell morphology and viability as long as the cells were maintained in the culture medium. Beyond mammalian and bacterial cell cultures, a growth temperature above 30 °C is often stressful and even lethal, notably for the widely utilized model organisms *D. melanogaster* (118) and *C. elegans* (119). Nevertheless, a temperature shift can potentially result in confounding effects in any cell type by affecting endogenous cold shock proteins. Research have shown that exposing mammalian cells to reduced temperatures can result in slowed cell cycle progression, reduction in translation and transcription, and delayed apoptosis (120, 121). It is therefore important to avoid the application of tsTEVp in natural or synthetic systems that are intrinsically affected by temperature, or used for shorter time-scales such that off-target effects can be minimized.

Considering that temperature has excellent tissue penetration, tsTEVp could find therapeutic applications. In contrast to small-molecule induced systems such as the split-TEVp, there will be no complex pharmacokinetic issues relating to tsTEVp induction *in vivo*. Besides, hypothermia (at 30 °C to 32 °C) has already been an established therapeutic method for diseases such as cardiac arrest (122), stroke (123), and brain injury (124). Future work will involve the engineering of appropriate TEVp substrates that could be implemented along with tsTEVp to elicit temperature-induced biological effects.

### 5.4 Conclusion

In this chapter, a truncation of TEVp at its N-terminus allowed its activity to be inducible by temperature in mammalian cells.

1. A TEVp specific sensor, TEVpSensor, was designed. TEVpSensor is capable of reporting TEVp activity by membrane to nuclear translocation.

2. A series of truncated variants of TEVp were made. Truncations after S15 displayed marked reduction in catalytic efficiency at 37 °C.
3. Lowering reaction temperature to 30 °C can restore activity of truncated variants. N18TEVp variant displayed greatest signal-to-background ratio and was selected as temperature-sensitive TEVp (tsTEVp). Temperature inducible activity tsTEVp was quantified in mammalian cells using real-time imaging of TEVpSensor cleavage.

Research objective 2 outlined in chapter 1 was addressed. An inducible TEVp, tsTEVp, was developed. tsTEVp was inactive at physiological temperature 37 °C, and can be conditionally activated in mammalian cells by 30 °C induction. The completion of this objective allows for the conditional manipulation of biosystems outlined in research objective 3.
6 Synthetic manipulation of biosystems using tsTEVp

6.1 Introduction and Aims

Living cells use complex molecular signaling circuits to monitor external and internal signals and translate them into appropriate physiological responses. These circuits are made up of a dynamic network of interacting proteins that executes specific tasks such as translocation, activation/inhibition and degradation. The growing field of synthetic biology aims to create protein interactions beyond those found in nature, ultimately using them to probe, control, or even replace existing biological systems. The irreversible modification of proteins by proteolytic cleavage represents a simple and ubiquitous mechanism for signal propagation in living cells. The development of tsTEVp in the previous chapter offers the possibility of rewiring multiple biosystems to respond to a temperature cue without complete redesign of each system.

The specific aims for the work in this chapter are:

1. Use tsTEVp to activate protein translocation
2. Apply tsTEVp in a multi-module synthetic system to influence cellular blebbing behavior based on a combination of multiple input parameters
3. Design a protein degradation system conditionally controlled by tsTEVp
4. Use tsTEVp for temperature-induced cell death through the activation of a caspase-7 chimera

Several fusion proteins were created in this chapter to demonstrate the versatility of tsTEVp in the design of synthetic systems (Figure 6.1).
Figure 6.1: Overview of fusion proteins employed for Chapter 6.

Schematic of the fusion proteins are shown along with their names used in the text. Variants of tsTEVp and TEVp containing different fluorescent tags were used depending on the context indicated alongside the names. Five variants for RFP-tevX-Deg-YFP were created (X = C, F, H, S, and Y) and were omitted from this figure for simplicity.

6.2 Results

6.2.1 Protein translocation from nucleus to membrane mediated by tsTEVp and DAGR

A target protein can be rerouted from the nucleus to the membrane by combining two synthetic inputs: 1) temperature-induced tsTEVp activity, and 2) PDBu (phorbol 12,13-dibutyrate)-induced translocation of DAGR (DAG reporter, Cys1 domain of PKC β–isoform) (104). The approach is illustrated in Figure 6.2A. Within the construct of CFP-NLS-tevS-DAGR-YFP-NES, YFP is the target protein for translocation. YFP will be originally confined within nucleus by the presence of NLS but can be moved to the plasma membrane by a two-step process. First, the temperature-induced cleavage at tevS by tsTEVp will result in the separation DAGR-YFP-NES from CFP-NLS and its release into the cytoplasm by a nuclear export signal (NES). Second,
addition of PDBu will allow recruitment of DAGR to membrane-bound DAG (diacylglycerol) molecules, thereby directing YFP to the plasma membrane.

Figure 6.2: tsTEVp and DAGR mediated protein translocation.

(A) Cartoon representation of how a target protein (YFP) can be released from nuclear confinement and rerouted to the plasma membrane. (B) Percent COS-7 cells co-expressing CFP-NLS-tevS-DAGR-YFP-NES and tsTEVp displaying differential localization of YFP and CFP after 4-hour incubation at indicated temperature (data are mean ± std, n = 3 experiments/20 cells each, *p < 0.01). (C) Representative images showing 30 °C-induced release of YFP into the cytoplasm and subsequent PDBu-induced membrane localization. (D) Representative images
showing nuclear-confined YFP in the absence of temperature induction, followed by ineffective PDBu-induction. Scale bars represent 30 µm. Cell outlines are marked by dashed lines in selected panels.

A 4-hour incubation at 30 °C resulted in 57 ± 13% of COS-7 cells co-expressing CFP-NLS-tevS-DAGR-YFP-NES and tsTEVp to display a differential localization of CFP and YFP between the nucleus and the cytoplasm (Figures 6.2B and 6.2C). Both CFP and YFP were nuclear localized in most of the cells maintained at 37 °C for the same duration. The addition of PDBu triggered the secondary translocation event where cytoplasmic DAGR-YFP migrated to the plasma membrane (PM), typically within 2 to 5 minutes (Figure 6.2C). This event was visible in all cells with a sufficient cytoplasmic concentration of DAGR-YFP-NES (i.e. most of the cells in 30 °C-induced cultures). Cells maintained at 37 °C did not display PM-localized DAGR-YFP after PDBu-induction. Instead, the entire construct CFP-NLS-tevS-DAGR-YFP-NES formed a perimeter around the nucleus, a phenomena likely caused by the binding to DAG molecules present on the nuclear membrane (Figure 6.2D).

6.2.2 Cellular blebbing mediated by tsTEVp-controlled localization of CaRQ

tsTEVp can be used in conjunction with other synthetic modules to allow cellular states to be determined by a combination of different physical and chemical cues. Here describes the integration of tsTEVp with CaRQ, a synthetic Ca²⁺-sensitive RhoA which uses Ca²⁺ signaling to control cell morphology and migration (35). An important feature of CaRQ is that it responds differentially to different spatial patterns of Ca²⁺ signaling. In other words, the cellular localization of CaRQ defines its ability to be activated by a local Ca²⁺ transient. This property of CaRQ is harnessed in the design illustrated in Figure 6.3A by creating a derivative of CaRQ, pLyn-tevS-CaRQ. In this construct, CaRQ will be expressed either on the PM (by virtue of pLyn) or the cytoplasm depending on TEVp activity. When pLyn-tevS-CaRQ was co-expressed with tsTEVp in HEK-293, the differential localization of CaRQ became temperature dependent.
Only 28.3 ± 10.4% of cells which experienced a 4-hour 30 °C-incubation displayed membrane-localized CaRQ compared to almost all of cells maintained at 37 °C (Figures 6.3B).

To verify if pLyn-tevS-CaRQ can be activated by Ca^{2+}, ionomycin was added to stimulate a period of high Ca^{2+} influx into the cytoplasm. pLyn-tevS-CaRQ was sensitive to ionomycin regardless of its temperature-dependent localization by displaying high occurrence of blebbing after both 30 °C and 37 °C inductions (data not shown). To test the sensitivity of pLyn-tevS-CaRQ towards localized Ca^{2+} signals under different temperature conditions, cells were stimulated using two Ca^{2+} signaling modalities, ATP and LOVS1K/Orai1.

**Figure 6.3: tsTEVp and CaRQ mediated control over cell blebbing.**

(A) Cartoon representation of using tsTEVp to alter the localization of CaRQ, thereby modulating its sensitivity towards local Ca^{2+} signals. (B) Percent HEK-293 cells co-expressing
pLyn-tevS-CaRQ and tsTEVp displaying membrane localized CaRQ after 4-hour incubation at indicated temperature (data are mean ± std, n = 3 experiments/20 cells each, *p < 0.001). (C) Percent cells with blebbing morphology under each expression and stimulation conditions (data are mean ± std, n = 3/15, *p < 0.01, #p < 0.001).

ATP creates a brief Ca\(^{2+}\) transient near the PM via the IP\(_3\) pathway (125). 26.7 ± 2.9% of the cells maintained at 37 °C responded to ATP induction by displaying blebbing morphology within 20 minutes (Figures 6.3C and 6.4A, Movie 6.1). This value is not significantly different (p > 0.5) from the positive control, pLyn-CaRQ, where CaRQ was PM-localized under all conditions. However, blebbing was significantly reduced to 6.7 ± 2.9 % in 30 °C-induced cells, coherent with the depletion of CaRQ from the PM following tsTEVp activity (Figures 6.3C and 6.4B). This value is not significantly different (p > 0.5) from the negative control, CaRQ.

![Image](image_url)

**Figure 6.4:** tsTEVp and CaRQ mediated control over cell blebbing.
(A) HEK-293 cells expressing pLyn-tevS-CaRQ (shown) and tsTEVp incubated at 37 °C. ATP was added at 4 h. (B) HEK-293 cells expressing pLyn-tevS-CaRQ (shown) and tsTEVp incubated at 30 °C. Cleavage of pLyn-tevS-CaRQ appeared to be complete at 2 h. ATP was added at 2 h. (C) HEK-293 cells expressing DualCMVLOVS1K/Orai1, pLyn-tevS-CaRQ (shown) and tsTEVp after incubation at 37 °C for 4 hours. Image acquired after 1 min of blue light stimulus, pulsed at 1 s every 10 s. (D and E) HEK-293 cells expressing DualCMVLOVS1K/Orai1, pLyn-tevS-CaRQ (shown) and tsTEVp after incubation at 30 °C for 4 hours. Image acquired after 1 min (D) and 10 min (E) of blue light stimulus, pulsed at 1 s every 10 s. All scale bars represent 30 µm. Arrows indicate visible blebs.

LOVS1K is a synthetic protein which can reversibly activate membrane-bound Orai1 channels following blue light stimulation (36). In the following experiments, LOVS1K and Orai1 were co-expressed using a single DualCMV vector (DualCMVLOVS1K/Orai1). One minute of blue light excitation (1 s exposure, 10 s intervals, total 7 1s-doses) resulted in blebbing morphology in a vast majority (91.7 ± 7.6%) of the HEK-293 cells expressing DualCMVLOVS1K/Orai1, tsTEVp and pLyn-tevS-CaRQ at 37 °C (Figures 6.3C and 6.4C). Only 13.3 ± 10.4% of the 30 °C-induced cells were blebbing after the same dosage of blue light (Figures 6.3C and 6.4D). However, prolonged blue light excitation (> 5 minutes at 1 s exposure, 10 s intervals) eventually led to blebbing in all cells regardless of temperature induction due to global accumulation of Ca^{2+} (Figure 6.4E, Movie 6.2).

6.2.3 tsTEVp mediated protein degradation by N-degron activation

Temperature-dependent protein degradation was achieved by using tsTEVp to activate a dormant N-degron. The N-degron is a class of degrons where the N-terminal sequence of a protein dictates its in vivo stability (83). In eukaryotes, the N-degron consists of a destabilizing N-terminal residue (Cys, Arg, His, Lys, Leu, Phe, Trp, Tyr or Ile) (83), followed by a strategically positioned internal Lys residue for polyubiquitination (85, 86). Since Met is not a destabilizing N-terminal residue, N-degrons can only be activated by post-translational modifications capable of placing a destabilizing residue at the N-terminus of the target protein. The enzymatic activity of TEVp is flexible towards the last residue (X) of its 7-residue recognition sequence (ENLYFQ-
X), which becomes the new N-terminal residue after proteolytic cleavage (68). By placing a destabilizing residue at X, the new downstream fragment following TEVp cleavage will become susceptible to N-degron mediated degradation, as first demonstrated in yeast by Taxis et al. (112).

```
/    * *
M...RFP...RHSTGAEFLYFOXASSGAWLLPVSSLVKRASLFKLVSKGEE...YFP...
```

**Green**: TEVp recognition site (tevX)
**Orange**: Spacer residues from cloning sites
**Blue**: Deg peptide
**/**: Location of TEVp cleavage
**X**: New N-terminal residue after cleavage
**/**: Possible Lys sites for polyubiquitination

**Figure 6.5: Protein sequence of RFP-tevX-Deg-YFP.**

For simplicity, the complete sequence of RFP and YFP are not shown. The key elements of the N-degron are the TEVp recognition site (green), a destabilizing residue at X and an unstructured peptide containing internal Lys residues (blue).

This study presents a conditional degradation mechanism determined by temperature-induced activity of tsTEVp. The general strategy is illustrated in Figure 6.7A. YFP was used as a target of degradation, N-terminally coupled to RFP-tevX-Deg, yielding a fluorescent reporter in the form of RFP-tevX-Deg-YFP (Figure 6.5 and 6.7A). tevX are TEVp recognition sites with alternative residues (X) in place of Ser of tevS, and Deg is an unstructured peptide sequence which harbors an internal Lys residue for polyubiquitination (112). The cleavage of the reporter at tevX creates two fragments, RFP and X-Deg-YFP, where X becomes the new N-terminal residue of the downstream fragment.
Figure 6.6: Characterization of tevX-Deg variants.

Unnormalized YFP/RFP ratios from the co-expression of RFP-tevX-Deg-YFP and TEVp. Cleavage by TEVp resulted in reduction in YFP/RFP in constructs containing destabilizing residues C, F, H and Y (p < 0.001, n = 48).

The susceptibility of X-Deg-YFP to degradation is dependent on the nature of its N-terminal residue (X). The extent of degradation of X-Deg-YFP can be estimated. Four variants of RFP-tevX-Deg-YFP were tested (where X = C, F, N, Y) with full-length TEVp in COS-7 cells. One day after transfection, comparable levels of YFP degradation was observed for each variant as shown by the difference in YFP/RFP ratios compared to experiments in the absence of TEVp or when the stable variant RFP-tevS-Deg-YFP was used (Figure 6.6).
Figure 6.7: Synthetic activation of protein degradation by tsTEVp.

(A) Cartoon representation of temperature dependent protein degradation following N-degron exposure. (B) Normalized YFP/RFP values of COS-7 cells co-expressing the indicated proteins under different temperature conditions temperature (data are mean ± std, n = 48 cells, *p < 0.001). The YFP/RFP values are normalized to the average YFP/RFP value when RFP-tevH-Deg-YFP or RFP-tevS-Deg-YFP was expressed on their own. (C) Representative images of cells co-expressing tsTEVp and RFP-tevH-Deg-YFP after a 4-hour 30 °C-induction. Scale bars represent 30 µm.
RFP-.tevH-Deg-YFP was selected to analyze the capability of tsTEVp to activate degradation. Likewise, the extent of degradation was estimated by the ratio between YFP and RFP quantified from fluorescence microscopy. The ratios were normalized to the YFP/RFP ratio of the fluorescent reporters (RFP-tevH-Deg-YFP and RFP-tevS-Deg-YFP) without co-transfection of the derivatives of TEVp. Therefore, a ratio of one represents either an uncleaved reporter, or a cleaved reporter without differential degradation between RFP and YFP components. There was significantly higher YFP/RFP ratio after 1-day at 30 °C than at 37 °C (p < 0.001) when RFP-tevH-Deg-YFP was co-transfected with tsTEVp (Figure 6.7B). When RFP-tevS-Deg-YFP was used, the YFP/RFP ratio remained unchanged (Figure 6.7B). This indicates that the stability of Deg-YFP was dependent on the nature of its N-terminal residue, which was conditionally exposed at 30 °C by tsTEVp activity. Time-lapse imaging of cells was used to monitor the depletion of YFP relative to RFP. Representative images of cells expressing RFP-tevH-Deg-YFP and tsTEVp at 30 °C over a course of eight hours are shown in Figure 6.7C.

6.2.4 Synthetic activation of caspase-7 by tsTEVp

tsTEVp can be used to activate caspase-7 by proteolytic removal of its prodomain. Caspase-7 is a member of the caspase family of proteases which are important participants in the process of apoptosis (93). Also known as an executioner caspase, caspase-7 is directly responsible for cleaving most of the substrates involved in apoptosis and the activation of caspase-7 itself is sufficient to induce cell death (8, 35, 93). Caspase-7 is translated as an inactive zymogen which can be activated by proteolytic removal of its prodomain at Asp24 and separation of its p20 and p11 subunits at Asp198 (97), a process typically performed by initiator caspasases in response to extrinsic signals or intrinsic cellular stress (93). However, it is possible to rewire the activation mechanism of caspase-7 by allowing exogenous proteases such as TEVp to perform the required proteolysis. This concept was recently demonstrated in the SNIPer method by Gray et al. (8). In SNIPer, a rapamycin-inducible split-TEVp was used to activate a caspase-7 chimera containing embedded tevS sites at Asp24 and Asp198. The activation of this chimera upon TEVp induction led to robust apoptosis, rigorously validated by standard biochemical assays (8).
Two tevS sites are inserted between residues P56 and T57 of full length human Caspase-7. For simplicity, RFP sequence is not shown.

Here, caspase-7 activation was rewired to the temperature-induced activity of tsTEVp. A similar caspase-7 chimera (named Casp7tevS) was created in which only one tevS insertion was made between residues Pro56 and Thr57 of caspase-7 (Figures 6.8 and 6.9A). This site of insertion was chosen because caspase-7 truncated at Thr57 (57casp7) has been previously shown to cleave its substrate DEVD and induced apoptosis when overexpressed in mammalian cells (35, 98). A membrane localized derivative, pLyn-Casp7tevS, was also created so that its cleavage by tsTEVp can be visually inspected (Figure 6.9A). When pLyn-Casp7tevS was co-expressed with tsTEVp in COS-7, temperature dependent differential localization of the cleaved portion, 57Casp7 was observed (Figures 6.9B). A 4-hour 30 °C induction resulted in cleaved pLyn-Casp7tevS in 70.0 ± 8.7% of the cells. These cells began to develop hallmarks of cell death (membrane blebbing and cell rounding) over the next 2 to 4 hours.

In the study summarized by Figure 6.9C, each combination of constructs were expressed in COS-7 cells as indicated. The cultures were then placed in two study groups (temperature-induced and non-induced) over a total duration of 8 hours. Cultures in the temperature-induced group were 30 °C-induced for the first 4 hours and then re-incubated at 37 °C for the remaining 4 hours. Cultures in the non-induced group were incubated at 37 °C for the entire duration. More than 70% of cells displayed apoptotic phenotypes when Casp7tevS was expressed with full length TEVp, indicating its potency as an inducer of cell death (Figure 6.9C). For cells
expressing tsTEVp and Casp7tevS, cell death was suppressed at 37 °C due to tsTEVp inactivity, but 30 °C-induction resulted in a significant increase in the number of apoptotic cells from 16 ± 6.5 % to 47 ± 13.5% (p < 0.01). A greater number of 30 °C-induced cells also stained positive for propidium iodide (PI), a dye used to mark late apoptosis or necrotic cells (Figure 6.10).

Expression of tsTEVp alone or with the uncleavable Casp7 did not result in an increase in apoptosis-like cells regardless of temperature induction. Expression of Casp7tevS with inactive tsTEVp(C151A) mutant shows that background activity from Casp7tevS itself is low.

Figure 6.9: Synthetic activation of a Caspase-7 chimera.

(A) Schematic of Caspase-7 chimeras used in this study. Fluorescent tags are omitted for simplicity. (B) Percent COS-7 cells co-expressing pLyn-Casp7tevS and tsTEVp displaying membrane localized Casp7tevS after 4-hour incubation at indicated temperature (data are mean ± std, n = 3 experiments/20 cells each, *p < 0.001). (C) Percent COS-7 cells under indicated expression and temperature conditions for 8 hours, displaying at least one apoptosis-like morphology (membrane blebbing, cell rounding or disintegration into apoptotic bodies) (data are mean ± std, n = 5 experiments/20 cells each, *p < 0.01).
Figure 6.10: Propidium iodide staining.

(A) Percent COS-7 cells expressing Casp7tevS and tsTEVp stained positive with propidium iodide (PI) after 1 day incubation at indicated temperature (data are mean ± std, n = 5/20, *p < 0.001). (B) Representative images of cells from (A) stained by propidium iodide (red). Transfected cells are shown in blue. Scale bars represent 30 µm.

A caspase-3/7 specific biosensor, CaspSensor was developed to allow activity of Casp7tevS to be monitored in real time (Figure 6.11A). Based on a previous design (126), CaspSensor reports the cleavage of DEVD (a well-known substrate of caspase-3/7) by the translocation of CFP from the membrane to the nucleus. CaspSensor responded reliably to both the constitutively active 57Casp7 and staurosporine (STS)-induced apoptosis by displaying membrane to nucleus translocation (Figures 6.11B and 6.11C).
Figure 6.11: Characterization of CaspSensor.

(A) CaspSensor is single fluorophore probe which can translocate from the membrane to the nucleus following cleavage at DEVD. (B) COS-7 cells co-expressing CaspSensor (blue) with pLyn-RFP (red) shows colocalization on membrane. (C) Co-expression of CaspSensor, pLyn-RFP and 57Casp7 resulted in COS-7 cells with apoptotic morphology and differential localization between CaspSensor (blue, nucleus) and pLyn-RFP (red, membrane). (D) COS-7 cells co-expressing CaspSensor (blue) and pLyn-RFP (red) stimulated with apoptosis inducing agent STS (2 μM) at 0 mins. Onset of apoptosis at 200 mins marked by translocation of CaspSensor and gradual cell shrinkage. All scale bars represent 30 μm.

The cartoon schematic in Figure 6.12A demonstrates how CaspSensor was employed to report tsTEVp mediated activation of Casp7tevS. CaspSensor was employed in a similar population study demonstrated in Figure 6.9C and found correlation between the appearance of apoptotic morphologies with CaspSensor cleavage (judged by lack of membrane localized CFP and/or presence of nuclear localized CFP) for most of the cells under each expression and temperature
conditions (Figure 6.9C and 6.12B). The entire Casp7tevS activation process within a representative cell expressing pLyn-Casp7tevS, tsTEVp and CaspSensor can be visualized in a time-lapse imaging experiment (Figure 6.12D, Movies 6.3 and 6.4). In this particular example, cleavage of pLyn-Casp7tevS was evident from its translocation from membrane to cytoplasm within the first hour of 30 °C incubation. This was followed by gradual cell shrinkage, and ultimately characteristic apoptotic morphologies of cell blebbing and rounding after 6 hours. CaspSensor cleavage was observed as CFP detached from the membrane and entered the nucleus after 4 hours of temperature induction (Figure 6.12D). Nuclear localization of CaspSensor could only be observed for a short duration of about 20 minutes, likely due to nuclear envelop breakdown. This process was reproducible, each with similar timings of the different apoptotic events of cell shrinking, cleavage of CaspSensor and cell blebbing. An interesting observation was that the completion of CaspSensor cleavage was not gradual (which one would expect after the activation of Casp7tevS), but rapid during a short window of time which also coincided with an increased rate of cell shrinking. This provides strong indication that endogenous regulatory processes were involved which inhibits activity Casp7tevS until a certain activation threshold was achieved.
Figure 6.12: Casp7tevS activity indicated by CaspSensor cleavage.

(A) Cartoon representation of how activation of tsTEVp leads to cleavage of pLyn-Casp7tevS followed by CaspSensor, resulting in a visible translocation event of CFP from membrane to
nucleus. (B) Percent COS-7 cells displaying cleaved CaspSensor under indicated expression and temperature conditions for 8 hours (data are mean ± std, n = 5/20, *p < 0.01). (C) Time course showing percentage of cells displaying cleaved CaspSensor under indicated expression and temperature conditions (data are mean ± std, n = 5/20). (D) Time-lapse images of a cell expressing pLyn-Casp7tevS (red), CaspSensor (blue) and tsTEVp incubated at 30 °C. The panels show representative milestones over the course of the experiment. At 1 h, cleavage of pLyn-Casp7tevS was evident from the lack of membrane ruffles compared to 0 h. At 4 h 20 mins, cell underwent significant shrinking, and cleavage of CaspSensor was indicated by its detachment from the membrane and translocation to the nucleus (indicated by arrows). At 6 h 10 mins, blebbing morphology was observed. Scale bars represent 30 µm.

6.3 Discussion

In nature, post-translational processes regulate biochemical events that are both temporally rapid and spatially resolved. However, development of synthetic tools for post-translational control has lagged behind those for controlling gene expression (15, 127). One of the reasons is that post-translational systems are difficult to engineer on an individual basis. tsTEVp offers the possibility of rewiring multiple systems to respond to a temperature cue without complete redesign of each system. The broad applicability of tsTEVp lies in the fact that proteolytic cleavage can regulate protein function in many ways. For example, the cleavage of self-inhibitory domains can induce constitutively activation of a protein (e.g. ROCK1 (95)). The truncation of proteins by cleavage can destabilize a protein and result in loss of function (e.g. ICAD (128)). Removal of localization motifs can result in redistribution of a protein to its active or inactive compartment, resulting in either gain or loss of function (e.g. Chk1 (129)). Here, each mode of protein regulation was demonstrated. Specifically, target proteins were activated (DAGR) or inactivated (CaRQ) by translocation, inactivated by removal of stabilizing fragments (N-degrons), and activated by removal of auto-inhibitory domains (Casp7tevS). Simple observable traits of localization, concentration and morphology were used to evaluate consequences of each cleavage. The results of each study are further discussed below.

Translocation of DAGR: Spatial localization is a key process in the regulation of protein function in eukaryotic cells. As such, modifying the location of a protein is a proven tool for rewiring
signaling pathways (130-134). In this study, two translocation events were programmed using a single fusion protein, CFP-NLS-tevS-DAGR-YFP-NES. Each translocation event highlights a specific theme in protein translocation as means to regulate function.

The first event is a temperature-induced nuclear to cytoplasmic translocation of DAGR triggered by tsTEVp cleavage. The nuclear membrane acts as a physical barrier for the passive exchange of nuclear and cytoplasmic proteins and associated molecular interactions to take place (135). In the context of CFP-NLS-tevS-DAGR-YFP-NES, DAGR is enclosed within the nucleus at 37 °C. The addition of PDBu could not stimulate the binding between DAGR and DAG before the translocation of DAGR out of the nucleus took place. The mechanism of this translocation event was mediated by a pair of dominant and secondary “cellular address” tags. In the absence of cleavage, the target protein DAGR will be localized in accordance with the dominant tag (NLS). After cleavage, the target protein will be released from the dominant tag and localize according to the secondary tag (NES). This demonstrates a general strategy for protease-mediated translocation which has unique advantages over the more commonly used chemical-induced dimerization (CID) methods (130-134). CIDs cannot be used to mediate translocation from a specific cellular address (e.g. nucleus) to non-specific cellular address (e.g. cytosol).

Translocation between two specific addresses (e.g. between nucleus and the plasma membrane) can also be difficult using CIDs because the affinity between the dimerization partners has to outcompete the affinity of the tags to their respective addresses. The second event is a PDBu-induced cytoplasm to membrane translocation, specifically involving the binding of DAGR to DAG. The lipid DAG is a component of biological membranes of all eukaryotic cells (136). The key function of DAG is to mediate protein translocation to membranes by its direct binding to their Cys1 domains (136). DAGR is the Cys1 domain of PKC β–isoform and is primarily a localization motif without other functionalities (104). The plasma membrane can also be considered as a good location for specific molecular interactions to take place even without a delimiting boundary. The transition from the three-dimensional space of the cytoplasm to the two-dimensional PM surface can result in an increase in concentration of at least one order of magnitude (137). For specific subdomains such as lipid rafts on the membrane, this increase can be even greater. This is particularly relevant for DAGs because they tend to accumulate in
specific regions on the membrane, such as the immunological synapse during T-lymphocyte activation (138, 139). In summary, this method allows target proteins to be rerouted from the nucleus to the membrane by combining two synthetic inputs: 1) temperature-induced tsTEVp activity, and 2) PDBu-induced translocation of DAGR to DAG.

**Induction of cellular blebbing using CaRQ:** This study demonstrates that tsTEVp can be built into a multi-module synthetic system, thereby enabling it to elicit differential output behaviors based on a combination of input parameters. This multi-module synthetic system was composed of three synthetic transgenes: tsTEVp, pLyn-tevS-CaRQ and LOVS1K/Orai1. CaRQ is a synthetic Ca\(^{2+}\)-sensitive RhoA which uses Ca\(^{2+}\) signaling to control cell morphology and migration (35). Several exogenous stimuli can be used to mobilize intracellular Ca\(^{2+}\) stores or mediate extracellular Ca\(^{2+}\) entry. These include chemicals such as ATP (125) and acetylcholine (140), and physical cues such as light (36, 141) and voltage (142). In this study, the two Ca\(^{2+}\) mobilizing modules employed were purinergic receptors (through ATP binding) and LOVS1K/Orai1 (activated by blue-light). A modification was made to CaRQ by attaching a TEVp cleavable membrane localization tag so that its localization is dependent on temperature-induced tsTEVp activity. The regulation of pLyn-tevS-CaRQ activity is another example of the localization theme described previously. Localization can regulate a protein of interest either by sequestering it away from its biological targets or away from its effectors. In the case of pLyn-tevS-CaRQ, both modes of regulation might be present. The change from the two-dimensional membrane surface to the three-dimensional space of the cytosol results in a concentration effect of at least one order of magnitude (137). Localization of CaRQ away from the membrane could prevent its access to the sites of Ca\(^{2+}\) entry through Orai1 channels and purinergic receptors (35, 36) while at the same time reduces its apparent concentration around its targets (e.g. ROCK) which are most likely membrane associated (92). However, the cytoplasmic localization of pLyn-tevS-CaRQ after tsTEVp cleavage did not have a significant effect on its sensitivity towards global Ca\(^{2+}\) increase induced by ionomycin or prolonged LOVS1K/Orai1 activation. This suggests that the localization of its effector (Ca\(^{2+}\) ions) has a greater contribution towards CaRQ regulation in the context of pLyn-tevS-CaRQ. While cytoplasmic CaRQ was still mildly sensitive to Ca\(^{2+}\) signals, it might be possible to further suppress its activity by directing it to a
membrane-confined compartment such as the nucleus using an NLS tag in the manner of TEVpSensor (chapter 5). The observable biological output of this multi-module system is cellular blebbing. The specific physiological role of blebs is still subject to debate. Most known for its role in apoptosis (93), cellular blebbing has been recently implicated as an important motility mechanism in the absence of the common lamellipodia-driven migration (92, 143). Notably, bleb-driven migration is essential for development in some cell types (144) and can be used by cancer cells for metastasis (145). While temperature-dependent tsTEVp activity does not induce blebs per se, it diminishes the cell’s susceptibility to CaRQ-mediated blebbing. In essence, tsTEVp provides an avenue for rapid inactivation or attenuation of other synthetic modules once they are no longer required. Taken together, this study is among the first to demonstrate how well defined synthetic tools (tsTEVp, CaRQ and LOVS1K/Orai1) can be pieced together to form more complex networks that is able to process multiplexed input stimulus.

N-degron mediated protein degradation: A conditional degradation mechanism determined by temperature-induced activity of tsTEVp was engineered. Activation of N-degron was achieved by tsTEVp cleavage, resulting in rapid degradation of a fluorescent tester construct within one day. The implementation of tsTEVp-mediated degradation is similar to the previously reported TIPI (TEVp induced protein inactivation) system (112, 146): 1) Both systems utilize a conserved degradation pathway known as the N-end rule, 2) Conditional degradation is achieved by removal of a blocker moiety, which exposes a destabilizing residue located at the N-terminus of the target protein, and 3) Removal of the blocker moiety is facilitated by cleavage of a TEVp substrate. The key difference is that TIPI is activated by induced expression of TEVp from the galactose responsive GAL1 promoter. This limits TIPI as a yeast specific technique. While in theory TIPI could be applied to mammalian cells using equivalent inducible expression systems such as the Tet-based induction system, they are slower and often suffer from leakiness (147, 148). In the construct RFP-tevH-Deg-YFP presented in this study, RFP was used as the blocker moiety so that quantitative comparisons with the target protein (YFP) could be made, although it is not necessary for N-degron function. The minimum required N-terminal attachment to the target protein is the 24-residue tevH-Deg peptide. Deg is a 17-residue truncated version of a 40-residue E. coli Lac repressor-derived sequence called e^K (88). The truncation was made to
minimize the length of the degron while still containing the crucial Lys residue at positions 15 and 17 for attachment of ubiquitin molecules. In essence, Deg is the first minimally sufficient N-degron sequence that has been shown to work in mammalian cells. Another temperature-induced N-degron system, simply known as the TS degron, was previously described (40, 57-60). TS degron functions between 23 °C and 37 °C, with 37 °C being the temperature that induces degradation. A great advantage of TS degron is that degradation occurs within 30 minutes after temperature shift in yeast, which makes it one of the fastest degradation systems available. Given that the N-end rule is conserved in all eukaryotes, TS degron should work in mammalian cells in principle. However, its activation temperature of 37 °C coincides with the homeostatic temperature of most mammals, which makes its actual application infeasible. To date, tsTEVp-mediated protein degradation is the first temperature-sensitive degradation system for mammalian cells.

**Synthetic activation of caspase-7:** tsTEVp was used to activate caspase-7 by proteolytic removal of its prodomain. In mammals, Caspase-7 is translated as an inactive zymogen that can be activated by separation of its p20 and p11 subunits at Asp198 followed proteolytic removal of its prodomain at Asp24 (97, 149). Gray et al. first demonstrated that it is possible to rewire the activation mechanism of caspase-7 by using exogenous proteases such as TEVp to mimic endogenous proteolysis at the two key sites (8). In their SNIPer method, a rapamycin-inducible split-TEVp was used to activate a caspase-7 chimera containing embedded tevS sites at Asp24 and Asp198. The activation of this chimera upon TEVp induction led to robust apoptosis (8). In this work, caspase-7 was activated solely by cleavage at the non-canonical cleavage site of Thr57. This mode of activation has never been tested before and relies on auto-proteolysis by the caspase itself at Asp198. The current model suggests that the N-terminal prodomain only serves as an inhibitor to caspase auto-proteolysis. Removal of the prodomain does not result in caspase activation per se as cleavage at Asp198 remains the key determinant of caspase activation (97, 150). Consequently, the cleavage at Asp24 alone could not result in caspase-7 activity in the SNIPer method (8). Nevertheless, studies have shown that the caspase-7 truncated at Thr57 (57Casp7) is active in vitro (98) and overexpression in mammalian cells causes apoptosis-like cell death (35, 151), although the exact mechanism of activation remains elusive. The presented
data in this study showed that the cleavage of the caspase-7 chimera Casp7tevS directly in front of Thr57 by tsTEVp led to apoptotic signatures such as shrinking, blebbing, and cleavage of CaspSensor within eight hours. However, these events did not occur until a few hours after cleavage as indicated by time lapse imaging experiments, as one would expect from the gradual accumulation of cleaved Casp7tevS. This has two implications: 1) The cleavage of caspase-7 before Thr57 generates an inactive fragment which takes time to self-process into its active form by auto-proteolysis, and 2) The cells have an innate ability to buffer transient caspase-7 activity before a certain threshold is reached. While little is known about the kinetics of caspase-7 auto-proteolysis when truncated at Thr57, there could be several endogenous inhibitory proteins such as XIAP which can bind and inhibit caspase-7 activity (152). The activity from the Casp7tevS transgene alone may not be sufficient to overcome endogenous regulatory mechanisms but prolonged feedback activation of endogenous caspases (such as caspase-9 (153)) by Casp7tevS may eventually surmount a certain activation threshold, after which results in rapid cleavage of caspase-7 specific targets responsible for the characteristic apoptotic morphology. This was exemplified by the rapid cleavage of CaspSensor observed during the same time frame as the morphological changes. The temporal accuracy of caspase-7 activation by tsTEVp could facilitate further studies of the timing of apoptotic events at the single cell level.

### 6.4 Conclusions

In this chapter, tsTEVp was used together with a catalogue of other synthetic tools to generate temperature-induced biological responses, namely protein localization, cellular blebbing, protein degradation, and cell death.

1. Two translocation events were programmed using a single fusion protein, CFP-NLS-tevS-DAGR-YFP-NES. tsTEVp cleavage mediated nuclear to cytoplasmic translocation of DAGR which is subsequently rerouted to the plasma membrane by PDBu-induced binding to DAG.
2. Cellular blebbing was programmed using three synthetic modules tsTEVp, pLyn-tevS-CaRQ and LOVS1K/Orai1. The localization of pLyn-tevS-CaRQ mediated by tsTEVp determined its sensitivity towards local Ca\(^{2+}\) ions mobilized by purinergic receptors (ATP) and LOVS1K/Orai1 (light).

3. A conditional degradation mechanism determined by temperature-induced activity of tsTEVp was engineered. This mechanism is based on exposure of destabilizing N-terminal residues resulting in degradation of target protein via the N-end rule.

4. tsTEVp was used to activate caspase-7 by proteolytic removal of its prodomain and subsequently resulted in apoptosis-like cell death.

Research objective 3 outlined in chapter 1 was addressed. The versatility of tsTEVp for synthetic activation or deactivation of proteins was demonstrated. In the process, cellular pathways were rewired to respond to the physical stimulus of temperature. Key observable cellular phenotypes including protein concentration, subcellular localization, and cell morphology, were successfully induced by tsTEVp activity.
7 Summary

The work discussed in this thesis harnessed the potential of the highly specific TEVp and demonstrated how it could be used to manipulate biosystems. The results of this work were presented in three chapters (chapters 4-6), addressing each of the three objectives laid out at the beginning of this study. Together, it fits into the long-term vision of engineering proteins and synthetic systems to reprogram cells for carrying out a set of well defined tasks, ultimately facilitating their widespread use for therapeutic applications.

Objective 1 and Chapter 4 - Using TEVp self-cleavage for the expression of multiple genes:
Borrowing from a theme in RNA viruses, a single-vector multiple gene expression strategy utilizing TEVp self-cleavage was created. The feasibility of this approach for robust expression of multiple genes was demonstrated in both bacterial (E. coli) and mammalian cells (COS-7, HeLa, and HEK-293). In E. coli cells, up to three fluorescent proteins were expressed and could be individually purified using respective purification tags. SH3 and its binding peptide CB1 were expressed simultaneously and purified as a complex using only one purification tag. In mammalian cells, up to three fluorescent-tagged genes were simultaneously expressed. Each gene was differentially targeted to subcellular compartments either constitutively by a localization motif (NLS), or conditionally by small-molecular induced interactions (DAGR recruitment to membrane-bound diglycerides). In addition, the stoichiometry of protein products was correlated with the frequency of appearance of their genes on the expression vector.

Objective 2 and Chapter 5 – Engineering a temperature-sensitive TEVp: A novel modification to TEVp in the form of N-terminal truncations was presented. Taking advantage of the inherent temperature sensitivity of TEVp, incremental N-terminal truncations were conducted to weaken its native activity until it becomes negligible at the unfavorable temperature (37 °C) while retaining sufficient activity at the favorable temperature (30 °C). This temperature-inducible TEVp, tsTEVp, was realized by truncating the first 17 residues of full-length TEVp. To assist in the monitoring of tsTEVp activity in live-cell imaging experiments, a novel translocating
biosensor, TEVpSensor, was developed to allow the extent of TEVp to be quantified. Temperature-induced activity of tsTEVp was validated in several mammalian cell cultures (COS-7, HEK-298, HeLa and CHO) and cleavage was rapid typically within 4 hours. Results also suggested that there was considerable cell-to-cell variability in cleavage efficiency and background activity due to different expression levels, which makes the system more suitable for population studies where such variability is less of an issue.

Objective 3 and Chapter 6 - Synthetic manipulation of biosystems using tsTEVp: tsTEVp was used together with a catalogue of other synthetic tools to generate temperature-induced biological responses. The versatility of tsTEVp was demonstrated by the number of ways it can be used to influence protein function. In the process, cellular pathways were rewired to respond to the physical stimulus of temperature, namely protein localization, cellular blebbing, protein degradation, and apoptosis. Key observable cellular phenotypes including protein concentration, subcellular localization, and cell morphology, were successfully induced by tsTEVp activity.

1. tsTEVp cleavage mediated nuclear to cytoplasmic translocation of DAGR which was subsequently rerouted to the plasma membrane by PDBu-induced binding to DAG.

2. Cellular blebbing was programmed using three synthetic modules tsTEVp, pLyn-tevS-CaRQ and LOVS1K/Orai1. The tsTEVp-activated plasma membrane to cytoplasm localization of pLyn-tevS-CaRQ rendered it insensitive to membrane localized Ca2+ signals.

3. Conditional protein degradation via the N-end rule was achieved by using tsTEVp to cleave and expose destabilizing N-terminal residues.

4. Caspase-7 was activated by tsTEVp by proteolytic removal of its prodomain and subsequently resulted in apoptosis-like cell death.
8 Future work

This chapter will describe several ways in which the technology developed in the previous chapters could be improved and broadened for more complex applications, and applied towards answering specific biological questions. First, further development of tsTEVp in itself will be discussed, including ways to improve its solubility, and how it could be supplemented with auxiliary technologies to expand its applicability. Second, further development within the theme of inducible proteolytic devices will be explored, including alternative ways of controlling TEVp activity, and the possibility of engineering other proteases to function as inducible proteolytic devices. Third, ways of using tsTEVp to answer specific biological questions will be discussed, specifically, how tsTEVp could be used to functionally characterize proteolytic events in the caspase cascade.

8.1 Further development of tsTEVp and associated applications

tsTEVp inherits intrinsic defects from full-length TEVp. One is that full-length TEVp is poorly expressed in *E. coli* and has low solubility (71, 75). While full-length TEVp has been shown to cleave substrates effectively in both *E. coli* and in solution, it owes very much to its high catalytic efficiency, something that the truncated variants lack. At this stage, application of tsTEVp is limited to mammalian cells. In order to expand the utility of tsTEVp to *E. coli*, and associated *in vitro* applications as a recombinant protein, its low solubility has to be overcome. Effort has been made in this regard in the past, including refolding of TEVp from inclusion bodies (154), coexpression with chaperone proteins (155), deletion of C-terminal residues (156), and addition of N-terminal fusion partners (157, 158). Several point mutations have also been identified to improve the solubility of full-length TEVp (75, 76, 159). It is however uncertain whether these mutations could improve the solubility of tsTEVp while concurrently preserving its temperature sensitivity. Further random mutagenesis studies could be conducted using YFP as a solubility screen to identify corresponding mutations specific to tsTEVp.
A major advantage of protein systems over genetic systems is that protein systems can be spatially localized. Spatial localization allows isolation of protein function either by limiting local concentration, or by physical barriers such as biomembranes. One of the next directions of tsTEVp development will focus on subcellular targeting of tsTEVp so that its activity will not only be temporally controlled, but also spatially confined. For instance, tsTEVp activity could be limited to only a subcellular domain (e.g. neuronal axons and dendrites), a subcellular compartment (e.g. nucleus), or in the vicinity of a scaffolding protein (e.g. Ste5 in the MAPK pathway (160)). Two general methods exist by which subcellular localization of tsTEVp can be controlled: 1) a functional tag can be permanently appended to tsTEVp, resulting in constitutive localization, or 2) chemical inducers of dimerization (CIDs) systems can be used for conditional localization (137). The functional tag could either be a localization motif, or another protein (e.g. either the TEVp substrate itself or a scaffolding protein) which tsTEVp could be recruited to.

Many localization motifs are readily available and can be used to target nucleus, plasma membrane, mitochondria, and endoplasmic reticulum to name a few (132). Conditional targeting can be achieved using CIDs such as rapamycin induced FKBP12/FRB heterodimerization (132-134). For example, FRB domain could be tethered to the C terminus of tsTEVp and the FKBP12 could be fused at its C terminus with the functional tag. Clustering of tsTEVp with its effector proteins could allow faster cleavage dynamics of its target only at the intended subcellular domain. Some effector proteins, such as caspase-8, display differential functions depending on spatial localization to drive distinct biological processes between proliferation and apoptosis (161, 162). As such, differentially localized tsTEVp could become a valuable tool to study spatial effects of protein activation.

Other technologies could be developed to facilitate the secretion or cell-to-cell transfer of tsTEVp so that tsTEVp could be extended beyond the cell it is expressed. For example, tsTEVp secretion could be engineered using currently available conditional protein secretion systems inducible by small molecules (163, 164) or light (165). Secreted tsTEVp could be used to mimic extracellular proteases such as matrix metalloproteinases, thrombins, and cathepsins (166), which could in turn be used to activate extracellular ligands such as latent TGF-β complexes (167). Gene transfer of tsTEVp into adjacent cells could be facilitated by co-expression of
vesicular stomatitis virus glycoprotein (VSV-G), which is capable of initiating membrane fusion at low pHs (168). As such, tsTEVp could be conditionally delivered and activated in cells present in low pH microenvironments (e.g. tumors).

8.2  Further development of inducible proteolytic devices

An inducible proteolytic device can be defined as a protein or a set of proteins that converts an input stimulus (such as physical cues, ligands, or cellular states) and translates it into a proteolytic output. In this thesis, a temperature-inducible TEV protease was engineered to control protein function and manipulate biosystems. Beyond this work, there is still considerable interest in engineering novel proteolytic devices to expand the repertoire of tools and to create diversity in the ways synthetic circuits could be designed. Different proteolytic devices can be optimized for different catalytic efficiency, for different substrate recognition, and to function in different cellular contexts. Future development of inducible proteolytic devices after tsTEVp will proceed in two general directions:

1. Engineering alternative methods for inducing TEVp activity

2. Engineering alternative proteolytic output using other orthogonal proteases

tsTEVp will not be applicable in organisms which cannot tolerate or sustain a temperature shift, or in synthetic systems where another temperature sensitive protein is utilized. Further development in creating alternative methods for TEVp induction could impart greater flexibility towards the design of complex synthetic systems. Three alternative methods are presented below:

**Light-sensitive TEVp**: Light-switchable proteins are an emerging class of synthetic tools that has been used to reversibly effect biological function with spatiotemporal precision. The blue-light induced conformational change of LOV2 domain from phototropin has been used to create novel photosensitive switches by fusing N-terminally to effector domains. The versatility of this technique is manifested in the successful creation of light-sensitive proteins such as Rac1 (157),
STIM1 (36), TrpR (169) and caspase-7 (35). N-terminal fusion of LOV2 domain to TEVp could potentially regulate its activity through two mechanisms: 1) steric occlusion of the TEVp active site, and 2) propagation of conformation strain which distorts the TEVp active site. The efficacy of the LOV2 domain in suppressing TEVp activity in the dark state will depend on the junction sequence between the LOV2 and TEVp. It may be necessary to engineer appropriate junction sequences in order to position LOV2 in the right spatial orientation for interfering with TEVp active site. Circular permutations of TEVp may uncover alternative N- and C- termini that might allow LOV2 domain better access to the active site of TEVp.

**Calcium-sensitive TEVp:** Several exogenous stimuli can be used to activate calcium sensitive proteins by mobilizing intracellular calcium stores or mediate extracellular calcium entry. These include chemicals such as ATP (125) and acetylcholine (140), and physical cues such as light (36, 141) and voltage (142). Calcium sensitive proteins such as calmodulin (CaM) could in turn be used to engineer unique calcium-responsive switches. The embedding of CaM-binding peptides within surface-exposed loops of proteins could allow CaM to regulate their activity based on local calcium concentration. This strategy has been successfully demonstrated for small GTPases RhoA (170), Cdc42 (170), and Rac1 (170). Likewise, a calcium-sensitive TEVp could potentially be engineered by identifying an appropriate CaM-binding peptide insertion site within one of five surface-exposed loops in the TEVp secondary structure (72, 114).

**Conditional splicing of TEVp:** In the split-TEVp system, reconstitution of TEVp activity is achieved by the non-covalent reassembly of split TEVp fragments with the assistance of dimerization partners (7). Conceptually similar is a process known as protein trans-splicing, where split-inteins mediate the covalent joining of two protein fragments before splicing themselves out (171). Split-inteins could potentially be incorporated into the split-TEVp system to provide two additional benefits: 1) the reconstituted fragments will be covalently linked, which might result in a more stable conformation resembling that of native TEVp, and 2) the reassembled TEVp will be liberated from the constraints of any dimerization partners. Most split-inteins, such as DnaE from *Nostoc punctiforme*, results in spontaneous trans-splicing, which is not applicable when conditional TEVp activity is desired (171). However, if both
fragments were to be placed under the control of orthogonal promoters, it might be possible to achieve conditional TEVp activity in a subset of cells or tissues expressing transcriptional factors for both promoters (172). Nevertheless, some split-inteins such as Sce VMA intein show very little spontaneous trans-splicing and conditional splicing variants inducible by light (173) and small molecules (174) have already been developed.

Output versatility could be engineered by creating inducible forms of other proteases with orthogonal substrate specificity. By using several orthogonally acting proteases, it might become possible in the future to construct fairly complex signaling circuits. For example, sequential proteolysis of several substrates either within the same protein, or within different proteins could be induced. Many pathways involve sequential proteolysis of differential signaling substrates, including the caspase cascade (93) and the coagulation cascade (175). Some processes also involve sequential cleavage of a single protein at different sites, such as the activation of SARS coronavirus (176) or the deactivation of E-cadherins (177). In the context of the work described in this thesis, having two orthogonal proteases could allow self-cleavage of multiple genes (chapter 4) and temperature-induced cleavage (chapter 5 and 6) to be simultaneously utilized. Even the applications developed in chapter 6 could be mixed and matched to create more complex systems. For example, a membrane to nucleus translocation could be first triggered, followed by N-degron mediated degradation.

In order to create such a system, one could look towards a close relative of TEVp, the tobacco vein mottling virus protease (TVMVp). TEVp has been widely used as a tool for proteolytic cleavage of engineered proteins because of its stringent sequence specificity. TVMVp can also be used for this purpose (73). Although they share a similar 3-dimensional fold and 52% sequence identity, the two proteases have distinct substrate specificities and have negligible cross activity towards each other’s substrates (70, 178). Like TEVp, TVMVp has a temperature activity profile which peaks at around 30 °C (73). An incremental truncation strategy could potentially be also applied towards engineering a temperature sensitive TVMVp due to aforementioned similarities. A temperature-sensitive TVMVp could be used in tandem with
other inducible forms TEVp discussed previously to induce cleavage of distinct substrates in a sequential manner using two orthogonal inputs.

8.3 Using tsTEVp as a tool for understanding biology

The proteome is a dynamic set of proteins that is constantly changing due to post-translational modifications. Proteolytic cleavages represent a central post-translational process which sculpts the proteome. Every protein, at some point of its life, will be affected by proteolysis, whether it's systematic shredding by the proteasome or precise cleavages involved in signal propagation (3). To understand the role of a protease in a biological process and its relationships with health and disease, it is necessary to identify and characterize the protease’s physiological substrates.

The development of high-throughput proteomics has recently uncovered thousands of substrates of the nearly 600 proteases present in the human genome (2, 179). However, due to the inherent complexity of proteolytic networks, these combined data sets of protease substrates have limited use for understanding how a single substrate can contribute to their respective signaling cascades and associated biological phenomena.

The technology developed in this work can be used to address this specific issue. The temporal activation of site-specific proteolysis provided by the tsTEVp can be used to answer several lingering questions about proteases and their substrates:

1. What is the functional effect of each cleavage in isolation?

2. Which cleavages are essential, which are just bystanders?

3. How long does the substrates remain stable?

4. Is it possible to fine-tune cellular behavior using a minimum subset of substrate proteolysis?

Of particular interest is the caspase signaling cascade. The caspase cascade ends in cell-wide proteolysis of over 400 proteins (180-183). The sum total of these proteolytic events is known to
define the cellular morphology known as apoptosis, but may also play a role in other cellular events such as differentiation and inflammation depending on cellular context (180, 184, 185). The cleavage of each substrate can have a gain-of-function, loss-of-function or change-of-localization effect on the protein, but some of them may just be bystander proteins with no functional impact (180, 183). To further the complexity of caspase signaling, the quantity of each cleavage may also influence the outcome (186). There is currently no easy way to identify functionally important substrates from proteolytic noise, and there is almost no knowledge of how much cleavage of any caspase substrate there should be to result in physiological outcome (180). tsTEVp provides the unique ability to imitate caspase cleavage while bypassing the canonical signaling cascade and keeping other substrates unperturbed. This method could be particularly useful in investigating gain-of-function cleavages which either directly contribute to an observable morphology or are part of a positive feedback loop which enhances the efficiency of apoptotic signaling.

Substrate candidates can be obtained from literature search or a caspase substrate database such as CASBAH (187) and MEROPS (188). TEVp recognition sequences would be engineered in place of confirmed or putative caspase cleavage sites within target substrates. TEVp has a short recognition sequence that can potentially be embedded within most proteins without perturbing their native structure. Below are a few signature substrates which could be investigated:

**Structural modifiers:** Cleavage of structural modifiers directly generates distinct morphologies that could be monitored by microscopy. The typical morphologies associated with apoptosis are membrane blebbing, cell rounding, chromatin condensation and nuclear fragmentation (93). Notable gain-of-function cleavages which result in the characteristic blebbing morphology include kinases ROCK1 (95) and PAK2 (189), and actin-binding protein gelsolin (190). PAK2, in particular, may be involved in multiple avenues of signaling, as it has displayed both nuclear and membrane targeting following caspase cleavage (189, 191). It would be interesting to study how PAK2 can trigger distinct signaling cascades from ambiguous subcellular targeting following cleavage. ROCK1 is also implicated in nuclear fragmentation (192), while cleavage of
ICAD liberates CAD nuclease that mediates apoptotic DNA fragmentation and chromatin condensation (94).

**Signaling intermediaries:** A wide variety of proteins involved in signal transduction are cleaved by caspases. A lot of these substrates either propagate apoptotic signaling, or function in a positive feedback loop to increase its robustness. The best examples are the cleavage of caspases themselves by their upstream family members (93). Cleavage of Bid generates an active fragment that induces the mitochondrial pathway of apoptosis (93). Several kinases could also be cleaved by caspases, resulting in gain-of-function kinase activity (193). For example, cleavage of kinases PKCδ, Abl and MST1 generates a constitutively active C-terminal fragment that translocates to the nucleus where it can phosphorylate pro-apoptotic targets (193, 194). However, the physiological consequence of these cleavages remains mostly unclear.

Many questions about the orchestration of cell death could be answered. Given the large number of substrates, there must be some that are more crucial than others, and some that are simply bystanders. The cleavage of certain substrates such as ROCK1 and ICAD affects typical apoptotic morphology, but are they sufficient on their own to drive the cell towards a point of no return? Is there a threshold amount of proteolysis which is required to drive a particular biological outcome? The existing way of answering these questions is by overexpressing the cleaved fragments. While this method is straightforward, the effects of timing and quantity of each cleavage cannot be assessed. tsTEVp induced cleavage could provide useful information in this regard. In addition, the half-lives and translocation patterns of each substrate after cleavage could be observed. Cleavage assays could be carried out in a 96-well format or at even higher throughputs, and multiplexed for parallel comparisons (195).
9 References


Appendices

The appendices contain sequences of key proteins used in this work, and additional experimental results that may be of interest to the reader.

Appendix A: DNA and protein sequences

Appendix B: List of oligonucleotide primers

Appendix C: Supplemental movie legends

Appendix D: Additional experiments on N-degrons
Appendix A: DNA and protein sequences

DNA and protein sequences of key proteins and short peptides used in this work are shown below. Protein sequences are represented by standard one letter amino acid code. Fluorescent tags are omitted for simplicity.

pLyn

DNA sequence:
ATGGGCTGCATCAAGAGCAAGGCAAGGACACGACGCCACTAGT

Protein sequence:
MGCIKSKGKDSATS

NLS

DNA sequence:
CGGATTCGAAAAAGcTACGG

Protein sequence:
RIRKKLR

NES

DNA sequence:
CTACAAAGAAGCTAGAGGAGCTGGAGCTAGATGAACAA

Protein sequence:
LQKKLEELDEQ
Deg

DNA sequence:
AGCGGCCTGGCTGCTGCCCTGGCTGCTGCC
Protein sequence:
SGAWLLPVSLV KRK

SH3

DNA sequence:
GGATCCGCAGAGTATGTGCG
Protein sequence:
GSAEYVRALFDFNGD LRI RDKPEEQWWNAE DSEGKRGMIPVYP YVEKY

CB1

DNA sequence:
CCGCGCCCGGCCCTGCGCCCAAGAAGAGG
Protein sequence:
PPPALPPK KR

TEVp

DNA sequence:
ATGGGCCTGACTAGTGGATCCGAAAGCTTTGTTAAGGGGCCTGGCTGATTACAACCGATATCGA GCACCATTTGTCATTTGACGAATGAACCTGATGGGCACACAACATCGTTGTATGGTATTGGATT TGTCCTCTCATACATTACAAAACAGCACTTGTATTTAGAAGAAAATATTGGAACACTGTTGGTC CAA TCACTACATGGGTATTCAAGGTCGAAGAACCACACGTCTTTGCAACAACACCTCATTGATGGGA GGGACATGATAAATTATTCGCATGCTAAGGATTTCCACCATTTCCTCAAAAGCTGAATTTAG
AGAGCCACAAAGGGAAGAGCATATATGTCTTTGTGACACACAAACTTCCAAAACCTAAAGAGCATGCTT
AGCATGGTGTCAGACACTAGTTGGCACAATCCCTCTCATTCTGATGGCATATTCTCGGAAAGATGGA
TTCAAAACCAAGGATGGGCGATGGCGAGTTCAATTAGTATCAACTAGAGATGGGTGTCATTGG
TATATACCTAGATGACTGTAATTTTACCACACAAACAAACTATATGTCACAAGCGTGGCAGAAACTTC
ATGGAATTGGTACAAATACAGGAGGCGAGCAGACGATGGGTTAGTCGATAAATGCTGACCTGATTGG
CATATTGGGAGGCGACAAAGGTAACACCTTCAAACTTCAACGAACTTTCCAAACACTTTCACCT
Protein sequence:
MGLTSGSesLKFGRDYNPSSTICTHLMNEDGHTSLYGIGFPIITNKHLFRRNNGTLLVQSLIGVF
KNTTLQQHLIDGRMIIRPKFPFPKLRKFREPFQEREICLVTTNQFTKMSMSVSDTSTCFPSSDGI
FWKHIIQTKDQCGSPLVSTRDGFIVGIHSASNFTNTNYFTSVPKNFELLTNQEAQQWVSWRLNAD
SVELWGHHKFMVKPEEPFPQPKVEATQLMNRRRR

tsTEVp

DNA sequence:
ATGGGCCTGACTAGTGATCCATTTTGTGACAAATCTGATGGGCACACAAACTCAGTGGGATTAAG
AGAATCTGGATTTGGATTTGGTCCCTCTCATCATTACAACACAGACTGGTATTAGAAGAAATAATTGG
AACACTCGTGTGGTCTCAATCACTACATGCTGATTCAAGGCGAAGCACACGGCGATTTGCAACAA
CACCTCATGATGGGAGGACATGATAATTATTGTCAGCTGCTATAGGATTCTCCACCAGATTCTCTC
AAAAGCTGAAATTTAGAAGCCAAAGGGAAGAGCCGACATATGTCTTTGGCAATCACCATTCTCCT
AACAGCACTGCTAGCTAGTGCAGACACTAGTTGGCACAATCCCTCACTGACATTGGCATAATGCT
TTCTGGAGACATTTGGATCAAACAAAGGATGGGCGATGGCGATCCATTAGTATTCACTAGAG
AGTGGGCATTAGTGGATTAACACTACAGCTGAAATTTACCAACACAAACAAATTATTCCACAAG
CGTTGCCAAAACACTTCATGGAGATTGTCGACAAATCCAGAGGCACGAGTGATGGTGGCGATATG
CGATTTACGTCTGACTGATTGGTGGGGGGCCATAAAAGTTTTTACTGTGGAACCTGAAGACGC
CTTTCAAGCAGTTAAGGAAGACGACTCAACTCATGAAATCGTGCTCGCCGTCGC

Protein sequence:
MGLTSGSICHLTNESDGHSTSYYGIFGPFIITNKHLFRRNNGTLLVQSLIGHVFKVNTTTLQQ
HLIDGRMIIRPKFPFPKLRKFREPFQEREICLVTTNQFTKMSMSVSDTSTCFPSSDGI
FWKHIIQTKDQCGSPLVSTRDGFIVGIHSASNFTNTNYFTSVPKNFELLTNQEAQQWVSWRLNAD
SVELWGHHKFMVKPEEPFPQPKVEATQLMNRRRR

Casp7

DNA sequence:
ATGGGCCTGACTAGTGGATCCATGCGAGATGTCAGGGCTGTCAGGCTTGATATTGAGAGCAAGAAGAGGGTGGT
ATTCAGCAATAAGAGATTCAGTGGATCATAAGCCAGCCGTCTCGTTGTACCTGCCTCGAT
CAGTAGAAAGAGAAATGTCACCATAGGATCCATCAGGGCTGTATTGAAGAGCAGGGGGTT
GAGGATTCAGCAAATGAAGATTCA

Protein sequence:
MGLTSGSMADDQGCIEEQGVEDSANEDSVDAKPDRESSVFPSLSFKKKNVTMRISKTRTDVPT

57Casp7
DNA sequence:
ATGGGCCTGACTAGTGGATCCATGCGAGATGTCAGGGCTGTCAGGCTTGATATTGAGAGCAAGAAGAGGGTGGT
ATTCAGCAATAAGAGATTCAGTGGATCATAAGCCAGCCGTCTCGTTGTACCTGCCTCGAT
CAGTAGAAAGAGAAATGTCACCATAGGATCCATCAGGGCTGTATTGAAGAGCAGGGGGTT
GAGGATTCAGCAAATGAAGATTCA

Protein sequence:
MGLTSGSMADDQGCIEEQGVEDSANEDSVDAKPDRESSVFPSLSFKKKNVTMRISKTRTDVPT
PKLFFIQACRGTELDDGIQADSGPINDDANPRYKIPVEADFLFAYSTVPGYYSWRSPGRGSWF VQALCSILEEHHKDLEIMQILTRVNDVARHFEQSQDDPHFHEKKQIPCVVSMLTKELYFSQ

Casp7tevS

DNA sequence:
ATGGGCTGACTAGTGGATCCATGGCAGA
TGATCAGGGCTGTATTGAAG
AGCAGGGGGTTGAGG
ATTCAGCAAATGAAGATTCA
GTGGATGCTAAGCCAGACCG
GTCCTCGTTTGTACCGTCCC
TCTT
CAGTAAGAAGAAATGTCACCATGCAGATCCATCAAGACCCGAGACTCCCTCTTCAGG
CTTGCCAGGACCAGCTTGAGTGACCGATCCAGCCGGCATCGGGCCCTCAATGACACAGA
TGCTAATCCGTCGATCAAGATCCCGCTCTCGCTCCTACAGCTCATCTCTTTAGAGAAACCACCTCTTCCTTCATCGAG
CTGGCCAGGACCAGCTTGAGTGACCGATCCAGCCGGCATCGGGCCCTCAATGACACAGA
TGCTAATCCGTCGATCAAGATCCCGCTCTCGCTCCTACAGCTCATCTCTTTAGAGAAACCACCTCTTCCTTCATCGAG
CTGGCCAGGACCAGCTTGAGTGACCGATCCAGCCGGCATCGGGCCCTCAATGACACAGA
ggctattac
TCGTGGAGGAGCCCAGGAAG
AGGCTCCTGGTTTGTGCAAG
CCCTCTGCTCCATCC
TGGAGGAGCACGGAAAAGAC
CGCCTGCTTCGCCTGCATCC
TCTT
AAGCCATGG
AGAAGAAAATGTAATTTATG
GGAAAGATGGTGTCACACCA
GCCCACTTTAGGGGGGATAG
CTGGAAATCATGCAGATCCT
CACCAGGGTGAATGACAGAG
TTGC
CAGGCACTTTGAGTCTCAGT
CTGATGACCCACACTTCCAT
GAGAAGAAGCAGATCCCCTG
TGTG
GTCTCCATGCTCAACAGGAACCTACTTCACTTGCAAA

Protein sequence:
MGLTSGMSADDQQGCIEEQGVEDSANEDSVDAPRSSFVSLSFKKKKNVTRSIKTTRDRVPA
SENLYFQSASENLONYQNFNFEKLKCIINNKDFKVGMVRGTDSDEALFKF
RSLGFDVIVYNSCSAKMQDLKKAEEHTANCAFCILLSHGEENVYIKGDKGVTPIKDLTAH
FRGDRCKTLLEKPRLFFIQACRGTELDDGIQADSGPINDDANPRYKIPVEADFLFAYSTVPGYYSWRSPGRGSWFVQALCSILEEHHKDLEIMQILTRVNDVARHFEQSQDDPHFHEKKQIPCVVSMLTKELYFSQ

pLyn-tevS-CarQ

DNA sequence:
ATGGGCTGCATCAAGAAGCAAGGCAAGCAGCCACTAGTGAAACCTGATTTCAGAGCG
CTAGTGAAAACCTGTATTTTCAGAGCGCTAGTGAGGATTGACAGATTGACAGATGTTCAAAAGAAGGCTTCTTC
ATTATCTGACAAGGATGCGGACCCACTACCAACAAAAAGGAACCTTGGCACCCTTTATGAGGTCCG
CTTGAGAACAAAAACCCAGGAGCAAGGAAATGCAGAATGATGACATTGACATGATGACAGATGATGACAGA
CTGGAAATCATGCAGATCCT
CACCAGGGTGAATGACAGAG
TTGC
CAGGCACTTTGAGTCTCAGT
CTGATGACCCACACTTCCAT
GAGAAGAAGCAGATCCCCTG
TGTG
GTCTCCATGCTCAACAGGAACCTACTTCACTTGCAAA
CGAAGAGGAAATCCGAGAAGCATTCGCAGTCTAAGGAGGACGCTAACATCAGCGCT
GCTGAATTACGTACGTCATGAACAACTCCGGGAGAGGAGTTGATGAAGATGAAT
TAGAAGGAAGAGCAGATACGATGGATGTCAGGGCAAGATGAAACTATGAGAGTT
GACAGCAAAAGGCTAGTAAAGACGTCGTAAGAGAAGAAACTTCTATTGGCGTCAGCCAGCG
TACAAGAGAGTCTCCAGCTCCGGGACATGGCTGCCATCCCGGAAAGAACTGGTGATGGT
ATGGAAGCCCTGTGGAAGACATGCTTTGCTCATAGTCTCGACAGGACAGTTCACAGGAGGT
TGCCACAGTGTTGGAAGAATATGTGGCAGATATCGAGGAGTGCTAGTGATCACCCGCTATT
ACTGTCAGGAATATGTCCAGGAGATCCAGCTGAGCTGTTATGCTAGATTCCCTCGCTAGTT
ATCGCTCGAGCCCTCTCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CCCTGATAGTTTAGAAAAACATCCCGAGAAAAAGTGAGCCCAAGAAGTCAAGATTTTCTGTCCC
ACGAGGCTCATCCTGTGTTTGGGAATAGAGAGATCTTCGGAAGAAACTGGTGATTGTT
ATGCGCTGAGGCCCTCCTCCTACCCAGATACCAGTGTATATCTGATGTTGTTTTCCCATCGACG
CGCTTTTGGGTACATGGAGTGTTCGACAAAGACGATCGTGGAGAGCTGCTCTGCA
Protein sequence:
MGCIKSKGKDSATSENLYFQSASENLYFQSAEQIAEFKEAFSLFDKDGDTITTKELGTVMRS
LGQNPTEAELQDMINEVDAGNGTIIYFPFELTMARKNDTDEEEREAFRVFDKGNGTISA
AEILRHVMNLGKLTDEEVMERADIDGDQVNYEFVQMMTAKSRRWKNFIALVSAANR
YKKISSSSALAAIRKVLIVGDGACGKTCLLIVFSDKQPVEVYPVTFVENYVADIEVDASGS
TVQRVYRGIQRARAYARFLASYDRLRFLSYPDTDVILMFSDPSLENIPEKWTPEVKHFCPN
VPIILVGNKDLIRNDEHTRRELAHKMKEPVKPEGRDMANRIRGAFCMYECSAKTDGVR

Appendix B: List of oligonucleotide primers

Oligonucleotide primers used to create key proteins and peptides in this work are shown here.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Oligonucleotide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEVp_53</td>
<td>CCGGATCCGAAAGCTTGTTGTTAAGGGGCCGC</td>
</tr>
<tr>
<td>TEVp_35</td>
<td>CTAGCTAGCCGACGGGCAGCGACGATTCAT</td>
</tr>
<tr>
<td>TEVpMut_53</td>
<td>GGTAGGGCAGGCTGTCAGCTCC</td>
</tr>
<tr>
<td>TEVpMut_35</td>
<td>GGACTGCCAGCCCTGCCCACCTCC</td>
</tr>
<tr>
<td>N12TEVp_53</td>
<td>CATGCCATGGGCATAGTAACCCGATATCGAGCACCAATTTC</td>
</tr>
<tr>
<td>P13TEVp_53</td>
<td>CATGCCATGGGCATAGTAATGAGCACCAATTTCACGAGTGGCGCGGCACCATATCAGAG</td>
</tr>
<tr>
<td>I14TEVp_53</td>
<td>CATGCCATGGGCATAGTAATGAGCACCAATTTCACGAGTGGCGCGGCACCATATCAGAG</td>
</tr>
<tr>
<td>S15TEVp_53</td>
<td>CATGCCATGGGCATAGTAATGAGCACCAATTTCACGAGTGGCGCGGCACCATATCAGAG</td>
</tr>
<tr>
<td>S16TEVp_53</td>
<td>CATGCCATGGGCATAGTAATGAGCACCAATTTCACGAGTGGCGCGGCACCATATCAGAG</td>
</tr>
<tr>
<td>T17TEVp_53</td>
<td>CATGCCATGGGCATAGTAATGAGCACCAATTTCACGAGTGGCGCGGCACCATATCAGAG</td>
</tr>
<tr>
<td>H20TEVp_53</td>
<td>CATGCCATGGGCATAGTAATGAGCACCAATTTCACGAGTGGCGCGGCACCATATCAGAG</td>
</tr>
<tr>
<td>L21TEVp_53</td>
<td>CATGCCATGGGCATAGTAATGAGCACCAATTTCACGAGTGGCGCGGCACCATATCAGAG</td>
</tr>
<tr>
<td>tevS_35</td>
<td>CTAGCTAGCCGCTCTGAAAATACAGGTGTTCACGAGTCAGTGCGCGGCACGACTCAG</td>
</tr>
<tr>
<td>tevH_35</td>
<td>CTAGCTAGCTAGCTGAAAATACAGGTGTTCACGAGTCAGTGCGCGGCACGACTCAG</td>
</tr>
<tr>
<td>tevF_35</td>
<td>CTAGCTAGCTAGCTGAAAATACAGGTGTTCACGAGTCAGTGCGCGGCACGACTCAG</td>
</tr>
<tr>
<td>tevY_35</td>
<td>CTAGCTAGCTAGCTGAAAATACAGGTGTTCACGAGTCAGTGCGCGGCACGACTCAG</td>
</tr>
<tr>
<td>tevC_35</td>
<td>CTAGCTAGCTAGCTGAAAATACAGGTGTTCACGAGTCAGTGCGCGGCACGACTCAG</td>
</tr>
<tr>
<td>Deg_35</td>
<td>CTAGCTAGCTAGCTGAAAATACAGGTGTTCACGAGTCAGTGCGCGGCACGACTCAG</td>
</tr>
<tr>
<td>Casp7_53</td>
<td>CGGGATCGCCATGGCAGATGAGTGATCAGGCTCGGAGCCAGCCAGCCAGGGCGCGCGCG</td>
</tr>
<tr>
<td>Casp7_35</td>
<td>CTAGCTAGCTAGCTGAAAATACAGGTGTTCACGAGTCAGTGCGCGGCACGACTCAG</td>
</tr>
<tr>
<td>57Casp7_53</td>
<td>CGGGATCGCCATGGCAGATGAGTGATCAGGCTCGGAGCCAGCCAGCCAGGGCGCGCGCG</td>
</tr>
<tr>
<td>proCasp7_35</td>
<td>CTAGCTAGCTAGCTGAAAATACAGGTGTTCACGAGTCAGTGCGCGGCACGACTCAG</td>
</tr>
</tbody>
</table>
Appendix C: Supplemental movie legends

This section describes the experimental conditions and timescales of the movies presented in this work.

Movie 4.1: PDBu stimulated migration of DAGRVen in COS-7 cells expressing TEVpCNDVR.

Cleaved components DAGRVen (green) and CeruNLS (cyan) were shown while mRFP was omitted for simplicity. PDBu was added at 0 s. Movie was created using Windows Movie Maker in .wmv format. Image dimensions are 190 µm by 150 µm. 1 s playback is equivalent to 50 s elapsed time.

Movie can be downloaded here: http://individual.utoronto.ca/xichent/Movie4.1.wmv

Movie 5.1: COS-7 cell expressing tsTEVp and TEVpSensor at 30 °C

TEVpSensor (pLyn-RFP-tevS-CFP-NLS) was shown. Blue represents CFP-NLS fragment while red represents pLyn-RFP fragment. Movie was created using Time Lapse Assembler in .mov format. Images dimensions are 110 µm by 110 µm. 1s playback is equivalent to 100 mins elapsed time (total 4 hours elapsed time).

Movie can be downloaded here: http://individual.utoronto.ca/xichent/Movie5.1.mov

Movie 5.2: COS-7 cell expressing tsTEVp and TEVpSensor at 37 °C

TEVpSensor (pLyn-RFP-tevS-CFP-NLS) was shown. Blue represents CFP-NLS fragment while red represents pLyn-RFP fragment. Movie was created using Time Lapse Assembler in .mov
format. Images dimensions are 110 µm by 110 µm. 1 s playback is equivalent to 100 mins elapsed time (total 4 hours elapsed time).

Movie can be downloaded here: http://individual.utoronto.ca/xichent/Movie5.2.mov

**Movie 6.1: HEK293 cells expressing tsTEVp/pLyn-tevS-CaRQ at 37 °C stimulated by ATP.**

pLyn-tevS-CaRQ is shown in grayscale. The movie was taken after 4 hours of incubation at 37 °C. ATP was added at 0 s of the movie. Movie was created using Time Lapse Assembler in .mov format. Images dimensions are 80 µm by 80 µm. 1 s playback is equivalent to 200 s elapsed time (total 16 mins elapsed time).

Movie can be downloaded here: http://individual.utoronto.ca/xichent/Movie6.1.mov

**Movie 6.2: HEK293 cells expressing tsTEVp/pLyn-tevS-CaRQ/DualCMVLOVS1K/Orai1 at 30 °C stimulated by light**

pLyn-tevS-CaRQ is shown in grayscale. The movie was taken after 4 hours of incubation at 30 °C. Blue light stimulation began at 0 s of the movie. Movie was created using Time Lapse Assembler in .mov format. Images dimensions are 80 µm by 80 µm. 1 s playback is equivalent to 200 s elapsed time (total 16 mins elapsed time).

Movie can be downloaded here: http://individual.utoronto.ca/xichent/Movie6.2.mov

**Movie 6.3: COS-7 cells expressing tsTEVp/pLyn-Casp7tevS/CaspSensor at 30 °C**

pLyn-Casp7tevS (red) and CaspSensor (blue) are shown. Movie was created using Time Lapse Assembler in .mov format. Images dimensions are 120 µm by 120 µm. 1 s playback is equivalent to 100 mins elapsed time (total elapsed time is 8 h).
Movie can be downloaded here: http://individual.utoronto.ca/xichent/Movie6.3.mov

**Movie 6.4: COS-7 cells expressing tsTEVp/pLyn-Casp7tevS/CaspSensor at 37 °C**

pLyn-Casp7tevS (red) and CaspSensor (blue) are shown. Movie was created using Time Lapse Assembler in .mov format. Images dimensions are 120 µm by 120 µm. 1 s playback is equivalent to 200 mins elapsed time (total elapsed time is 8 h).

Movie can be downloaded here: http://individual.utoronto.ca/xichent/Movie6.4.mov
Appendix D: Additional experiments on N-degrons

TEVp can be used to activate degradation of itself through N-degrons. This is particularly useful in TEVp self-cleavage gene expression for the clearance of residual TEVp. This appendix summarizes the experiments done to demonstrate N-degron mediated degradation of TEVp after self-cleavage. It also establishes the foundation for N-degron mediated degradation by tsTEVp in Chapter 6.

A set of constructs in the form of pLyn-Ceru-tevX-TEVp-Ven was constructed to assess the cleavage efficiency of non-canonical substrates tevX (X is each of the four N-degron residues C, F, H, and Y. When transfected into COS-7 cells, each construct was cleaved efficiently at tevX. The pLyn-Ceru and X-TEVp-Ven modules localized to the plasma membrane and cytoplasm respectively without any visible trace of co-localization (data not shown).

The constructs Ceru-tevX-TEVp-Ven, along with the Ceru-tevS-TEVp-Ven control, were used to characterize degradation after tevX cleavage. 41 cells were randomly sampled from cells expressing each construct two days after transfection. The extent of degradation in each cell was measured by the relative ratio of total Ven fluorescence and total Ceru fluorescence (Ven/Ceru) two days after transfection (Figure D1). The distribution of Ven/Ceru values for each construct was normal and both colors were still visible in all the cells. The differences in the mean Ven/Ceru values between each pair was insignificant (p > 0.05). Despite bearing an N-terminal destabilizing residue, X-TEVp-Ven did not show any more degradation than the stable S-TEVp-Ven. This shows that the N-terminal destabilizing residue itself is insufficient to function as an N-degron.
Figure D1: Assessment of tevX degradation.

Distribution of Ven/Ceru values for Ceru-tevS-TEVp-Ven and Ceru-tevX-TEVp-Ven (total of 41 cells for each construct) on day 2 after transfection. All Ven/Ceru values were normalized to the Ceru-tevS-TEVp-Ven control. There is no significant difference in Ven/Ceru between any two constructs (p > 0.05)
**Figure D2: Assessment of tevX-Deg degradation.**

(A) Percentage of cells (out of 60) without visible Ven on day 1 (black) and day 2 (grey) after transfection of each mRFP-tevX-Deg-TEVp-Ven construct. Experiments were conducted in triplicates. Error bars represents s.d. (n = 3). The day 2 values are significantly greater than day 1 values for each construct (p < 0.001). (B) Distribution of Ven/mRFP values for one of the experimental replicates (total of 60 cells for each construct). All Ven/mRFP values were normalized to the mRFP-tevS-Deg-TEVp-Ven control. The Ven/mRFP values of mRFP-tevX-Deg-TEVp-Ven are significantly lower compared to the control, in addition to the high number of counts where Ven is invisible (Ven/mRFP = 0).

To construct functional N-degrons, each tevX was supplemented with a 16-residue unstructured peptide Deg (ASSGAWLLPVSLVKRK). The constructs mRFP-tevX-Deg-TEVp-Ven, along with the mRFP-tevS-Deg-TEVp-Ven control, were used to characterize degradation after tevX cleavage. The degradation of X-Deg-TEVp-Ven in cells expressing mRFP-tevX-Deg-TEVp-Ven was apparent one day after transfection (Figure D2A). A total of 60 cells were randomly sampled for each construct. Ven was not visible in at least 27% of cells expressing mRFP-tevX-Deg-TEVp-Ven but was visible in all of the cells expressing the mRFP-tevS-Deg-TEVp-Ven control. Two days after transfection, Ven remained visible in all of the cells expressing mRFP-tevS-Deg-TEVp-Ven, but was almost completely removed from the mRFP-tevX-Deg-TEVp-Ven constructs (Figure D2A). To quantify the extent of degradation, the relative ratio of total mRFP fluorescence and total Ven fluorescence (Ven/mRFP) was measured for each cell (Figure D2B). This ratio was not significantly different between day 1 and 2 for the mRFP-tevS-Deg-TEVp-Ven control (p > 0.05), indicating that the stability of S-Deg-TEVp-Ven was not affected by the addition of Deg. On the other hand, Ven/mRFP values for mRFP-tevX-Deg-TEVp-Ven are significantly lower (Figure D2B) going from day 1 to day 2. This suggests that both the N-terminal residue X and the N-terminal peptide Deg are necessary elements of an N-degron. From the data, it is also evident that the rate of ubiquitin dependent proteolysis varied from cell to cell, and it is less than ideal that degradation was incomplete over the 2-day period, although this is most likely due to compensation by protein production.