Engineering robust enzymes via head-to-tail cyclization using spontaneous isopeptide bond formation



Christopher Schoene

Department of Biochemistry
Lady Margaret Hall
University of Oxford
Trinity 2016

Thesis for a Doctorate in Philosophy.

Abstract

Enzymes are selective, operate efficiently in mild conditions and are biodegradable. Hence, enzymes have become an attractive option for a wide range of academic and industrial processes. However, enzymes are often labile at higher temperatures which can limit the operational range. Current enzyme stabilization approaches can be time-intensive and often have only a minor effect. Enzyme cyclization has been suggested as an alternative and more general approach. SpyTag/SpyCatcher, SnoopTag/SnoopCatcher and Isopeptag/Pilin-C are peptide/protein pairs capable of forming an irreversible isopeptide bond upon reconstitution. By genetically encoding SpyTag, SnoopTag and Pilin-C on the N-terminus and SpyCatcher, SnoopCatcher and Isopeptag on the C-terminus, I developed a method for the simple *in vivo* cyclization of enzymes. TEM-1 β-lactamase, which is prone to irreversible aggregation above 37 °C, remained soluble at temperatures up to 100 °C after isopeptide-mediated cyclization and upon cooling retained high catalytic activity. Differential scanning calorimetry showed that isopeptide-mediated cyclization had little effect on the Tm of the enzyme. Instead, isopeptide-mediated cyclization increased the reversibility of the denaturation. SpyTag/SpyCatcher was found to be the most effective pair for conferring thermal resilience and was also used to cyclize and increase the thermal resilience of dihydrofolate reductase and phytase. SpyTag/SpyCatcher-mediated cyclization had only minor effects on the thermal resilience of glucose oxidase and luciferase. Therefore, there is a limit to how general the effect of isopeptide-mediated cyclization is. To start addressing the limitations of the current approach, I developed a phage display-based strategy for the directed evolution of SpyTag/SpyCatcher reactivity. Isopeptide-mediated cyclization is a potentially powerful approach for the enhancement of protein resilience.

Authorship declaration

I declare that all work performed in this thesis is my own. The thesis has not been submitted for any other degree at the University of Oxford or any other institution.

Publications arising from this thesis are detailed below:

- Schoene, C., Fierer, J.O., Bennett, S.P., and Howarth, M. (2014). SpyTag/SpyCatcher Cyclization Confers Resilience to Boiling on a Mesophilic Enzyme. Angew. Chem. 53, 6101–6104.

- Schoene, C., Bennett, S.P., and Howarth, M. (2016a). SpyRing interrogation: analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries. Sci. Rep. 6, 21151.

- Schoene, C., Bennett, S.P., and Howarth, M. (2016b). SpyRings Declassified: A Blueprint for Using Isopeptide-Mediated Cyclization to Enhance Enzyme Thermal Resilience. B.-M. in Enzymology, ed. (Academic Press), p.

Acknowledgements

I would like to thank Prof. Mark Howarth for always being available to provide good guidance and advice throughout my degree. I am grateful for the funding provided by the BBSRC and Sekisui Diagnostics. I would like to express my gratitude to Dr. Nick Major and Dr. Paul Bennett at Sekisui diagnostics for their time, resources and knowledge.

I would like to thank the members of the Sherratt lab for their help and support. I would like to give a special mention to Rachel Baker, Katarzyna Zawadzka and Dr. Katarzyna Ginda for always being so kind and awesome. I would like to thank the members of the Whitby group as well, in particular Dr. Carl Morrow and Manisha Jalan for always providing a good laugh, especially after 7 PM. I would also like to thank all the awesome people from the Oxford University Sul Ki Do club for all the support they provided and the good memories they have left me with.

I thank Dr. Michael Fairhead for his almost bottomless reservoir of knowledge, Dr. Jacob Fierer for his dad-like advice, for always being there to listen to me even when he is overseas and for his hopefully lifelong friendship, Khalil Davis, Dejan Mesner and Raphael Gayet for the nice memories during their short but memorable stays and Dr. Samuel Reddington and Can Buldun for their laddish antics which always brightened up the lab.

I would like to give a very special thank you to Dr. Gianluca Veggiani and Eric Liang. I honestly do not know what I would have done without either of you and I cannot begin to express the gratitude I have for having met you both.

Finally, I would like to thank my family. In particular, I would like to thank my parents for their support and unconditional love. To my father, who survived the horrors of World War II and fled East Germany, and my mother who survived a kidnapping in Colombia by the ELN; thank you for teaching me how to survive, no matter what life throws at us.

Abbreviations

AP	acceptor peptide
atm	atmosphere
ATP	adenosine triphosphate
BCA	bicinchoninic acid
BFP	blue fluorescent protein
BG	O(6)-benzylguanine
BirA	biotin ligase
BLA	TEM-1 β-lactamase
BMMH	buffered methanol minimal media
BMMY	buffered methanol complex media
bp	base pair
BPTI	bovine pancreatic trypsin inhibitor
BSA	bovine serum albumin
CD	circular dichroism
cfu	colony forming unit
CLEA	cross-linked enzyme aggregate
CLEC	cross-linked enzyme crystal
CPEC	circular polymerase extension cloning
Ct	threshold cycle
Da	Dalton
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DHF	dihydrofolic acid
DHFR	dihydrofolate reductase
DMSO	dimethyl sulfoxide

DSC	differential scanning calorimetry
DLS	dynamic light scattering
DR5	death receptor 5
DSF	differential scanning fluorimetry
dTMP	thymidine monophosphate
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
EDC	1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide
EDTA	ethylenediaminetetraacetic acid
eGFP	enhanced GFP
EGFR	epidermal growth factor receptor
ELP	elastin-like polypeptide
epPCR	error-prone PCR
ESI-MS	electrospray ionization mass spectrometry
FACS	fluorescence-activated cell-sorting
FAD(H_2)	flavin adenine dinucleotide
FLuc	firefly luciferase
FRET	Förster resonance energy transfer
Ftr	formylmethanofuran:tetrahydromethanopterin formyltransferase
GCSF-3	granulocyte colony-stimulating factor 3
GFP	green fluorescent protein
GOx	glucose oxidase
h	hour
hAGT	human O(6)-alkylguanine-DNA alkyltransferase
IFNα2	interferon α2
IL-1Ra	interleukin-1 receptor
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl-β-D-thiogalactopyranoside

ISM	iterative saturation mutagenesis
k_{cat}	catalytic rate constant of an enzyme
K_d	equilibrium dissociation constant
K_M	Michaelis constant
LB	Luria Bertani broth
MBP	maltose binding protein
min	minute
MS	mass spectrometry
MSA	multiple sequence alignment
NADP(H)	nicotinamide adenine dinucleotide phosphate
NCL	native chemical ligation
NHS	N-hydroxysuccinimide
Ni-NTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
OD_{280}	optical density at 280 nm
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEI	polyethyleneimine
PfMBP	*Pyrococcus furiosus* maltodextrin-binding protein
PHA	polyhydroxyalkanoate
PhyC	phytase
PMSF	phenylmethylsulfonylfluoride
PTFE	polytetrafluoroethylene
PVDF	polyvinylidene fluoride
QM/MM	quantum mechanics/molecular mechanics
qPCR	quantitative PCR

RBS	ribosome binding site
R_H	hydrodynamic radius
RT	room temperature
s	second
SAXS	small-angle x-ray scattering
SDS	sodium dodecyl sulfate
SDSM	site directed saturation mutagenesis
SH3	Src homology 3
SLIM	site-directed ligase-independent mutagenesis
SpyCatcher	protein that can form an isopeptide bond with SpyTag
SpyCatcher EQ	unreactive SpyCatcher variant with a E→Q mutation
SpyTag	peptide that can form an isopeptide bond with SpyCatcher
SpyTag DA	unreactive SpyTag variant with a D→A mutation
Sulfo-LC-SPDP	sulfosuccinimidyl 6-[3'-(2-pyridyldithio)propionamido]hexanoate
T_{50}	temperature at which enzyme activity is 50 %
TBS	tris buffered saline
TCA	trichloroacetic acid
TEV	tobacco etch virus
TG2	transglutaminase 2
Tm	temperature at which 50 % of the protein has unfolded
TOF	time-of-flight
UV	ultraviolet
VLP	virus-like particle
WT	wild-type
YPD	yeast extract peptone dextrose

Table of contents

Abstract ..... **ii**

Authorship declaration ..... **iii**

Acknowledgements ..... **iv**

Abbreviations ..... **v**

Table of contents ..... **ix**

Chapter 1: Introduction **2**

1.1 Enzymes: Challenges and opportunities 2

1.1.1 Evolvability of an enzyme 2

1.1.2 Improving reaction efficiency 3

1.1.3 Improving enzyme yield 5

1.1.4 Improving an enzyme's operational stability range 6

1.2 Thermal stabilization of enzymes for industrial applications 8

1.2.1 Rational design approaches 8

1.2.2 Directed evolution approaches 10

1.2.3 Disulfide bond engineering 12

1.2.4 CLEAs, CLECs and immobilization-based approaches 13

1.3 Cyclization techniques as a method for stabilization 16

1.3.1 Carbodiimide, native chemical ligation and cyanocysteine-mediated cyclization 16

1.3.2 Sortase, Butelase 1 and Intein-mediated cyclization 19

1.3.3 Cyclization-mediated improvements in thermal resilience 22

1.4 Spontaneous isopeptide forming Tags and Catchers 24

1.4.1 Spontaneous isopeptide bond formation by bacteria 24

1.4.2 The Isopeptag/Pilin-C system 26

1.4.3 The SpyTag/SpyCatcher system 27

1.4.4 The SnoopTag/SnoopCatcher system 29

1.5 Phage display 29

1.5.1 Introduction to phage display 29

1.5.2 Use of phage display to improve non-covalent interactions 30

1.5.3 Use of Phage display to improve covalent interactions 31

1.6 Thesis aims 33

Chapter 2: Methods **36**

2.1 Cloning 36

2.1.1 Insert amplification 36
2.1.2 Vector amplification, SLIM, Inverse PCR and QuikChange 37
2.1.3 Overlap extension PCR and DNAWorks 37
2.1.4 CPEC, Gibson assembly and ligation 38
2.1.5 Cloning of BLA constructs 39
2.1.6 Cloning of DHFR, PhyC, FLuc and GOx constructs 43
2.1.7 Cloning of phage display constructs 45
2.2 Protein expression and purification 48
2.2.1 Protein expression in E. coli 48
2.2.2 Protein expression in P. pastoris 49
2.2.3 Ni-NTA Affinity Chromatography 50
2.2.4 Anion exchange chromatography 51
2.2.5 Gel filtration 52
2.2.6 Determination of protein concentration 52
2.3 Polyacrylamide Gel Electrophoresis 52
2.3.1 SDS-PAGE 52
2.3.2 Temperature-dependent solubility assay 53
2.3.3 TEV protease cleavage of SpyTag (DA)-BLA-TEV-SpyCatcher 53
2.4 Enzyme activity 54
2.4.1 BLA thermal resilience enzymatic assay 54
2.4.2 BLA kinetics assay 54
2.4.3 DHFR thermal resilience enzymatic assay 55
2.4.4 FLuc thermal resilience enzymatic assay 56
2.4.5 PhyC thermal resilience enzymatic assay 56
2.4.6 PhyC specific activity 57
2.4.7 GOx thermal resilience enzymatic assay 58
2.4.8 Effect of additives to GOx thermal resilience 58
2.5 Biophysical techniques 59
2.5.1 Electrospray Ionization Mass spectrometry (ESI-MS) 59
2.5.2 Circular Dichroism (CD) 59
2.5.3 Differential Scanning Calorimetry (DSC) 60
2.5.4 Dynamic Light Scattering (DLS) 60
2.6 Phage display 61
2.6.1 Conjugating SpyCatcher on magnetic beads 61
2.6.2 Producing pFab5cHis-derived phagemid particles 61
2.6.3 Biotinylation of AP-tagged proteins 62

2.6.4 α-pIII and α-His Western Blots 63
2.6.5 Streptavidin-HRP Western Blot 63
2.6.6 Phage viability test 64
2.6.7 Panning of the PhD-12 library (Rounds 1-5) 64
2.6.8 Panning of the PhD-12 library (Rounds 6-7) 67
2.6.9 qPCR phage kinetics assay 67
2.6.10 Error-prone PCR SpyCatcher library 68
2.7 Tag/Catcher reactions 69
2.7.1 Tag-MBP variants reconstitution reactions 69
2.7.2 Reconstitution of SpyCatcher and SnoopCatcher with SpyTag and SnoopTag peptide 69
Chapter 3: Exploring isopeptide-mediated cyclization approaches 72
3.1 Utilization of split isopeptide-containing domains to cyclize TEM-1 β-lactamase 74
3.1.1 Design of cyclized constructs 74
3.1.2 Isopeptide-mediated cyclization of BLA led to a gel shift when compared to linear mutants 75
3.1.3 TEV protease cleavage of cyclized enzyme compared to the non-cyclized linear control 76
3.1.4 Mass Spectra provides further evidence for successful cyclization 77
3.2 Effects of cyclization on aggregation and recovered activity after heating 80
3.2.1 Effect of cyclization on aggregation 80
3.2.2 Effects of cyclization on recovered activity after heating 82
3.3 Importance of protein domains and the isopeptide bond (comparison to point mutants and fusions like BLA-CnaB2) 83
3.3.1 Aggregation and recovered activity after heating of SnoopTag/SnoopCatcher and Isopeptag/Pilin-C linear points mutants 83
3.3.2 The importance of the isopeptide bond and cyclization on the thermal resilience of SpyTag-BLA-SpyCatcher 85
3.3.3 Cyclization vs. thermostable fusion – effects on aggregation resistance 86
3.3.4 Cyclization vs. thermostable fusion – effects on recovered activity after heating 87
3.4 Studying the increase in thermal resilience using Differential Scanning Calorimetry (DSC) 88
3.4.1 DSC data of cyclized constructs and point mutant controls 88
3.4.2 Studying the effects of cyclization on the refolding of BLA using DSC 91
3.5 Discussion 92

Chapter 4: Cyclization as a general approach to enhance thermal resilience 99

4.1 Cyclization of DHFR 101

4.1.1 Introduction of DHFR 101

4.1.2 MS and gel shift of cyclized DHFR 102

4.1.3 Effect of cyclization on protein aggregation of DHFR 103

4.1.4 Effect of cyclization on recovered activity after heating 104

4.1.5 DSC of DHFR constructs 106

4.2 Cyclization of Firefly Luciferase (FLuc) 107

4.2.1 Introduction of FLuc 107

4.2.2 Effects of cyclization on FLuc 108

4.3 Cyclization of Phytase (PhyC) 110

4.3.1 Introduction of PhyC 110

4.3.2 MS and gel shift of cyclized PhyC 111

4.3.3 Effect of cyclization on protein aggregation of PhyC 113

4.3.4 Cyclization improved the recovered activity after heating 114

4.3.5 DSC of PhyC constructs 116

4.3.6 DLS of PhyC showed little change in the size of cyclized PhyC after heating 117

4.3.7 CD of PhyC constructs provided further evidence that cyclization improved the ability to refold 119

4.4 Cyclization of Glucose Oxidase (GOx) 122

4.4.1 Introduction of GOx 122

4.4.2 Effects of sandwiching GOx between SpyTag/SpyCatcher 123

4.5 Discussion 127

Chapter 5: Phage display as an approach to improve the SpyTag/SpyCatcher interaction 131

5.1 Design of the phage display selection approach 133

5.1.1 Introduction to the phage display approach 133

5.1.2 Design of panning to select for better Tags 133

5.2 Verification of phage display as a potential approach for the selection of improved peptides 135

5.2.1 Western Blot verified the formation of a covalent bond between the SpyTag-displaying phages and SpyCatcher 135

5.2.2 The wash protocol was optimized to improve phage viability 138

5.2.3 A qPCR approach was developed for the rapid screening of selected phages 140

5.3 Panning the NEB PhD-12 library 142
5.3.1 Panning the PhD-12 library led to an enrichment of covalent-bond forming clones 142
5.3.2 Selected clones provided insight into the peptide sequence needed to be able to react with SpyCatcher 144
5.3.3 Initial screens showed that none of the clones were faster-reacting than SpyTag 144
5.3.4 Selected clones were fused to MBP and shown to react slower than SpyTag 145
5.3.5 SallyTag was able to cyclize BLA and confer an increase in thermal resilience 148
5.4 Development of SpyCatcher-displaying phages 149
5.4.1 Strategy devised to make a SpyCatcher library 149
5.4.2 Western blot showed that SpyCatcher was not being displayed properly on the phage 150
5.5 Discussion 152
Chapter 6: Summary and general discussion 157
Bibliography 166
Appendix 188
A.1 Migration of cyclized construct is acrylamide percentage-dependent 188
A.2 Cyclization did not substantially alter the activity of BLA. 189
A.3 Cyclization did not substantially alter the specific activity of PhyC. 190

Chapter 1: Introduction

Chapter 1: Introduction

1.1 Enzymes: Challenges and opportunities

Enzymes are biological catalysts and are usually polypeptides. The majority of biochemical processes in all living organisms are catalysed by enzymes. Humanity has been taking advantage of enzymatic processes since the invention of the fermentation process over 7,000 years ago (Chambers and Pretorius, 2010). However, it was not until 1874 that the first industrially purified enzyme mix was made available commercially. Rennet, a mixture of enzymes including chymosin, was purified from dried calves' stomachs by Christian Hansen. The mixture was used for the coagulation of milk during the production of cheeses (Hui and Sherkat, 2005).

Biotechnology has become an important industry in modern society (Bevan and Franssen, 2006). Nowadays, enzymes are present in a number of industries. These include the: paper and pulp industry, textile industry, laundry detergents, food processing (dairy, baking, juice, etc.), agricultural industry, organic synthesis industry and the cosmetics industry (Li et al., 2012). Enzymes are also vitally important in the healthcare industry as diagnostic devices (Vigneshvar et al., 2016) or as therapeutic drugs (Vellard, 2003).

The utility of enzymes can be attributed to their biodegradability, regio-/chemo-/enantio-selectivity, their ability to function effectively in mild conditions and the ability to chemically and genetically modify enzymes to alter their characteristics (Bastioli, 2005). While enzymes have proven to be effective in the aforementioned industries, some challenges still remain.

1.1.1 Evolvability of an enzyme

The evolvability of an enzyme can be linked to three factors: the enzyme's promiscuity, mutational robustness and thermostability (O'Loughlin et al., 2006). Enzymes that require few mutations to obtain a novel function are considered to be

more likely to survive change. An antibody, for example, can be considered a highly evolvable scaffold due to the immune system's inherent need of a molecule that can be adapted to bind to new external threats (O'Loughlin et al., 2006). Mutational robustness is determined by the amount and distribution of intramolecular bonds (O'Loughlin et al., 2006). Previous work has attempted to quantify molecular robustness through an "x factor" which was defined as the probability that a random mutation will inactivate an enzyme (Guo et al., 2004). Thermostability has been linked to increased evolvability. The effect of thermostability was demonstrated by comparing the amount of P450 mutants with novel catalytic functions derived from a library where the parent sequence was thermostable or marginally stable. The library derived from the thermostable P450 variant contained more clones with novel catalytic activity than the library derived from the marginally stable variant (Bloom et al., 2006).

Molecular robustness, enzyme promiscuity and thermostability can be difficult aspects to study and define. Therefore, in most cases, the evolvability of an enzyme is not taken into consideration when deciding on an enzyme target for a directed evolution or rational design campaign (O'Loughlin et al., 2006).

1.1.2 Improving reaction efficiency

One of the factors limiting the uses of some enzymes is a low catalytic rate. For example, the computational design of novel enzymes with novel functions is currently hampered by the low catalytic rates of the designed enzymes (Damborsky and Brezovsky, 2014).

Enzyme catalytic efficiency can range from a turnover rate of 600,000 s^{-1} for carbonic anhydrase to 1 s^{-1} for tyrosinase (Robinson, 2015). Carbonic anhydrase is considered a diffusion-limited enzyme because the rate of reaction is limited by the diffusion rate. The k_{cat}/K_M of carbonic anhydrase is 10^7 M^{-1} s^{-1} (Williamson, 2011). Most enzymes are not diffusion-limited and instead are limited by other factors such as the formation of the enzyme-substrate complex, formation of intermediates, stabilization

of the transition state, the release of a leaving group or the release of the product from the enzyme and return of the enzyme to the ground state (Linder, 2012).

An enzyme's catalytic activity can be improved via rational approaches, random approaches or a combination of both approaches. Rational approaches utilize biochemical data, enzyme structures and computational methods to identify beneficial mutations (Steiner and Schwab, 2012). For example, the 3DM database combines multiple sequence alignment information with known mutations, protein-ligand contact information, conserved residues and protein-protein contact information of a protein superfamily (Kuipers et al., 2009). The database has been successfully used to improve the catalytic activity of an esterase from *Pseudomonas fluorescens* by 240-fold (Jochens and Bornscheuer, 2010).

A random approach (directed evolution) is the dominant approach (Zinchenko et al., 2014). Directed evolution requires a large library size generated via random mutagenesis (e.g. error-prone PCR, epPCR) or random recombination (e.g. gene shuffling) (Steiner and Schwab, 2012). The important factor in using a directed evolution approach is the selection and/or screening method. Screening a library of $>10^3$ clones cannot be accomplished without a degree of automation. Fluorescence-activated cell-sorting (FACS) has been used to screen for highly active clones (Yang and Withers, 2009). However, the substrate needs to be able to penetrate the cell membrane or the enzyme needs to be displayed on the surface. One of the reasons why microfluidics has been gaining popularity for the purpose of enzyme selection is because the method does not require the substrate to penetrate a cell membrane (Levin and Aharoni, 2012).

Many metabolic and natural degradation processes are reliant on multiple enzymes. For example, the degradation of lignocellulose, a complex branched polymer, requires β-glucosidase, endoglucanase and laccase. The mixture of the three enzymes can have a synergistic effect on the degradation of beechwood xylan (Scharf et al., 2011). In the cases where a product requires multiple enzymes to form, it has

been found that the rate of product formation can be increased by bringing the enzymes in close proximity to one another and in the correct order through the use of a DNA scaffold and zinc finger fusions (Conrado et al., 2012). Bringing enzymes closer together by using an engineered scaffold could also be used *in vivo;* such as for the increase in production of terpenoid indole alkaloids (Masakapalli et al., 2014) or for the study of the effects of compartmentalization on the metabolic flux of a system (Sweetlove et al., 2014).

1.1.3 Improving enzyme yield

An enzyme's yield has to surpass a threshold in order to become viable. Increasing the yield of enzymes expressed by an organism can be achieved using the following approaches: codon optimization, optimization of the ribosome binding site (RBS), optimization of transcription speed, optimization of transport route and optimization of fermentation conditions (i.e. to increase the total cell mass) (Liu et al., 2013). Other important factors affecting an enzyme's yield are post-translational modifications (Sahdev et al., 2008) and enzyme misfolding upon expression (Morra et al., 2016).

Codon optimization can have a large effect on the yield of protein. The most simple strategy was the one amino acid one codon approach which used the most abundant codon for each amino acid (Fuglsang, 2003). However, a gene constructed using the one codon one amino acid principle does not reflect the tRNA pool present in the organism, creating an imbalance which could deplete the tRNA and decrease the growth and production of the organism (Gong et al., 2006). Therefore, the most popular codon optimization approach takes into account the codon frequency in each organism and will randomly pick a codon for each amino acid based on a probability weighting (Jayaraj et al., 2005). Other optimizations that have gone into codon optimization software include GC content optimization (Gustafsson, 2009), avoiding unfavourable codon context (Villalobos et al., 2006), avoiding mRNA sequence motifs

which can target the mRNA for degradation (Pertzev and Nicholson, 2006) and using rare codons at the N-terminus of a gene (Goodman et al., 2013).

The sequence of the promoter can affect the rate of transcription by affecting the binding strength between the RNA polymerase complex and the promoter (Davis et al., 2011). RBSs modulate the rate of translation by affecting the strength of the interaction between the 16S ribosomal RNA, the mRNA and the forymylmethionine-tRNA in the translation initiation complex (directly or through the use of mRNA secondary structures) (Salis et al., 2009).

In some cases, optimizing the quantity of tRNA can have a positive effect on the yield of protein. For example, the expression of a native-sized glycine-rich (45 %) spider silk protein in *E. coli* was improved through the co-expression of extra glycyl-tRNA (Xia et al., 2010).

In certain cases, the yield of a protein could be improved by optimizing the transportation route. By choosing to secrete the protein into the medium, one can skip a purification step, since the protein can be easily separated from the cytosolic proteins via centrifugation. Secretion is the preferred strategy when working with *Pichia pastoris* (Liu et al., 2013) and the yield of product can be dependent on the signal sequence (Lin-Cereghino et al., 2013).

Fermentation conditions have a major effect on the cell density and protein yield. Fermentation processes vary depending on the organism (Liu et al., 2013). Moreover, there appears to also be a target protein dependency on the optimum fermentation condition, which makes it difficult elucidate any general rules (Sivashanmugam et al., 2009). Central factors are the main carbon source and medium, the amount of dissolved oxygen, the pH and the temperature.

1.1.4 Improving an enzyme's operational stability range

Enzymes tend to work in a narrow range of conditions. Broadening the range of conditions at which an enzyme is able to operate can make new applications

feasible. The conditions which need to be broadened for some applications include pH and tolerance to organic solvents and oxidizing agents.

The operational pH range can be a limiting factor for an enzyme's application. For example, one might want a fungal laccase with an active pH range of 2.0 - 4.0 to operate at pH 7.0 for pulp biobleaching or enzymatic cofactor regeneration (Torres-Salas et al., 2013). In order to obtain a laccase which could operate at a more alkali pH researchers have used a directed evolution approach. The approach was based on selecting active mutants in pH 5.0 buffer from an epPCR- and gene shuffling-derived library. A clone was selected which had a k_{cat}/K_M value at neutral pH which was 31-fold higher than the parental strain (Torres-Salas et al., 2013). Using a similar approach, researchers discovered a mutant α-amylase with 5-fold higher activity at pH 10 for use in laundry detergents (Bessler et al., 2003).

Micro-organism derived endoglucanases tend to have acidic pH optima (Fukumori et al., 1988) which can also be problematic for the use of cellulases in laundry detergents (Ito et al., 1989). Researchers used an epPCR library of endoglucanases to discover an endoglucanase mutant that can operate at more alkali conditions (pH optimum shifted from 4.8 to 5.6) (Wang et al., 2005).

Other than directed evolution approaches, bioprospecting can be a good way to find enzymes that operate well in acidic or alkali conditions. Acidophiles and alkaliphiles have evolved to survive in alkali or acidic conditions and therefore researchers have been able to find low and high pH-tolerant enzymes by studying these extremophiles (Fujinami and Fujisawa, 2010; Sharma et al., 2012).

An enzyme's ability to operate in organic solvents can also be highly desirable. Organic solvents can increase the solubility of hydrophobic products and can allow phase separation of the product from the side-products and substrate. Organic solvents are detrimental to hydrolysis and can therefore shift a thermodynamic equilibrium or avoid the formation of water-dependent side-products (Carrea and Riva, 2000). However, the majority of enzymes have evolved to operate in aqueous solution

(Stepankova et al., 2013). Some organisms have evolved to survive in environments containing organic solvents. By studying these organisms, we can discover enzymes which might be tolerant to organic solvents. For example, an organic solvent-tolerant protease was discovered in a strain of *Pseudomonas aeruginosa* (Ogino et al., 1999).

Other than through bioprospecting, tolerance to organic solvents can be increased via a multitude of approaches. These approaches include the use of media and supports to protect the enzymes from the organic solvent. Adsorption of the enzyme to a surface, covalent binding of the enzyme to a surface, chemical cross-linking, entrapment in a matrix addition of surfactants to encapsulate the enzyme and including ionic liquids in the buffer are all methods that have been used to shield the enzyme from the organic solvent (Stepankova et al., 2013). Alternatively, enzyme engineering through rational, semi-rational or random approaches have also been successful in increasing the stability of enzymes in organic solvents (Stepankova et al., 2013).

Laundry detergents are not only alkali, but they also contain oxidizing agents like bleach. Enzymes that are likely to be used in detergents such as proteases, amylases and lipases therefore need to be resistant to oxidizing agents (Fujinami and Fujisawa, 2010). Oxidant-resistant proteases and lipases have been previously found (Bisht et al., 2014; Joo et al., 2003; Xu et al., 2015). Encapsulation has been used to make an amylase resistant to surfactants and oxidizing agents (Amid et al., 2014).

1.2 Thermal stabilization of enzymes for industrial applications

1.2.1 Rational design approaches

Rational design approaches are data-driven and require structural and biochemical data to be effective. An advantage of rational design is that there is a higher probability for beneficial mutations, thereby limiting the number of clones that need to be screened (Steiner and Schwab, 2012). Moreover, limiting the amount of

clones can sometimes remove the need to set-up a high-throughput screening approach which usually rely on the use of surrogate substrates (fluorescent or chromophoric analogues) which might not be truly indicative of an enzyme's ability to react with the natural substrate (Goldsmith and Tawfik, 2013).

Rational design requires information to be effective. One way to obtain information is through the sequence consensus approach. The sequence consensus approach utilizes structural information like homology modelling (Vyas et al., 2012) or protein threading (Xu et al., 2008) to build protein structure models and/or consensus sequence information obtained through multiple sequence alignments (MSAs) with homologues. The conserved residues discovered through the MSAs tend to be evolutionarily favourable and have been linked to increases in thermal resilience. The sequence consensus approach has been used to increase the Tm of a fungal phytase by 28 °C (Lehmann et al., 2002). Not all consensus mutations are beneficial which is why using protein threading and homology modelling to obtain models for protein structure can be important for informing the researcher which residues might be important for structural stability (Steiner and Schwab, 2012).

Similarly to the sequence consensus approach, building phylogenetic trees from available consensus sequences can help researchers infer ancestral protein sequences which tend be more thermostable (Gaucher et al., 2008). Ancestral protein sequences appear to be more thermostable because the Earth was 30 °C hotter 3.5 billion years ago (Gaucher et al., 2008). This approach has been used to find mutations which increased the Tm of a 3-isopropylmalate dehydrogenase from *Thermus thermophilus* (Watanabe et al., 2006).

Software has not only been developed to predict structures but also to predict the effects that a mutation might have on said structure. Computational tools such as FoldX (Guerois et al., 2002), Rosetta (Kellogg et al., 2011), ERIS (Yin et al., 2007) and CUPSAT (Parthiban et al., 2006) allow us to predict whether a mutation will be energetically beneficial or not. However, these tools can only predict the effect of one

mutation at a time and sometimes these predictions are incorrect (Bednar et al., 2015). FireProt has been developed to combine the information obtained from evolution and energy-based prediction tools, to predict stable multiple-point mutants thereby decreasing the amount of variants that need to be tested (Bednar et al., 2015). FireProt has been used to increase the Tm of haloalkane dehalogenase by 24 °C and γ-hexachlorocyclohexane dehydrochlorinase by 21 °C (Bednar et al., 2015).

1.2.2 Directed evolution approaches

Directed evolution aims to mimic the natural process of evolution under two conditions: an increased rate of mutation and a selection pressure that can be chosen by the researcher. The increased rate of mutation can be accomplished through the use of epPCR (Chen and Arnold, 1993), DNA shuffling (Stemmer, 1994), site directed saturation mutagenesis (SDSM) (Miyazaki and Arnold, 1999) and iterative saturation mutagenesis (ISM) (Reetz and Carballeira, 2007). The selection pressure that needs to be implemented in many cases is high temperature. Since most organisms used for enzyme production have difficulties surviving high temperatures, cell-based screening and selection methods cannot be used to select for enzymes with high thermostability.

Finding a high-throughput selection and screening method for thermostable enzymes remains a challenge. Work has been carried out to create a microfluidic device that can analyze the stability of proteins, but that work was limited to measuring the fluorescence of green fluorescent protein (GFP) (Sagar et al., 2013). Differential scanning fluorimetry (DSF) remains the most high-throughput approach for the study of protein thermostability (Niesen et al., 2007) but is hampered by the need for a dye which can alter the transition temperature of a protein (Shi et al., 2013). Moreover, it is a challenge to study more than 10^3 – 10^4 clones using DSF.

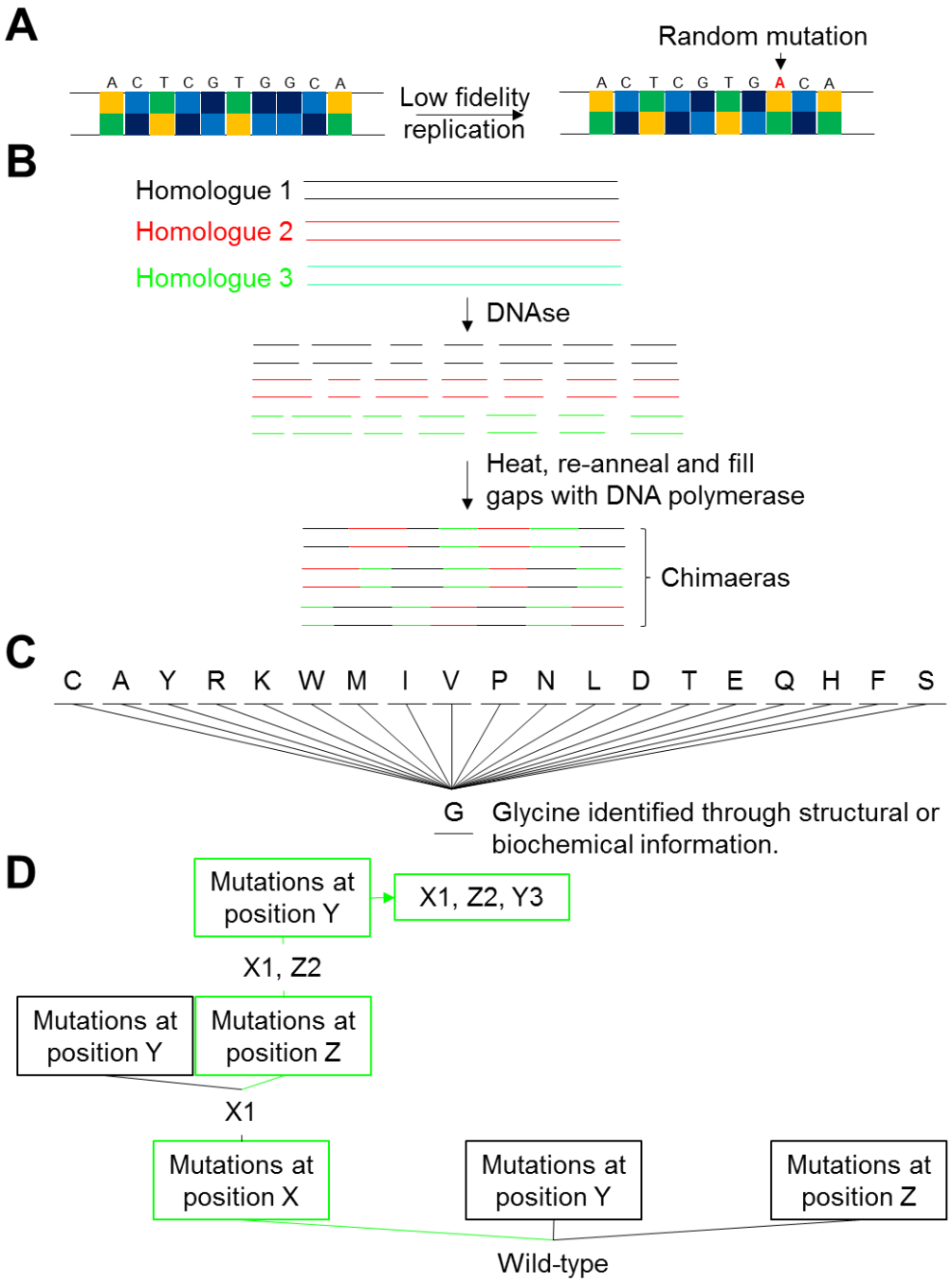


Figure 1.1: Methods for introducing diversity in a library. A) epPCR inserts random mutations into a coding sequence using a replication method with low fidelity. B) DNA shuffling uses homologous sequences to create chimaeric sequences. The homologues are cut randomly by a nuclease after which the fragments are melted, allowed to re-anneal and gaps are filled by a polymerase. C) SDSM involves mutating a residue of interest to all the other amino acids to find a beneficial mutation. D) Diagram of a typical ISM experiment. Structurally important positions in the protein sequence are identified and SDSM libraries are made in each position. The most stable clone is then selected and used to build SDSM libraries mutating the other positions. The process is repeated several times.

Even with the limitations imposed by a relatively low throughput, successes have been achieved. Through the use of epPCR to discover beneficial mutations and the use of SDSM to combine the mutations to have an overall additive effect, the T_{50} (temperature at which the enzyme is only 50 % active) of phosphite dehydrogenase was increased by 20 °C (Johannes et al., 2005). In a semi-rational approach which subjected the residues with the highest B-values in a crystal structure to ISM, the T_{50} value of a lipase from *Bacillus subtilis* was increased from 50 °C to >100 °C (Reetz et al., 2006).

1.2.3 Disulfide bond engineering

Disulfide bonds are formed between the two thiol groups of two cysteine residues in oxidizing environments when the distance of the two C_β residues is ≤ 5.5 Å. The ideal geometry of the disulfide bond is such that the C_α-C_β-S_γ angle is ~114°, the C_β-S_γ-S_γ is ~105° and the torsion angle between the two C_β atoms around the S-S bond is ~-87° and ~97°. Disulfide bonds can form using slightly different angles but tend to be weaker than in the ideal case (Dombkowski et al., 2014). The disulfide bond can be disrupted through the addition of a reducing agent or in a reducing cellular environment. The stabilizing effect of disulfide bonds has been known for a long time. Initial experiments on the T_1 Ribonuclease showed that removal of the two native disulfide bonds decreased the Tm of the protein from 59.3 °C to 16.6 – 27.2 °C (Pace et al., 1988). Introduction of a disulfide bond between C97 and the mutated I3C in T4 lysozyme increased the kinetic stability of the protein at 67 °C (Perry and Wetzel, 1984).

Currently, there are two theories to explain why the introduction of some disulfide bonds are beneficial while others are neutral or harmful to the stability of the protein. T4 lysozyme, barnase and dihydrofolate reductase (DHFR) mutant constructs with engineered disulphide bonds were crystallized and their structures elucidated.

The disulphide bonds either stabilized or destabilized the proteins depending on where they had been introduced. Simulations on the crystal structures indicated that stabilization of regions which unfold early is beneficial whereas stabilization of regions that unfold late can be detrimental (Ramakrishnan et al., 2012). Molecular simulations of the unfolding of a phytase identified unfolding regions that could be stabilized through the engineering of disulfide bonds. The introduction of three disulfide bonds increased the Tm of the protein by 12.1 °C and the introduction of 2-3 disulfide bonds had a positive effect on the reversibility of inactivation (Sanchez-Romero et al., 2013).

The second theory involves the use of disulfide bonds to cross-link flexible regions. Work on engineering disulfide bonds into T4 lysozyme led to the hypothesis that an engineered disulfide bond is only effective if the bond location includes a flexible region and also the disulfide bond has to create a large loop size (Matsumura et al., 1989). It is thought that the extra covalent bond decreases the conformational enthalpy of the unfolded states (Flory, 1956). B-factors have been used as a reference point to determine where to insert a disulfide bond. This approach was used to stabilize lipase B from *Candida antarctica,* increasing the T_{50} by 8.5 °C (Le et al., 2012).

Prediction tools exist to determine which residues in a protein could be mutated to cysteines and form a disulfide bond, e.g. Disulfide by Design 2.0 (Craig and Dombkowski, 2013).

1.2.4 CLEAs, CLECs and immobilization-based approaches

Enzyme immobilization has been linked to an increase in thermal resilience. One method to immobilize an enzyme is to cross-link the enzyme. Enzyme crosslinking through the addition of glutaraldehyde has been known since the 1960s (Sheldon, 2011). Cross-linked enzyme aggregates (CLEAs) were developed by inducing protein aggregation and then chemically cross-linking the aggregated protein (Cao et al., 2000). CLEAs can have a larger operational stability than their soluble enzyme counterparts in terms of both thermal stability and resistance to organic

solvents. Papain from *Carica papaya* was found to be more stable after the enzyme had undergone the CLEA process. The soluble papain lost >60 % of its activity after only a week whereas the CLEA papain was still ~90 % active after 5 weeks at 50 °C (Sheldon, 2011). Using a CLEA-based immobilization approach, the optimum temperature of subtilisin was increased from 55 to 70 °C and the kinetic stability at 60 °C was also increased (Sangeetha and Emilia Abraham, 2008).

The CLEA process can be further improved through the addition of other proteins and additives. For example, the addition of bovine serum albumin (BSA) and starch to phenylalanine ammonia lyase during the CLEA process further increased the stability of the enzyme (Cui et al., 2013). A disadvantage of using CLEAs is that the method is most effective with proteins which contain a lot of exposed amines. Carboxyl-CLEAs (CLEAs cross-linked using PEI) can be a good strategy for stabilizing enzymes with few surface amine groups Treating a lipase from *Candida rugose* with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) and N-hydroxysuccinimide (NHS) and then cross-linking with 1.3 kDa polyethyleneimine (PEI) increased the thermal stability of the enzyme more than using glutaraldehyde (Velasco-Lozano et al., 2014).

An alternative approach to CLEAs are cross-linked enzyme crystals (CLECs). CLECs are made by first crystallizing the enzyme and then cross-linking the crystal (St. Clair and Navia, 1992). These crystals are robust and can increase the thermal stability of an enzyme. A lipase from *Burkholderia cepacia* was stabilized after crystallization and cross-linking. The soluble enzyme was quickly inactivated at 70 °C whereas the CLEC lipase was still >80 % active after 6 h (Rajan and Abraham, 2007). The main disadvantage of using CLECs is that one requires a high concentration of pure enzyme in order to form a crystal and this can sometimes be too expensive to be practical (Brady et al., 2004). Moreover, some proteins can be difficult if not impossible to crystallize. The isoelectric point, average hydropathy, presence of transmembrane

helices and protein disorder can all determine whether an enzyme can or cannot be crystallized (Slabinski et al., 2007).

Instead of enzyme cross-linking, enzymes can also be stabilized by immobilizing them on a support or surface. It is thought that the surface provides a rigid steric backbone which can keep the enzyme folded. For example, adsorption of a lipase onto a mesoporous silica nanoparticles was able to increase the kinetic stability of the enzyme (Jiang et al., 2016). Immobilization of an enzyme onto a surface can be done through three approaches: physical adsorption, entrapment and covalent attachment (Figure 1.2) (Mohamad et al., 2015). Physical adsorption can be susceptible to enzyme leaching (Mohamad et al., 2015). Enzyme leaching can be avoided by covalently attaching the enzyme to the surface or by trapping the enzyme on the surface through the addition of a silicone coat (Sheldon and van Pelt, 2013). However, when encapsulating or entrapping an enzyme, one needs to be cautious with the mass transfer limitation that the coating could inflict (Popat et al., 2011).

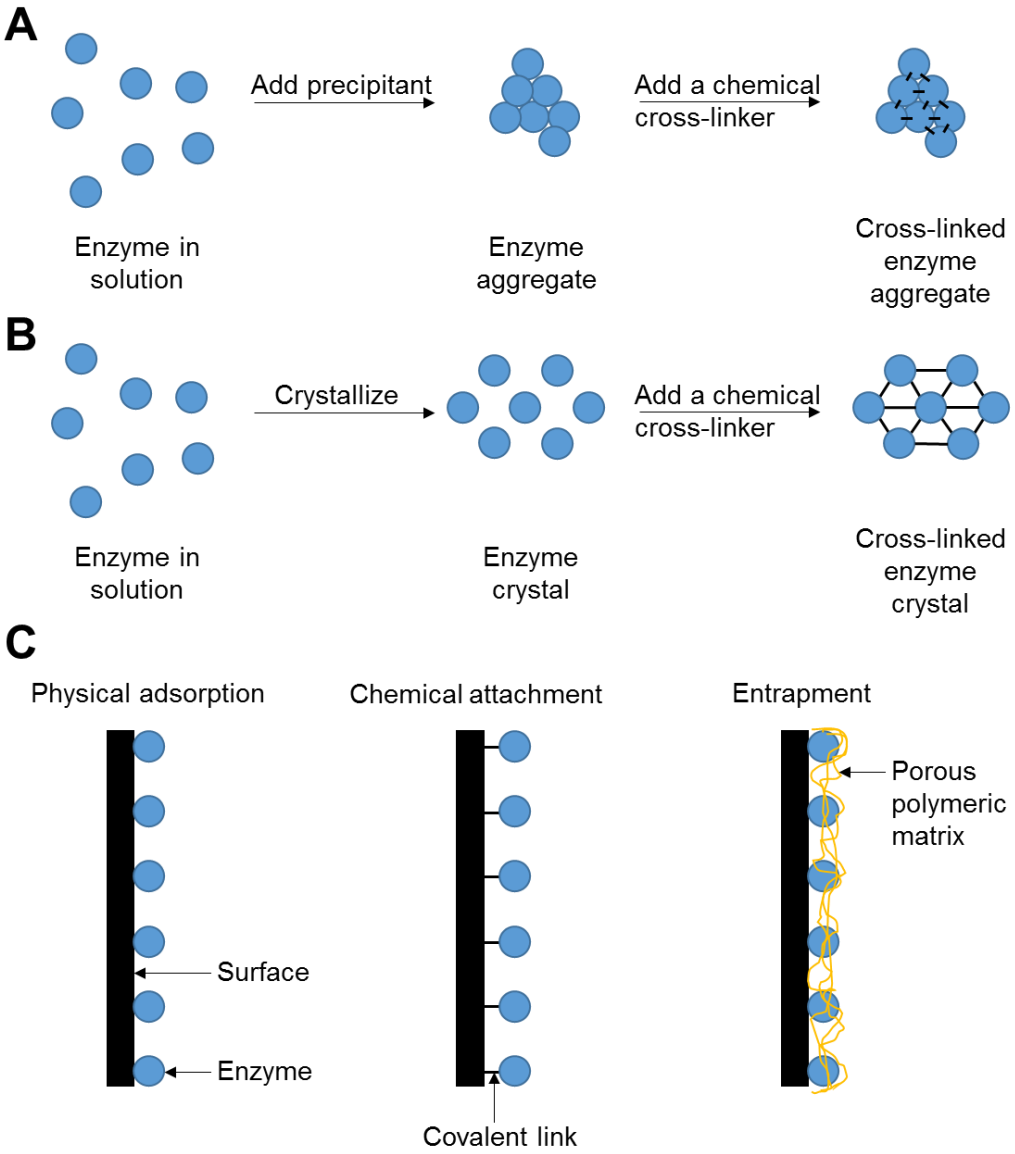


Figure 1.2: Examples of immobilization methods. A) Outline of method to create CLEAs. B) Outlines of method to create CLECs. C) The three ways to immobilize an enzyme onto a surface. Physical adsorption relies on the non-covalent interactions between the enzyme and the surface. Interactions can include van der Waals forces, hydrophobic interactions or ionic interactions. Chemical attachment uses a chemical cross-linker to bind the enzyme to a surface. Entrapment uses physical adsorption to attach the enzyme to the surface and then coats the enzyme with a porous polymeric matrix.

1.3 Cyclization techniques as a method for stabilization

1.3.1 Carbodiimide, native chemical ligation and cyanocysteine-mediated cyclization

The first reported instance of a cyclized protein used EDC, a carbodiimide, to link the N-terminus and C-terminus of bovine pancreatic trypsin inhibitor (BPTI), a 58 amino acid protein. The main interest in cyclizing BPTI was to study the folding intermediates and their relationship to the formation of the three disulfide bonds present in the protein (Goldenberg and Creighton, 1983).

The linkage was made possible by the short distance between the N- and C-terminus (8.3 Å, PDB: 4PTI) (Marquart et al., 1983). Less than 50% of the total protein was successfully cyclized through this approach. Cyclization increased the mobility of the protein on an SDS-PAGE. This increase in mobility was countered by trypsin-induced linearization of the protein (Goldenberg and Creighton, 1983).

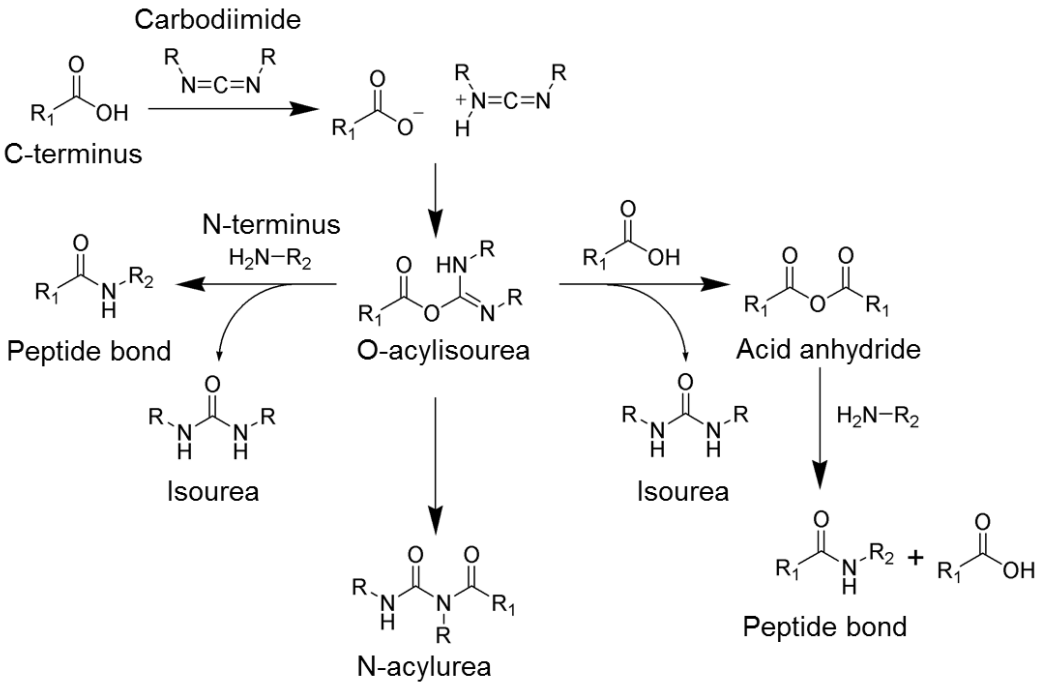


Figure 1.3: Carbodiimide-induced cyclization of a protein (adapted from: username: "~K" (Chemical mechanism of amide formation using a carbodiimide) [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0/)], via Wikimedia Commons). Carbodiimide can undergo one undesired rearrangement to form a stable N-acylurea. The unstable O-acylisourea can also be hydrolyzed by water to yield an unmodified carboxylic acid group or the O-acylisourea can react with an amine group to form a peptide bond or another carboxylic group to form an acid anhydride which can then react with an amine group to form a peptide bond.

The main disadvantage of the carbodiimide approach was that the approach has the potential of forming a lot of unwanted side-products, both through the formation of N-acylurea (Goldenberg and Creighton, 1983) as well as the formation of bonds between exposed amine and carboxylic acid-containing amino acid side-chains (Figure 1.3).

Additionally, native chemical ligation (NCL) has also been used to cyclize polypeptides. NCL was developed by Professor Stephen Kent and his colleagues and was initially used in order to chemically synthesize larger polypeptides (Dawson et al., 1994). More recently, NCL has been used to synthesize cyclotides and sunflower trypsin inhibitor –I as well as cyclizing canonotoxins (Clark and Craik, 2010). The advantages of synthesizing and/or cyclizing these small proteins (14-29 amino acids long) (Jia et al., 2014) is an increase in stability towards proteolysis, thermal stresses and chemical stresses (Colgrave and Craik, 2004). The mechanism for NCL-mediated cyclization is illustrated in Figure 1.4A. So far, the use of NCL to cyclize polypeptides has been limited to peptides and small proteins (Takahashi et al., 2007).

In order to cyclize a larger polypeptide through a chemical biology approach, one can use cyanocysteine-mediated cyclization. Cyanocysteine-mediated cyclization (Figure 1.4B) has been applied in the cyclization of a methionine- and cysteine-free dihydrofolate reductase (DHFR) mutant called ANLYF. The authors wanted to stabilize an enzyme with DHFR activity because of DHFR's clinical importance (target of antifolate drugs such as methotrexate) (Takahashi et al., 2007) and ability to produce leucovorin, an anti-cancer drug (Uwajima et al., 1990). Cyanocysteine-mediated cyclization was ~50 % efficient and increased the mobility of ANLYF on SDS-PAGE. Cyclized ANLYF showed increased resistance to proteolysis by carboxypeptidase (Takahashi et al., 2007).

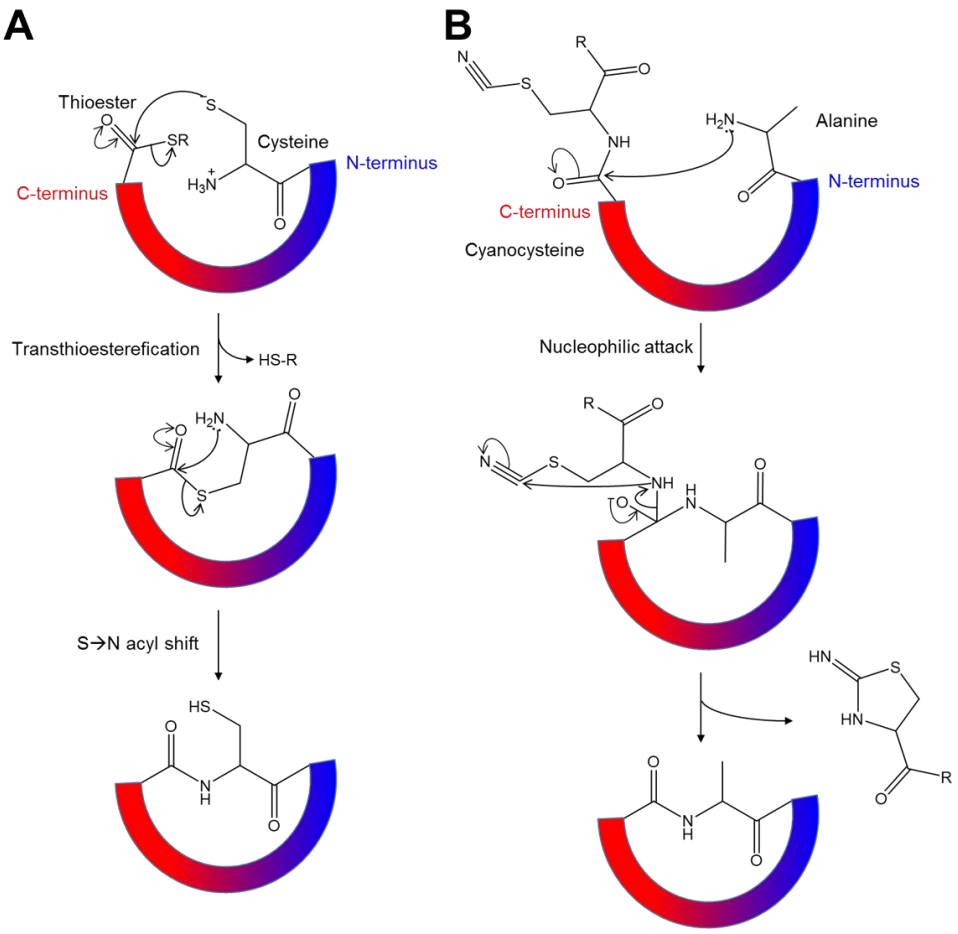


Figure 1.4: Diagrams of NCL and cyanocysteine-mediated cyclization. A) NCL-mediated cyclization of a linear polypeptide. NCL requires an N-terminal cysteine residue and a C-terminal thioester group. The sulfur group on the cysteine nucleophilically attacks the carbonyl group in a transthioesterification step. Then the amine group nucleophilically attacks the carbonyl group in an S→N acyl shift. B) Cyanocysteine-mediated cyclization of a linear polypeptide. A cysteine was engineered close to the C-terminus and alanine was engineered on the N-terminus. The cysteine was cyanylated through the use of 2-nitro-5-thiocyanobenzoic acid. In an alkaline solution, the cyanocysteine residue is excised through a nucleophilic attack from the N-terminal amine and a peptide bond forms between the N- and C-terminus (Jacobson et al., 1973; Takahashi et al., 2007).

1.3.2 Sortase, Butelase 1 and Intein-mediated cyclization

Sortase A is a transpeptidase commonly found in Gram-positive bacteria. The main function of sortase A is to catalyse the condensation reaction that anchors polypeptides to the cell wall (Marraffini et al., 2006). Sortase A recognizes a C-terminal LPXTG motif and form an acyl-intermediate which can then be nucleophilically attacked by a polyglycyl amine group (Figure 1.5A) (Parthasarathy et al., 2007). By

incorporating a string of glycines on the N-terminus and LPXTG on the C-terminus, researchers have managed to use sortase A to cyclize proteins such as enhanced GFP (Antos et al., 2009; Parthasarathy et al., 2007). Cyclization increased the mobility of eGFP on SDS-PAGE and the reaction reached a yield of ~50 % cyclized protein after incubating at 37 °C for 24 h and a final yield of >90 % (Antos et al., 2009). Cyclization of Cre recombinase was >90 % efficient and increased the mobility of the protein (Antos et al., 2009). Sortase A has also been used to stabilize cytokines to improve the stability and circulatory half-life of these therapeutic proteins. In all the cases, the cytokines were cyclized and cyclization led to an increased in mobility on SDS-PAGE (Popp et al., 2011).

Disadvantages of using sortase A are that sortase A requires a long incubation time (overnight) and a large amount of sortase A (0.1 – 10 moles of Sortase per mole of target) (Nguyen et al., 2015). Butelase I is the first reported asparagine/aspartate protein ligase which requires a C-terminal Asn/Asp-His-Val motif to function (Figure 1.5B) (Nguyen et al., 2014). Butelase I is fast (k_{cat} = 17 s^{-1}) and efficient (yield = >95 %) (Nguyen et al., 2014) and was able to cyclize GFP ~20,000 times faster than sortase A (Nguyen et al., 2015). Butelase I was also used to cyclize the therapeutic proteins interleukin-1 receptor agonist (IL-1Ra) and human growth hormone (somatropin). In all butelase I cyclization events, the mobility of the proteins were increased on SDS-PAGE (Nguyen et al., 2015).

Inteins (*int*ervening prot*eins*) are protein domains found in immature proteins which participate in protein splicing. Inteins will excise themselves from the coding sequence and then ligate the two extein fragments back together (Shah and Muir, 2014). Intein-mediated cyclization was first successfully attempted on TEM-1 β-lactamases (BLA) (Iwai and Plückthun, 1999), maltose binding protein (MBP), thioredoxin (Evans et al., 1999a), *E. coli* DHFR (Scott et al., 1999) and a Src homology 3 (SH3) domain (Camarero and Muir, 1999).

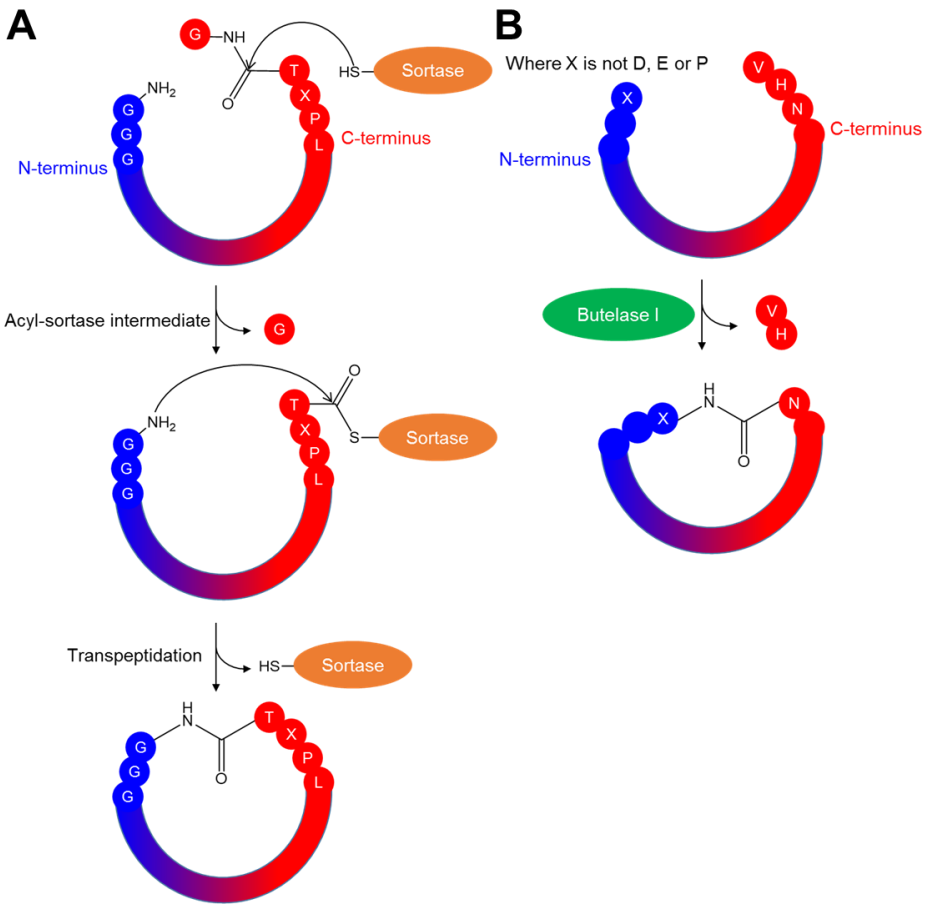


Figure 1.5: Sortase A- and Butelase I-mediated cyclization. A) Diagram of sortase A-mediated cyclization. Sortase A utilizes a catalytic cysteine residue to create a thioacyl-sortase intermediate which is then excised through a transpeptidation reaction. B) Diagram of butelase I-mediated cyclization. The N-terminal amino residue cannot be aspartate, glutamate or proline. The C-terminus needs to have the Asn/Asp (Asx)-His-Val motif to be recognized by butelase I. Butelase I then cleaves between the Asx and the His residues and links the Asx residue to the N-terminal amino acid (Nguyen et al., 2014).

There are two main approaches that can be used to cyclize a protein using intein technology. The first approach requires a cysteine on the N-terminus and fuses a mutant intein domain to the C-terminus of the protein to be cyclized. The mutant intein contains a N→A mutation which removes C-terminal splicing activity (Chong et al., 1996). A thiolate nucleophile is needed for the transthioesterification reaction which releases the C-terminus of the protein from the intein and creates a C-terminal thioester group that can react with an N-terminal cysteine (Iwai and Plückthun, 1999). The second approach uses split inteins. $Intein_N$ is fused to the C-terminus of the protein and $intein_C$ is fused to the N-terminus. When the two fragments associate, the two

fragments are able to catalyse the excision of the two fragments and the formation of a peptide bond between the N- and C-terminus (Figure 1.6) (Scott et al., 1999).

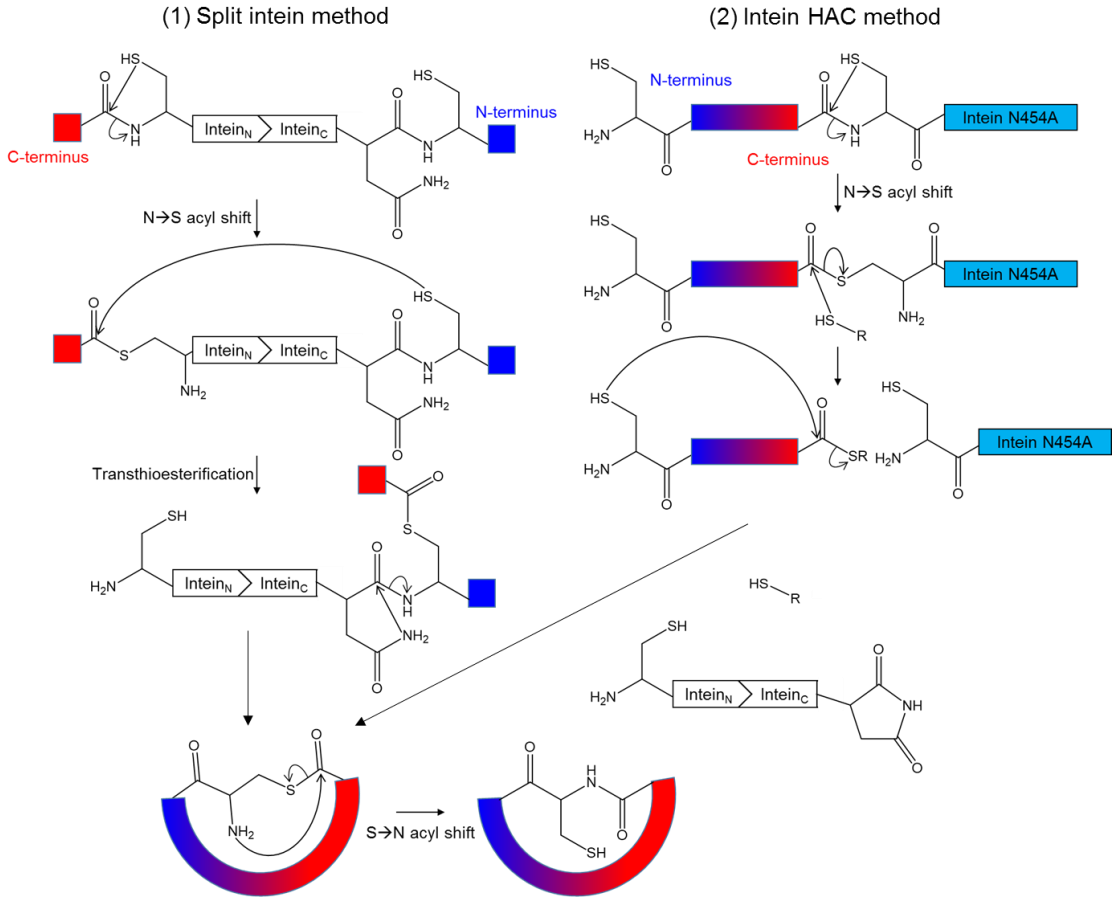


Figure 1.6: Intein-mediated cyclization. This diagram highlights the mechanism of the (1) split intein method (note: the association step of the two intein fragments has not been depicted) and the (2) mutant intein method described in the text.

A disadvantage of using inteins for protein cyclization is that the method requires the genetic fusion of the intein or split intein domains onto the protein to be cyclized which might impair the folding and solubility of the construct (Wood and Camarero, 2014).

1.3.3 Cyclization-mediated improvements in thermal resilience

Cyclization is thought to increase protein stability by constraining the protein backbone (van 't Hof et al., 2015). The constraint leads to decreased conformational entropy when the protein is denatured (Flory, 1956). Increases in thermal resilience

have been reported using intein-, cyanocysteine-, butelase I- and sortase-mediated cyclization.

Intein-mediated cyclization is the most prominent cyclization approach and has been reported to increase the thermal resilience of several proteins. Intein-mediated cyclization of *E. coli* DHFR improved the kinetic stability of the enzyme. The kinetic stability was studied by incubating the enzyme at 65 °C and measuring the recovered activity after cooling the sample to room temperature at different time-points (Scott et al., 1999). Cyclization of the N-terminal domain of *E. coli* DnaB helicase led to an increase in Tm of 13.6 °C as measured through differential scanning calorimetry (DSC) (Williams et al., 2002). Cyclizing a subunit from the *E. coli* glucose transporter led to an increase in Tm of 5 °C, measured by monitoring the change in ellipticity at 222 nm (Siebold and Erni, 2002).Intein-mediated cyclization of GFP led to an increase in solubility at higher temperatures. The temperature at which 50 % of the protein aggregated was increased from 73 to 77 °C (Zhao et al., 2010).

Intein-mediated cyclization has also been used to stabilize the industrially important enzyme, endoglucanase (Yennamalli et al., 2013). Intein-mediated cyclization of an endoglucanase increased the kinetic stability at 65 °C. Moreover, DSC and fluorescence spectroscopy determined that the cyclization had led to an increase in thermodynamic stability as well. Interestingly, a cyclized variant with a 14 amino acid connecting loop was more kinetically stable than a cyclized variant with a 20 amino acid connecting loop (van Lieshout et al., 2012).

Cyanocysteine-mediated cyclization is complicated by the fact that the protein should not contain any sulphur groups to avoid non-specific cyanylation. However, the approach led to an increase in thermodynamic stability of a mutant DHFR. Absorption spectra were used to determine that cyclization had increased the Tm by 9.3 °C. Moreover, the cyclized protein remained soluble after a 30 min incubation at 80 °C, whereas the linear protein aggregated (Takahashi et al., 2007).

Butelase I appears to be a promising and effective approach for enzyme cyclization. However, the technique is still quite novel so not much cyclization data is currently available. Butelase I-mediated cyclization of IL-1Ra was shown to increase the thermodynamic stability of the protein. DSF was used to determine that the Tm of the linear protein was 53 °C and the Tm of the cyclized protein was 57 °C (Nguyen et al., 2015).

While butelase I might have some advantages in terms of reaction speed and efficiency over sortase, sortase-mediated cyclization has also been shown to increase the thermal resilience of proteins. Sortase-mediated cyclization of cytokines led to an increase in thermodynamic stability with an increase in Tm of 3.6 °C for human interferon α2 (IFNα2) and an increase in Tm of 4.9 °C for human granulocyte colony-stimulating factor 3 (GCSF-3) (measured using DSF) (Popp et al., 2011). Enhanced GFP has also been cyclized using sortase and this led to improvement in refolding capability over time after a 5 min incubation step at 90 °C (Antos et al., 2009).

1.4 Spontaneous isopeptide forming Tags and Catchers

1.4.1 Spontaneous isopeptide bond formation by bacteria

The major pilin protein, Spy0128, from *Streptococcus pyogenes* was crystallized and its structure elucidated in 2007. Analysis of the electron densities in the structure determined that Spy0128 contained two intramolecular isopeptide bonds formed between the side-chains of a Lys and Asn (Kang et al., 2007). By studying the published structures of other Gram positive immunoglobulin-like fold pilin proteins and adhesins, Kang et. al discovered that the Cna protein from *Staphylococcus aureus* and the minor pilin (GBS52) protein from *Streptococcus agalactiae* also contain isopeptide bonds. Moreover, sequence alignments showed that Spy0125 might also contain an isopeptide bond (Kang et al., 2007). Analysis of the crystal structure of the

Streptococcus pyogenes-derived CnaB2 domain from the fibronectin-binding (FbaB) protein determined that the protein contained an isopeptide bond between a Lys and Asp (Oke et al., 2010).

The isopeptide bond is formed spontaneously without the need of any cofactor or enzyme. The isopeptide bond formation is mediated by a Glu residue. All three residues are buried in a hydrophobic pocket surrounded by aromatic residues. The hydrophobic environment is thought to keep the Glu in a protonated state and the Lys in an unprotonated state by increasing the pK_a of Glu and decreasing the pK_a of Lys. Computational quantum mechanics/molecular mechanics (QM/MM) simulations revealed that the reaction mechanism might have two stages. The first stage involves the direct nucleophilic attack of the Cγ group of the Asx (Asn/Asp) residue by the Lys ε-amino group forming a tetrahedral intermediate. The second stage then involves two coordinated proton transfers which releases either a water (if Asp) or ammonia (if Asn) molecule (Figure 1.7) (Hu et al., 2011).

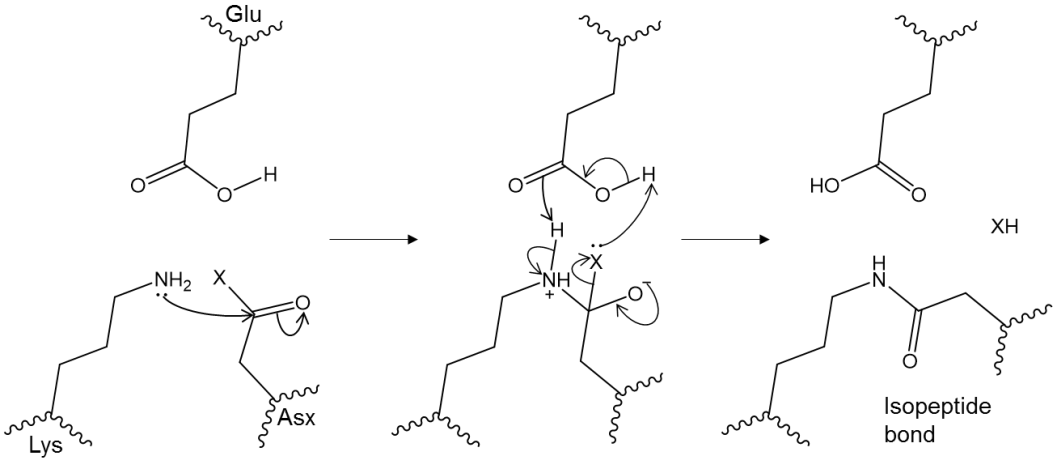


Figure 1.7: Formation of an intramolecular isopeptide bond. The proposed mechanism is based on the QM/MM findings (Hu et al., 2011).

The isopeptide bond confers a resistance against mechanical stress (800 pN), low pH (pH 2), temperature (Tm of 85 °C as determined by circular dichroism) and proteolysis (Kang and Baker, 2009, 2011). Gram-positive bacteria are thought to often lack the machinery which Gram-negative bacteria use to cross-link pilins and adhesins with disulfide bonds and therefore might rely on Sortase A and intramolecular

isopeptide bonds to confer resistance to the aforementioned stresses (Kang et al., 2007).

1.4.2 The Isopeptag/Pilin-C system

Previous work with split-protein systems such as the split DHFR (Pelletier et al., 1999), β-galactosidase (Rossi et al., 1997), BLA (Galarneau et al., 2002), and yellow fluorescent protein (Hu et al., 2002) provided evidence that a genetically split protein can associate into a functional bimolecular protein. Therefore, a split Spy0128 protein where the catalytic Asn and Lys are separated might be able to form a covalent bond upon the association of the two components.

Isopeptag and Pilin-C were developed by splitting the final C-terminal β-strand of the Spy0128 domain (Figure 1.8). Isopeptag is a 16 amino acid peptide containing the reactive Asn, whereas Pilin-C is a 32 kDa protein which contains the reactive Lys and Glu necessary to form the isopeptide bond. Upon mixing of the two molecules, an intermolecular isopeptide bond was formed (Zakeri and Howarth, 2010).

The Isopeptag/Pilin-C system has been used to create hyperthin fibrils. The system was engineered to encode the Isopeptag sequence on the N-terminus of Pilin-C and a protein cap was engineered into the binding pocket with a disulfide bond which would block the Lys from reacting with the Isopeptag on the N-terminus. Upon the addition of a reducing agent, the engineered protein would polymerize and form fibrils (Matsunaga et al., 2013).

By fusing blue fluorescent protein (BFP) to the N-terminus of Pilin-C, it was demonstrated that the efficiency of Förster resonance energy transfer (FRET) after isopeptide bond formation was dependent on whether the Isopeptag was fused to the N- or C-terminus of GFP. Isopeptag/Pilin-C is therefore a powerful protein scaffold that allows for the spatial arrangement of the proteins that the scaffold is joining (Abe et al., 2013).

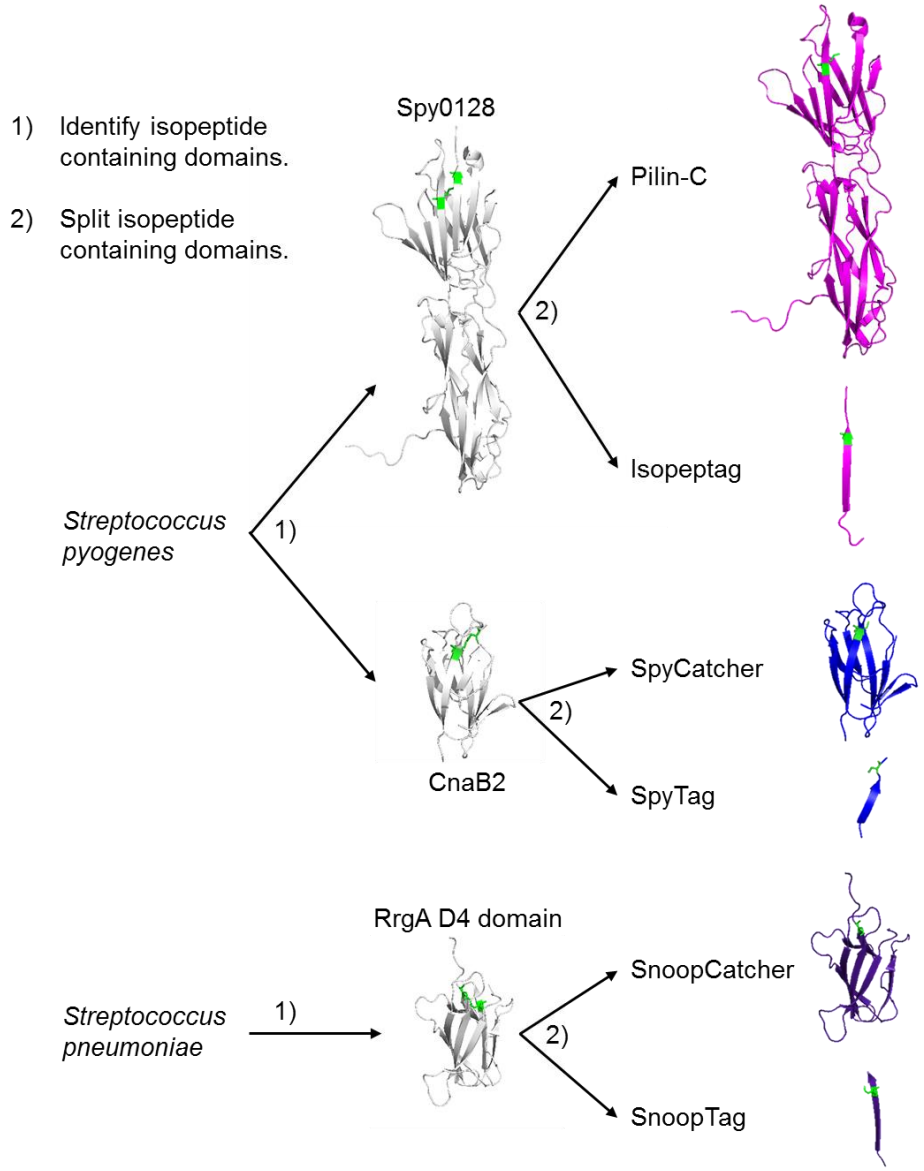


Figure 1.8: Forming Tag/Catcher pairs. Spy0128, CnaB2 and RrgA were identified as proteins containing intramolecular bonds. The proteins were split to create the three Tag/Catcher pairs: Isopeptag/Pilin-C, SpyTag/SpyCatcher and SnoopTag/SnoopCatcher.

1.4.3 The SpyTag/SpyCatcher system

The Isopeptag/Pilin-C system had several disadvantages. The reaction was slow and inefficient, reaching a yield of ~60 % in 72 h. Moreover, the Pilin-C protein is quite large (32 kDa) (Zakeri and Howarth, 2010). To find a better Tag/Catcher system, Zakeri et. al split the CnaB2 domain to make the SpyTag/SpyCatcher system (Figure 1.8). The final C-terminal β-sheet became a 13 amino acid peptide (contains catalytic

Asp) that could form an irreversible isopeptide bond with a 12 kDa protein called SpyCatcher (contains catalytic Lys and Glu). The reaction was fast reaching ~80 % completion in 10 min with a rate constant of 1.4 x 10^3 M^{-1} s^{-1} (Zakeri et al., 2012).

SpyTag/SpyCatcher is a versatile system which has been used in a lot of applications. One of the applications involves the manufacture of vaccine candidates. SpyTag/SpyCatcher can be used to covalently bind an antigen to a dendritic cell-targeting antibody (Liu et al., 2014). Fusing SpyCatcher to the coat protein of a virus-like particle (VLP) allowed researchers to decorate the VLP with SpyTag-antigen fusions (Brune et al., 2016; Thrane et al., 2016).

SpyTag/SpyCatcher has also been used for the assembly of large protein complexes. By combining the SpyTag/SpyCatcher with Streptavidin/biotin, Fairhead et. al was able to assemble an eicosamer capable of binding 12 biotinylated targets (Fairhead et al., 2014).

Biomaterials is another field where SpyTag/SpyCatcher has proven to be useful. By fusing SpyTag/SpyCatcher to elastin-like polypeptides (ELPs), researchers were able to form functionalizable hydrogels which could be used for cell culture and stem cell programming (Sun et al., 2014). Curli fibers are another potential biomaterial which could be used with conjunction with SpyCatcher. Curli fibers are an extracellular amyloid material found in the biofilm of some strains of *E. coli*. By fusing one of the monomeric subunits of the curli fiber (CsgA) with SpyTag, researchers were able to decorate the fibers with quantum dot-bound SpyCatcher (Chen et al., 2014). SpyTag fused to CsgA was also used to functionalize curli fibers with α-amylase fused to SpyCatcher. Immobilization of the amylase to the curli fibers led to an increase in amylase stability towards pH, organic solvents and inactivation over time (Botyanszki et al., 2015).

1.4.4 The SnoopTag/SnoopCatcher system

Elucidation of the crystal structure of domain D4 of a pilus-related adhesin (RrgA) from *Streptococcus pneumoniae* revealed the existence of an intramolecular isopeptide bond between a side-chain Lys and Asn (Izoré et al., 2010). In order to develop an orthogonal pair to SpyTag/SpyCatcher, Veggiani et. al split the RrgA D4 domain after the first N-terminal β-strand (Figure 1.8). SnoopTag is a 12 amino acid peptide that contains the reactive Lys residue, whereas SnoopCatcher is a 13 kDa protein which contains the catalytic Asn and Glu residue. SnoopCatcher would not be able to react with SpyTag because the catalytic Lys is missing (Veggiani et al., 2016).

By fusing SpyCatcher to the N-terminus of SnoopCatcher, Veggiani et. al were able to perform solid-phase protein chain synthesis. By combining nanobodies that target death receptor 5 (DR5) with affibodies targeting epidermal growth factor receptor (EGFR), Veggiani et. al were able to enhance killing of a cancer cell-line (Veggiani et al., 2016).

1.5 Phage display

1.5.1 Introduction to phage display

Phage display is an *in vitro* selection method first developed in 1985 (Smith, 1985). Phage display involves the use of a filamentous phage where the protein of interest is expressed on the surface, fused to one of the coat proteins. Since a virus is composed of DNA or RNA and protein, any selected phage can be directly identified via the genetic code that is contained within. Selection usually occurs via protein-protein interactions or other non-covalent interactions which will select for the properties of interest. This tool is an incredibly effective high-throughput screening method, with uses in protein engineering via directed evolution thanks to the phage's relatively fast replication cycle (Clackson and Lowman, 2004).

The phages used in most phage display assays are either the M13, f1 or fd phage. These phages are 98% genetically identical and are therefore referred to collectively as Ff phages. These phages are lysogenic- the phage progeny is assembled and later secreted, rather than lysing the cell. They tend to have a fixed diameter of 6.5 nm and their length is determined by the size of their genome. The main constituent of the phage coat is the pVIII protein, which multimerizes into a flexible tube like structure around the single-stranded circular DNA. The hydrophobic interactions between the pVIII multimers cause the monomers to be oriented at 20° from the vertical axis like the metal plates on scale armour. However, unlike scale armour, the monomers multimerize to form a right-handed helix (Clackson and Lowman, 2004).

The two ends of the phage capsid can be distinguished by their shapes under an electron microscope. One end is flat and is composed of 3-5 copies each of pVII and pIX. The flat end is the end on which the other components are built during phage assembly and the end where the packaging sequence (a hairpin structure found in the ssDNA genome) is orientated. The other end of the phage is pointed and is composed of 5 copies of pVI and 5 copies of pIII. The function of pVI is not well studied. pIII is involved in the initial infection process. The remaining proteins encoded in the genome are used by the phage for replication (pII, pX and pV) or during phage assembly (pI, pXI and pIV) (Clackson and Lowman, 2004).

1.5.2 Use of phage display to improve non-covalent interactions

The main use of phage display has focused on the selection of strong non-covalent binding. One of the major uses of phage display has been the selection of antibodies that bind tightly to an antigen (Bazan et al., 2012), which was first developed in 1990 (Huse et al., 1989; McCafferty et al., 1990). Phage display peptide libraries have also been important for finding therapeutic peptides for drug manufacture (Bazan et al., 2012) as well as for the discovery of peptides that can bind to membrane

proteins, finding enzyme inhibitors, studying protein-protein interactions, selection of peptides that bind to non-protein targets or the selection of cell-specific peptides (Wu et al., 2016).

1.5.3 Use of Phage display to improve covalent interactions

Phage display has been used for the development of covalent bond-forming technologies. For example, SNAP-tag was developed using phage display. SNAP-tag is a 10 kDa mutant enzyme derived from human O(6)-alkylguanine-DNA alkyltransferase (hAGT). hAGT can react with O(6)-benzylguanine (BG) by nucleophilically attacking the molecule with a reactive cysteine and forming an irreversible covalent bond with the molecule. Derivatives of BG can be made with fluorophores and other molecules which are useful for labelling. SNAP-Tag was obtained from hAGT via phage display by fusing the enzyme to the pIII protein and mutating four key residues using the NNK method (Reetz et al., 2008) to select for an hAGT mutant with better activity towards BG (Juillerat et al., 2003). Phage display has also been used to improve the reaction efficiency of subtiligase. Subtiligase catalyzes the ligation of C-terminally ester-activated peptides to the N-terminal α-amine of a second peptide. In 1999, Atwell and Wells selected for subtiligase mutants with increased ligation efficiency, by expressing subtiligase with an extended N-terminus on pIII of filamentous phage and then incubating the phage with peptides which are known substrates of subtiligase. Atwell and Wells managed to select for two mutants that exhibited 2.5-fold and 2.7-fold increased activity (Atwell and Wells, 1999).

Other than increasing reaction rates, phage display has also been used to make sortase's specificity less stringent. A penta-glycine tail was added on the N-terminus of the sortase-pIII fusion (G5-sortase) and then the G5-sortase was allowed to react with a biotin molecule containing the LPXTG motif on the C-terminus. A sortase mutant was isolated which was less selective at the leucine position of the motif (Piotukh et al., 2011).

Phage display is an effective technique for the selection of peptide sequences with which proteins are able to interact covalently. Phage display has been used for the selection of orthogonal Sfp (*Bacillus subtilis*) and AcpS (*E. coli*) substrate recognition sequences. Sfp and AcpS are phosphopantetheinyl transferases that covalently attach a pantetheinyl group onto a serine side chain using coenzyme A as a substrate. Zhou et al. obtained two 12 amino acid long peptides that were labeled by either AcpS or Sfp, by first depleting each library for non-orthogonal peptides and then selecting for peptides that could still react with AcpS or Sfp (Zhou et al., 2007).

Phage display can also be used to study the substrate motif of proteins. Polyak et al. fused a known biotin domain from yeast pyruvate carboxylase to the pIII protein and used error prone PCR to build their library of mutants. Polyak et al. were able to select for biotinylated phage which could be sequenced to elucidate the motif necessary for *E. coli* biotin ligase-mediated biotinylation (Polyak et al., 2001). Moreover, phage display can be used for the selection of good peptide substrates. In 2005, Keresztessy et al. used phage display in order to select for good transglutaminase 2 (TG2) substrates. TG2 functions by first creating a covalent bond with a Gln residue by using a Cys residue to create an enzyme-substrate intermediate and then releases it when an amine donor is available (usually the amino group of the K side chain), creating a isopeptide bond between the two substrates (Keresztessy et al., 2006).

In the future phage display might be used to improve the reaction efficiency of split inteins. Garbe et al. have shown that it is possible to fuse one of the segments of *Ssp* DnaB split intein to a truncated pIII on a M13 phage without impairing its activity. The researchers are looking into using this to improve the splicing efficiency of the split intein system (Garbe et al., 2010).

1.6 Thesis aims

The main goal of this thesis was to develop a novel cyclization approach which was (i) quick and simple to execute, (ii) efficient and produced a good yield of enzyme, (iii) effective at increasing the thermal resilience of enzymes and (iv) a general approach which could be used on more than one enzyme. A cyclization approach that would fulfil the four criteria listed would be invaluable in many industries as a simple method to increase the operational range of many enzymes remains elusive.

Genetically encoding Tag/Catcher pairs at the N- and C-termini of a model enzyme would first allow us to test the feasibility of the approach as well as gaining an understanding of the underlying mechanism of thermal resilience. Determining which Tag/Catcher pair works well and why was crucial for understanding how the system might work and for hypothesising how the system might be improved in the future.

Once the best Tag/Catcher pair was discovered, the generality of the approach would have to be tested on various enzymes. How well isopeptide-mediated cyclization works on each enzyme might help us determine the limitations of the system and to think about alternative ways that the system could be used instead to confer resilience.

Using phage display to evolve Tag/Catcher pairs could provide us with faster-reacting pairs that would increase the cyclization efficiency, increase enzyme solubility and yield and/or could confer a greater thermal resilience which might eliminate some of the limitations of the current system and make isopeptide-mediated cyclization even more general and effective.

Chapter 2: Methods

Chapter 2: Methods

2.1 Cloning

Polymerase Chain Reaction (PCR) was performed in a C1000 Thermal Cycler (Bio-Rad). KOD Hot Start DNA Polymerase (Roche) was used. DNA constructs were transformed into chemically competent *E. coli* DH5α (Life Technologies). All primers were synthesized by Life Technologies. All cloned constructs were verified by Sanger sequencing. DNA was quantified using a ND-1000 NanoDrop (Thermo Scientific) OD_{260}.

2.1.1 Insert amplification

A typical PCR reaction contained 1 X KOD Hot Start Buffer, 0.2 mM mixed dNTPs, 0.2 µM forward and reverse primers, 2.5 mM $MgSO_4$, 30 ng template DNA, 0.02 U/µl KOD Hot Start DNA polymerase, with or without 3 % DMSO (v/v) and double deionised water to a volume of 25 µl. Typical cycling conditions were 95 °C for 3 min, 20 cycles of 95 °C for 30 s followed by 55 or 62 °C for 30 s and 68 °C for 1 min. Samples were then cooled to 10 °C. For most amplifications, 4 conditions were tested: annealing temperature at 55 or 62 °C and 0 or 3 % DMSO (v/v). The length of the elongation step at 68 °C was varied depending on the length of the insert to be amplified.

PCR products were analysed on a 0.75 - 2 % agarose gel (dependent on the size of the DNA fragment) containing 0.01 % (v/v) ethidium bromide (Sigma). 5 µl of sample was mixed with 1 µl of 6 X Gel loading dye (1.2 M sucrose, 0.7 mM bromophenol blue, 0.9 mM xylene cyanol FF) and loaded on the gel. The gel was run at 130 V for 45 min. Gels were imaged using a ChemiDoc XRS+ (Bio-Rad) and analysed using Image Lab 3.0 software (Bio-Rad).

If necessary, 1 µl of DpnI was added to the remaining 20 µl of PCR sample and incubated at 37 °C for 3 h to remove the template DNA. Inserts were then purified using the QIAquick Gel Extraction Kit (Qiagen).

2.1.2 Vector amplification, SLIM, Inverse PCR and QuikChange

A typical Inverse PCR (Kaluz et al., 1999), vector amplification PCR or QuikChange (Stratagene) reaction contained 1 X KOD Hot Start Buffer, 0.2 mM mixed dNTPs, 0.2 µM forward and reverse primers, 2.5 mM $MgSO_4$, 30 ng template DNA, 0.02 U/µl KOD Hot Start DNA polymerase, with or without 3 % DMSO (v/v) and double deionised water to a volume of 25 µl. For site-directed ligase-independent mutagenesis (SLIM) PCR (Chiu et al., 2004), 4 primers were used at 0.2 µM. Typical cycling conditions were 95 °C for 3 min, 20 cycles of 95 °C for 30 s followed by 55 or 62 °C for 30 s and 68 °C for 4 min. Samples were then cooled to 10 °C. For most amplifications, 4 conditions were tested: annealing temperature at 55 or 62 °C and 0 or 3 % DMSO (v/v). The PCR product was digested with DpnI, purified using the QIAquick Gel Extraction Kit and then treated with T4 Polynucleotide Kinase (NEB). The DNA was then self-ligated.

PCR products were analysed on 0.75 % agarose gels and then the methylated template DNA was digested with DpnI. For amplified vector and Inverse PCR, products were purified using the QIAquick Gel Extraction Kit. QuikChange and SLIM PCR products were transformed directly into *E. coli* DH5α.

2.1.3 Overlap extension PCR and DNAWorks

A typical Overlap Extension PCR (Bryksin and Matsumura, 2010) reaction contained 1 X KOD Hot Start Buffer, 0.2 mM mixed dNTPs, 0.2 µM forward and reverse primers, 2.5 mM $MgSO_4$, equimolar quantities of 2 or more DNA fragments (mixed to a final weight of ~20 ng), 0.02 U/µl KOD Hot Start DNA polymerase, with or without 3 % DMSO (v/v) and double deionised water to a volume of 50 µl. Typical

cycling conditions were 95 °C for 3 min, 10 cycles of 95 °C for 30 s followed by 55 °C for 30 s and 68 °C for 1 min. Samples were then cooled to 10 °C. For most overlap extension PCRs, 4 conditions were tested: annealing temperature at 55 or 62 °C and 0 or 3 % DMSO (v/v).

DNAWorks (Hoover and Lubkowski, 2002) was used to design the 60 bp overlapping primers used to assemble the PhyC gene. A typical assembly reaction contained 1 X KOD Hot Start Buffer, 0.2 mM mixed dNTPs, 0.4 µM flanking forward and flanking reverse primers, 40 nM of each primer in-between the flanking reverse or forward primer, 2.5 mM $MgSO_4$, 0.02 U/µl KOD Hot Start DNA polymerase and double deionised water to a volume of 50 µl. Cycling conditions were 95 °C for 3 min, 25 cycles of 95 °C for 30 s followed by 55 or 62 °C for 30 s and 68 °C for 1 min and 1 final incubation step at 68 °C for 2 min 20 s. Samples were then cooled to 10 °C. The correct fragment was purified by gel extraction using the QIAquick Gel Extraction Kit.

2.1.4 CPEC, Gibson assembly and ligation

A typical circular polymerase extension cloning (CPEC) (Quan and Tian, 2009) reaction contained 1 X KOD Hot Start Buffer, 0.4 mM mixed dNTPs, 2.25 mM $MgSO_4$, 250-500 ng of vector and equimolar amounts of insert, 0.02 U/µl KOD Hot Start DNA polymerase and double deionised water to a volume of 50 µl. Cycling conditions were 95 °C for 3 min, 10 cycles of 95 °C for 30 s followed by 62 °C for 30 s and 68 °C for 4 min. 3 µl of CPEC reaction was transformed into chemically competent *E. coli* DH5α.

Gibson assembly (Gibson et al., 2009) was performed using a Gibson Assembly Master Mix (NEB) and by following the recommended conditions. 50 ng of vector were mixed with a 3-fold molar excess of insert and double deionised water was added to a volume of 10 µl. 10 µl of Gibson Assembly Master Mix (2X) was added and the reaction was incubated at 50 °C for 60 min. 3 µl of Gibson reaction was transformed into chemically competent *E. coli* DH5α.

Ligations were performed by digesting inserts and vectors using the appropriate restriction enzymes. The insert was purified using the QIAquick Gel Extraction Kit or the QIAquick PCR Purification Kit (Qiagen) and the vector was purified using the QIAquick Gel Extraction Kit. 25 ng of vector was mixed with a 3-fold molar excess of insert, 1 X T4 DNA Ligase Buffer (NEB), 20 U/µl T4 DNA Ligase (NEB) and double deionised water up to 20 µl final volume. The reaction was incubated at 20 °C for 1 h and 3 µl was transformed into chemically competent *E. coli* DH5α.

2.1.5 Cloning of BLA constructs

pET28a SpyTag-β-lactamase-SpyCatcher (SpyTag-BLA-SpyCatcher) and pET28a SpyTag DA-BLA-SpyCatcher were cloned by Dr. Jacob O. Fierer in the Howarth laboratory. The SpyTag-BLA fragment was amplified from the pDEST14 (Life Technologies) vector using primers 5'-AGGACATATGGGAGCCCACATCGTGATGGTGGACGCCTACAAGCCGACGAAG GGTTCAGGGGGTTCCGGT and 5'-TCCGCTGCCACCACTCCCCCAATGCTTAATCAGTGAGGCACCTATCTCAG. The SpyTag DA-BLA fragment was amplified using primers 5'-AGGACATATGGGAGCCCACATCGTGATGGTGGCCGCCTACAAGCCGACGAAG GGTTCAGGGGGTTCCGGT and 5'-TCCGCTGCCACCACTCCCCCAATGCTTAATCAGTGAGGCACCTATCTCAG. The SpyCatcher fragment was amplified from pDEST14 SpyCatcher (Zakeri et al., 2012) using primers 5'-GGGAGTGGTGGCAGCGGAGGCGCCATGGTTGATACCTTATCAGG and 5'-TTTAAAGCTTTCATTAAATATGAGCGTCACCTTTAGTTGCTTTGCC. The fragments were then assembled using Overlap Extension PCR. The SpyTag-BLA-SpyCatcher insert was amplified using primers 5'-AGGACATATGGGAGCCCACATCGTGATGGTGGACGCCTACAAGCCGACGAAG GGTTCAGGGGGTTCCGGT and 5'-

TTTAAAGCTTTCATTAAATATGAGCGTCACCTTTAGTTGCTTTGCC. The SpyTag DA-BLA-SpyCatcher insert was amplified using primers 5'-AGGACATATGGGAGCCCACATCGTGATGGTGGCCGCCTACAAGCCGACGAAG GGTTCAGGGGGTTCCGGT and 5'-TTTAAAGCTTTCATTAAATATGAGCGTCACCTTTAGTTGCTTTGCC. The inserts and pET28a (Novagen) vector were digested using NdeI (NEB) and HindIII (NEB) and then ligated.

pET28a SpyTag-BLA was cloned by Dr. Jacob O. Fierer. The insert was amplified from pDEST14 using primers 5'-GGTTCAGGGGGTTCCGGTCACCCAGAAACGCTGGTGAAAGTAAAAGATG and 5'-TTTAAAGCTTTCATTACCAATGCTTAATCAGTGAGGCACCTATCTCAG. The insert was then amplified again using primers 5'-AGGACATATGGGAGCCCACATCGTGATGGTGGACGCCTACAAGCCGACGAAG GGTTCAGGGGGTTCCGGT and 5'-TTTAAAGCTTTCATTACCAATGCTTAATCAGTGAGGCACCTATCTCAG. The insert and pET28a (Novagen) vector were digested using NdeI (NEB) and HindIII (NEB) and then ligated.

pET28a BLA was generated using Inverse PCR (Kaluz et al., 1999) from pET28a SpyTag-BLA using primers 5'-CACCCAGAAACGCTGGTGAAAG and 5'-TCCCATATGGCTGCCGCGC. The PCR product was digested with DpnI and then re-ligated using T4 Polynucleotide Kinase and T4 DNA Ligase.

pET28a SpyTag-BLA-SpyCatcher EQ was generated via QuikChange using pET28a SpyTag-BLA-SpyCatcher as a template and primers 5'-GGAAAATATACATTTGTCCAAACCGCAGCACCAGACG and 5'-CGTCTGGTGCTGCGGTTTGGACAAATGTATATTTTCC.

pET28a SpyTag-BLA-TEV-SpyCatcher and pET28a SpyTag DA-BLA-TEV-SpyCatcher (containing a TEV protease cleavage site between the BLA and SpyCatcher) were generated via SLIM PCR (Chiu et al., 2004) using pET28a SpyTag-

BLA-SpyCatcher or pET28a SpyTag DA-BLA-SpyCatcher as templates and primers 5'-GGCGCCATGGTTGATACCTTATCAG, 5'-GAAAACCTGTATTTTCAGGGCGGCGCCATGGTTGATACCTTATCAG, 5'-TCCGCTGCCACCACTCCCC and 5'-GCCCTGAAAATACAGGTTTTCTCCGCTGCCACCACTCCCC.

pET28a BLA-SpyCatcher was generated using Inverse PCR (Kaluz et al., 1999) from pET28a SpyTag-BLA-SpyCatcher using primers 5'-CACCCAGAAACGCTGGTGAAAG and 5'-TCCCATATGGCTGCCGCGC. The PCR product was digested with DpnI and then re-ligated using T4 Polynucleotide Kinase and T4 DNA Ligase.

pET28a BLA-CnaB2 was generated via SLIM PCR (Chiu et al., 2004) using pET28a BLA-SpyCatcher as a template and primers 5'-TAATGAAAGCTTGCGGCCGCACTC, 5'-GTGATGGTGGATGCGTACAAGCCGACGAAGTAATGAAAGCTTGCGGCCGCACTC, 5'-AATATGAGCGTCACCTTTAGTTGCTTTGC and 5'-CTTCGTCGGCTTGTACGCATCCACCATCACAATATGAGCGTCACCTTTAGTTGCTTTGC.

pET28a SnoopTag-BLA-SnoopCatcher was generated via Overlap Extension PCR (Bryksin and Matsumura, 2010). The SnoopTag-BLA fragment was amplified from pET28a SpyTag-BLA-SpyCatcher using primers 5′-ATTACATATGGGAAAACTGGGGGACATCGAATTCATCAAAGTAAACAAAGGTTCAGGGGGTTCCGGTCACC and 5′-CTAAACACGGCACCACGCAGCGGCTTTCCACTGCCACCACTCCCCCAAT. The BLA-SnoopCatcher fragment was amplified from pET28a SnoopCatcher T842G (Veggiani et al., 2016) using primers 5′- AAGCCGCTGCGTGGTGCCG and 5′-ATTAAAGCTTTCATTATTTCGGCGGTATCGGTTCATTGGTG. After assembly, the SnoopTag-BLA-SnoopCatcher insert was amplified using primers 5′-ATTACATATGGGAAAACTGGGGGACATCGAATTCATCAAAGTAAACAAAGGTTC

AGGGGGTTCCGGTCACC and 5′-ATTAAAGCTTTCATTATTTCGGCGGTATCGGTTCATTGGTG, digested using NdeI and HindIII and ligated into pET28a.

pET28a SnoopTag KA-BLA-SnoopCatcher was generated via QuikChange using pET28a SnoopTag-BLA-SnoopCatcher as a template and primers 5′-CCCTGAACCTTTGTTTACTGCGATGAATTCGATGTCCCC and 5’-GGGGACATCGAATTCATCGCAGTAAACAAAGGTTCAGGG.

pET28a Pilin-C-BLA-Isopeptag was generated using Gibson assembly (Gibson et al., 2009). The vector was amplified using pET28a SpyTag-BLA-SpyCatcher and primers 5′- TCCCATATGGCTGCCGCGCGG and 5′-AAGCTTGCGGCCGCACTCGAGC. The Pilin-C fragment was amplified from pET28a Pilin-C (Zakeri and Howarth, 2010) using primers 5′-CCGCGCGGCAGCCATATGGGAGCTACAACAGTTCACGGGGAGAC and 5′-GGTGACCGGAACCCCCTGAACCAATGGTCATATCTTTATCAGTAGATGTCTCTTG. The BLA-Isopeptag fragment was generated using pET28a SpyTag-BLA-SpyCatcher as a template and primers 5′- GGTTCAGGGGGTTCCGGTCAC and 5′-GCTCGAGTGCGGCCGCAAGCTTTCATTATTCCGCATCTTTTTTGTTGGTAAAGGTAATGGTCATATCTTTATCGGTTCCACTGCCACCACTCCCCC. The three fragments were combined in a single Gibson assembly reaction (Gibson et al., 2009).

pET28a Pilin-C KA-BLA-Isopeptag was generated via QuikChange using primers 5′- TCTACTACATTAACGGTGAAGGCAAAAGTTTCAGGTACCGGTGG and 5’- CCACCGGTACCTGAAACTTTTGCCTTCACCGTTAATGTAGTAGA.

pET28a SallyTag-BLA-SpyCatcher was generated via SLIM PCR (Chiu et al., 2004) using pET28a SpyTag-BLA-SpyCatcher as a template and primers 5’-GGTTCAGGGGGTTCCGGTCACC, 5’-AGTTATCAGATTAGTATGGTTGATACGAAGACGGAGGGTTCAGGGGGTTCCGGTCACC, 5’- TCCCATATGGCTGCCGCGCGG and 5’-

CTCCGTCTTCGTATCAACCATACTAATCTGATAACTTCCCATATGGCTGCCGCG
CGG.

pET28a SallyTag DA-BLA-SpyCatcher was generated via QuikChange using pET28a SallyTag-BLA-SpyCatcher as a template and primers 5'-GTTATCAGATTAGTATGGTTGCGACGAAGACGGAGGGTTCAG and 5'-CTGAACCCTCCGTCTTCGTCGCAACCATACTAATCTGATAAC.

pET28a Spy0128 was cloned by Dr. Bijan Zakeri in the Howarth laboratory and the cloning method was described previously (Zakeri, 2011).

2.1.6 Cloning of DHFR, PhyC, FLuc and GOx constructs

pET28a SpyTag-dihydrofolate reductase-SpyCatcher (SpyTag-DHFR-SpyCatcher) and pET28a SpyTag DA-DHFR-SpyCatcher were generated via CPEC (Quan and Tian, 2009). The vector fragment was generated by using pET28a SpyTag-BLA-SpyCatcher or pET28a SpyTag DA-BLA-SpyCatcher as templates and primers 5'-GGGAGTGGTGGCAGCGGAG and 5'-ACCGGAACCCCCTGAACCCTTC. The DHFR insert was generated by amplifying the *FolA* gene from the genome of *E. coli* BL21 DE3 RIPL (Stratagene) using primers 5'-CCGACGAAGGGTTCAGGGGGTTCCGGTATGATCAGTCTGATTGCGGCGTTAGC and 5'-GCGCCTCCGCTGCCACCACTCCCCCGCCGCTCCAGAATCTCAAAGC.

pET28a DHFR was generated using pET28a SpyTag-DHFR-SpyCatcher as a template and primers 5'-ATTACATATGGGAATGATCAGTCTGATTGCGGCGTTAGC and 5'-AAAAAAGCTTTTACCGCCGCTCCAGAATCTCAAAG. The insert was then digested using NdeI and HindIII and then ligated into pET28a.

pET28a SpyTag-PhyC-SpyCatcher and pET28a SpyTag DA-PhyC-SpyCatcher were generated via CPEC (Quan and Tian, 2009). The vector fragment was generated by using pET28a SpyTag-BLA-SpyCatcher or pET28a SpyTag DA-BLA-SpyCatcher as templates and primers 5'-GGGAGTGGTGGCAGCGGAG and 5'-ACCGGAACCCCCTGAACCCTTC. The PhyC insert (NCBI reference sequence:

WP_013352583.1) was generated by assembling the gene using DNAWorks (Hoover and Lubkowski, 2002) and then using primers 5′-CCGACGAAGGGTTCAGGGGGTTCCGGTAAACATAAACTGAGCGATCCGTATCATTTCAC and 5′-GCGCCTCCGCTGCCACCACTCCCTTTGCCCGAACGATCGGTCAATTTACG.

pET28a PhyC was generated using pET28a SpyTag-PhyC-SpyCatcher as a template and primers 5′-TATACATATGGGAAAACATAAACTGAGCGATCCGTATCATTTCAC and 5′-TTTTAAGCTTTCATTATTTGCCCGAACGATCGGTCAATTTACG. The insert was then digested using NdeI and HindIII and then ligated into pET28a.

pET28a SpyTag-firefly luciferase-SpyCatcher (SpyTag-FLuc-SpyCatcher) and pET28a SpyTag DA-FLuc-SpyCatcher were generated via CPEC (Quan and Tian, 2009). The vector fragment was generated by using pET28a SpyTag-BLA-SpyCatcher or pET28a SpyTag DA-BLA-SpyCatcher as templates and primers 5'-GGGAGTGGTGGCAGCGGAG and 5'-ACCGGAACCCCCTGAACCCTTC. The FLuc insert was generated using pGL2-Basic (Promega) as a template and primers 5′-CCGACGAAGGGTTCAGGGGGTTCCGGTGAAGACGCCAAAAACATAAAGAAAGGCCC and 5′-GCGCCTCCGCTGCCACCACTCCCCAATTTGGACTTTCCGCCCTTCTTGG.

pPICLIT SpyTag-glucose oxidase (SpyTag-GOx, NCBI reference sequence: P13006.1) was generated via Gibson assembly (Gibson et al., 2009). The insert was synthesized as a GeneArt String DNA fragment (Life Technologies) and amplified using primers 5’-GGCTGAAGCTGCAGGAATTCCTGCAGGTGCTCACATCGTTATGGTTGACGC and 5’-CAGATCCTCTTCTGAGATGAGTTTTTGTTCTAGACCTTAACCAGATCCACCGGAACCTTGCATAG. The vector was amplified from pPICLIT (derived from pPICZα, Life

Technologies) using primers 5'-TCTAGAACAAAAACTCATCTCAGAAGAGGATCTG and 5'-TGCAGGAATTCCTGCAGCTTCAGCC.

pPICLIT GOx was generated via Gibson assembly (Gibson et al., 2009). The insert was generated using pPICLIT SpyTag-GOx as a template and primers 5'-GGCTGAAGCTGCAGGAATTCCTGCATCCAACGGTATTGAAGCTTCCTTGTTGAC and 5'-CAGATCCTCTTCTGAGATGAGTTTTTGTTCTAGACCTTATTGCATAGAAGCGTAG TCCTCCAAAATAGC. The vector was amplified from pPICLIT SpyTag-GOx using primers 5'-TCTAGAACAAAAACTCATCTCAGAAGAGGATCTG and 5'-TGCAGGAATTCCTGCAGCTTCAGCC.

pPICLIT SpyTag-GOx-SpyCatcher was generated via Gibson assembly (Gibson et al., 2009). The SpyCatcher insert was generated using pET28a SpyTag-PhyC-SpyCatcher as a template and primers 5'-TTCTATGCAAGGTTCCGGTGGATCTGGTGGCGCCATGGTTGATACCTTATCAG and 5'-CAGATCCTCTTCTGAGATGAGTTTTTGTTCTAGACCTTAAATATGAGCGTCACCT TTAGTTGCTTTGC. The vector was amplified from pPICLIT SpyTag-GOx using primers 5'-TCTAGAACAAAAACTCATCTCAGAAGAGGATCTG and 5'-ACCAGATCCACCGGAACCTTGCATAG.

2.1.7 Cloning of phage display constructs

pDEST14 Cys-SpyCatcher and pDEST14 Cys-SpyCatcher EQ were cloned by Dr. Bijan Zakeri in the Howarth laboratory and the cloning method has been described previously (Zakeri, 2011). pET9d TEV protease was kindly supplied by Prof. Kim Nasmyth (University of Oxford) (Haering et al., 2008).

pDEST14 acceptor peptide-SpyCatcher (AP-SpyCatcher) (Fairhead and Howarth, 2015) and pDEST14 AP-SpyCatcher EQ were generated via SLIM PCR (Chiu et al., 2004) using pDEST14 SpyCatcher or pDEST14 SpyCatcher EQ (Zakeri

et al., 2012) and primers 5'-GATTACGACATCCCAACGACCGAAAACCTG, 5'-GCCTGAACGATATTTTTGAAGCGCAGAAAATTGAATGGCATGAAGGCGATTACG ACATCCCAACGACCGAAAACCTG, 5'-GTGATGGTGATGGTGATGGTAGTACGACATATG and 5'-TGCCATTCAATTTTCTGCGCTTCAAAAATATCGTTCAGGCCGCTGCCGTGATGG TGATGGTGATGGTAGTACGACATATG.

pET28a AP-SpyTag-MBP and pET28a AP-SpyTag DA-MBP were generated by using pET28a SpyTag-MBP or pET28a SpyTag DA-MBP (Zakeri et al., 2012) as a template and primers 5'-ATTACATATGGGTCTGAATGATATTTTCGAAGCGCAGAAAATTGAATGGCATGAA GGTAGCGGAGCCCACATCGTGATGGTG and 5'-GGGGAAGCTTTTACGAGCTCGAATTAGTCTG. The insert was digested with HindIII and NdeI and ligated into pET28a.

pET28a AKTE SpyTag-MBP (Figure 5.10) was generated using Inverse PCR (Kaluz et al., 1999). pET28a SpyTag-MBP was used as a template and primers 5'-CTCCGTCTTGGCGTCCACCATCACGATG and 5'-GGTAGTGGTGAAAGTGGTAAAATCG were used. The PCR product was digested with DpnI and then re-ligated using T4 Polynucleotide Kinase and T4 DNA Ligase.

pET28a TKTE SpyTag-MBP (Figure 5.10) was generated using Inverse PCR (Kaluz et al., 1999). pET28a SpyTag-MBP was used as a template and primers 5'-GGTAGTGGTGAAAGTGGTAAAATCG and 5'-CTCCGTCTTCGTGTCCACCATCACGATGTGGG were used. The PCR product was digested with DpnI and then re-ligated using T4 Polynucleotide Kinase and T4 DNA Ligase.

pET28a P3C1-MBP, pET28a P4C8-MBP and pET28a P4C9-MBP (Figure 5.10) were generated using CPEC (Quan and Tian, 2009). The vector was amplified using pET28a SpyTag-MBP as a template and primers 5'-GAGGTTCGGCCGAAGGTAGTGGTGAAAGTGGTAAAATC and 5'-

AGGTACCACTAAAGGTCCCATATGGCTGCCGCG. The inserts were amplified from M13KE P3C1, M13KE P4C8 and M13KE P4C9 phage genomic DNA using primers 5'-GGCAGCCATATGGGACCTTTAGTGGTACCTTTCTATTCTCAC and 5'-CTTTCACCACTACCTTCGGCCGAACCTCCACC.

pET28a PhD12C SpyTag-MBP (Figure 5.10) was generated via SLIM PCR (Chiu et al., 2004) using pET28a P3C1-MBP as a template and primers 5'-GGTGGAGGTTCGGCCGAAGG, 5'-ATCGTGATGGTGGACGCCTACAAGCCGACGAAGGGTGGAGGTTCGGCCGAAGG, 5'-AGAGTGAGAATAGAAAGGTACCACTAAAGG and 5'-CGGCTTGTAGGCGTCCACCATCACGATGTGGGCAGAGTGAGAATAGAAAGGTACCACTAAAGG.

pET28a SallyTag-MBP was generated via QuikChange using pET28a P3C1-MBP as a template and primers 5'-CTCTAGTTATCAGATTAGTATGGTTGATACGAAGACGGAGGGTGGAGGTTC and 5'-GAACCTCCACCCTCCGTCTTCGTATCAACCATACTAATCTGATAACTAGAG.

pET28a P4C8 ND-MBP was generated via QuikChange using pET28a P4C8-MBP as a template and primers 5'-GTTGGTTTTGAGCTGACTATGATTGATAAGCAGAGGGGTGGAGGTTCGG and 5'-CCGAACCTCCACCCCTCTGCTTATCAATCATAGTCAGCTCAAAACCAAC.

To generate a phage that could display SpyTag (DA) monovalently on the surface, SpyTag (DA) was inserted into the pFab5cHis phagemid (Engberg et al., 2000) by Prof. Mark R. Howarth. To generate pFab5cHis SpyTag, pFab5cHis was digested with XhoI (NEB) and NotI (NEB). Primers 5'-TCGAGGGCGGCGCCCACATCGTGATGGTGGACGCCTACAAGCCGACGAAGGGCGC and 5'-GGCCGCCTTCGTCGGCTTGTAGGCGTCCACCATCACGATGTGGGCGCCGCCC were annealed and ligated into pFab5cHis. To generate pFab5cHis SpyTag DA, pFab5cHis was digested with XhoI (NEB) and NotI (NEB). Primers 5'-

TCGAGGGCGGCGCCCACATCGTGATGGTGGCCGCCTACAAGCCGACGAAGGG CGC and 5'-GGCCGCCTTCGTCGGCTTGTAGGCGGCCACCATCACGATGTGGGCGCCGCCC were annealed and ligated into pFab5cHis.

To generate a phage that could display 5 copies of SpyTag on the surface, SpyTag was inserted into the M13KE (NEB) vector by using SLIM PCR (Chiu et al., 2004). M13KE P3C1 was used as a template and primers 5'-AGAGTGAGAATAGAAAGGTACCACTAAAGG, 5'-TCGGCTTGTAGGCGTCCACCATCACGATGTGGGCAGAGTGAGAATAGAAAGGT ACCACTAAAGG, 5'-GGTGGAGGTTCGGCCGAAAC and 5'-CATCGTGATGGTGGACGCCTACAAGCCGACGAAGGGTGGAGGTTCGGCCGAA AC were used.

To generate a phage that could display SpyCatcher monovalently on the surface, SpyCatcher was inserted into the pFab5cHis phagemid (Engberg et al., 2000). SpyCatcher was amplified from pDEST14 SpyCatcher using primers 5'-taatCTCGAGATCAGGGCGCCATGGTTGATACCTTATC and 5'-atatGCGGCCGCTCCACTCCCCTGGAAGTAGAGGTTTTC. The insert and vector were digested using XhoI and NotI and then ligated.

2.2 Protein expression and purification

2.2.1 Protein expression in E. coli

pET28a constructs were expressed in *E. coli* BL21 DE3 RIPL, with the exception of Pilin-C-BLA-Isopeptag and Pilin-C KA-BLA-Isopeptag which were expressed in *E. coli* B834 DE3 (Novagen). pDEST14 constructs were expressed in *E. coli* BL21 DE3 pLysS (Stratagene). pET9d constructs were expressed in *E. coli* B834 DE3. Colonies were picked and grown overnight in LB containing 0.8 % glucose (v/v) and the appropriate antibiotic (50 µg/ml kanamycin for pET28a and pET9d; 100 µg/ml

ampicillin for pDEST14). Cultures were diluted 1:100 into 750 ml of LB containing 0.8 % glucose (v/v) and the appropriate antibiotic and grown at 37 °C, 200 RPM to an OD_{600} of ~0.5 and then induced with 0.4 mM IPTG. Typically, cultures were incubated at 30 °C, 200 RPM for 4 h. Cyclized enzyme constructs and controls (including non-tagged enzyme controls such as pET28a BLA) were grown at 18 °C, 200 RPM overnight. TEV protease was expressed at 22 °C, 200 RPM overnight. For MS samples SpyTag (DA)-BLA-SpyCatcher (EQ), SnoopTag (KA)-BLA-SnoopCatcher, Pilin-C (KA)-BLA-Isopeptag, SallyTag (DA)-BLA-SpyCatcher, SpyTag (DA)-PhyC-SpyCatcher overnight cultures were grown in LB without glucose and grown at 30 °C, 200 RPM for 3 h after induction to avoid gluconylation (Geoghegan et al., 1999). Cells were spun at 4,000g for 20 min and the pellets were harvested and stored at -80 °C.

2.2.2 Protein expression in P. pastoris

2-5 µg of pPICLIT constructs were linearized using SacI (NEB) and PCR purified using a QIAquick PCR Purification Kit. DNA was electroporated into electrocompetent *P. pastoris* X-33 cells (Life Technologies). Cells were rescued using a 1:1 solution of yeast extract peptone dextrose (YPD; 10 g/l yeast extract, 20 g/l bacteriological peptone, 20 g/l glucose) and 2 M sorbitol, allowed to recover for 2 h at 30 °C and plated on YPD plates containing 50 µg/ml zeocin. Plates were incubated at 30 °C for 2-3 days. Colonies were inoculated into YPD and grown at 30 °C, 200 RPM for 24 h. Cultures were diluted 1:33 into buffered methanol minimal (BMMH; 0.004 % histidine (w/v), 1.34 % yeast nitrogen base (w/v), 4 x 10^{-5} % biotin (w/v), 0.5 % methanol (v/v), 100 mM potassium phosphate pH 6.0) medium or buffered methanol complex (BMMY; 1 % yeast extract (w/v), 2 % peptone (w/v), 1.34 % yeast nitrogen base (w/v), 4 x 10^{-5} % biotin (w/v), 0.5 % methanol (v/v), 100 mM potassium phosphate pH 6.0) medium and grown at 30 °C, 200 RPM. Every 24 h a sample was removed for analysis by SDS-PAGE and additional methanol was added so that the final concentration of methanol in the medium was maintained at ~ 1 % (v/v). The cells

were harvested after 48 h and spun at 4,000g for 20 min. The pellet was discarded and the supernatant was stored at -80 °C.

2.2.3 Ni-NTA Affinity Chromatography

His-tagged proteins were purified using standard Ni-NTA purification procedures taking caution to maintain the samples at 4 °C throughout the procedure. Cell pellets obtained from a 750 ml culture were resuspended in 8 ml NiNTA buffer (50 mM Tris-HCl pH 7.8 containing 300 mM NaCl) containing 10 mM phenylmethylsulfonylfluoride (PMSF, Roche) and 1× EDTA-free mixed protease inhibitors (Roche, 1 tablet Resuspended in 1.5 ml double deionised H_2O was 100× stock). Cells were sonicated 3 times in a Sonics Vibra-Cell sonicator (Ultrasonic Processor) for 60 s at 50 % duty cycle, micro-tip limit set to '4' with the 'pulse' setting 'on'. Samples were spun at 17,000g, 4 °C for 20 min. Supernatant was collected and loaded onto Ni-NTA resin (Qiagen) in a 15 ml tube which had been washed with water and equilibrated with Ni-NTA buffer. The resin was incubated at 4 °C for 30 min on a rotor to keep the resin suspended. The resin was then pelleted at 4,000g for 10 s and the supernatant was discarded. The resin was washed twice with 10 column volumes of Ni-NTA buffer containing 10 mM imidazole and then loaded onto a Polyprep column (Bio-Rad). The resin was then washed with 5 column volumes of Ni-NTA buffer containing 30 mM imidazole. Typically, proteins were eluted using 3 column volumes of Ni-NTA buffer containing 100 mM imidazole and collecting 0.5 ml fractions. For cyclized enzymes and controls elution was done using Ni-NTA buffer containing 75 mM imidazole to enrich for monomeric species (Howarth and Ting, 2008). Protein purity in each fraction was assessed using SDS-PAGE.

High purity fractions were selected for dialysis. Typically proteins were dialyzed using dialysis tubing with a 3.5 kDa cut-off (Fisher Scientific). 1 ml of sample was dialyzed into 1000 ml of buffer 3 times. The dialysis buffer was replaced with fresh buffer every 4 hours. Typically proteins were dialyzed into PBS (140 mM NaCl, 2.7

mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4 pH 7.4). PhyC constructs used for DSC, aggregation studies and activity studies were dialyzed into 50 mM Tris-HCl containing 2 mM $CaCl_2$. TEV protease was not dialyzed at all. PhyC constructs used for CD and DLS were dialyzed into 10 mM Tris-HCl pH 8.0 containing 2 mM $CaCl_2$ to be further purified using anion-exchange. Samples used for MS were dialyzed into 10 mM ammonium acetate.

2.2.4 Anion exchange chromatography

For purification of GOx and SpyTag-GOx-SpyCatcher, samples were dialyzed into 10 mM Tris-HCl pH 8.0 three times. 0.6 ml packed resin volume of Q Sepharose High Performance resin (GE Healthcare) was washed with 10 column volumes of water and then equilibrated with 10 column volumes of 10 mM Tris-HCl pH 8.0. Resin and protein sample was mixed in a 50 ml tube and left on a rotor at 4 °C for 30 min to keep the resin suspended. The sample was loaded onto a Polyprep column and washed with 10 column volumes of 10 mM Tris-HCl pH 8.0. 5 column volumes of 10 mM Tris-HCl pH 8.0 containing 50, 100, 150, 200, 250, 300 or 700 mM NaCl was added sequentially, slowly increasing the concentration of NaCl. 1 ml fractions were collected and analysed by SDS-PAGE. Samples containing protein were pooled and concentrated using a Vivaspin 6 centrifugal concentrator with a 5 kDa cut-off (Sartorius). Protein samples were then dialyzed into PBS 3 times.

SpyTag (DA)-PhyC-SpyCatcher and PhyC for CD and DLS analysis were further polished using anion exchange. Anion-exchange chromatography was performed on an ÄKTA purifier (GE Healthcare) using a 1 ml HiTrap HP Q column (GE Healthcare). Sample buffer was 10 mM Tris-HCl pH 8.0 containing 2 mM $CaCl_2$ and a linear gradient was applied using 10 mM Tris-HCl pH 8.0 containing 2 mM $CaCl_2$ and 1 M NaCl at a flow rate of 4 ml/min. OD_{280} was used to select the fractions to analyze using SDS-PAGE. Samples containing high purity monomeric protein were pooled and concentrated using a Vivaspin 6 centrifugal concentrator with a 5 kDa cut-off. Samples

to be analyzed by DLS were dialyzed 3 times into 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$. Samples analyzed by CD were dialyzed 3 times into 10 mM Tris-SO_4 (Tris base with pH adjusted using H_2SO_4) pH 7.0 containing 2 mM $CaSO_4$.

2.2.5 Gel filtration

Samples used to determine kinetic parameters or specific activity were purified with gel filtration after the Ni-NTA Affinity Chromatography purification step, to remove polymeric species in the cyclized constructs. Gel filtration was performed on an ÄKTA purifier using a Superdex 200 10/300 GL column (GE Healthcare) using a flow rate of 0.4 ml/min. For BLA constructs, PBS was used as the running buffer. For PhyC constructs, 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$ was used as the running buffer. OD_{280} was used to select the fractions to analyze using SDS-PAGE. Samples containing high purity monomeric protein were pooled and concentrated using a Vivaspin 6 centrifugal concentrator with a 5 kDa cut-off.

2.2.6 Determination of protein concentration

Protein concentration was determined using a Bicinchoninic Acid (BCA) assay kit (Thermo Scientific) using the manufacturer's instructions or by measuring the OD_{280} using a ND-1000 NanoDrop and calculating the molar extinction coefficient using ProtParam (Gasteiger et al., 2005).

2.3 Polyacrylamide Gel Electrophoresis

2.3.1 SDS-PAGE

SDS-PAGE was run in an XCell SureLock module (Life Technologies) at 200 V for 1 h and then stained with InstantBlue Coomassie (Expedeon). For BLA, FLuc, DHFR and GOx constructs, a final concentration of 100 mM dithiothreitol (DTT, Sigma) was added with the exception of the GOx aggregation experiments. For phytase

constructs, a final concentration of 100 mM ethylenediaminetetraacetic acid (EDTA) was added to remove tightly-bound divalent cations affecting phytase mobility. Samples were mixed with 6× SDS-PAGE loading buffer (0.23 M Tris-HCl, 0.24% glycerol, 6.7% SDS, and 12 mM bromophenol blue) and heated at 95 °C for 7 min in a C1000 Thermal cycler. Gels were imaged on a ChemiDoc XRS+ and analyzed using Image Lab 3.0 software (Bio-Rad).

2.3.2 Temperature-dependent solubility assay

20 µL 25 µM BLA construct in PBS containing 100 mM DTT was kept at 10, 25, 37, 55, 75, 90 or 100 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler (ramp rate was 3 °C/s). 20 µL 20 µM DHFR construct in PBS containing 100 mM DTT was kept at 10, 25, 37, 55, 75, 90 or 100 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler. 20 µL 5 µM FLuc construct in PBS containing 100 mM DTT was kept at 10, 25, 37, 55, 75 or 90 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler. 20 µL 5 µM GOx construct in PBS was kept at 10, 25, 37, 55, 75, 90 or 100 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler. 20 µL 10 µM PhyC construct in 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$ was kept at 10, 25, 55, 75, 90 or 100 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler. Insoluble aggregates were separated by centrifugation at 17,000g at 4 °C for 30 min and the supernatant was analysed by SDS-PAGE. The samples stored at 10 °C were used as loading controls and was defined as 100 % soluble fraction. Image Lab 3.0 software was used to quantify the bands via densitometry.

2.3.3 TEV protease cleavage of SpyTag (DA)-BLA-TEV-SpyCatcher

10 µM of SpyTag-BLA-TEV-SpyCatcher or SpyTag DA-BLA-TEV-SpyCatcher was incubated with 10 µM TEV protease in 50 mM Tris-HCl pH 8.0 containing 0.5 mM EDTA and 1 mM DTT for 2 h at 34 °C. The reaction was stopped through the addition

of DTT to a final concentration of 100 mM and 6× SDS-PAGE loading buffer. Samples were heated for 7 min at 95 °C on a C1000 Thermal Cycler.

2.4 Enzyme activity

2.4.1 BLA thermal resilience enzymatic assay

A 96 well clear flat-bottom polystyrene plate (Greiner) was pre-blocked to avoid adsorption of the enzyme to the plastic with 200 µL PBS containing 3% bovine serum albumin (BSA, ≥ 98%, Sigma) at 37 °C for 2 h. The pre-blocked plate was washed two times using 0.1 M NaH_2PO_4 pH 7.0 containing 1 mM EDTA. 20 µL 25 µM BLA construct in PBS containing 100 mM DTT was kept at 10, 25, 37, 55, 75, 90 or 100 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler (ramp rate was 3 °C/s). The samples were diluted 1 in 400 using 0.1 M NaH_2PO_4 pH 7.0 containing 1 mM EDTA and 1% BSA. A final concentration of 3 nM of BLA construct was mixed with 100 µM nitrocefin (Merck) in 0.1 M NaH_2PO_4 pH 7.0 containing 1 mM EDTA at 20 °C. The reaction was stopped at different time-points through the addition of a suicide inhibitor, potassium clavulanate (Hugonnet and Blanchard, 2007). An equal volume of 0.1 M NaH_2PO_4 pH 7.0 containing 1 mM EDTA and 500 µM potassium clavulanate (Sigma). Plates were protected from light to avoid loss of signal. Half of the volume was reverse pipetted into a new 96 well clear flat-bottom polystyrene plate to avoid forming bubbles that could interfere with the light pathway in the plate reader. OD_{486} was measured using a SpectraMax M3 microplate reader (Molecular Devices). Absorbance values were multiplied by two to take into account the dilution caused by the addition of the quenching reagent (potassium clavulanate). Blanks consisted of all the components bar the enzyme.

2.4.2 BLA kinetics assay

2.0 ml sample cups (Kartell Labware) were pre-blocked to avoid adsorption of the enzyme to the plastic with 50 mM NaH_2PO_4 pH 7.0 containing 3% BSA at 37 °C for 2 h. The pre-blocked sample cups were washed two times with 50 mM NaH_2PO_4 pH 7.0. Enzymes were diluted with 50 mM NaH_2PO_4 pH 7.0 containing 0.1% BSA to a concentration of 0.25 nM. The diluted enzyme was then mixed with 50 mM NaH_2PO_4 pH 7.0 containing 100, 50, 30, 20, 15, or 10 µM Nitrocefin (Oxoid) in a Gallery Plus Automated Photometric Analyzer (Thermo Scientific) at 30 °C. The Photometric analyser was set to take a measurement every 9 s at 30 °C, 500 nm ($\Delta\varepsilon_{500}$ = 15,900 $M^{-1}cm^{-1}$) (Chen and Herzberg, 1999). The initial rate of reaction was determined using the linear portion of the curve. Kinetic parameters were extrapolated from the data by using the Excel 2013 (Microsoft) Solver add-in, a non-linear least squares curve fitting program. The K_M and Vmax values were calculated for each individual run and an average and standard deviation were calculated from the three values.

2.4.3 DHFR thermal resilience enzymatic assay

A 96 well clear flat-bottom polystyrene plate (Greiner) was pre-blocked to avoid adsorption of the enzyme to the plastic with 200 µL PBS containing 3% BSA at 37 °C for 2 h. The pre-blocked plate was washed two times using 1× assay buffer (Dihydrofolate Reductase Assay Kit, Sigma). 20 µL 20 µM DHFR construct in PBS containing 100 mM DTT was kept at 10, 25, 37, 55, 75, 90 or 100 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler (ramp rate was 3 °C/s). 1× assay buffer was used to dilute the protein to a concentration of 1 µM in the pre-blocked plate. A final concentration of 100 nM of DHFR construct was mixed with 100 µM dihydrofolic acid (DHF, Sigma) and 100 µM β-nicotinamide adenine dinucleotide phosphate reduced (NADPH, Sigma) in 1× assay buffer at 25 °C. The reaction was monitored by measuring the changes in OD_{340} in a SpectraMax M3 microplate reader.

Samples were blanked against a control containing 1× assay buffer and 100 nM DHFR construct.

2.4.4 FLuc thermal resilience enzymatic assay

A 96 well clear flat-bottom polystyrene plate was pre-blocked, to avoid adsorption of the enzyme to the plastic, with 200 µL PBS containing 3% bovine serum albumin BSA at 37 °C for 2 h. The pre-blocked plate was washed two times using 50 mM Tris-acetate pH 7.8 containing 10 mM $MgSO_4$ and 2 mM EDTA. 20 µL 5 µM FLuc construct in PBS containing 100 mM DTT was kept at 25, 37, 55, 75 or 90 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler (ramp rate was 3 °C/s). The protein was diluted 1 in 40 using 50 mM Tris-acetate pH 7.8 containing 10 mM $MgSO_4$ and 2 mM EDTA. 12.5 nM of FLuc construct was allowed to react with 4 mM adenosine triphosphate (ATP, Sigma) and 300 µM luciferin (Life Technologies) in 50 mM Tris-acetate pH 7.8 containing 10 mM $MgSO_4$, 2 mM EDTA and 2 mM DTT at 25 °C in a 96-well black solid plate (Corning). After mixing the enzyme with the substrate there was a 15 s delay before the luminescence was measured on a Pherastar FS (BMG Labtech) plate-reader with the gain set to 3,600. Activity was defined as the maximum initial luminescence value obtained. 100 % activity was defined as the activity obtained after incubating the FLuc construct at 25 °C. Samples were blanked against a control lacking enzyme.

2.4.5 PhyC thermal resilience enzymatic assay

A 96 well clear flat-bottom polystyrene plate was pre-blocked with 200 µl 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$ and 3% BSA at 37 °C for 2 h to avoid adsorption of the enzyme to the plastic. The pre-blocked plate was washed two times using 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$. 20 µl of 10 µM PhyC construct in 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$ was kept at 10, 55, 75, 90 or 100 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler (ramp rate was 3 °C/s).

The enzyme was diluted to 385 nM using 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$, 10% glycerol and 1% BSA in the preblocked plate. The final reaction concentrations were 77 nM PhyC construct reacting with 2 mM phytic acid (Sigma) in 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$ and 10% glycerol at 50 °C. The reaction was stopped at different time points through the addition of an equal volume of 20% trichloroacetic acid (TCA). The colour mix solution was prepared fresh daily and contained 10 mM ammonium heptamolybdate (Sigma), 3.4% H_2SO_4 and 36 mM $FeSO_4$ (Scientific Laboratory Supplies). The colour mix solution was added 1:2 to the quenched reaction solution. Plates were centrifuged to remove precipitated BSA and colour pigments at 3,500 g at 20 °C in a Multifuge X3R (Thermo Scientific) for 10 min. 60% of the supernatant (90 µl of 150 µl) was pipetted into a new 96 well clear flat-bottom polystyrene plate and a OD_{700} reading was taken within 10 min using a SpectraMax M3 microplate reader. Blanks consisted of all the components bar the enzyme.

2.4.6 PhyC specific activity

40 nM PhyC construct reacted with 1.6 mM phytic acid at 37 °C in 100 mM Tris-HCl pH 7.5 containing 1 mM $CaCl_2$. The reaction was stopped after 1, 2, 3, 4 or 5 min through the addition of an equal volume of 20% trichloroacetic acid. Centrifugation was performed to remove any precipitated protein at 17,000g, 20 °C for 10 min. The supernatant was mixed at a 2:1 ratio with a solution containing 10 mM ammonium heptamolybdate, 3.4% H_2SO_4 and 36 mM $FeSO_4$ (colour mix solution, prepared fresh daily). The samples were kept at 25 °C for 20 min to develop the colour. OD_{700} was measured using a Varian Cary 50 UV-Vis Spectrophotometer (Agilent). The blank contained all the components of the sample but the components were mixed in a different order. The enzyme was mixed with the 20 % TCA solution before mixing in the phytic acid. Mass of phosphate released was obtained using a standard curve where 20 to 100 µM NaH_2PO_4 was mixed with TCA and then the colour mix solution

and an OD_{700} value was obtained. The initial rate was calculated by fitting a linear line of best fit. Each time-point was done in triplicate. Three separate initial rates were calculated and an average and standard deviation value was obtained.

2.4.7 GOx thermal resilience enzymatic assay

A 96 well clear flat-bottom polystyrene plate (Greiner) was pre-blocked to avoid adsorption of the enzyme to the plastic with 200 µL PBS containing 3% BSA at 37 °C for 2 h. The pre-blocked plate was washed two times using 50 mM sodium acetate pH 5.1. 20 µL 5 µM GOx construct in PBS was kept at 10, 25, 37, 55, 75, 90 or 100 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler (ramp rate was 3 °C/s). 50 mM sodium acetate pH 5.1 was used to dilute the protein to a concentration of 50 nM in the pre-blocked plate. A final concentration of 2.5 nM of GOx construct was mixed with 160 µM o-dianisidine hydrochloride (Sigma), 90 mM glucose and 2 U/ml horseradish peroxidase (Sigma) in 50 mM sodium acetate pH 5.1. The reaction was monitored by measuring the changes in OD_{485} in a SpectraFluor Plus microplate reader (Tecan) every 20 s at 35 °C. Initial rates could be calculated using the linear portion of the enzyme plots. The initial activity after incubating the enzyme at 25 °C was defined as 100 % activity. Blanks consisted of all the components bar the GOx construct.

2.4.8 Effect of additives to GOx thermal resilience

A 96 well clear flat-bottom polystyrene plate (Greiner) was pre-blocked to avoid adsorption of the enzyme to the plastic with 200 µL PBS containing 3% BSA at 37 °C for 2 h. The pre-blocked plate was washed two times using 50 mM sodium acetate pH 5.1. 20 µL 5 µM GOx construct was incubated at 25, 55 or 75 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler (ramp rate was 3 °C/s) in either PBS, PBS containing 1 mM flavin adenine nucleotide (FAD, Sigma) and 0.2 M K_2SO_4 (Sigma), PBS containing 0.6 M trehalose (Fisher Scientific) or PBS containing 4 M xylitol

(Sigma). 50 mM sodium acetate pH 5.1 was used to dilute the protein to a concentration of 50 nM in the pre-blocked plate. A final concentration of 2.5 nM of GOx construct was mixed with 160 µM o-dianisidine hydrochloride (Sigma), 90 mM glucose and 2 U/ml horseradish peroxidase (Sigma) in 50 mM sodium acetate pH 5.1. The reaction was monitored by measuring the changes in OD_{485} in a SpectraFluor Plus microplate reader (Tecan) every 20 s at 35 °C. Blanks consisted of all the components bar the GOx construct. Initial rates could be calculated using the linear portion of the enzyme plots. The initial activity after incubating the enzyme at 25 °C was defined as 100 % activity.

2.5 Biophysical techniques

2.5.1 Electrospray Ionization Mass spectrometry (ESI-MS)

A Micromass LCT time-of-flight (TOF) ESI-MS (Micromass UK) was used to obtain an m/z (mass/charge) spectrum for each sample. The m/z spectrum was converted to a mass spectrum using a Maximum Entropy algorithm on the MassLynx V 4.00.00 (Water Corporation) software. Predicted masses were obtained by analysing the amino acid sequences on ProtParam (Gasteiger et al., 2005) taking into account the loss of water or ammonia mass upon the formation of an isopeptide bond (Izoré et al., 2010; Kang et al., 2007; Zakeri et al., 2012).

2.5.2 Circular Dichroism (CD)

A J-815 Circular Dichroism Spectrometer (Jasco) was used to determine the ellipticity of the samples. Samples were filtered through a Millex-GV PVDF 0.22 µm syringe filter unit (Merck) and then analysed in a 1 mm path-length Spectrosil quartz cuvette (Starna Scientific) sealed with a polytetrafluoroethylene (PTFE) stopper. Sample concentration was 0.15 mg/ml in 10 mM Tris-SO_4 containing 2 mM $CaSO_4$. Blanks were obtained by just scanning the buffer (passed through a 0.22 µm filter) in

the absence of the protein. The sample temperature was regulated using a CDF-426S Peltier unit (Jasco). Samples were heated to 25, 55, 75 or 90 °C, incubated at the temperature for 3 min and then the Far-UV spectra was determined at the aforementioned temperatures (the “During heating” spectra). Samples were then cooled to 10 °C, incubated at the temperature for 3 min and then the Far-UV spectra was determined at 10 °C (the “After cooling” spectra). Far-UV spectra were determined by scanning from 260 – 185 nm at 50 nm/min. Four scans were obtained for each independent sample and averaged. Four scans were obtained for the blank and averaged. The averaged blank spectrum was subtracted from the averaged sample spectrum. The CD spectra were smoothened using the Savitzky-Golay method (Savitzky and Golay, 1964) in the SpectraManager 2 software (Jasco). Ellipticity values were converted to molar ellipticity values using the SpectraManager 2 software.

2.5.3 Differential Scanning Calorimetry (DSC)

PhyC samples for DSC analysis were dissolved in 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$. All other samples for DSC analysis were dissolved in PBS. All samples contained 20 µM protein. DSC profiles were measured on a VP Cap DSC (GE Healthcare). Scans were performed from 20-110 °C at 1 °C/min and 3 atm of pressure. DSC profiles were normalized for volume and protein concentration, blank traces (buffer in the absence of protein) were subtracted from the sample traces and the baseline was subtracted using the MicroCal DSC Origin Pro 7.0 software (GE Healthcare). The reversibility of denaturation of the BLA domains in the SpyTag (DA)-BLA-SpyCatcher and BLA samples was determined by scanning from 20-60 °C at 2 °C/min and 3 atm of pressure, cooling the samples to 20 °C and then repeating the scan from 20-60 °C to obtain a second DSC trace.

2.5.4 Dynamic Light Scattering (DLS)

50 µl of 10 µM PhyC construct in 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$ was kept at 25, 55, 75, 90 or 100 °C for 10 min and then cooled to 10 °C on a C1000 Thermal Cycler (ramp rate was 3 °C/s). The sample was filtered using a Millex-GV PVDF 0.22 µm syringe filter unit. Samples were analysed using a Viscotek 802 DLS instrument (Malvern). Measurements were taken at 20 °C. Data was then analysed using OmniSize 3.0 software (Malvern).

2.6 Phage display

2.6.1 Conjugating SpyCatcher on magnetic beads

Dynabeads M-270 Amine (Invitrogen) magnetic beads were coated with Cys-SpyCatcher (EQ) using sulfosuccinimidyl 6-[3'-(2-pyridyldithio)propionamido]hexanoate (Sulfo-LC-SPDP, Thermo Scientific). 100 µl of these Dynabeads were washed two times with PBS containing 0.1 M $NaHCO_3$ pH 8.3. A MagRack 6 (GE Healthcare) magnetic rack was used to capture the beads. The beads were then resuspended in 890 µl PBS containing 0.1 M $NaHCO_3$ pH 8.3. 110 µl 2 mg/ml sulfo-LC-SPDP was added. The beads were incubated on a shaker at 25 °C and 1,400 RPM for 1 h. The beads were then washed once more with PBS containing 0.1 M $NaHCO_3$ pH 8.3 and then resuspended in 70 µl PBS. 30 µl 200 µM SpyCatcher (EQ) was added to the beads and incubated on a shaker at 25 °C and 1,400 RPM for 16 h. The reaction was monitored by reading the absorbance of the supernatant at a wavelength of 363 nm to determine the pyridine-2-thione concentration. The beads were then washed 4 times with PBS containing 0.5% BSA and 0.1% Triton X-100 and then resuspended in 200 µl of PBS containing 3% BSA and+ 0.05% sodium azide. The beads were incubated at 25 °C and 1,400 RPM for 1 h and then stored at 4 °C.

2.6.2 Producing pFab5cHis-derived phagemid particles

The phagemids were transformed into *E. coli* ER2738 (NEB). An overnight culture was grown in LB containing ampicillin (100 µg/ml) and tetracycline (10 µg/ml) (LB-AMP/TET) medium. 100 µl of overnight culture was used to inoculate 200 ml of pre-warmed LB-AMP/TET medium. The culture was grown at 37 °C until it reached an OD_{600} of 0.5. $1x10^{12}$ R408 helper phages (Stratagene) were added to the culture and then the culture was shaken gently for 20 min at 37 °C. Cultures were induced with 0.4 mM IPTG and grown at 25 °C for 10-16 h. Phage was harvested by pelleting the cells at 15,000 g at 4 °C for 10 min. The supernatant was moved to a clean container and phage precipitation buffer [20% (w/v) polyethylene glycol-8000, 2.5 M NaCl] was added in a supernatant:buffer ratio (v/v) of 4:1. The solution was incubated on ice for 1 h before pelleting the phage particles. The supernatant was removed and the pellet was resuspended in PBS. A clearing spin at 15,000g and 4 °C was done and phages were then precipitated again as above. A second clearing spin was done and then the phages were precipitated and resuspended in PBS containing 50% glycerol and stored at -80 °C.

2.6.3 Biotinylation of AP-tagged proteins

Purified AP-SpyCatcher (EQ) or AP-SpyTag (DA)-MBP was biotinylated in PBS containing 5 mM $MgCl_2$ and 1 mM ATP. 1 µM GST-BirA was added (expressed and purified by Dr. Michael Fairhead, Howarth laboratory) for every 10 µM of AP-tagged protein. 50 µM of biotin was added for every 1 µM of GST-BirA and 1 µM GST-BirA. The reaction was incubated for 1 h at room temperature. After 1 h incubation, the same amount GST-BirA was added to the reaction again and the reaction was incubated for 1 h at room temperature. GST-BirA was removed by mixing the reaction with Hi-Cap Glutathione matrix (Qiagen) for 1 h at 4 °C. The supernatant was collected and the matrix was discarded. To test the effectiveness of the biotinylation, a gel-shift assay was performed by allowing the biotinylated AP-tagged protein to bind with Streptavidin and analysing the samples on a SDS-PAGE (Fairhead and Howarth,

2015). Biotinylated AP-SpyCatcher was incubated with 1, 5, 10, 20, 30 and 40 µM of Streptavidin for 1 h at room temperature. 6x SDS loading buffer was added and the samples were loaded onto a 14% SDS-PAGE gel and run on an X-Cell SureLock at 200 V for 1 h on ice. The gel was stained with InstantBlue (Expedeon) Coomassie stain.

2.6.4 α-pIII and α-His Western Blots

10 µl ~10^{11} cfu/ml of phage was added to 5 µl 432 µM SpyCatcher or 5 µl 238 µM SpyTag-MBP in PBS. The reaction was incubated at room temperature for 30 min. 6x SDS loading buffer was added and the samples were boiled at 95 °C for 7 min in a C1000 Thermal Cycler. Samples were loaded onto an 8 % (SpyCatcher phagemid) or 12% (SpyTag phagemid) SDS-PAGE and run at 200 V for 1 h on a XCell SureLock (Invitrogen). The gels were then transferred to a polyvinylidene fluoride (PVDF, Merck) membrane in Transfer buffer (12 mM Tris-base, 96 mM glycine, 20% methanol) for 2 h at 233 mA in an XCell SureLock. Membranes were blocked overnight at 4 °C using 40 ml of PBS containing 12.5% fat-free milk. The membranes were washed 4 times using PBS containing 0.05% Tween-20 (PBST). The membrane was incubated with either with 1 µg/ml of anti-M13 pIII antibody (NEB) or 0.33 µg/ml of anti-Penta-His monoclonal antibody (Qiagen) in PBST containing 2.5% fat-free milk for 1 h at room temperature. Membranes were washed 4 times with PBST and incubated with a 0.4 µg/ml solution of anti-mouse IgG-HRP secondary antibody (Sigma) in PBST for 40 min on an orbital shaker at room temperature. The membranes were washed 4 times with PBST on an orbital shaker and then incubated with 1 ml of SuperSignal West Pico (Thermo Pierce) for 5 min before developing the blot on a Chemidoc XRS+ at 30 s intervals for 10 min.

2.6.5 Streptavidin-HRP Western Blot

10 µl ~10^{11} cfu/ml of phage was added to 5 µl 10 µM biotinylated AP-SpyCatcher or 1 µl 8 µM biotinylated AP-SpyTag (DA)-MBP. The reaction was incubated at room temperature for 30 min. Samples were loaded onto an 8% SDS-PAGE and transferred to a PVDF membrane as above. Western blot was performed as previously described, but 1 µg/ml Streptavidin-HRP (Thermo Scientific) was used to detect Biotinylated AP-SpyCatcher. SpyTag-displaying phage blots were allowed to develop onto an Amersham Hyperfilm High performance autoradiography film (GE Healthcare) for 5 s. Films were developed on a Compact X4 (Xograph Imaging System) using Devalex M solutions (Champion Photochemistry). SpyCatcher-displaying phage blots were developed on the Chemidoc XRS+ as described in section 2.6.4.

2.6.6 Phage viability test

~8 x 10^9 cfu/ml SpyTag-displaying phagemids were incubated in PBS or 0.2 M glycine-HCl pH 2.2 buffer for 10 min on a C1000 Thermal Cycler at 25, 40, 55 or 65 °C. The sample was then cooled to 10 °C. Phage was serially diluted in PBS. A culture of ER2738 cells was grown to an OD_{600} of ~0.5. 200 µl of ER2738 cells were mixed with 10 µl of diluted phage. Mixture was incubated at 37 °C for 15 min and then plated onto LB-agar plates containing ampicillins (100 µg/ml). Plates containing <100 colonies were counted and the phage titre was calculated.

2.6.7 Panning of the PhD-12 library (Rounds 1-5)

Panning of the PhD.-12 library (NEB) was done by adding 1-2 x 10^{11} phages to 10 µl coated SpyCatcher beads and letting the phages bind to the beads. The beads were then washed under increasingly stringent conditions at each round of selection (Table 2.1). Phage titration, library amplification, plaque amplification and DNA isolation were done according to the Ph.D. Phage Display Libraries (NEB) manufacturer's protocol.

Panning round	Protocol followed
1	1. Wash beads 3 times with 100 µl TBS + 0.1% Tween-20. 2. Resuspend in 90 µl of TBS +0.1% Tween-20. Add 10 µl of phage library (~1.5 x 10^{11} phages). 3. Incubate at RT for 18 h at 25 °C and 1,200 RPM. 4. Wash beads 3 times with 100 µl of TBS +0.5% pH 7.5. 5. Wash beads 1 time with TBS +1% Tween-20 pH 7.5. 6. Wash 1 time with PBS +0.9M NaCl pH 7.5. 7. Wash 1 time with 0.2 M Glycine-HCl pH 2.2. 8. Wash 3 times with 100 µl TBS + 1% Tween-20 pH 7.5. 9. Wash once with PBS. 10. Elute with 100 µl 38 µM TEV protease in 50 mM Tris-HCl buffer pH 8.0 + 0.5 mM EDTA at 34 °C for 2 h at 1,200 RPM.
2	1. Wash beads 3 times with 100 µl TBS + 0.1% Tween-20. 2. Resuspend in 99 µl of TBS +0.1% Tween-20. Add 1 µl phage (~10^{11} phages). 3. Incubate at RT for 18 h at 25 °C and 1,200 RPM. 4. Wash beads 3 times with 100 µl of TBS +0.5% pH 7.5. 5. Wash beads 1 time with TBS +1% Tween-20 pH 7.5. 6. Wash 1 time with PBS +0.9M NaCl pH 7.5. 7. Wash 1 time with 0.2 M Glycine-HCl pH 2.2. 8. Wash 3 times with 100 µl TBS + 1% Tween-20 pH 7.5. 9. Wash once with PBS. 10. Elute with 100 µl 38 µM TEV protease in 50 mM Tris-HCl buffer pH 8.0 + 0.5 mM EDTA at 34 °C for 2 h at 1,200 RPM.
3	1. Wash beads 3 times with 100 µl TBS + 0.1% Tween-20. 2. Resuspend in 99 µl of TBS +0.1% Tween-20. Add 1 µl phage (~10^{11} phages). 3. Incubate at RT for 2 h at 25 °C and 1,200 RPM. 4. Wash beads 3 times with 100 µl of TBS +0.5% pH 7.5. 5. Wash beads 1 time with TBS +1% Tween-20 pH 7.5. Incubate for 10 min at 25 °C at 1,200 RPM. 6. Wash 1 time with PBS +0.9 M NaCl pH 7.5. Incubate for 10 min at 25 °C at 1,200 RPM.

	7. Wash 1 time with 0.2 M Glycine-HCl pH 2.2. Incubate for 10 min at 25 °C at 1,200 RPM. 8. Wash 3 times with 100 µl TBS + 1% Tween-20 pH 7.5. 9. Wash once with PBS. 10. Elute with 100 µl 38 µM TEV protease in 50 mM Tris-HCl buffer pH 8.0 + 0.5 mM EDTA at 34 °C for 2 h at 1,200 RPM.
4	1. Wash beads 3 times with 100 µl TBS + 0.1% Tween-20. 2. Resuspend in 99 µl of TBS +0.1% Tween-20. Add 1 µl phage (~10^{11} phages). 3. Incubate at RT for 2 h at 25 °C and 1,200 RPM. 4. Wash beads 3 times with 100 µl of TBS +0.5% pH 7.5. 5. Wash beads 1 time with TBS +1% Tween-20 pH 7.5. Incubate for 10 min at 65 °C at 1,200 RPM. 6. Wash 1 time with PBS +0.9 M NaCl pH 7.5. Incubate for 10 min at 65 °C at 1,200 RPM. 7. Wash 1 time with 0.2 M Glycine-HCl pH 2.2. Incubate for 10 min at 65 °C at 1,200 RPM. 8. Wash 3 times with 100 µl TBS + 1% Tween-20 pH 7.5. 9. Wash once with PBS. 10. Elute with 100 µl 38 µM TEV protease in 50 mM Tris-HCl buffer pH 8.0 + 0.5 mM EDTA at 34 °C for 2 h at 1,200 RPM.
5	1. Wash beads 3 times with 100 µl TBS + 0.1% Tween-20. 2. Resuspend in 99 µl of TBS +0.1% Tween-20. Add 1 µl phage (~10^{11} phages). 3. Incubate at RT for 1 min. 4. Wash beads 3 times with 100 µl of TBS +0.5% pH 7.5. 5. Wash beads 1 time with TBS +1% Tween-20 pH 7.5. 6. Wash 1 time with PBS +0.9 M NaCl pH 7.5. 7. Wash 1 time with 0.2 M Glycine-HCl pH 2.2. 8. Wash 3 times with 100 µl TBS + 1% Tween-20 pH 7.5. 9. Wash once with PBS. 10. Elute with 100 µl 38 µM TEV protease in 50 mM Tris-HCl buffer pH 8.0 + 0.5 mM EDTA at 34 °C for 2 h at 1,200 RPM.

Table 2.1: Conditions used for the first 5 rounds of panning of the PhD-12 library.

2.6.8 Panning of the PhD-12 library (Rounds 6-7)

8 x 10^{11} phages were incubated with 50 nM Biotinylated AP-SpyCatcher in 80 µl volume containing 2.4 mg/ml MCF7 cell line lysate (prepared by Dr. Gianluca Veggiani, Howarth lab) in PBS (to simulate the crowding effect in a cell and to remove any peptides that also bind to proteins in the cell) at room temperature for 2 min. To quench the reaction an excess of SpyCatcher was added to the solution and incubated at room temperature for 30 min. Phage precipitation buffer was added and the sample was incubated on ice for 40 min. The phage was pelleted, supernatant removed and the phage was resuspended in PBS. 25 µl of Dynabeads Biotin Binder (Invitrogen) was washed 3 times with TBS (50 mM Tris-HCl pH 7.5 containing 150 mM NaCl) containing 0.1% Tween-20 (TBST). Beads were blocked with a TBST solution containing 0.1% BSA. Beads were incubated at 1,200 RPM and 25 °C for 1 h. After 1 h, 10 µl phage solution was added and the mixture was incubated at 25 °C, 1,200 RPM for 1 h. Beads were washed once with 100 µl 0.2 M Glycine-HCl pH 2.2 followed by 10 washes with 100 µl PBST containing 0.1% BSA. Phage was eluted with 100 µl 38 µM TEV protease in 50 mM Tris-HCl buffer pH 8.0 containing 0.5 mM EDTA. Sample was incubated at 34 °C for 2 h at 1,200 RPM. Panning round 7 was performed as above, but half of the biotinylated AP-SpyCatcher concentration was used.

2.6.9 qPCR phage kinetics assay

~10^{7} phages were incubated in PBS containing 630 nM biotinylated AP-SpyCatcher and 2.5 % BSA to a final volume of 8 µl at room temperature. The reaction was quenched at different time points through the addition of 10 µl of PBS containing 70 µM SpyCatcher. The reaction was incubated at room temperature for 30 min. The sample was then diluted through the addition of 7 µl of PBS. 6.3 µl of phage precipitation buffer was added and the sample was incubated for 40 min on ice. Samples were spun at 15,000g for 30 min at 4 °C and the supernatant was discarded. The pellet was resuspended in 200 µl of PBS. A 96 well clear flat-bottom polystyrene

plate (Greiner) was pre-blocked to avoid adsorption of the phage to the plastic with 200 µL PBS containing 3% BSA at 37 °C for 2 h. The pre-blocked plate was washed 4 times with PBS containing 0.05 % Tween-20 (PBST). 25 µl of Dynabeads Biotin Binder were pipetted into each well and washed 3 times with TBST using a Magnabot 96 magnetic separation device (Promega) to capture the beads. The beads were then resuspended in 130 µl of the reacted phage solution prepared previously. 20 µl of TBS containing 1 % Tween-20 was added and the plate was incubated at 25 °C, 1,000 RPM for 1 h. The beads were washed once with 0.2 M glycine-HCl pH 2.2 and then 4 times with TBS containing 0.5 % Tween-20. The beads were resuspended in 150 µl of PBS and 50 µl of the resuspended beads was heated at 95 °C for 7 min on a C1000 Thermal Cycler. Beads were captured with a magnet and the supernatant was analyzed via quantitative PCR (qPCR) using primers 5'-GTCTGACCTGCCTCAACCTC and 5'-TCACCGGAACCAGAGCCAC. SensiMix SYBR Hi-ROX Kit (Bioline Reagents) was used in combination with a Rotor-Gene 6000 RT-PCR machine (Corbett Life Science) to quantify the amount of phages in the sample. qPCR cycling conditions consisted of an initial 10 min 95 °C step followed by 45 cycles of 95 °C for 10 s, 60 °C for 15 s and 72 °C for 15 s.

2.6.10 Error-prone PCR SpyCatcher library

SpyCatcher insert was amplified using primers 5'-CCTCGAGATCAGGGCGCCATGG and 5'-GAAGTAGAGGTTTTCGCTACCACTGGATC using pFab5cHis SpyCatcher as a template. An error-prone PCR reaction contained 0.2 mM dATP, dTTP and dCTP, 0.24 mM dGTP, 0.1 U/µl Taq polymerase (NEB), 0.2 µM forward and reverse primers, 50 ng template DNA, 1× ThermoPol reaction buffer (NEB) and 0.6 mM $MnCl_2$ (Sigma). Cycling conditions were initial denaturation step at 95 °C for 5 min followed by 25 cycles of 95 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min and a final extension cycle of 72 °C for 4 min. The vector to make the SpyCatcher library was amplified using the

standard vector amplification protocol using primers 5'-GGATCCAGTGGTAGCGAAAACC and 5'-AACCATGGCGCCCTGATCTCG and pFab5cHis SpyCatcher as a template. 3 µg of vector was combined with 760 ng of insert in a final volume of 300 µl using Gibson assembly (Gibson et al., 2009). The DNA was purified using a QIAquick PCR Purification Kit and eluted to a final volume of 30 µl. 100 ng was transformed into each vial of electrocompetent ER2738 cells (Lucigen). 10 transformations yielded ~10^9 transformants.

2.7 Tag/Catcher reactions

2.7.1 Tag-MBP variants reconstitution reactions

10 µM SpyCatcher was reacted with 10 µM Tag-MBP variants in 40 mM Na_2HPO_4 with 20 mM sodium citrate pH 5.0 at 25 °C. Reactions were quenched after 1 or 10 min through the addition of 6 x SDS loading buffer. The samples were heated to 95 °C for 7 min in a C1000 Thermal Cycler and the samples were loaded onto a 12 % SDS-PAGE. The bands were quantified using the Image Lab 3.0 software.

2.7.2 Reconstitution of SpyCatcher and SnoopCatcher with SpyTag and SnoopTag peptide

SpyCatcher was reacted with 4-fold molar excess of SpyTag peptide (GAHIVMVDAYKPTK, Insight Biotechnology, >95 % purity) in PBS for 2 h and SnoopCatcher was reacted with a 3-fold molar excess of SnoopTag peptide (GKLGDIEFIKVNKGY, Insight Biotechnology, >95 % purity) in PBS for 20 h. Samples were then dialyzed 3 times into PBS (for DSC experiments) or 10 mM Tris-SO_4 containing 2 mM $CaSO_4$ (for CD experiments). The samples were run on a 20 % SDS-PAGE to observe whether there had been a gel shift after the reaction with the peptide.

Chapter 3: Exploring isopeptide-mediated cyclization approaches

Work in this chapter has been published in the manuscripts (Schoene et al., 2014 and Schoene et al., 2016a) provided in the appendix.

Chapter 3: Exploring isopeptide-mediated cyclization approaches

Enzymes are biological catalysts that have several distinct advantages when compared to other catalysts. Enzymes exhibit exquisite regioselectivity, are able to operate under relatively mild conditions and are biodegradable (de Regil and Sandoval, 2013). However, enzymes can be unstable and inactivate at high temperatures (Kintses et al., 2006). Enzymes tend to denature irreversibly at elevated temperatures (Daniel and Danson, 2013).

Enzyme thermostability has been improved via several approaches. Rational design approaches utilize structural and biochemical data to determine which amino acid mutations, deletions and/or additions might lead to an increase in thermostability (Yang et al., 2015). Directed evolution approaches require the use of large libraries whereby certain or all amino acids are mutated randomly (Packer and Liu, 2015). The library then has to be selected for the enzyme which exhibits the highest stability (Eijsink et al., 2005) or activity (Kintses et al., 2012). Both of these approaches can be time-consuming and labour-intensive and sometimes do not lead to large improvements in thermostability. For example, after 700 mutations, the Tm of T4 lysozyme was only increased by 8 °C at the cost of the enzyme's activity (Baase et al., 2010).

Enzyme cyclization, covalently ligating the N-terminus of an enzyme to its C-terminus, could be a more general approach to increase the thermal resilience of enzymes (Iwai and Plückthun, 1999). Enzyme cyclization has previously been achieved via carbodiimide cross-linking (Goldenberg and Creighton, 1983), sortase-mediated cyclization (Parthasarathy et al., 2007) or intein-mediated cyclization (Camarero and Muir, 1999; Evans et al., 1999b; Iwai and Plückthun, 1999; Scott et al., 1999).

To further develop and study protein cyclization, we sought to develop an alternative cyclization approach. We hypothesized that SpyTag/SpyCatcher (Zakeri et al., 2012), Isopeptag/Pilin-C (Zakeri and Howarth, 2010) and SnoopTag/SnoopCatcher (Veggiani et al., 2016) could be used to cyclize enzymes *in vivo* by genetically encoding the peptide and Catcher-protein at the N- and C-termini of an enzyme. Once the gene would be expressed, the peptide and Catcher-protein would associate and form an irreversible amide bond between the reactive lysine side-chain and the reactive asparagine/aspartic acid side chain.

The three peptide-protein pairs were utilized to cyclize TEM-1 β-lactamase (BLA). BLA was chosen as a model system because it is well studied, the activity assay is simple and the N- and C-termini are close to each other in the crystal structure (7 Å). I used Mass Spectrometry (MS) and TEV protease cleavage to confirm successful cyclization. I tested the effects of cyclization on aggregation resistance and recovered enzyme activity. I included different point mutants and fusion controls in my aggregation and activity experiments to determine the importance of the isopeptide bond. Finally, I characterized the cyclized proteins using Differential Scanning Calorimetry (DSC) to better understand the mechanism leading to the increase in enzyme resilience.

3.1 Utilization of split isopeptide-containing domains to cyclize TEM-1 β-lactamase

3.1.1 Design of cyclized constructs

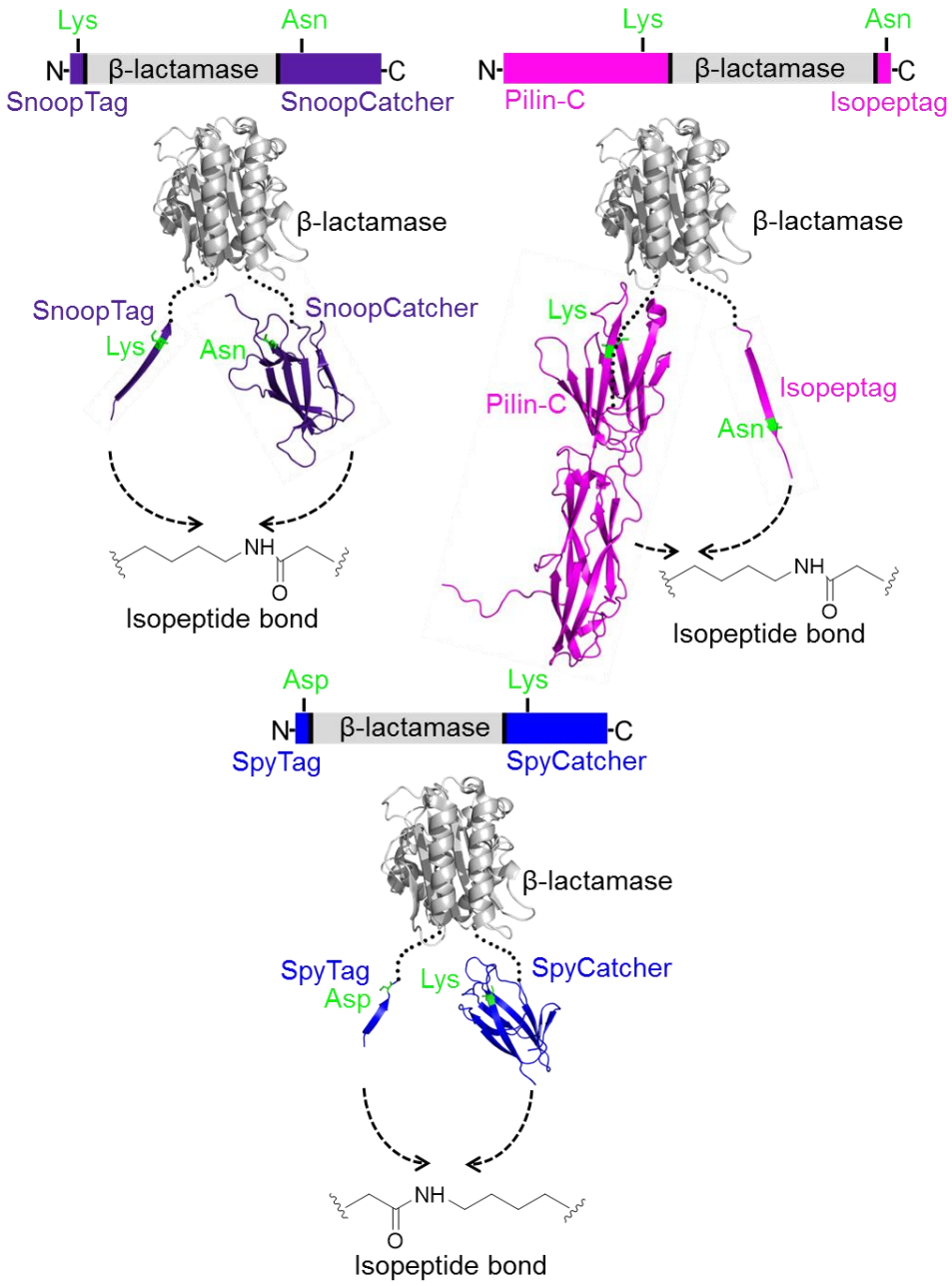


Figure 3.1: Design of cyclized BLA constructs. BLA was cyclized by genetically encoding SnoopTag, Pilin-C or SpyTag on the N-terminus and the orthogonal pair was genetically encoded on the C-terminus. Cartoons are based on Protein Data Bank files 1ZG4 (Stec et al., 2005), 3B2M (Kang et al., 2007), 2WW8 (Izoré et al., 2010) and 2X5P (Oke et al., 2010). Parts of this figure have been previously published in (Schoene et al., 2016a).

Figure 3.1 is a cartoon representation of the different constructs designed to test the effects of isopeptide-mediated cyclization on the thermal resilience of BLA. A flexible $(GSG)_2$ linker was chosen to favour the formation of an isopeptide bond and to avoid the Tag-Catcher pair affecting the structure and activity of BLA. The N- and C-termini of BLA are 7 Å apart and are far away from the active site. I postulated that these constructs would lead to the spontaneous *in vivo* cyclization of BLA.

3.1.2 Isopeptide-mediated cyclization of BLA led to a gel shift when compared to linear mutants

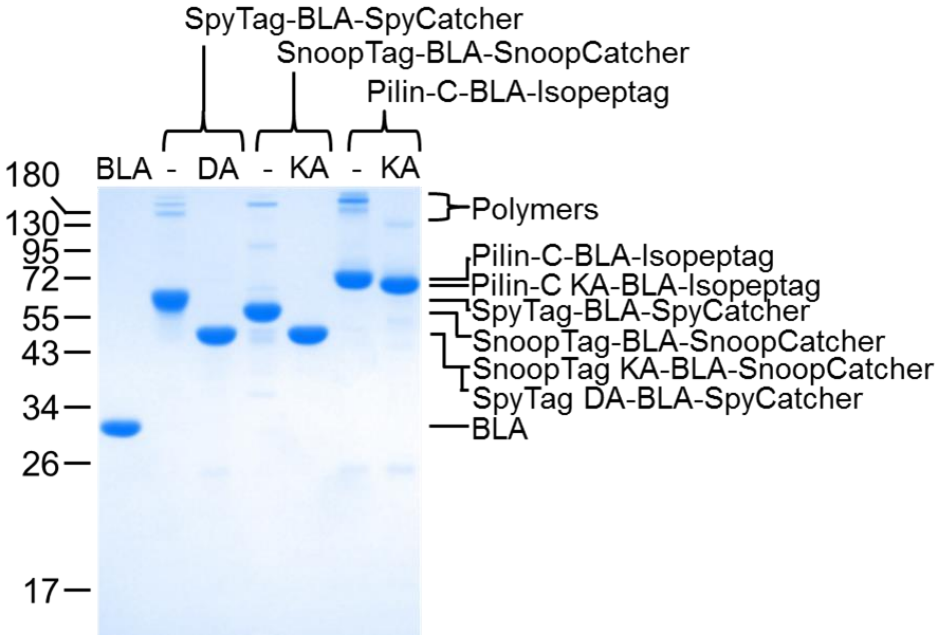


Figure 3.2: Cyclization reduced the mobility of the BLA constructs on a SDS-PAGE. Reduced SDS-PAGE gel (14 %) of the different cyclized constructs, their point mutant linear controls and wild-type BLA. Proteins were expressed with an N-terminal His-tag in *E. coli* and purified using IMAC. The SpyTag D→A, SnoopTag K→A and Pilin-C K→A mutations remove one of the three amino acids needed to form the isopeptide bond and therefore were not cyclized. Data were previously published in (Schoene et al., 2016a).

Expression of SpyTag-BLA-SpyCatcher, SnoopTag-BLA-SnoopCatcher and Pilin-C-BLA-Isopeptag led to the formation of a monomeric species and several polymeric species of higher molecular weight *in vivo*. The monomeric species migrated slower than the point mutant controls (Figure 3.2). The difference in migration between

the cyclized and linear constructs was also found to be acrylamide percentage-dependent (Figure 3.2, 3.3B, 3.6A and A.1).

3.1.3 TEV protease cleavage of cyclized enzyme compared to the non-cyclized linear control

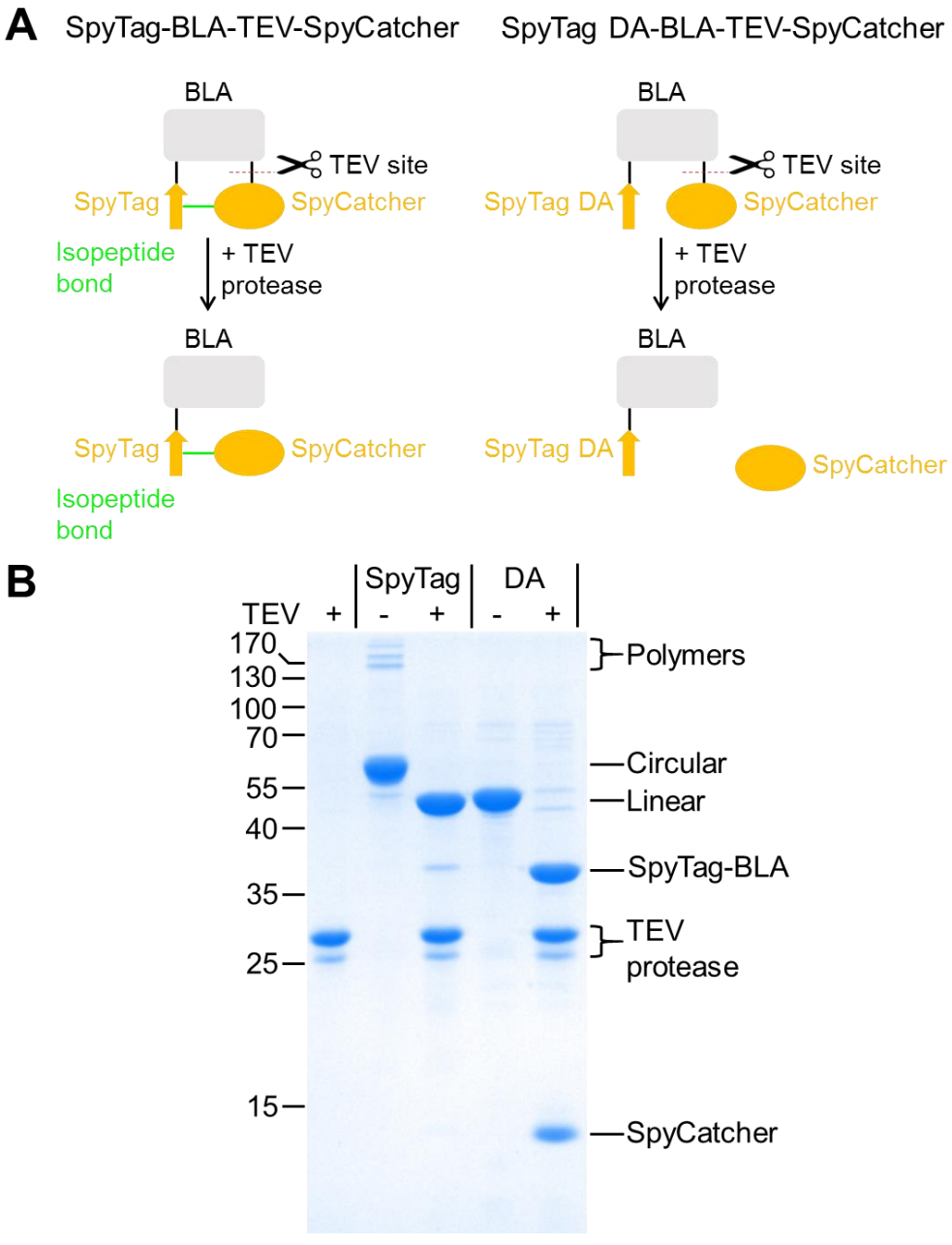


Figure 3.3: TEV protease cleavage of cyclized and non-cyclized BLA. A) Diagram of expected products obtained from TEV protease cleavage. B) Reduced SDS-PAGE gel (14 %) of SpyTag-BLA-TEV-SpyCatcher and SpyTag DA-BLA-TEV-SpyCatcher in the presence and absence of TEV protease. Data were previously published in (Schoene et al., 2014).

To test whether BLA had been successfully cyclized by SpyTag/SpyCatcher a TEV (Tobacco Etch Virus) protease cleavage site was inserted between BLA and SpyCatcher to make the constructs SpyTag-BLA-TEV-SpyCatcher and SpyTag DA-BLA-TEV-SpyCatcher.

As we can see in Figure 3.3A, cleavage of a cyclized protein will yield a single linear product whereas cleavage of a linear protein will yield two products. Cleavage of SpyTag-BLA-TEV-SpyCatcher migrated nearly the same distance as the point mutant control. Moreover, I observed the disappearance of the polymer bands. I saw the appearance of a faint band of SpyTag-BLA in the sample which might originate from the cleavage of some of the polymeric species or from some non-cyclized species in the sample (Figure 3.3B).

In the case of SpyTag DA-BLA-TEV-SpyCatcher, after cleaving with TEV protease I saw the disappearance of the major band in the sample and the appearance of two new bands which are likely to be SpyTag DA-BLA and SpyCatcher (Figure 3.3B). This was in agreement with the expected experimental outcome outlined in Figure 3.3A.

3.1.4 Mass Spectra provides further evidence for successful cyclization

Mass spectrometry was used to determine the molecular weight of the proteins. If the protein was cyclized I expected to see a molecular weight that is 18 Da below the ProtParam-predicted molecular weight, due to the loss of a water molecule upon formation of the isopeptide bond (Figure 3.4A). As expected, the SpyCatcher E→Q mutation and the SpyTag D→A mutation led to the control constructs not forming an isopeptide bond (Figure 3.4B-C)

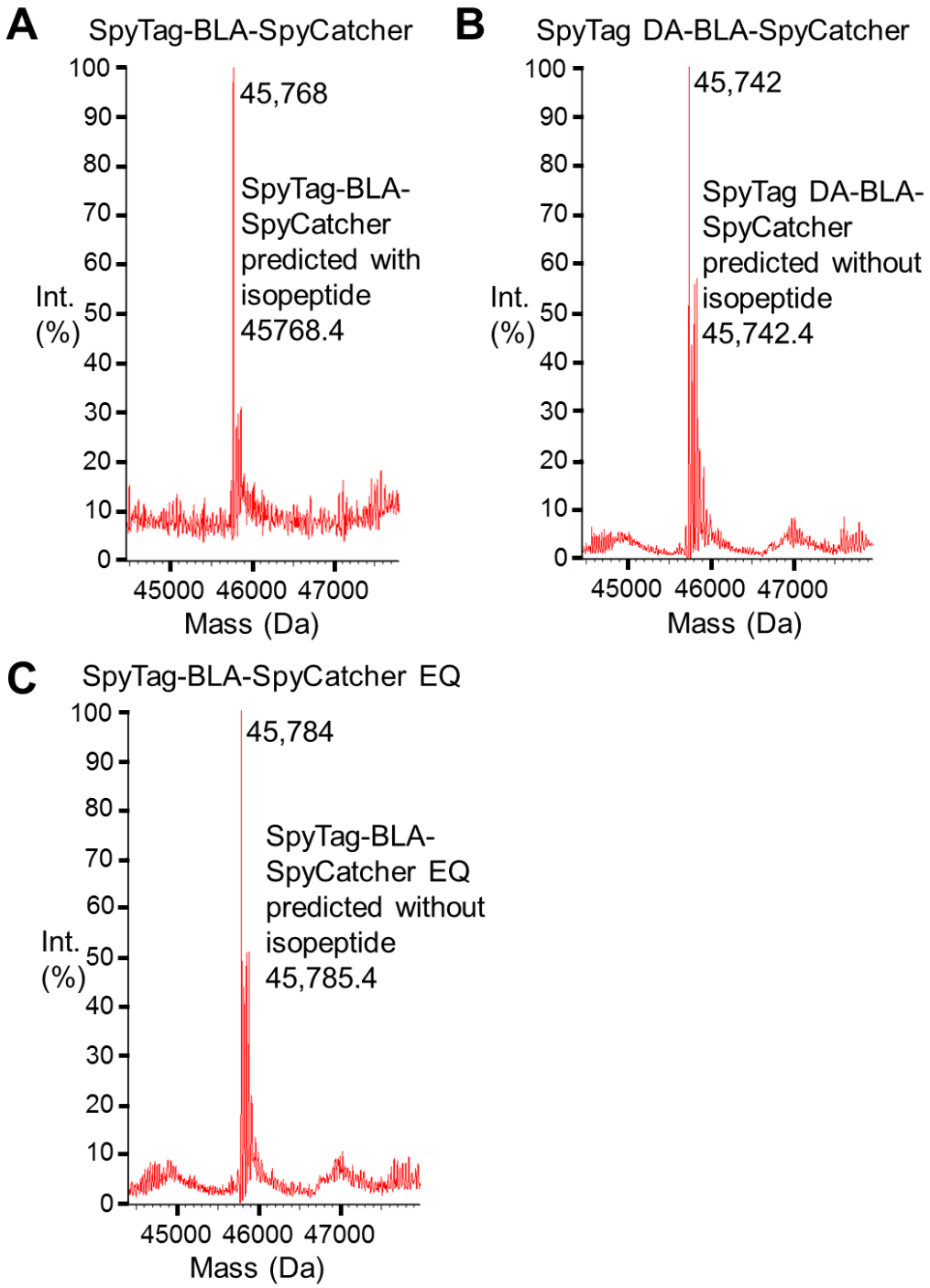


Figure 3.4: MS was used to verify the loss of a water molecule after the cyclization of SpyTag-BLA-SpyCatcher. A) MS of SpyTag-BLA-SpyCatcher. Predicted molecular weight was 45,786.4 – 18 = 45768.4 Da. B) MS of SpyTag DA-BLA-SpyCatcher where the catalytic aspartate residue has been mutated to an alanine. Predicted molecular weight was 45,742.4 Da. C) MS of SpyTag-BLA-SpyCatcher EQ where the catalytic glutamate residue has been mutated to a glutamine. Predicted molecular weight was 45,785.4 Da. Data were previously published in (Schoene et al., 2014).

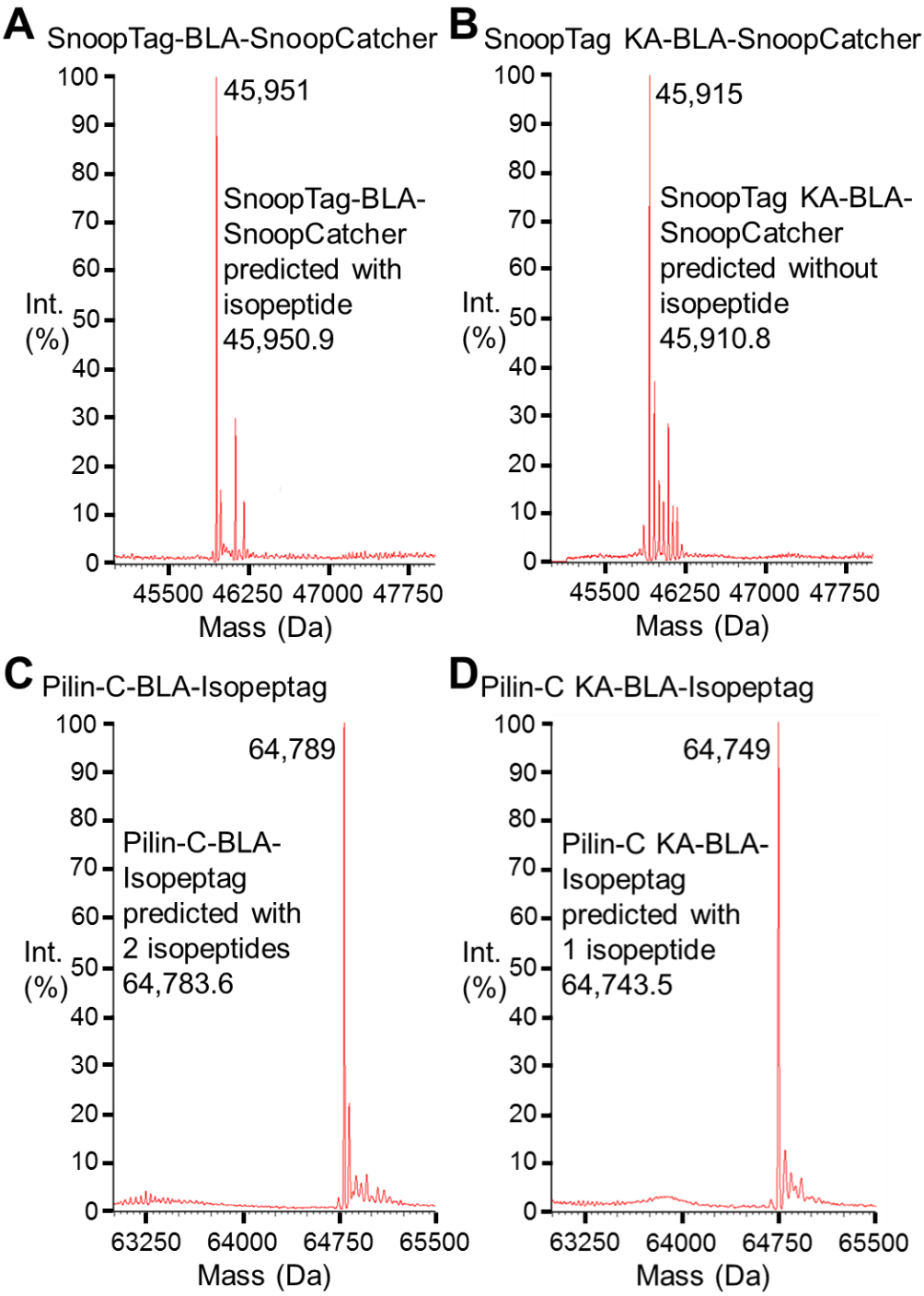


Figure 3.5: MS was used to verify the loss of an ammonia molecule after the cyclization of SnoopTag-BLA-SnoopCatcher and Pilin-C-BLA-Isopeptag. A) MS of SnoopTag-BLA-SnoopCatcher. Predicted molecular weight was 45,967.9 – 17 = 45,950.9 Da. B) MS of SnoopTag KA-BLA-SnoopCatcher. Predicted molecular weight was 45,910.8 Da. C) MS of Pilin-C-BLA-Isopeptag. Predicted molecular weight was 64,817.6 – 34 = 64,783.6 Da. D) MS of Pilin-C KA-BLA-Isopeptag. Predicted molecular weight was 64,760.5 – 17 = 64,743.5 Da. Data were previously published in (Schoene et al., 2016a).

In the cases of Isopeptag/Pilin-C and SnoopTag/SnoopCatcher, we expected to see a molecular weight that is 17 Da smaller due to the loss of an ammonia molecule (Figure 3.5A,C). As expected, the SnoopTag K→A mutation and the Pilin-C K→A mutation led to the control constructs not forming an isopeptide bond (Figure 3.5B,D).

3.2 Effects of cyclization on aggregation and recovered activity after heating

3.2.1 Effect of cyclization on aggregation

BLA was shown to start aggregating at 37 °C with almost complete loss of protein at and above 55 °C (Figure 3.6A,D). On the other hand, sandwiching BLA between SpyTag/SpyCatcher and Isopeptag/Pilin-C led to almost complete resistance to the formation of insoluble aggregates up to 100 °C (Figure 3.6A,C-D). Sandwiching BLA between SnoopTag/SnoopCatcher also led to a large improvement in aggregation resistance when compared to BLA (Figure 3.6A-B,D). However, incubating SnoopTag-BLA-SnoopCatcher at 75 °C caused a substantial amount of protein to form insoluble aggregates. Surprisingly, SnoopTag-BLA-SnoopCatcher was able to remain mostly soluble at 90 and 100 °C (Figure 3.6B,D). This result may indicate the existence of an energetic trap in the folding landscape of SnoopTag-BLA-SnoopCatcher where the protein is more likely to form insoluble aggregates at 75 °C than at any other temperature tested.

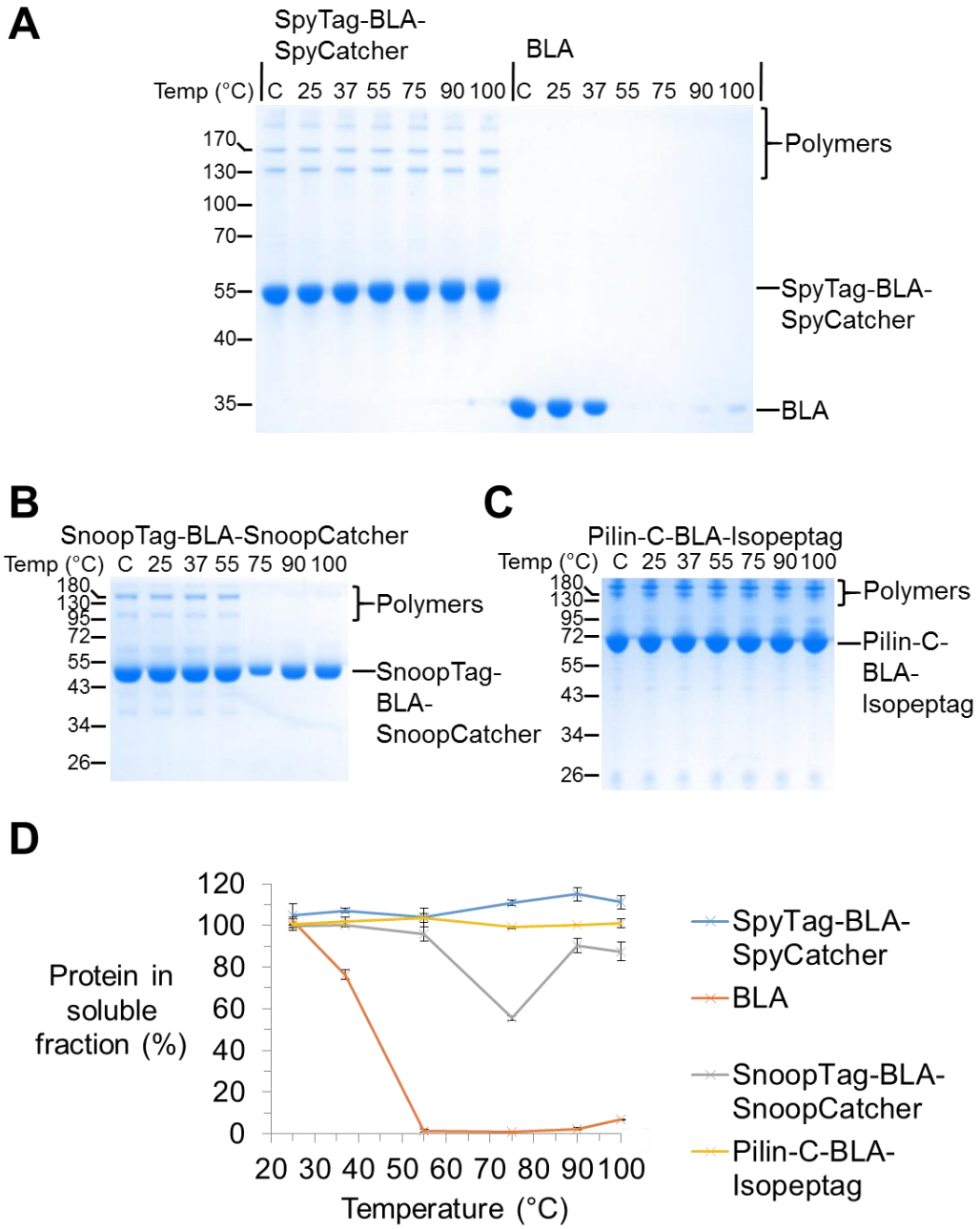


Figure 3.6: Effect of SpyTag/SpyCatcher, SnoopTag/SnoopCatcher and Isopeptag/Pilin-C-mediated cyclization on the aggregation resistance of BLA. A-C) Demonstrative gel of amount of soluble A) BLA and SpyTag-BLA-SpyCatcher (was run on a 10 % SDS-PAGE gel), B) SnoopTag-BLA-SnoopCatcher and, C) Pilin-C-BLA-Isopeptag in the sample after heating 25 µM enzyme to the indicated temperature for 10 min in PBS containing 100 mM DTT, cooling to 10 °C and then spinning the sample at 17,000g for 30 min. D) Densitometry-based quantification of bands (mean of triplicate ± 1 standard deviation). Data were previously published in (Schoene et al., 2014, 2016a).

3.2.2 Effects of cyclization on recovered activity after heating

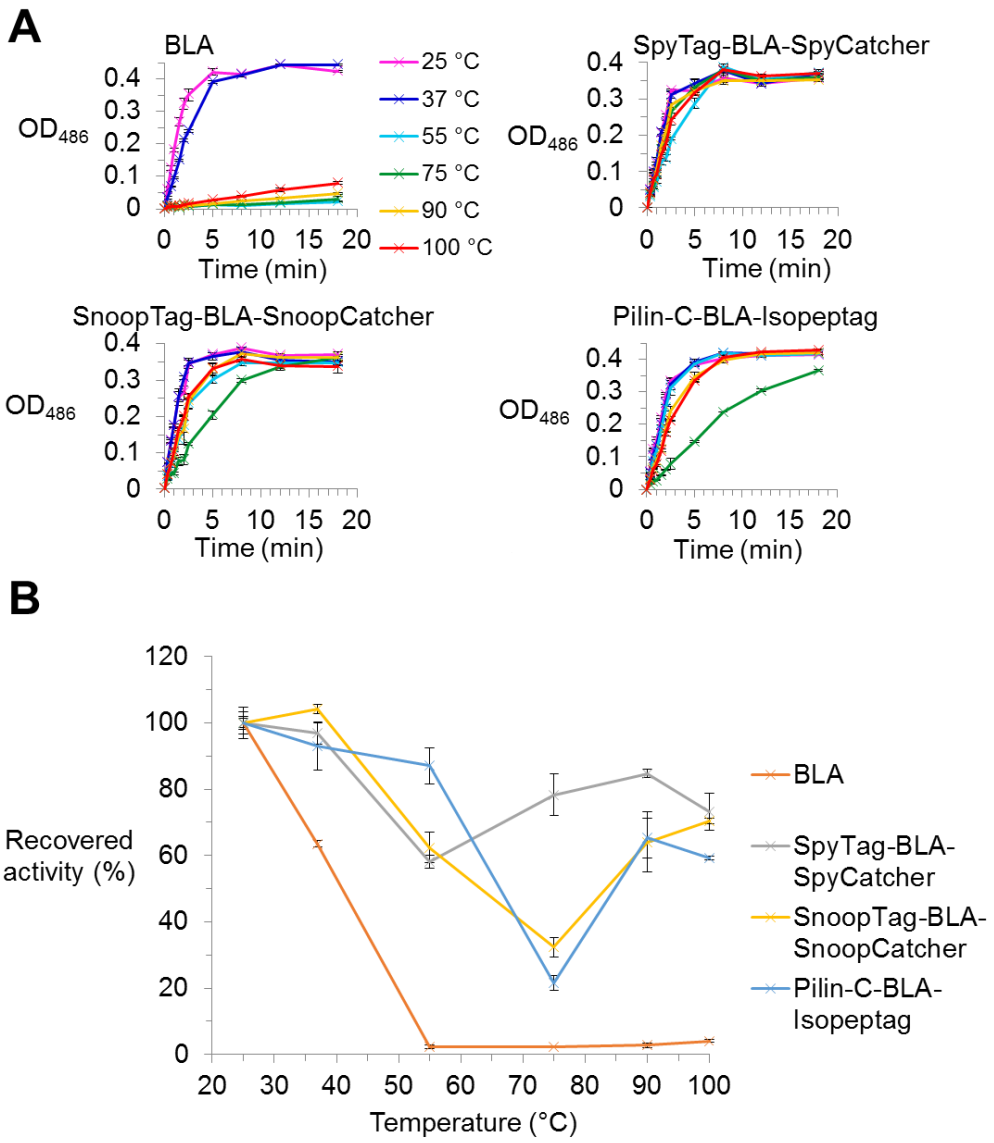


Figure 3.7: Effect of SpyTag/SpyCatcher, SnoopTag/SnoopCatcher and Isopeptag/Pilin-C-mediated cyclization on the activity of BLA. A) 25 µM BLA, SnoopTag-BLA-SnoopCatcher, Pilin-C-BLA-Isopeptag or SpyTag-BLA-SpyCatcher were incubated for 10 min at the indicated temperature in PBS containing 100 mM DTT then cooled to 10 °C, before returning to RT and then running a nitrocefin colorimetric activity assay where 3 nM of enzyme was allowed to react with 100 µM nitrocefin in 0.1 M phosphate buffer pH 7.0. The reaction was quenched at different time points through the addition of clavulunate (mean of triplicate ± 1 standard deviation). B) Recovered activity was calculated by obtaining the initial activity values of the curves in A) and then dividing the initial activity values of all the temperatures by the initial activity of each individual enzyme at 25 °C to obtain a percentage value (mean of triplicate ± 1 standard deviation). Data were previously published in (Schoene et al., 2016a).

As we can see in Figure 3.7, BLA lost some of its activity after a 10 min incubation at 37 °C with nearly complete loss of activity at 55 °C and little to no improvement in recovery at higher temperatures. On the other hand, all cyclized constructs improved upon the recovered activity after heating. Interestingly, the activity did not decay in a steady fashion with increasing temperatures. Instead every construct recovered differently depending on the temperature. For example, 55 °C was the most detrimental temperature for SpyTag-BLA-SpyCatcher whereas 75 °C was the most detrimental temperature for both Pilin-C-BLA-Isopeptag and SnoopTag-BLA-SnoopCatcher. Therefore, the cyclized BLA proteins appeared to exhibit non-Arrhenius behaviour (Wang and Roberts, 2013).

3.3 Importance of protein domains and the isopeptide bond (comparison to point mutants and fusions like BLA-CnaB2)

3.3.1 Aggregation and recovered activity after heating of SnoopTag/SnoopCatcher and Isopeptag/Pilin-C linear points mutants

When comparing the point mutant controls in Figure 3.8 with the cyclized proteins in Figure 3.6 and Figure 3.7 one can begin to elucidate the importance of the isopeptide bond for the perceived increase in thermal resilience. SnoopTag KA-BLA-SnoopCatcher exhibited a slight improvement in aggregation resistance and recovered activity at 37 °C when compared to the untagged BLA construct. However, at 55 °C SnoopTag KA-BLA-SnoopCatcher lost most of its activity. Therefore, the isopeptide bond was essential for the thermal resilience conferred by the SnoopTag/SnoopCatcher pair.

Pilin-C KA-BLA-Isopeptag fared better than SnoopTag KA-BLA-SnoopCatcher. The protein remained mostly in solution even at higher temperatures, unlike the untagged BLA construct. Moreover, Pilin-C KA-BLA-Isopeptag was able to retain some of its activity as well. However, the thermal resilience of Pilin-C-BLA-

Isopeptag was better than that of Pilin-C KA-BLA-Isopeptag. Pilin-C-BLA-Isopeptag did not form insoluble aggregates when heated and the recovered activity was higher after heating to 55, 75, 90 or 100 °C than that of Pilin-C KA-BLA-Isopeptag. Therefore, the isopeptide bond was also important for the improvement in thermal resilience obtained when using the Isopeptag/Pilin-C pair.

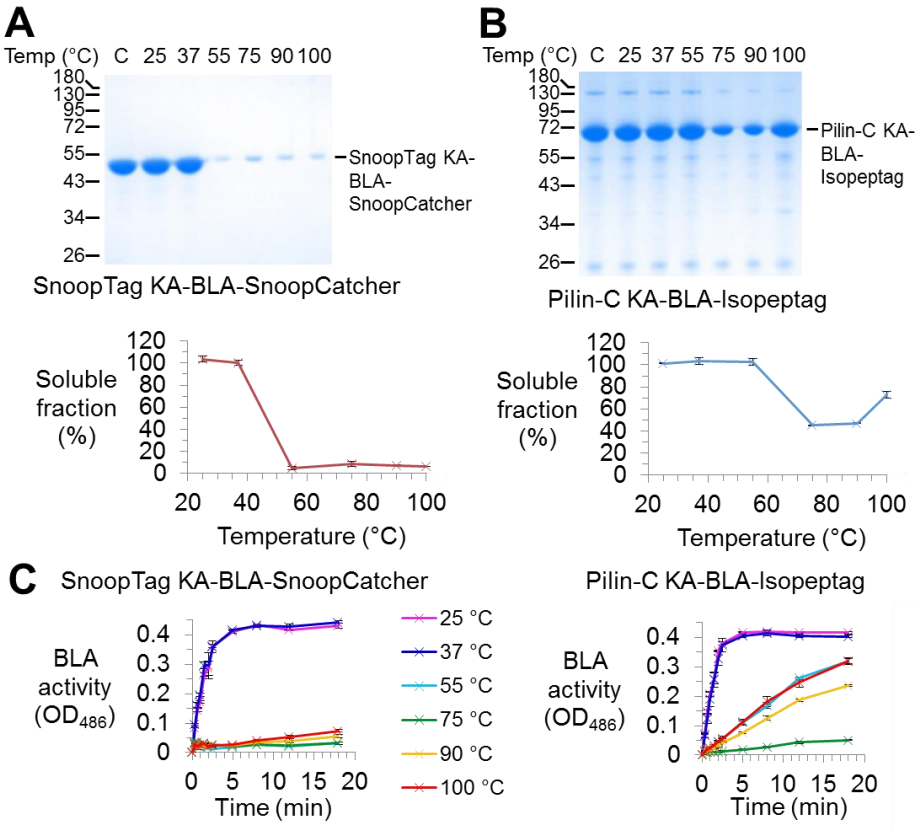


Figure 3.8: Thermal resilience of linear SnoopTag/SnoopCatcher and Isopeptag/Pilin-C point mutants. A-B) Demonstrative gel of amount of soluble A) SnoopTag KA-BLA-SnoopCatcher and B) Pilin-C KA-BLA-Isopeptag in the sample after heating 25 µM enzyme to the indicated temperature for 10 min in PBS containing 100 mM DTT, cooling to 10 °C and then spinning the sample at 17,000g for 30 min and quantification done by densitometry (mean of triplicate ± 1 standard deviation). C) 25 µM SnoopTag KA-BLA-SnoopCatcher or Pilin-C KA-BLA-Isopeptag were incubated for 10 min at the indicated temperature in PBS containing 100 mM DTT then cooled to 10 °C, before returning to RT and then running a nitrocefin colorimetric activity assay where 3 nM of enzyme was allowed to react with 100 µM nitrocefin in 0.1 M phosphate buffer pH 7.0. The reaction was quenched at different time points through the addition of clavulunate (mean of triplicate ± 1 standard deviation). Data were previously published in (Schoene et al., 2016a).

3.3.2 The importance of the isopeptide bond and cyclization on the thermal resilience of SpyTag-BLA-SpyCatcher



Figure 3.9: Amino acid sequence of constructs used to determine the importance of cyclization. Aligned amino acid sequence of SpyTag-BLA-SpyCatcher, SpyTag DA-BLA-SpyCatcher, SpyTag-BLA-SpyCatcher EQ and BLA-CnaB2. Sequence alignment was done using Clustal Omega (Sievers et al., 2011). N-terminal methionine was not included in the alignment as the methionine is removed in *E. coli* (Liao et al., 2004).

To test the importance of the isopeptide bond in the increase in thermal resilience I witnessed for SpyTag-BLA-SpyCatcher, I also tested a SpyTag D→A mutant and a SpyCatcher E→Q mutant which are both unable to form an isopeptide bond (Figure 3.9). However, removing the isopeptide bond has been shown to decrease the stability of the CnaB2 domain (Hagan et al., 2010). Therefore, the two linear controls alone were not adequate to determine whether the thermal resilience

was due to cyclization and an increase in backbone rigidity (Schumann et al., 2015) or due to the effect of the BLA domain being fused to a thermostable CnaB2-like domain (SpyTag + SpyCatcher) which could act like a refolding “scaffold” (Kim et al., 2009). Therefore, BLA-CnaB2 was used as a control to determine the thermal resilience one could achieve by just fusing BLA to a thermostable protein.

3.3.3 Cyclization vs. thermostable fusion – effects on aggregation resistance

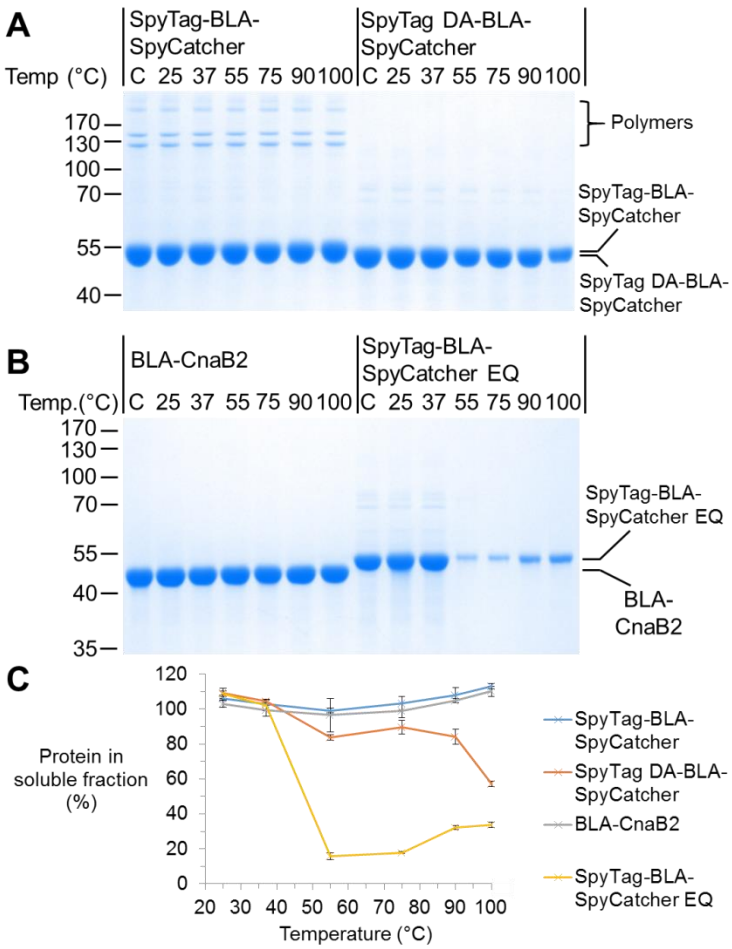


Figure 3.10: Aggregation resistance can be achieved via a fusion to CnaB2. A-B) Demonstrative gel of amount of soluble A) SpyTag-BLA-SpyCatcher and SpyTag DA-BLA-SpyCatcher and, B) BLA-CnaB2 and SpyTag-BLA-SpyCatcher EQ in the sample after heating 25 µM of enzyme to the indicated temperature for 10 min in PBS containing 100 mM DTT, cooling to 10 °C and then spinning the sample at 17,000g for 30 min. C) Densitometry based quantification of bands (mean of triplicate ± 1 standard deviation). Data were previously published in (Schoene et al., 2014).

By comparing the aggregation resistance of SpyTag-BLA-SpyCatcher with the DA and EQ point mutant, I determined that the addition of the SpyTag DA and SpyCatcher components were enough to confer resistance to aggregation. However, a small percentage of SpyTag DA-BLA-SpyCatcher still formed insoluble aggregates (Figure 3.10A,C). Most of SpyTag-BLA-SpyCatcher EQ formed insoluble aggregates at temperatures of 55 °C and above (Figure 3.10B-C).

By comparing the aggregation resistance of BLA-CnaB2 to that of SpyTag-BLA-SpyCatcher (Figure 3.10A-C), I determined that the stability and solubility of the CnaB2 domain was enough to minimize the formation of insoluble aggregates. I could not measure any BLA-CnaB2 or SpyTag-BLA-SpyCatcher falling out of solution. Therefore, the SpyTag/SpyCatcher domain acting like a refolding "scaffold" was enough to confer resistance to irreversible aggregation.

3.3.4 Cyclization vs. thermostable fusion – effects on recovered activity after heating

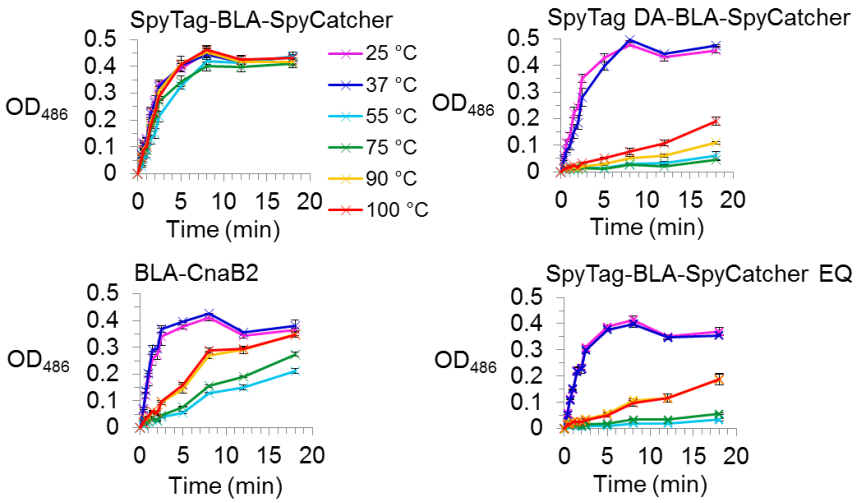


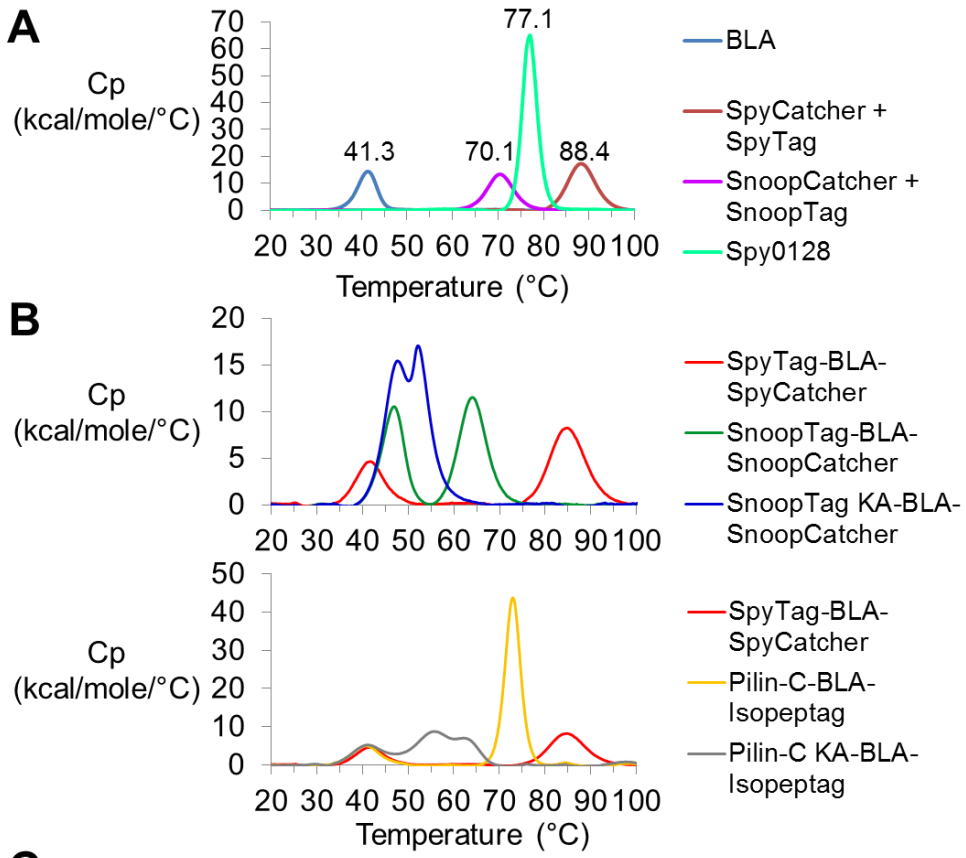
Figure 3.11: Cyclized BLA was the most resilient construct. 25 µM SpyTag-BLA-SpyCatcher, SpyTag DA-BLA-SpyCatcher, BLA-CnaB2 or SpyTag-BLA-SpyCatcher EQ were incubated for 10 min at the indicated temperature in PBS containing 100 mM DTT then cooled to 10 °C, before returning to RT and then running a nitrocefin colorimetric activity assay where 3 nM of enzyme was allowed to react with 100 µM nitrocefin in 0.1 M phosphate buffer pH 7.0. The reaction was quenched at different time points through the addition of clavulunate (mean of triplicate ± 1 standard deviation). Data were previously published in (Schoene et al., 2014).

Similar to the results in Figure 3.7 and 3.8, the isopeptide bond was essential in conferring thermal resilience to BLA using SpyTag/SpyCatcher (Figure 3.11A). Both SpyTag DA-BLA-SpyCatcher and SpyTag-BLA-SpyCatcher EQ lost most of their activity after incubating at and above 55 °C. In the case of SpyTag-BLA-SpyCatcher EQ, the results could be correlated with the formation of insoluble aggregates (Figure 3.10B-C). In the case of SpyTag DA-BLA-SpyCatcher, the protein was either not folding correctly after cooling and/or it formed soluble aggregates.

While BLA-CnaB2 did not form insoluble aggregates like SpyTag-BLA-SpyCatcher (Figure 3.10A-C), the recovered activity after heating was not as high as that obtained for SpyTag-BLA-SpyCatcher (Figure 3.11). However, the CnaB2 domain appears to function as a refolding “scaffold” since the recovered activity of the BLA-CnaB2 construct is a lot higher than that of BLA (Figure 3.7A and 3.11). Therefore, I postulated that the thermal resilience of SpyTag-BLA-SpyCatcher might be due to two factors. The first factor is the solubility and thermal resilience that can be obtained by fusing the protein of interest to a thermostable protein which can act like a refolding “scaffold” (de Marco et al., 2004) which can be observed by looking at the difference in recovered activity between BLA-CnaB2 and BLA. The second factor is the increased rigidity of the protein backbone caused by cyclization (Schumann et al., 2015) which can be observed by looking at the difference in recovered activity between BLA-CnaB2 and SpyTag-BLA-SpyCatcher.

3.4 Studying the increase in thermal resilience using Differential Scanning Calorimetry (DSC)

3.4.1 DSC data of cyclized constructs and point mutant controls



C

Sample	Peak	Temperature (°C)
SpyTag-BLA-SpyCatcher	Tm1	41.6
	Tm2	85.0
SnoopTag-BLA-SnoopCatcher	Tm1	47.0
	Tm2	64.1
SnoopTag KA-BLA-SnoopCatcher	Tm1	47.6
	Tm2	52.2
Pilin-C-BLA-Isopeptag	Tm1	40.9
	Tm2	72.9
Pilin-C KA-BLA-Isopeptag	Tm1	41.0
	Tm2	55.9
	Tm3	62.4

Figure 3.12: DSC of the cyclized BLA constructs, point mutant controls and the domains involved. A) DSC traces of 20 µM BLA, SpyCatcher + SpyTag peptide, SnoopCatcher + SnoopTag peptide and Spy0128 obtained from scanning the proteins from 20-110 °C at a rate of 1 °C/min in PBS. B) DSC traces of 20 µM SpyTag-BLA-SpyCatcher, SnoopTag (KA)-BLA-SnoopCatcher and Pilin-C (KA)-BLA-Isopeptag obtained from scanning the proteins from 20-110 °C at a rate of 1 °C/min in PBS. C) Tm values of the DSC traces in B). Data were previously published in (Schoene et al., 2016a).

The increases in thermal resilience that I obtained via isopeptide-mediated cyclization could be explained by an increase of the Tm of BLA. Cyclization could lead

to an increase in Tm which would allow BLA to survive higher temperatures. However, in the case of SpyTag-BLA-SpyCatcher and Pilin-C-BLA-Isopeptag I measured almost no difference in the Tm of the BLA domains when compared to the untagged BLA construct (Figure 3.12A-C). In the case of SnoopTag-BLA-SnoopCatcher, I measured an increase in Tm of 5.7 °C (Figure 3.12A-C). However, the increase in Tm was not due to cyclization, as the Tm of the BLA domain in SnoopTag KA-BLA-SnoopCatcher was similar to that of SnoopTag-BLA-SnoopCatcher. Either way, an increase in Tm of 5.7 °C could not account for the increase in thermal resilience I had previously measured (Figure 3.6 and 3.7).

I measured 2 or 3 peaks for the tagged constructs. In the case of SpyTag-BLA-SpyCatcher, I measured one peak at 41.6 °C which corresponded to the Tm of BLA and a second peak with a Tm of 85.0 °C which was close to the Tm of SpyCatcher + SpyTag peptide. SnoopTag-BLA-SnoopCatcher had two peaks in its DSC trace. One had a Tm of 47.0 °C and could have been a stabilized BLA domain while the second peak had a Tm of 64.1 °C which was close to the Tm of SnoopCatcher + SnoopTag peptide. Finally, I measured two peaks for Pilin-C-BLA-Isopeptag. One peak had a Tm of 40.9 °C and the second had a Tm of 72.9 °C, which was close to the measured Tm of the Spy0128 from which the Isopeptag/Pilin-C system originated (Zakeri and Howarth, 2010).

Removal of the isopeptide bond also decreased the Tm of the cyclization domain in all cases (Figure 3.12B-C and 3.13A). Therefore, removal of the isopeptide bond had two detrimental effects. The first is that the covalent bond was no longer there to rigidify the protein backbone and the second effect is that the removal of the isopeptide bond destabilizes the cyclization domains, thereby removing the effects associated with having a thermostable fusion partner. Moreover, the DSC data in Figure 3.12 is a measure of thermodynamic stability whereas the data in Figure 3.7B reflects the kinetic stability of the constructs. Therefore, isopeptide-mediated

cyclization appears to have a larger effect on kinetic stability rather than the thermodynamic stability.

3.4.2 Studying the effects of cyclization on the refolding of BLA using DSC

While the BLA domain unfolds at around 41 °C (Figure 3.12A), the isopeptide containing domains remain folded until higher temperatures. By remaining folded, the domains might have decreased the chance of aggregation occurring and might have made a refolding event of the BLA domain more likely.

To study the ability of the constructs to refold we utilized DSC. In the refold experiment I scanned the proteins, allowed them to cool, and then scanned them again. In the case of a protein being able to refold, one would see a peak in the second scan. In my experiment, only SpyTag-BLA-SpyCatcher had a measurable peak corresponding to the BLA domain in the second scan (Figure 3.13B). Therefore, we postulated that isopeptide-mediated cyclization improves BLA's likelihood to refold, instead of increasing the Tm of the domain.

SpyTag-BLA-SpyCatcher's peak in the second scan in Figure 3.13B is smaller than one would expect from the data obtained in Figure 3.6A and 3.7B. The second peak might be lower because the DSC experiment uses a slower heating and cooling ramp rate and therefore the protein is exposed to higher temperatures for longer (Houde and Berkowitz, 2014).

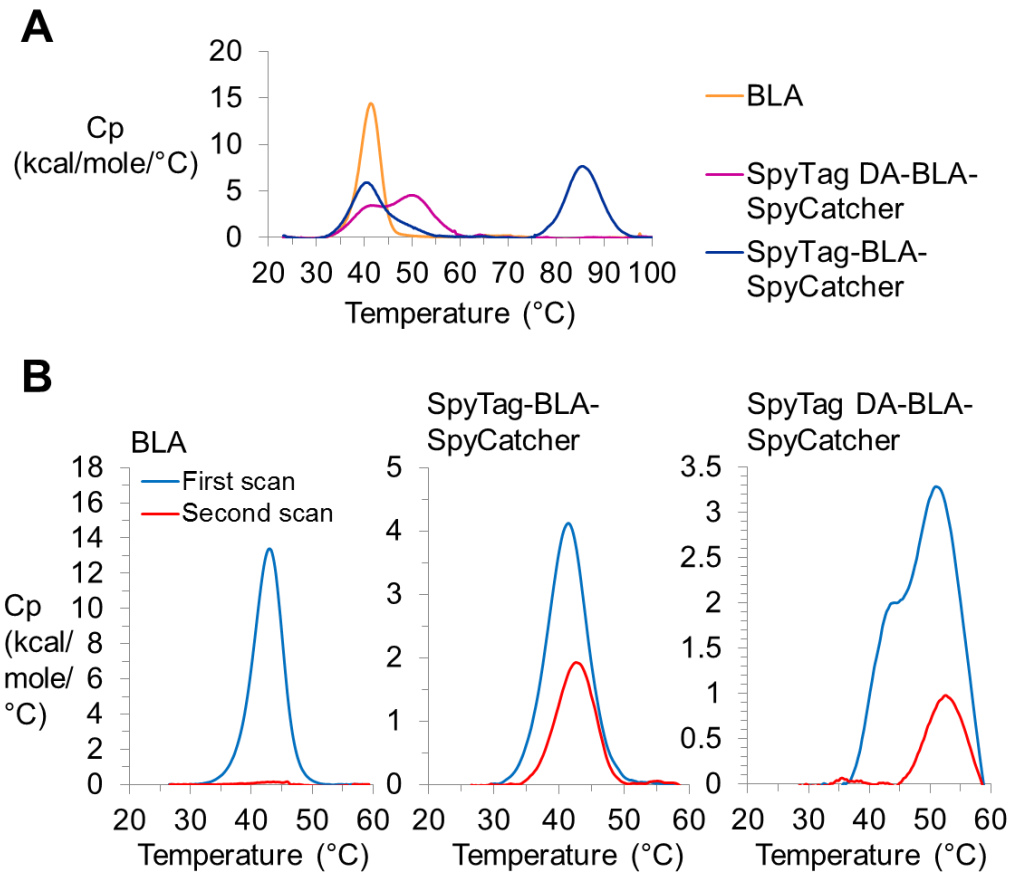


Figure 3.13: Studying the effects of cyclization on refolding. A) DSC traces of 20 µM BLA, SpyTag-BLA-SpyCatcher and SpyTag DA-BLA-SpyCatcher obtained from scanning the proteins from 20-110 °C at a rate of 1 °C/min in PBS. B) 20 µM BLA, SpyTag-BLA-SpyCatcher and SpyTag DA-BLA-SpyCatcher was scanned from 20-60 °C at a rate of 2 °C/min in PBS, cooled and then the scan was repeated. Data were previously published in (Schoene et al., 2014).

3.5 Discussion

Sortase and intein-based approaches to cyclize proteins are more commonly documented in the literature. Sortase can be used to cyclize peptides efficiently and has been used to cyclize larger proteins (van 't Hof et al., 2015). However, enhancements in thermal resilience or thermostability for cytokines have been small (Popp et al., 2011).

Previous work on BLA cyclization utilized an engineered intein on the C-terminus of BLA to react with a cysteine on the N-terminus (Iwai and Plückthun, 1999).

This led to the formation of a cyclized BLA with a residual cysteine residue at the cyclization site. The effect on thermal resilience was tested with an aggregation assay similar to what I have used. The temperature at which 50 % of the BLA aggregated was only increased by 5 °C (Iwai and Plückthun, 1999).

Previous work on increasing the thermal resilience of proteins via fusions with thermophilic proteins led to the discovery that an insertional fusion of BLA into *Pyrococcus furiosus* Maltodextrin-binding protein (PfMBP) increased the thermal resilience of BLA (Pierre et al., 2011). In the study, the authors found that insertion of BLA into PfMBP either did not change or slightly decreased the apparent Tm of BLA (obtained by DSC) depending on where the BLA was inserted. Further characterization included a kinetic stability study at 50 °C that showed an improvement when using certain insertional fusions. However, this approach did not manage to change the irreversible nature of BLA's denaturation (Pierre et al., 2015).

While cyclization approaches have already existed for quite some time, the improvements in thermal resilience were never impressive. Therefore, I researched whether isopeptide-mediated cyclization could lead to greater improvements in thermal resilience.

Isopeptide-mediated cyclization relies on the use of split isopeptide containing proteins that have been engineered to improve the reaction rate between the Tag component and the Catcher component. Isopeptag/Pilin-C contains 2 isopeptide bonds after reconstitution and is 33 kDa in size. Isopeptag reacts slowly and inefficiently with Pilin-C only reaching ~60 % completion in 72 h (Zakeri and Howarth, 2010). SpyTag/SpyCatcher contains 1 isopeptide bond after reconstitution and is 14 kDa in size. SpyTag/SpyCatcher is able to reach 80 % reconstitution completion after only 10 min (Zakeri et al., 2012). SnoopTag/SnoopCatcher contains 1 isopeptide bond after reconstitution and is 14 kDa in size. The reaction rate of SnoopTag and SnoopCatcher lies somewhere between that of SpyTag/SpyCatcher and Isopeptag-

Pilin-C (Veggiani et al., 2016). Therefore, by utilizing these three orthogonal pairs I was testing the importance of the reconstitution rate of the three proteins, the importance of the size of the domain, whether 2 isopeptides are better than 1 and what the importance of the Tm of the domains are.

Unlike intein and sortase-mediated cyclization, isopeptide-mediated cyclization was not as seamless as the components used to cyclize the enzyme remained fused to the protein backbone (Figure 3.1). It has been previously shown that fusing a protein to a thermostable domain can lead to an increase in thermal resilience. For example, this technique has been exploited to simplify the purification of vaccine components (de Marco et al., 2004). Therefore, we postulated that the Tag/Catcher domain might contribute to the thermal resilience of the construct not only via the tethering of the N- and C-termini but also via the aforementioned phenomenon. Moreover, the cyclization event occurred *in vivo* and so the purified product was already cyclized and therefore did not require an extra reaction step after purification (Figure 3.2). Interestingly, all three Tag/Catcher pairs led to efficient cyclization which could be verified by a decrease in gel mobility (Figure 3.2).

The decrease in gel mobility (Figure 3.2) was surprising because the presence of disulphide bonds in a sample usually increases the mobility of a protein (Selimova et al., 1981). Therefore, further evidence was necessary to confirm successful cyclization. The evidence was obtained by using an alternative construct which contained a TEV site between BLA and SpyCatcher (Figure 3.3) and with ESI-MS (Figure 3.4 and 3.5).

Cyclization increased the solubility of BLA at higher temperatures (Figure 3.6). Differences in insoluble aggregation propensity between the different Tag/Catcher's could be due to a function of stability or solubility. Surprisingly, different temperatures had variable detrimental effects on the three cyclized constructs (Figure 3.6 and 3.7). Our current hypothesis for the non-linear decline in activity is that the proteins are at a different unfolded state at every temperature tested and some of these unfolded states

are more productive than others when refolding. This phenomenon appears to be similar to previous work published where it was found that the refolding yields of lysozyme were three times higher at 98 °C than at 20 °C (Sakamoto et al., 2004).

I found that two isopeptide bonds are not necessarily better than one since Pilin-C-BLA-Isopeptag did not perform better than SpyTag-BLA-SpyCatcher (Figure 3.6). SpyTag-BLA-SpyCatcher performed best overall in terms of recovered activity after heating and aggregation resistance (Figure 3.6 and 3.7). A possible explanation for this might be that SpyTag/SpyCatcher was found to have the highest Tm by DSC (Figure 3.12). Therefore, the Tm of the isopeptide-containing domain might be important for the increase in thermal resilience conferred to BLA.

Point mutants did not exhibit the same extent of thermal resilience seen in the cyclized constructs (Figure 3.8, 3.10 and 3.11). Therefore, the isopeptide bond was important in conferring the improvement in thermal resilience that we observed. Removal of the isopeptide bond had two effects. The first effect was that it lowered the Tm of the Tag/Catcher domain (Figure 3.12 and 3.13) and the second effect was that the termini were left untethered. To provide evidence that untethering of the N- and C-termini had an effect on the construct as a whole, we cloned BLA-CnaB2. In this construct, the isopeptide bond is contained within the Tag/Catcher component and the N- and C-termini of BLA are left untethered. The Tm of the CnaB2-like domain should be similar to that of the SpyTag/SpyCatcher domain. Also, both domains have a similar sequence and should therefore function equally well as refolding “scaffolds”. While BLA-CnaB2 did not form insoluble aggregates (Figure 3.10), the recovered activity after heating was lower than that of SpyTag-BLA-SpyCatcher (Figure 3.11). Therefore, we postulated that isopeptide-mediated cyclization provided a greater thermal resilience than intein-mediated cyclization (Iwai and Plückthun, 1999) of BLA and insertional fusions of BLA into PfMBP (Pierre et al., 2011) because isopeptide-mediated cyclization is essentially a combination of the two aforementioned approaches.

DSC analysis allowed me to gain a better insight into the underlying mechanism causing the increase in thermal resilience. The unfolding mechanism had two transitions indicating that the unfolding of the BLA domain and the cyclization Tag/Catcher domain occurred sequentially. The Tag/Catcher's high Tm and the addition of the isopeptide bond protected BLA from irreversibly inactivating even at temperatures where the Tag/Catcher domain itself had melted (Figure 3.12).

The denaturation of BLA had been shown to be an irreversible event where the protein unfolds and forms insoluble aggregates (Iwai and Plückthun, 1999; Pierre et al., 2015). Therefore, an explanation for the increase in thermal resilience of the cyclized constructs might be that the Tag/Catcher domain and the tethering of the N- and C-termini of BLA introduced a degree of reversibility to the thermal denaturation of BLA. SpyTag/SpyCatcher-mediated cyclization is capable of making a completely irreversible system somewhat reversible.

Chapter 4: Cyclization as a general approach to enhance thermal resilience.

Some of the work in this chapter has been published in the manuscripts (Schoene et al., 2014, Schoene et al., 2016a and Schoene et al., 2016b) provided in the appendix.

Chapter 4: Cyclization as a general approach to enhance thermal resilience

To determine whether the enhancement in thermal resilience to BLA in Chapter 3 was generally applicable to other enzymes, I developed a simple cloning approach to insert genes of interest into the SpyTag/SpyCatcher construct. Previous work with intein cyclization had shown improvement in the stability of BLA (Iwai and Plückthun, 1999), GFP (Iwai et al., 2001), an endoglucanase (van Lieshout et al., 2012), a xylanase (Waldhauer et al., 2015) and a phytase (Ping et al., 2010). Therefore, there is precedence in the literature which would indicate that isopeptide-mediated cyclization could enhance the thermal resilience of more than one enzyme.

There currently is no study determining what characteristics a protein must have for cyclization to have a positive effect on the thermal resilience. A logical criterion was to select proteins whose termini are close together, since close termini should be easier to link. However, selecting the right linker length and linker type might make it possible to bridge further distances between termini (Wang et al., 2008). Moreover, the linkers chosen for the BLA study in Chapter 3 were six amino acids long and designed to be flexible (Chen et al., 2013). On top of that, SpyCatcher has an unstructured section at the N-terminus (Li et al., 2013) which might provide further flexibility.

The two main criteria I used when selecting the enzymes to work on were: (i) cyclization or a circular permutation had previously been shown to increase the thermal resilience of the enzyme, or (ii) there is a commercial need for a more stable version of said enzyme. Therefore, I chose dihydrofolate reductase (DHFR) from *E. coli* because of previous work which improved DHFR's thermal resilience via cyanocysteine-mediated cyclization (Takahashi et al., 2007). Luciferase was chosen because Luciferase is used in a wide range of applications but is an unstable enzyme prone to aggregation and irreversible denaturation (Ebrahimi et al., 2012). Phytase is

primarily used in the manufacture of feed stock to increase the nutritional value of the feed (Lei et al., 2013). Part of the pelleting process of the feed stock requires a high temperature incubation step, which would require a phytase with a higher thermal resilience. Finally, Glucose Oxidase (GOx) was chosen because GOx is a high value enzyme used in the production of *in vitro* diagnostic devices (Yoo and Lee, 2010). Sekisui Diagnostics is a major manufacturer of GOx and are interested in developing improved GOx which might decrease manufacturing costs and increase shelf life. A collaboration was struck between Sekisui Diagnostics, the University of Oxford and the BBSRC and I carried out the research.

4.1 Cyclization of DHFR

4.1.1 Introduction of DHFR

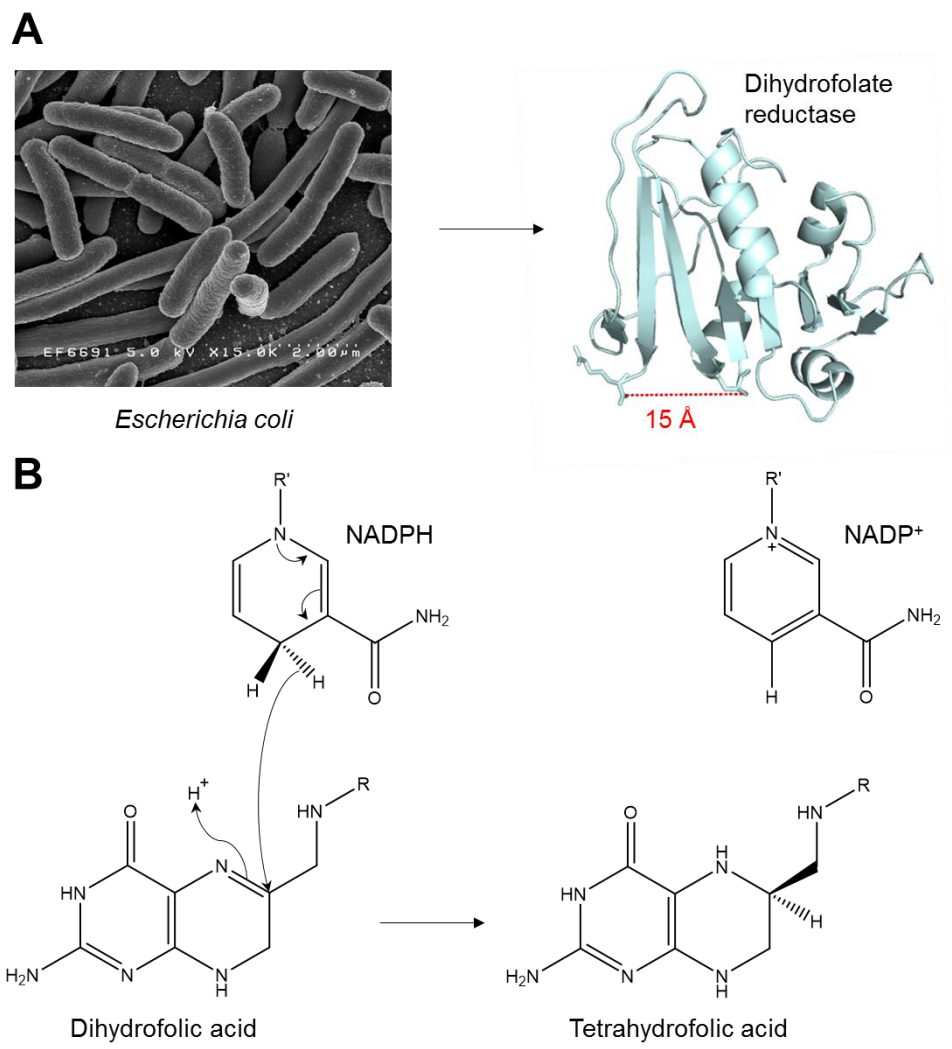


Figure 4.1: Structure and reaction mechanism of DHFR. A) Scanning electron micrograph of *E. coli* (credit: Rocky Mountain Laboratories, NIAID, NIH [Public domain], via Wikimedia Commons) and crystal structure of DHFR (PDB: 1RX2) (Sawaya and Kraut, 1997). B) Reaction mechanism of DHFR with its substrate, dihydrofolic acid.

DHFR was chosen because previous work has shown that backbone cyclization of a DHFR mutant lacking cysteines and methionines increased the Tm of the variant by 9.3 °C. The increase in Tm was measured by looking at the change in the absorption spectra between 250 and 310 nm. Moreover, the authors improved the reversibility of denaturation using their approach (Takahashi et al., 2007). Therefore, I

postulated that isopeptide-mediated cyclization could be used to improve the thermal resilience of DHFR. The advantage of using isopeptide-mediated cyclization instead of the cyanocysteine approach was that I could apply my approach to study the wild-type DHFR which contains cysteines and methionines.

In Figure 4.1A, I have illustrated the crystal structure of DHFR from *E. coli*. DHFR is classified as an α/β protein and the direct distance between the N- and C-terminus was measured to be 15 Å (Sawaya and Kraut, 1997). The activity of DHFR is dependent on its coenzyme, nicotinamide adenine dinucleotide phosphate (NADPH). In Figure 4.1B I have drawn a reaction mechanism for the reduction of DHFR's substrate, dihydrofolic acid. The reduction of dihydrofolic acid to tetrahydrofolic acid is essential for the production of methionine and thymidine monophosphate (dTMP). Hence, FolA (the gene encoding DHFR) is an essential gene for *E. coli* (Schnell et al., 2004).

4.1.2 MS and gel shift of cyclized DHFR

Cyclization of DHFR was efficient and caused the protein to migrate faster than the DA point mutant linear control (Figure 4.2A). This was surprising because SpyTag/SpyCatcher-mediated cyclization of BLA led to a decrease in gel mobility (Figure 3.2). Our current hypothesis is that the migration of the cyclized construct is dependent on two factors: the size of the insert between the SpyTag and SpyCatcher as well as the acrylamide percentage of the SDS-PAGE (Figure A.1). In the case of SpyTag-BLA-SpyCatcher, the insert was 263 amino acids long. SpyTag-DHFR-SpyCatcher had an insert length of 159 amino acids. Therefore, between 159 to 263 amino acids there is a shift from faster mobility (smaller insert) to slower mobility (larger insert) when compared to the linear point mutant.

MS provided further evidence for the efficient cyclization of the SpyTag-DHFR-SpyCatcher construct (Figure 4.2B-C). The second peak was attributed to a gluconylated species (Geoghegan et al., 1999).

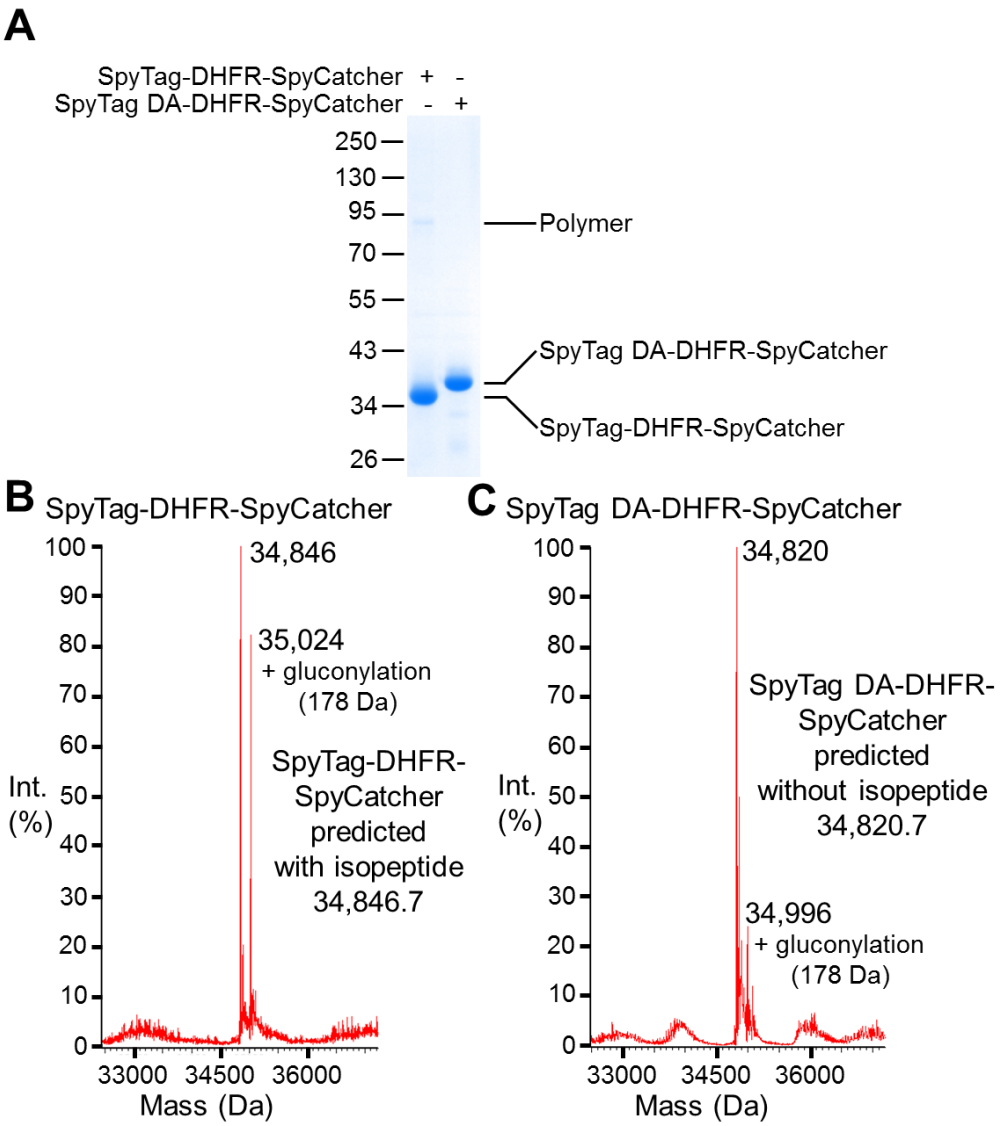


Figure 4.2: Gel shift and MS provide evidence for the successful cyclization of DHFR. A) Reduced SDS-PAGE of SpyTag-DHFR-SpyCatcher and the linear point mutant control, SpyTag DA-DHFR-SpyCatcher. B) MS of SpyTag-DHFR-SpyCatcher. Predicted molecular weight was 34,864.7 – 18 = 34,846.7 Da. C) MS of SpyTag DA-DHFR-SpyCatcher. Predicted molecular weight was 34,820.7 Da. Data were previously published in (Schoene et al., 2014).

4.1.3 Effect of cyclization on protein aggregation of DHFR

The untagged DHFR mutant was shown to be highly resistant to the formation of insoluble aggregates with only a small loss of soluble protein after incubation at 100 °C. SpyTag DA-DHFR-SpyCatcher and SpyTag-DHFR-SpyCatcher did not form insoluble aggregates.

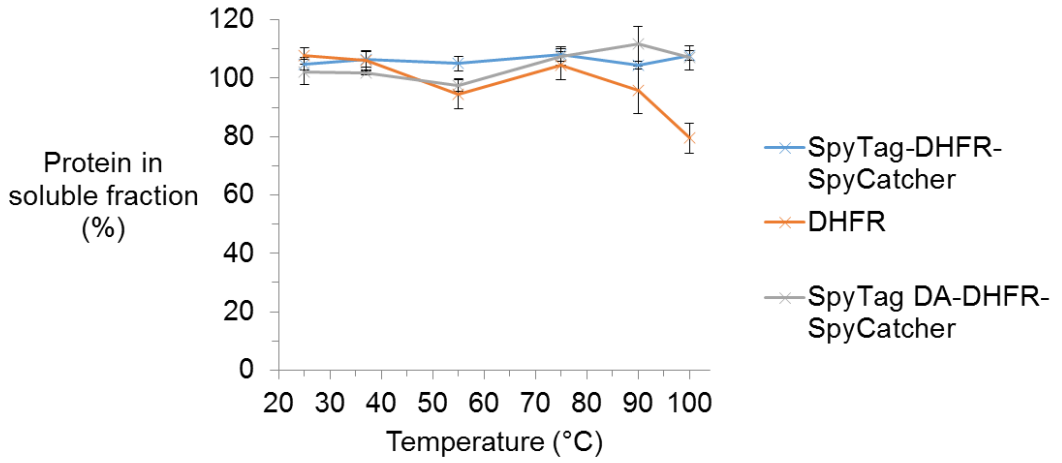


Figure 4.3: DHFR is highly soluble at high temperatures. Densitometry based quantification of amount of soluble SpyTag-DHFR-SpyCatcher, SpyTag DA-DHFR-SpyCatcher and DHFR in the supernatant after heating 20 µM of enzyme to the indicated temperature for 10 min in PBS containing 100 mM DTT, cooling to 10 °C and then spinning the sample at 17,000g for 30 min (mean of triplicate ± 1 standard deviation). Some data were previously published in (Schoene et al., 2014).

4.1.4 Effect of cyclization on recovered activity after heating

The activity of DHFR could be followed by measuring the absorbance at 340 nm which is the wavelength at which NADPH absorbs the light. In the presence of dihydrofolic acid, NADPH is oxidized to $NADP^+$ which does not absorb at 340 nm. While the effects of cyclization on the protein solubility was minor (Figure 4.3), the recovered activity after heating was much improved (Figure 4.4). The untagged DHFR lost most of its activity at 55 °C and the activity was not recovered at higher temperatures. SpyTag DA-DHFR-SpyCatcher lost even more activity at 55 °C, indicating that the addition of the SpyTag DA/SpyCatcher moiety decreased the stability of DHFR. This could be correlated with the decrease in Tm I measured using DSC (Figure 4.5). However, SpyTag-DHFR-SpyCatcher retained most of its activity after incubating the protein for 10 min up to 100 °C (Figure 4.4).

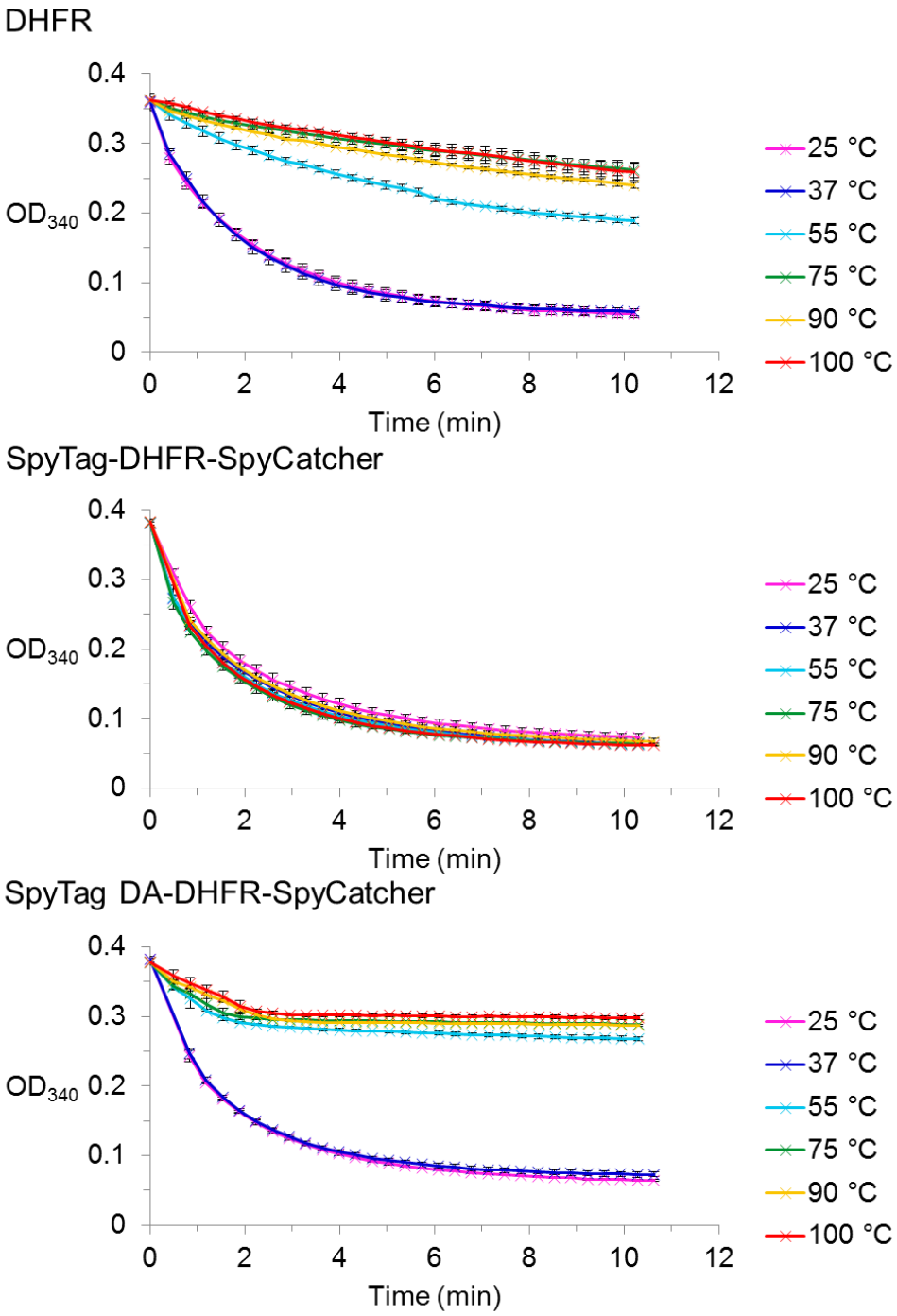


Figure 4.4: Cyclization improves the thermal resilience of DHFR. 20 µM DHFR, SpyTag-DHFR-SpyCatcher or SpyTag DA-DHFR-SpyCatcher were incubated for 10 min at the indicated temperature then cooled to 10 °C, before returning to RT and then monitoring the oxidation of NADPH at 340 nm by allowing 100 nM enzyme to react with 100 µM DHF in the presence of 100 µM NADPH (mean of triplicate ± 1 standard deviation). Some data were previously published in (Schoene et al., 2014).

4.1.5 DSC of DHFR constructs

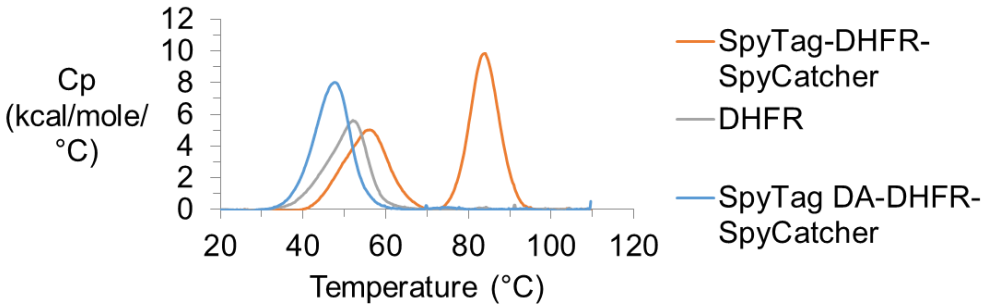


Figure 4.5: DSC indicated that DHFR was not stabilized by an increase in Tm. DSC traces of 20 µM DHFR, SpyTag-DHFR-SpyCatcher and SpyTag DA-DHFR-SpyCatcher obtained from scanning the proteins from 20-110 °C at 1 °C/min in PBS.

In Chapter 3 Figure 3.12 and 3.13 I discussed that the mechanism for thermal resilience in the case of BLA appeared to be due to an improvement of reversibility, rather than an improvement on the Tm of the domain. In the case of SpyTag-DHFR-SpyCatcher, the mechanism conferring thermal resilience appeared to be similar. I obtained a Tm of 52.1 °C for the untagged DHFR sample which is similar to the Tm values obtained previously for DHFR (Takahashi et al., 2007).

I obtained a Tm of 48.0 °C for the SpyTag DA-DHFR-SpyCatcher sample indicating that the SpyTag DA/SpyCatcher moiety destabilized the DHFR domain (Figure 4.5). Interestingly, I did not obtain a second peak for the SpyTag DA/SpyCatcher moiety as seen with SpyTag DA-BLA-SpyCatcher (Figure 3.13). In the case of SpyTag-DHFR-SpyCatcher, I obtained two peaks in the trace. The Tm's of the protein were 56.3 °C (DHFR domain) and 84.1 °C (SpyTag/SpyCatcher domain). Unlike SpyTag-BLA-SpyCatcher, the Tm of the DHFR domain in SpyTag-DHFR-SpyCatcher increased by 4.2 °C (Figure 4.5). However, the increase does not account for the improvement in thermal resilience I observed (Figure 4.4). Therefore, I hypothesize that the refold capability of DHFR was improved using a similar mechanism to that of BLA.

4.2 Cyclization of Firefly Luciferase (FLuc)

4.2.1 Introduction of FLuc

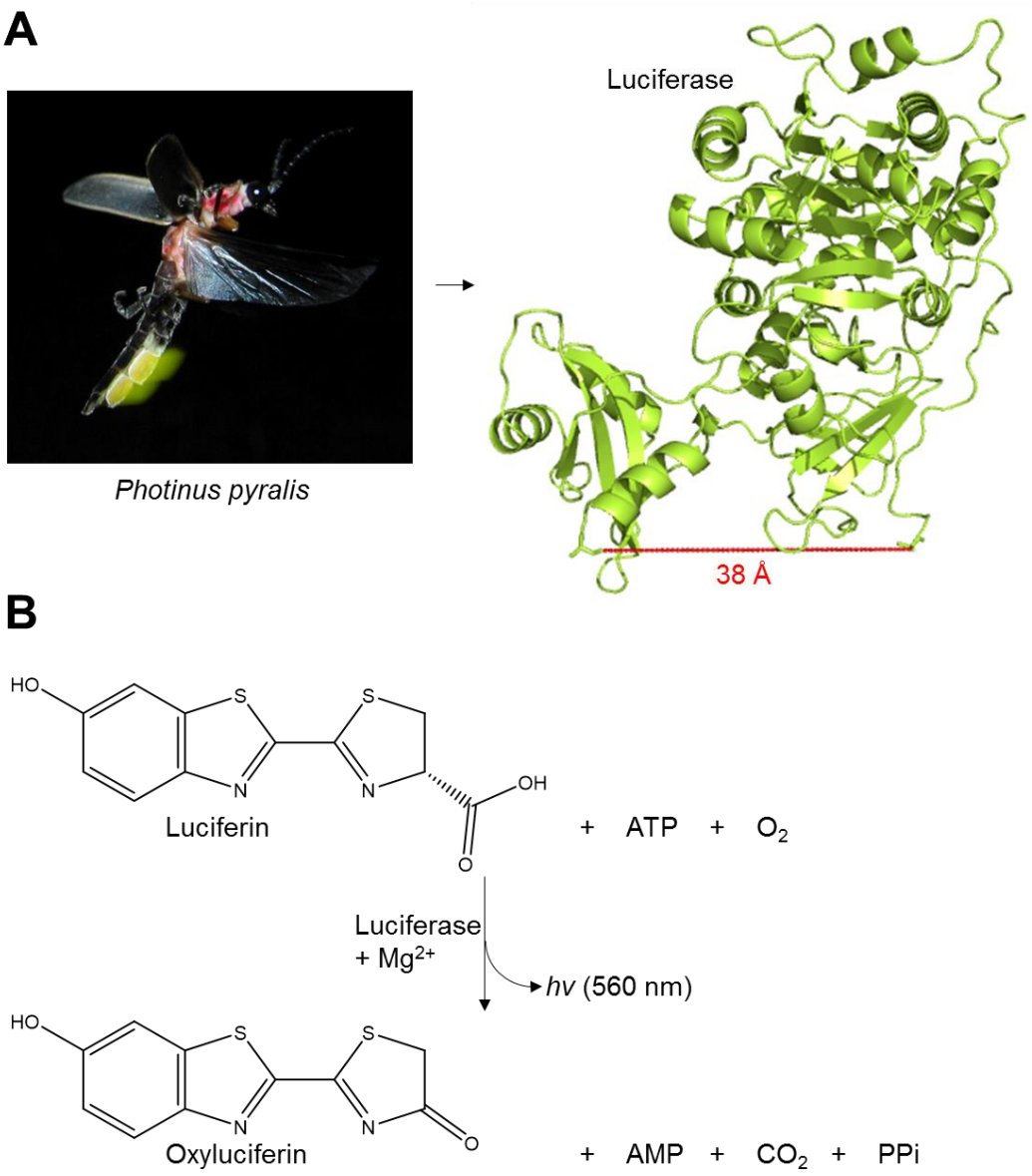


Figure 4.6: Structure and catalysis of Luciferase. A) Image of *Photinus pyralis* luminescing in flight (credit: Terry Priest (firefly w/ glow) [CC BY-SA 2.0 (http://creativecommons.org/licenses/by-sa/2.0)], via Flickr) and crystal structure of FLuc (PDB: 1BA3) (Franks et al., 1998). B) Catalysis of luciferin to oxyluciferin by Luciferase.

Firefly Luciferase (FLuc) is an important tool in the field of biotechnology and is primarily used as a reporter. FLuc has been used for the production of biosensors (Fan et al., 2008), studying gene expression (Gould and Subramani, 1988) and

protein-protein interactions (Paulmurugan et al., 2002), coupled to enzyme reactions to study the activity of other enzymes (Lundin et al., 1976) as well as for DNA sequencing (Ronaghi et al., 1996). However, one of the issues limiting FLuc's applicability is the enzyme's low stability (Ebrahimi et al., 2012). Therefore, FLuc was chosen as a target for cyclization due to the potential utility of a more stable FLuc.

FLuc originates from *P. pyralis*. The enzyme is produced in the light organ of males. The light organ can be clearly seen on the abdomen of the exemplar in Figure 4.6A. The fold of FLuc has been classified as an α/β fold. The enzyme is 550 amino acids long and the distance between the two termini in the crystal structure (Figure 4.6A) is 38 Å (Franks et al., 1998). FLuc catalyses the conversion of luciferin to oxyluciferin in the presence of ATP and oxygen, using magnesium as a cofactor. After the conversion to oxyluciferin, the oxyluciferin decays from an excited state to a ground state thereby releasing a photon of light (Marques and Esteves da Silva, 2009).

4.2.2 Effects of cyclization on FLuc

SpyTag-FLuc-SpyCatcher and SpyTag DA-FLuc-SpyCatcher were successfully produced in *E. coli* and purified using IMAC. When running the SpyTag DA-FLuc-SpyCatcher on SDS-PAGE, the protein migrated at around the expected molecular weight of ~77 kDa. However, SpyTag-FLuc-SpyCatcher migrated at an apparent molecular weight of ~120 kDa (Figure 4.7A). Therefore, FLuc was either successfully cyclized as a monomer or possibly as a dimer.

After I saw the promising results on the SDS-PAGE, I decided to test the solubility of the constructs after incubating them at different temperatures for 10 min. In Figure 4.7B we can see that there is a large improvement to the solubility of the cyclized construct when compared to the linear point mutant control. Then, I tested the residual activity of the FLuc after incubating the constructs at a different range of

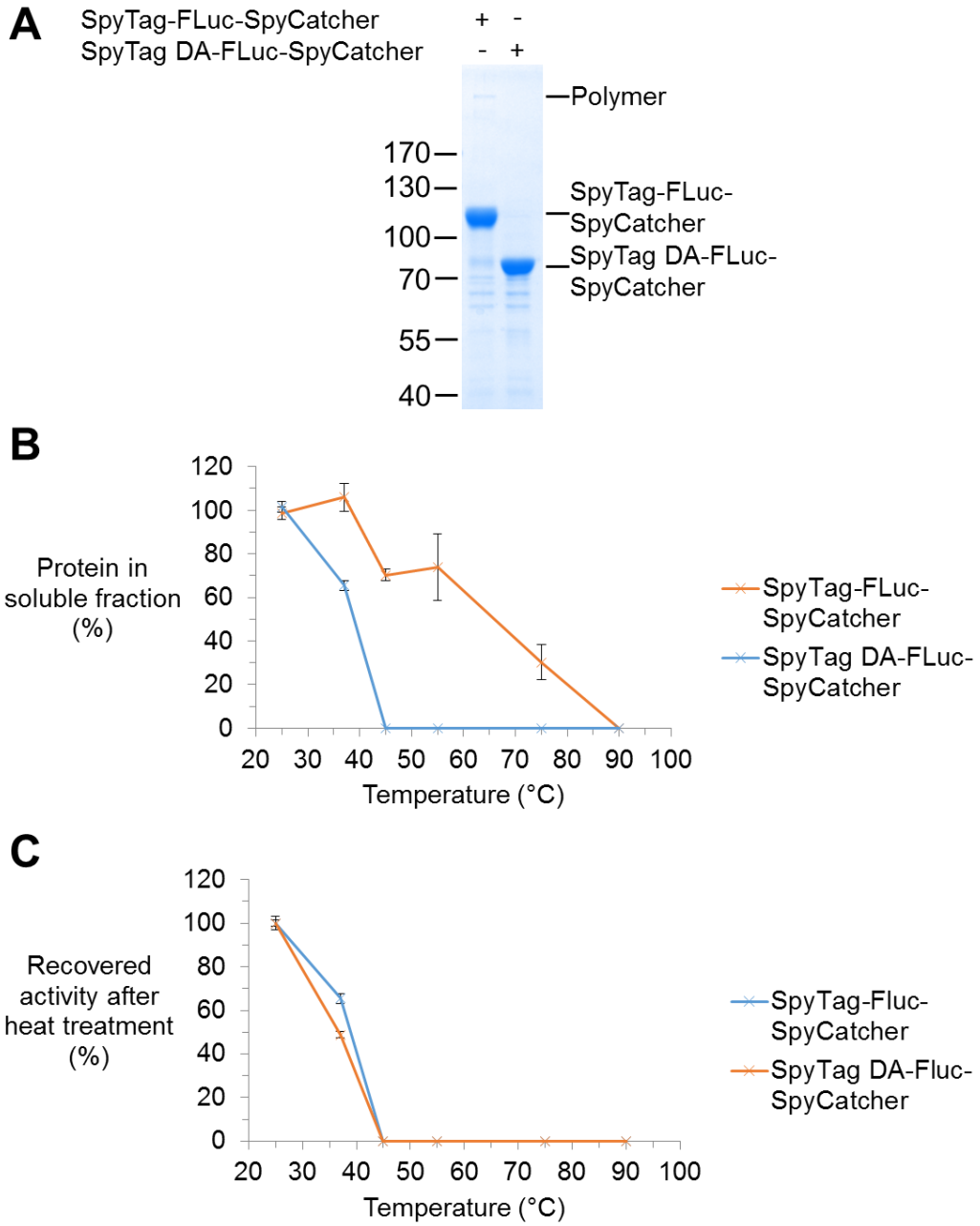


Figure 4.7: Sandwiching FLuc between SpyTag/SpyCatcher led to a small improvement in thermal resilience. A) Reduced SDS-PAGE of SpyTag-FLuc-SpyCatcher and the linear point mutant control, SpyTag DA-FLuc-SpyCatcher. B) Densitometry-based quantification of amount of soluble SpyTag-FLuc-SpyCatcher and SpyTag DA-FLuc-SpyCatcher in the supernatant after heating 5 µM enzyme to the indicated temperature for 10 min in PBS containing 100 mM DTT, cooling to 10 °C and then spinning the sample at 17,000g for 30 min. C) 5 µM SpyTag-FLuc-SpyCatcher or SpyTag DA-FLuc-SpyCatcher were incubated for 10 min at the indicated temperature in PBS containing 100 mM DTT, then cooled to 10 °C, before returning to RT and measuring the residual activity (mean of triplicate ± 1 standard deviation) by allowing 12.5 nM of enzyme to react with 300 µM luciferin in the presence of 4 mM ATP and 10 mM Mg. Some data were previously published in (Schoene et al., 2016b).

temperatures. I was only able to measure a small improvement to the residual activity after incubating the enzyme at 37 °C for 10 min (Figure 4.7C).

4.3 Cyclization of Phytase (PhyC)

4.3.1 Introduction of PhyC

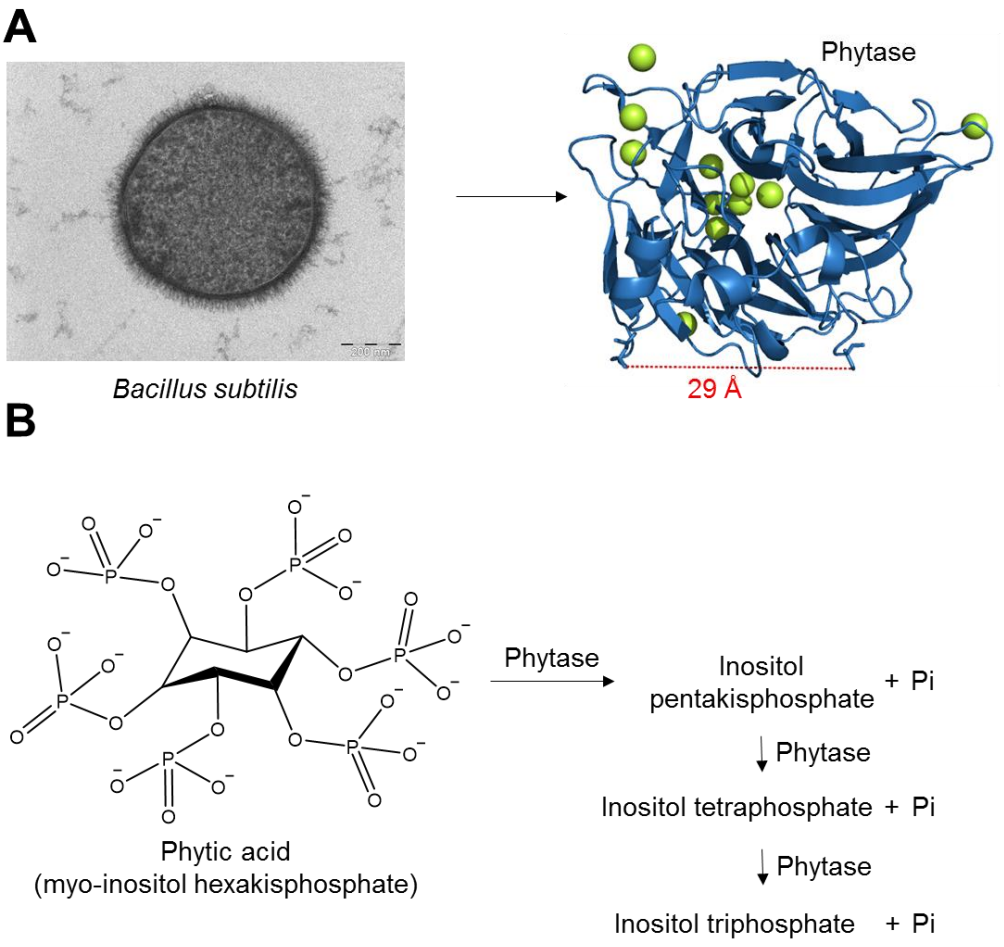


Figure 4.8: Structure and reaction mechanism of PhyC. A) Transmission electron micrograph of *B. subtilis* (credit: Allonweiner at en.wikipedia [Public domain], from Wikimedia Commons) and crystal structure of PhyC (PDB: 3AMR) (Zeng et al., 2011). B) Sequential hydrolysis of phytic acid catalysed by PhyC.

Phytic acid is a known anti-nutrient capable of chelating essential nutrients such as iron, magnesium and calcium (Bohlke et al., 2005; Selle et al., 2006). Moreover, phytic acid is the storage method of phosphates for many species of plants and cannot be digested by non-ruminant organisms. Therefore, it has also been shown

that including phytase, an enzyme capable of removing the inorganic phosphate groups from the phytic acid (Figure 4.8B), in animal feed stock increases the nutritional value. However, the process of feed stock production requires a high temperature incubation step which also rids the feed of pathogens such as *Salmonella* (Lei et al., 2013). Therefore, a phytase with a high thermal resilience would be beneficial for improving food security and so we chose to attempt to make a resilient phytase.

Previous work has shown that one can increase the thermal resilience of PhyC from *B. subtilis* using split inteins (Ping et al., 2010). Therefore, PhyC was chosen for isopeptide-mediated cyclization. PhyC is a β-propeller phytase which operates at an optimum pH of 7.0. PhyC relies on calcium both for its activity as well as its fold (Figure 4.8A) (Kerovuo et al., 1998). The distance between the termini in the crystal structure is 29 Å (Zeng et al., 2011) and it is therefore impressive that split intein-mediated cyclization yielded an active and resilient enzyme (Ping et al., 2010).

4.3.2 MS and gel shift of cyclized PhyC

Since PhyC is larger than BLA, I expected that SpyTag-PhyC-SpyCatcher would migrate slower than the non-cyclized point mutant on an SDS-PAGE. I was able to verify this hypothesis in Figure 4.9A. I also utilized MS to provide further evidence for successful cyclization by measuring the loss of mass of one water molecule for the cyclized construct (Figure 4.9B). SpyTag DA-PhyC-SpyCatcher did not lose a water molecule, based on the ProtParam-predicted mass (Figure 4.9C).

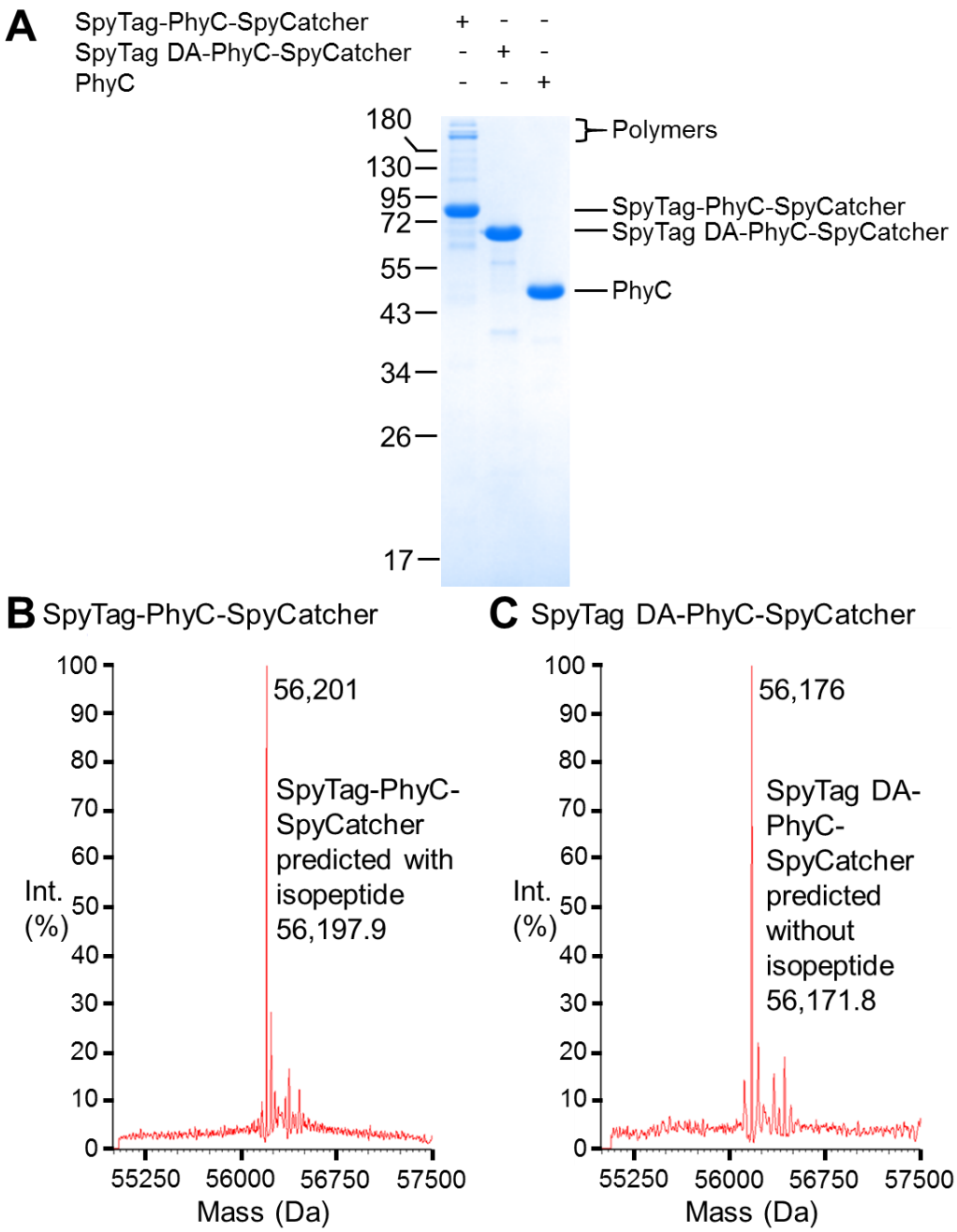


Figure 4.9: Gel shift and MS provide evidence for the successful cyclization of PhyC. A) SDS-PAGE of SpyTag-PhyC-SpyCatcher, SpyTag DA-PhyC-SpyCatcher and PhyC. B) MS of SpyTag-PhyC-SpyCatcher. Predicted molecular weight was 56,215.9 – 18 = 56,197.9 Da. C) MS of SpyTag DA-PhyC-SpyCatcher. Predicted molecular weight was 56,171.8 Da. Data were previously published in (Schoene et al., 2016a).

4.3.3 Effect of cyclization on protein aggregation of PhyC

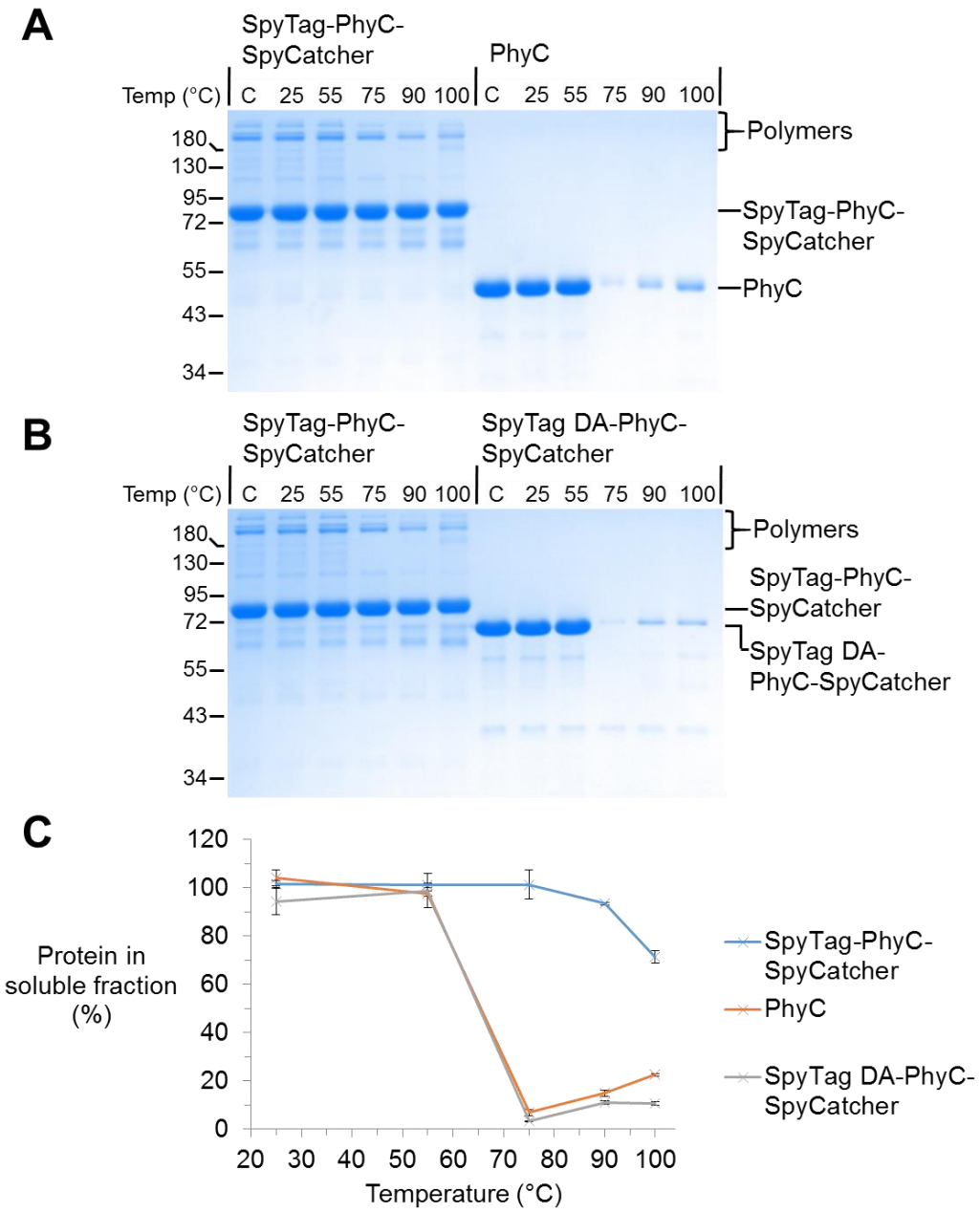


Figure 4.10: PhyC remains soluble at high temperatures when cyclized. A-B) Demonstrative gel of amount of soluble A) SpyTag-PhyC-SpyCatcher and PhyC and B) SpyTag-PhyC-SpyCatcher and SpyTag DA-PhyC-SpyCatcher in the supernatant after heating 10 µM enzyme to the indicated temperature for 10 min in 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$, cooling to 10 °C and then spinning the sample at 17,000g for 30 min. C) Densitometry based quantification of bands from A) and B) (mean of triplicate ± 1 standard deviation). Data were previously published in (Schoene et al., 2016a).

About 90 % of PhyC and SpyTag DA-PhyC-SpyCatcher formed insoluble aggregates at 75 °C. The amount of soluble protein was slightly higher at 90 and 100 °C for the aforementioned constructs. SpyTag-PhyC-SpyCatcher did not form insoluble aggregates at 75 °C. However, 5-10% of the soluble protein was lost after incubating SpyTag-PhyC-SpyCatcher at 90 °C and ~30% was lost after incubating at 100 °C. Since the pelleting process for feed stock production usually incubates the feed at 60-90 °C (Lei et al., 2013) for 5 min or less (Garrett et al., 2004), the results in Figure 4.10 were promising.

4.3.4 Cyclization improved the recovered activity after heating

PhyC is able to release inorganic phosphate from phytic acid. Therefore, the most common method to assay for phytase activity is to measure the amount of phosphate released by the phytase. Phosphate can be measured through a combination of ammonium molybdate, ferrous sulphate and sulphuric acid, which will turn the solution blue in the presence of phosphate (Sumner, 1944).

Cyclization led to a major improvement in recovered activity after heating of PhyC (Figure 4.11). While PhyC and SpyTag DA-PhyC-SpyCatcher lost most of their activity at 75, 90 and 100 °C, SpyTag-PhyC-SpyCatcher was still active at all these temperatures. Interestingly, SpyTag-PhyC-SpyCatcher exhibited a more gradual decay in activity than had been seen with the cyclized BLA constructs (Figure 3.7): the activity decreased slowly as the temperature increased.

The data in Figure 4.11 is encouraging since our data showed that cyclized PhyC could be a useful alternative to other phytases which need to be applied after the pelleting process (Lei et al., 2013).

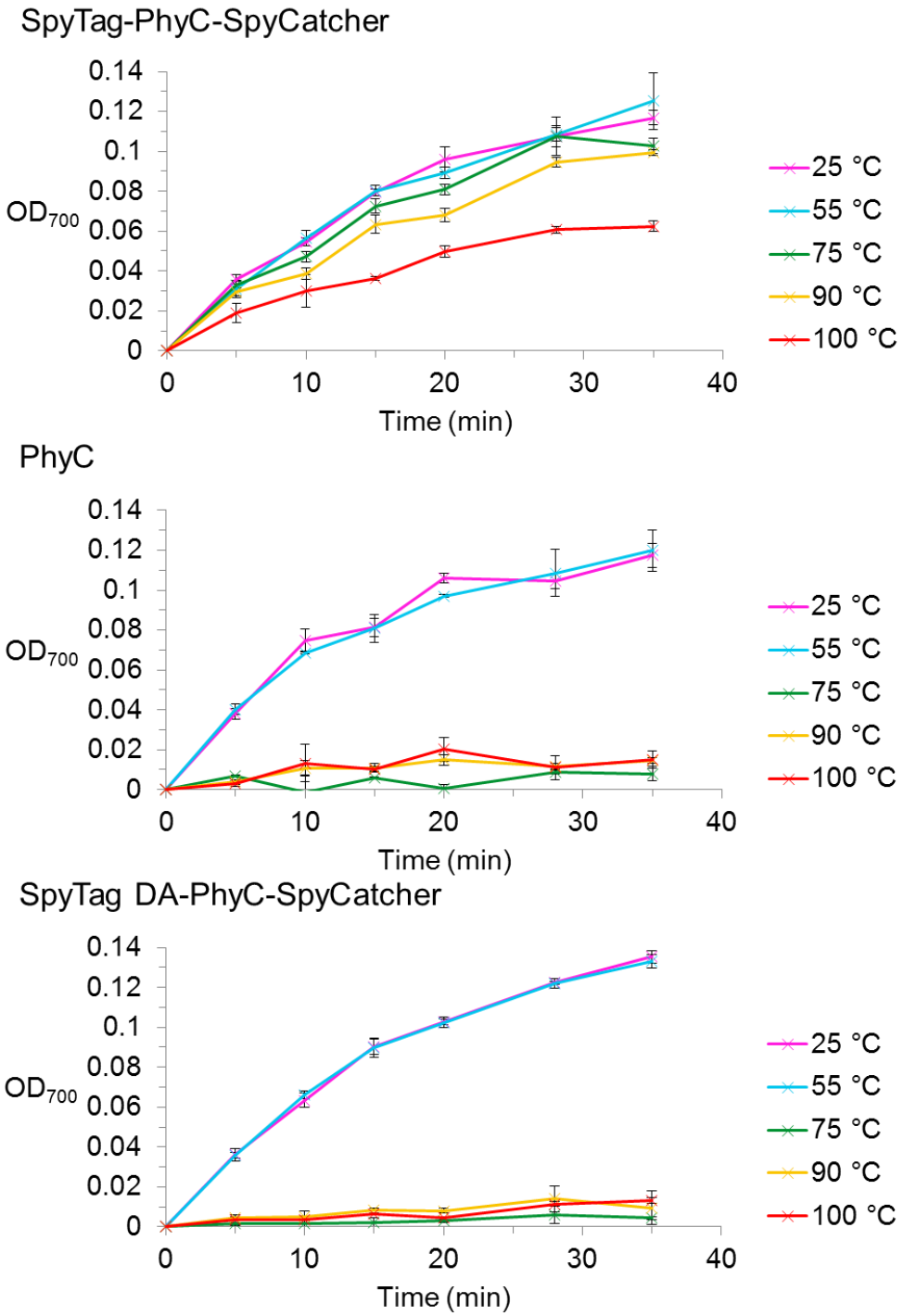


Figure 4.11: Isopeptide-mediated cyclization improves the thermal resilience of PhyC. 10 µM PhyC, SpyTag-PhyC-SpyCatcher or SpyTag DA-PhyC-SpyCatcher were incubated for 10 min at the indicated temperature in 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$, cooled to 10 °C and then an absorbance-based activity assay was run at RT using the molybdenum-blue method to quantify the amount of free phosphate (mean of triplicate ± 1 standard deviation). Data were previously published in (Schoene et al., 2016a).

4.3.5 DSC of PhyC constructs

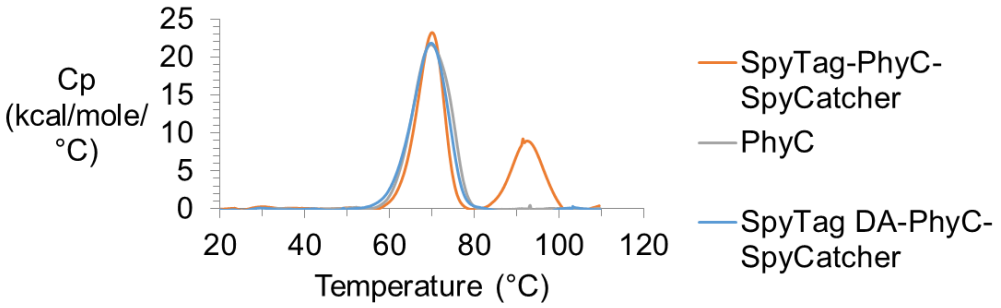


Figure 4.12: DSC indicated that PhyC was not stabilized by an increase in Tm. DSC traces of 20 µM PhyC, SpyTag-PhyC-SpyCatcher and SpyTag DA-PhyC-SpyCatcher obtained from scanning the proteins from 20-110 °C at 1 °C/min in 50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$. Data were previously published in (Schoene et al., 2016a).

Both PhyC and SpyTag DA-PhyC-SpyCatcher were shown to have a single Tm of 69.8 and 69.7 °C respectively (Figure 4.12). Like with DHFR (Figure 4.5), SpyTag DA-PhyC-SpyCatcher did not have a second Tm to account for the SpyTag DA/SpyCatcher moiety melting. SpyTag-PhyC-SpyCatcher was shown to have two unfolding transitions. The first accounts for the unfolding of the PhyC portion of the construct with a Tm of 70.0 °C. The second transition with a Tm of 91.5 °C was attributed to the unfolding of the SpyTag/SpyCatcher domain.

The Tm of the SpyTag/SpyCatcher domain was higher in the SpyTag-PhyC-SpyCatcher than for both the SpyTag-BLA-SpyCatcher (85.0 °C) and SpyTag-DHFR-SpyCatcher (84.1 °C) constructs. The difference might have been due to the different buffer (50 mM Tris-HCl pH 7.0 containing 2 mM $CaCl_2$) that I used to analyse SpyTag-PhyC-SpyCatcher or it might have been due to the PhyC domain having a stabilizing effect on the SpyTag/SpyCatcher domain.

4.3.6 DLS of PhyC showed little change in the size of cyclized PhyC after heating

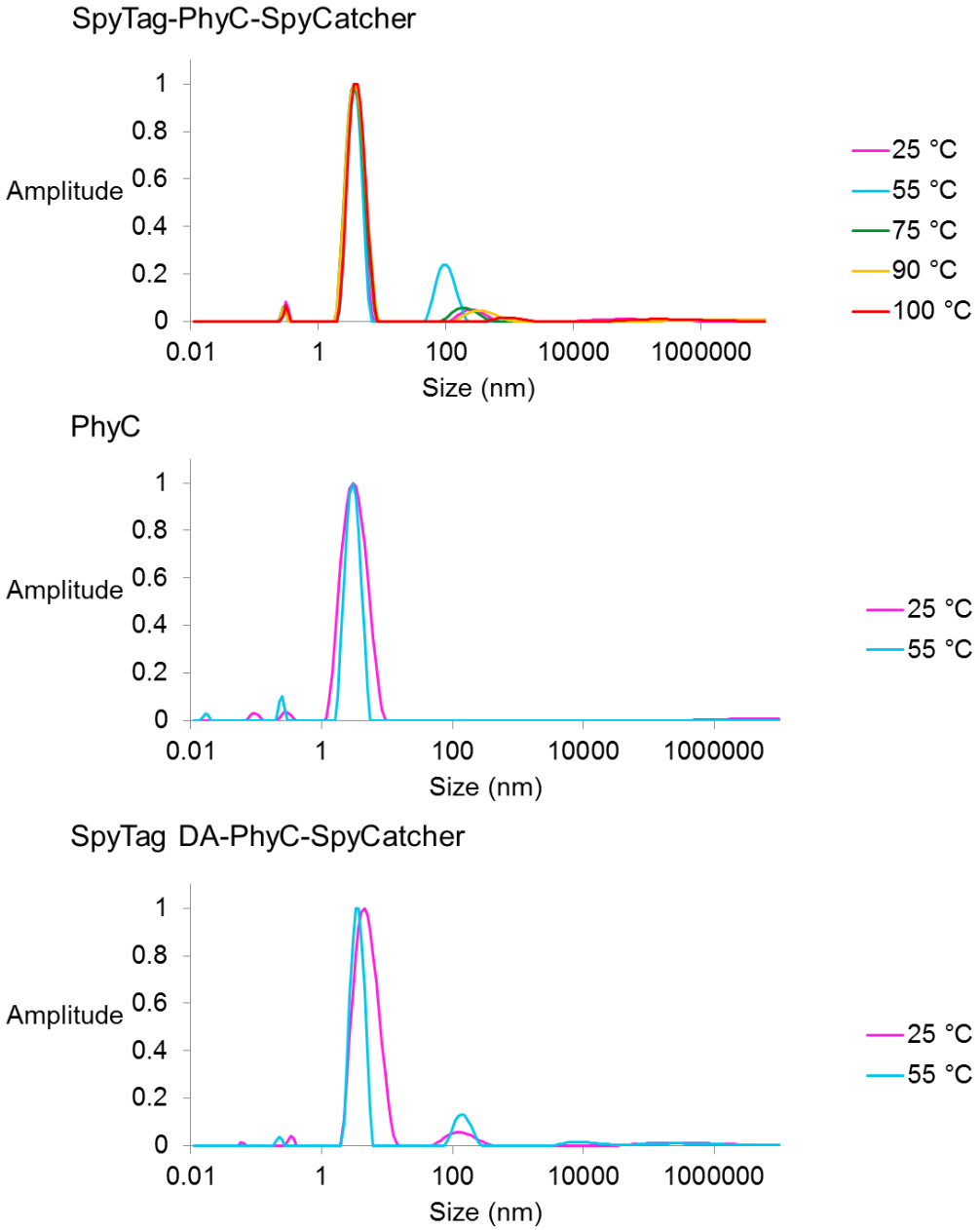


Figure 4.13: DLS showed little change in particle size after heating SpyTag-PhyC-SpyCatcher. Hydrodynamic radius (R_H) of SpyTag-PhyC-SpyCatcher, PhyC, SpyTag DA-PhyC-SpyCatcher after heating at the indicated temperatures for 10 min, cooling to 10 °C, filtering the sample through a 0.2 µm filter and measuring the light scatter at 20 °C.

The hydrodynamic radius (R_H) of PhyC and SpyCatcher + SpyTag peptide was predicted to be 1.7 nm and 2.7 nm respectively using HYDROPRO (Ortega et al., 2011) based on the atomic coordinates of PDB files 3AMR (Zeng et al., 2011) and

4MLI (Li et al., 2013). To ensure that I was measuring the size of the cyclized monomer, I purified the samples with ion exchange after the usual IMAC purification to remove the multimers in the SpyTag-PhyC-SpyCatcher sample.

Sample	Temperature (°C)	Size (nm)	%RSD
SpyTag-PhyC-SpyCatcher	25	3.85	24.9
	55	3.78	23.1
	75	3.90	29.3
	90	4.02	30.7
	100	4.13	27.6
SpyTag DA-PhyC-SpyCatcher	25	5.16	40.1
	55	3.60	21.8
	75	N/A	N/A
	90	N/A	N/A
	100	N/A	N/A
PhyC	25	3.41	41.4
	55	3.08	24.3
	75	N/A	N/A
	90	N/A	N/A
	100	N/A	N/A

Table 4.1: R_H size and polydispersity percentage values obtained from the data in Figure 4.13. RSD = N/A means that the sample aggregated and could not be analysed using DLS.

The value for the R_H of PhyC was variable but still close to the predicted size of the particle (Table 4.1), 2.7 nm. The R_H of SpyTag-PhyC-SpyCatcher was between 3.8 and 4.1 nm and the sample was quite monodispersed (23-31% RSD). Increasing the temperature had little effect on the polydispersity or the R_H of SpyTag-PhyC-SpyCatcher. Surprisingly, SpyTag DA-PhyC-SpyCatcher was larger and more polydisperse than SpyTag-PhyC-SpyCatcher at 25 °C. I postulated that SpyTag DA-PhyC-SpyCatcher might be forming multimers as the Tag and Catcher of multiple SpyTag DA-PhyC-SpyCatcher might have still been able to make transient complexes. The size of SpyTag DA-PhyC-SpyCatcher decreased from 5.2 nm to 3.6 nm after incubating the protein at 55 °C. The decrease in size could be attributed to the

irreversible denaturation of the SpyTag DA/SpyCatcher complex which I had previously shown to have a Tm of 54.0 °C (Figure 3.13A). The size of PhyC and SpyTag DA-PhyC-SpyCatcher could not be obtained at 75, 90 or 100 °C because the protein formed aggregates that could not pass through the 0.2 µm filter.

4.3.7 CD of PhyC constructs provided further evidence that cyclization improved the ability to refold

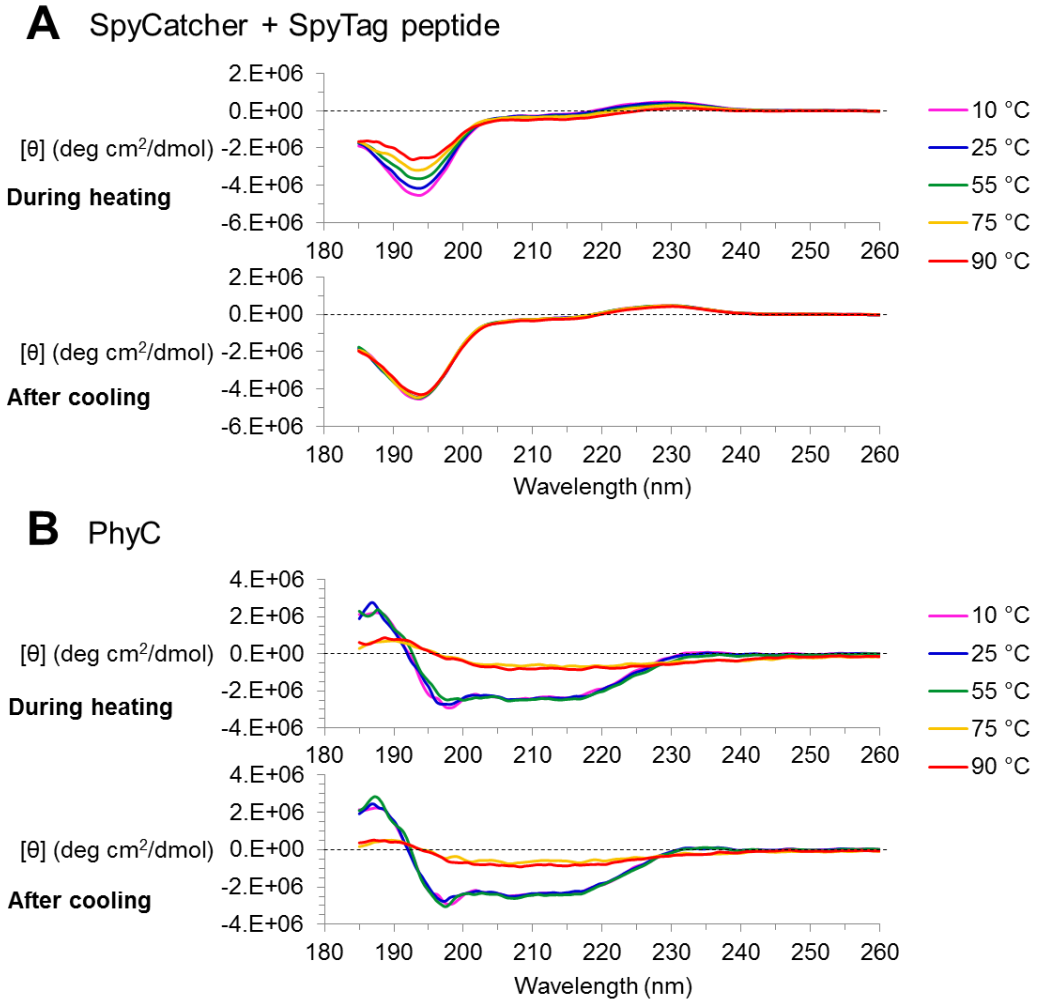


Figure 4.14: CD showed high refolding capability for SpyCatcher + SpyTag peptide but not for PhyC. Far-UV spectra of A) SpyCatcher + SpyTag peptide and, B) PhyC at the indicated temperature after 3 min of equilibration (During heating) or at 10 °C after the protein had been incubated at the indicated temperature for 10 min (After cooling). Some data were previously published in (Schoene et al., 2016a).

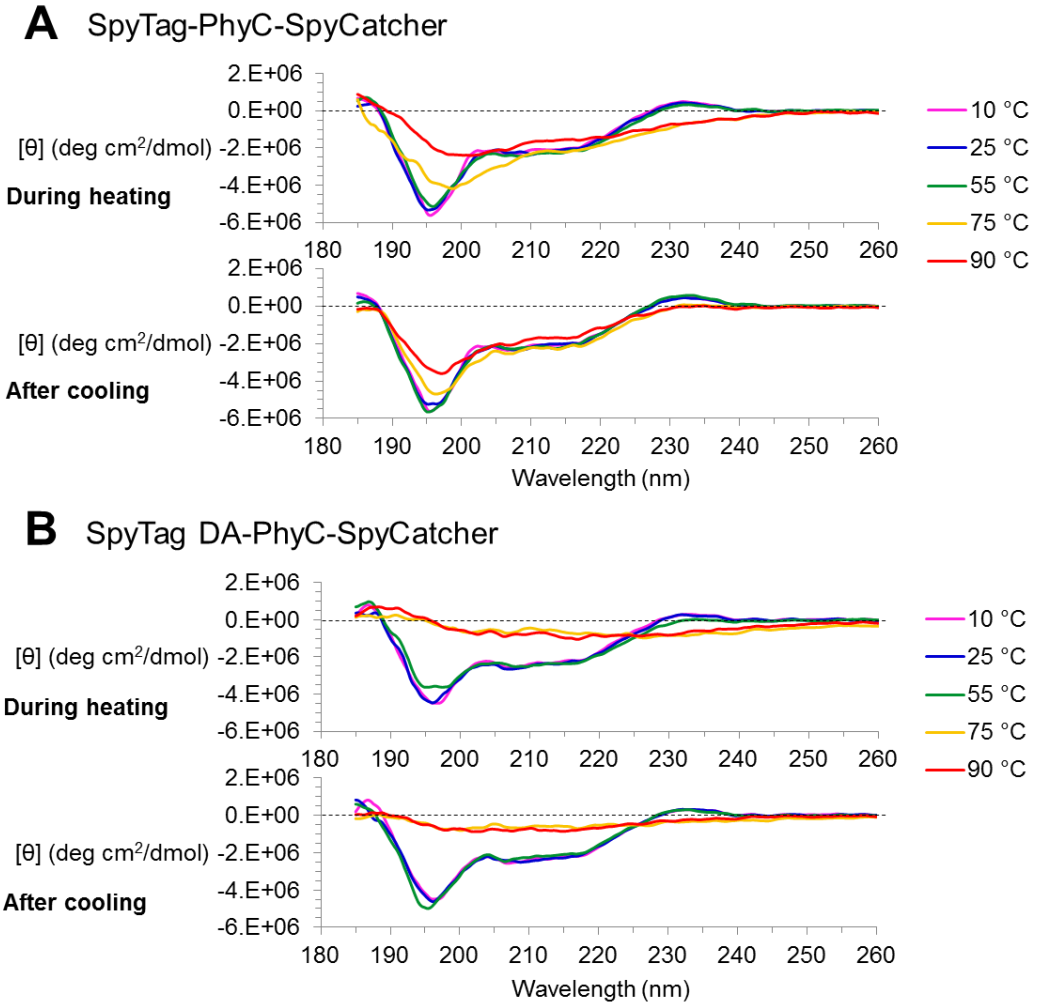


Figure 4.15: CD spectra showed refolding capability for SpyTag-PhyC-SpyCatcher but not for SpyTag DA-PhyC-SpyCatcher. Far-UV spectra of A) SpyTag-PhyC-SpyCatcher and B) SpyTag DA-PhyC-SpyCatcher at the indicated temperature after 3 min of equilibration (During heating), or at 10 °C after the protein had been incubated at the indicated temperature for 10 min (After cooling). Data were previously published in (Schoene et al., 2016a).

To gain insight on the effects of cyclization on the secondary structure, I used CD. I obtained the spectra of PhyC and a sample of SpyCatcher which had been allowed to react to completion with SpyTag peptide (Figure 4.14). The spectrum of SpyCatcher + SpyTag peptide was characterized by a trough with a local minimum at 195 nm which shifted upwards as the temperature increased. The spectrum did not contain the characteristics of a well-defined α-helix or β-sheet structure. Most

impressive was the protein's ability to regain its original secondary structure even after 90 °C incubation for 10 min (Figure 4.14A).

PhyC's spectrum was that of a well-folded protein containing α-helices and β-sheets. PhyC's spectrum remained unchanged at 25 or 55 °C. However, at 75 °C and 90 °C the signal decreased and PhyC was not able to recover, after cooling the protein to 10 °C (Figure 4.14B). I also observed aggregates in the cuvette after obtaining the spectra at 75 and 90 °C so I attributed this loss of signal to the PhyC aggregating.

The far-UV spectra of SpyTag-PhyC-SpyCatcher and SpyTag DA-PhyC-SpyCatcher were similar (Figure 4.15) with the exception of a more negative local minimum at 195 nm for SpyTag-PhyC-SpyCatcher. The spectra of SpyTag-PhyC-SpyCatcher appeared to be a combination of the characteristics obtained for the PhyC and SpyCatcher + SpyTag peptide spectra in Figure 4.14. SpyTag DA-PhyC-SpyCatcher behaved similarly to PhyC at 75 and 90 °C (Figure 4.14B and Figure 4.15A). However, SpyTag-PhyC-SpyCatcher was shown to unfold and 75 and 90 °C and then recover after cooling to 10 °C (Figure 4.15A).

4.4 Cyclization of Glucose Oxidase (GOx)

4.4.1 Introduction of GOx

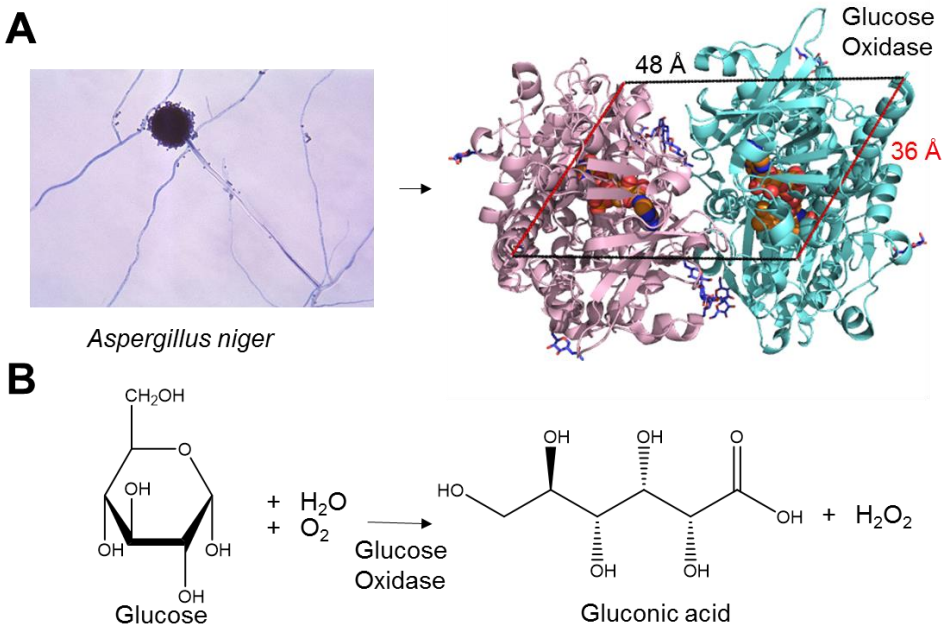


Figure 4.16: Structure and reaction mechanism of GOx. A) Photomicrograph of *Aspergillus niger* (credit: CDC/Lucille K. Georg, HHS, NIH [Public domain], via Public Health Image Library) and crystal structure of GOx (PDB: 3QVP) (Kommoju et al., 2011). GOx is a homodimer and the monomers have been coloured cyan and pink. Glycosylation is represented in stick form and the FAD is represented in sphere form. The red dashed line measures the intramolecular distance between the N and C-terminus and the black dashed line measures the intermolecular distance between the N and C-termini of the two monomeric units. B) Oxidation of glucose catalysed by GOx.

GOx is an industrially important enzyme most commonly known for its use in glucose biosensors (Heller and Feldman, 2008). Less commonly known, however, is the use of GOx in the production of low alcohol wines (Pickering et al., 1999). Sekisui Diagnostics, the company partly funding my research, is one of the leading manufacturers of GOx. There was interest in producing a more thermostable GOx that might have a longer shelf life than what is currently available on the market. Therefore, I decided to test the effects of isopeptide-mediated cyclization on GOx.

The most common GOx on the market originates from *A. niger* (Ferri et al., 2011). *A. niger* is a species of fungus which is known to reproduce by forming a

complex of cells called a conidial head which can be seen in Figure 4.16A. GOx is a glycoprotein which is known to be homodimeric (Zoldák et al., 2004). SpyTag/SpyCatcher could either react between the dimers (48 Å) or cyclize the monomer (36 Å) (Kommoju et al., 2011).

GOx utilizes flavin adenine dinucleotide (FAD) to oxidize β-D-Glucose to δ-gluconolactone thereby releasing H_2O_2. The δ-gluconolactone is then spontaneously hydrolysed to form D-gluconic acid. A simplified reaction scheme can be found in Figure 4.16B.

4.4.2 Effects of sandwiching GOx between SpyTag/SpyCatcher

I expressed GOx and SpyTag-GOx-SpyCatcher in *Pichia pastoris*, secreted using an α-factor secretion signal sequence (Meng et al., 2014). I tested two different media for induction, buffered methanol minimal medium (BMMH) and buffered methanol complex medium (BMMY). Both media led to the expression of GOx and SpyTag-GOx-SpyCatcher (Figure 4.17). However, we decided to work with BMMH because expression of GOx and SpyTag-GOx-SpyCatcher in BMMY led to a chromophore impurity in the sample which was difficult to get rid of (Moore et al., 2010).

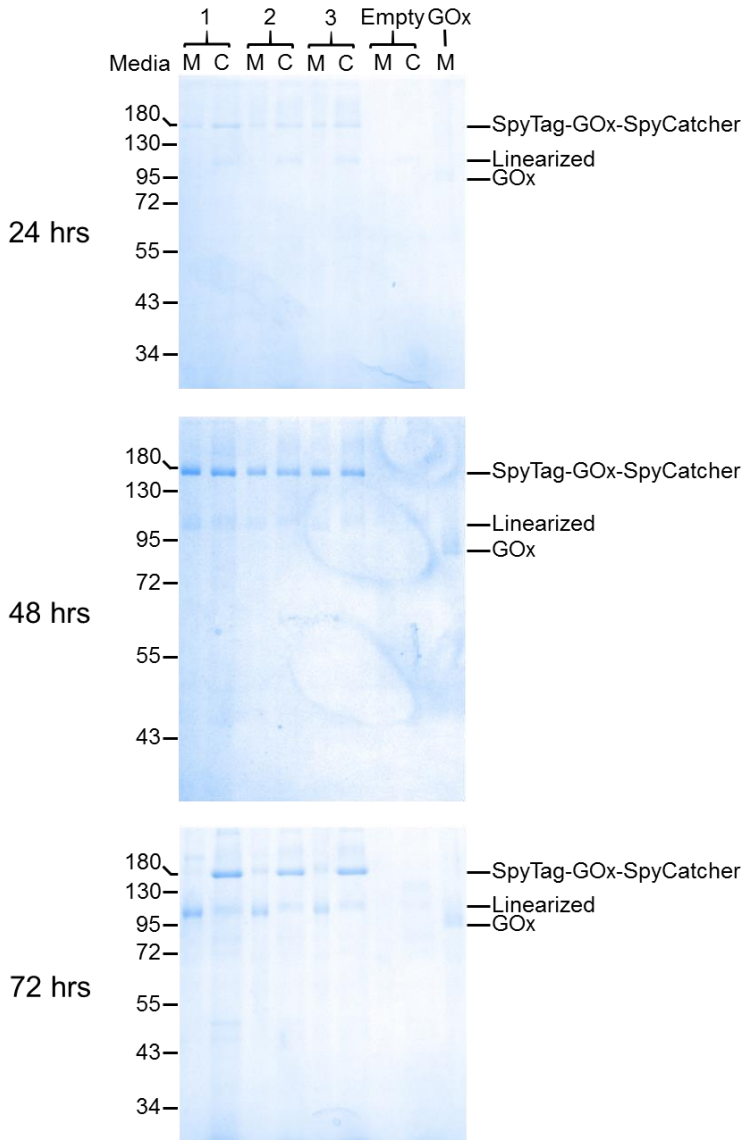


Figure 4.17: Expression of GOx and SpyTag-GOx-SpyCatcher in *P. pastoris*. 3 separate clones containing the SpyTag-GOx-SpyCatcher gene, 1 clone transformed with an empty vector and 1 clone containing the GOx gene were expressed. Induction was done in BMMH (M) or BMMY (C) media. Samples were spun at different time points to pellet the cells and the supernatant was removed and analysed by SDS-PAGE.

Untagged GOx had an apparent molecular weight of ~90 kDa on SDS-PAGE (Figure 4.18A). SpyTag-GOx-SpyCatcher produced two bands that could be purified from the media. The lower band had an apparent molecular weight of ~105 kDa and the higher band had an apparent molecular weight of ~140 kDa. During the second and third day of methanol-induced induction the species running at ~140 kDa disappeared, leaving only the species at a molecular weight of ~105 kDa (Figure 4.17). Since it has been shown that the concentration of proteases in the media increases

between 48 and 72 h (Sinha et al., 2005), I suggest that the species at ~105 kDa was a cleaved and linearized SpyTag-GOx-SpyCatcher and the species at ~140 kDa was the cyclized SpyTag-GOx-SpyCatcher construct. This hypothesis would need to be confirmed using MS.

Since GOx's termini are quite far apart from one another, a concern was that cyclization might drastically decrease activity. In a preliminary test of catalytic activity by measuring the initial rates of GOx and SpyTag-GOx-SpyCatcher under the same conditions I found that SpyTag-GOx-SpyCatcher was about 25% less active than GOx (Figure 4.18B). The loss in activity might be due to cyclization or the proteolytic cleavage of SpyTag-GOx-SpyCatcher.

Unfortunately, I found that both the aggregation resistance (Figure 4.18C) and the recovered activity after heating (Figure 4.18D) was largely unaffected by the inclusion of the SpyTag/SpyCatcher domain. Even adding additives to the buffer such as K_2SO_4 (Gouda et al., 2003), xylitol (Cioci and Lavecchia, 1994) and trehalose (Paz-Alfaro et al., 2009) was not more beneficial to the stability of SpyTag-GOx-SpyCatcher than to the stability of GOx (Figure 4.19). Adding FAD to the buffer as an additive also did not seem to improve the reversibility of SpyTag-GOx-SpyCatcher's denaturation in comparison to that of GOx (Figure 4.19).

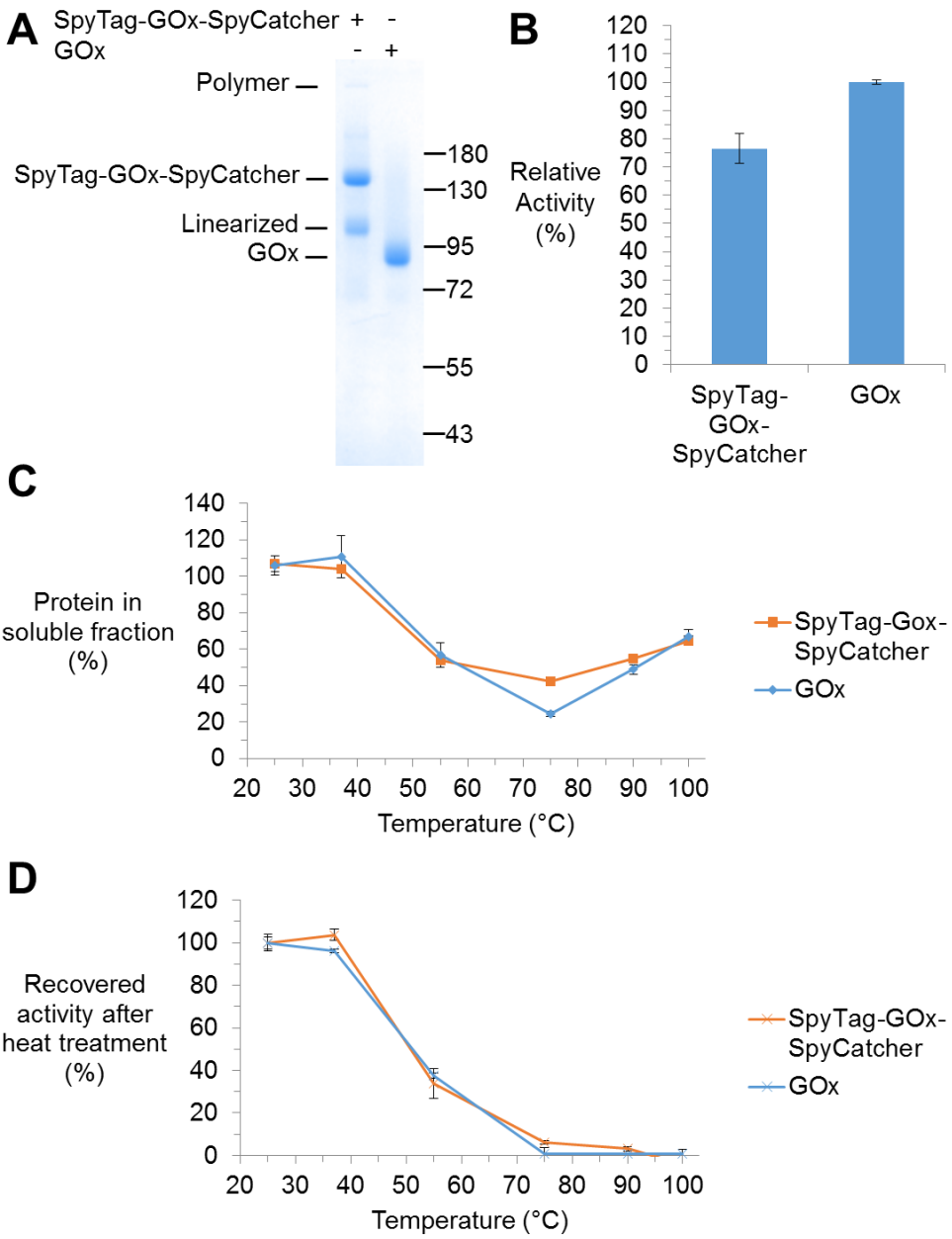


Figure 4.18: Sandwiching GOx between SpyTag/SpyCatcher led to no noticeable improvement in thermal resilience. A) Reduced SDS-PAGE of SpyTag-GOx-SpyCatcher and the untagged GOx. B) Relative initial activity of SpyTag-GOx-SpyCatcher in comparison to GOx at 34 °C. The initial activity of GOx was used as the 100% value. C) Densitometry-based quantification of amount of non-aggregated SpyTag-GOx-SpyCatcher and GOx in the sample after heating to the indicated temperature for 10 min, cooling to 10 °C and then spinning the sample at 17,000g for 30 min. D) SpyTag-GOx-SpyCatcher and GOx were incubated for 10 min at the indicated temperature and then cooled to 10 °C, before returning to RT and then measuring the residual initial activity. 100% was defined as the initial activity of each enzyme after incubating at 25 °C (mean of triplicate ± 1 standard deviation).

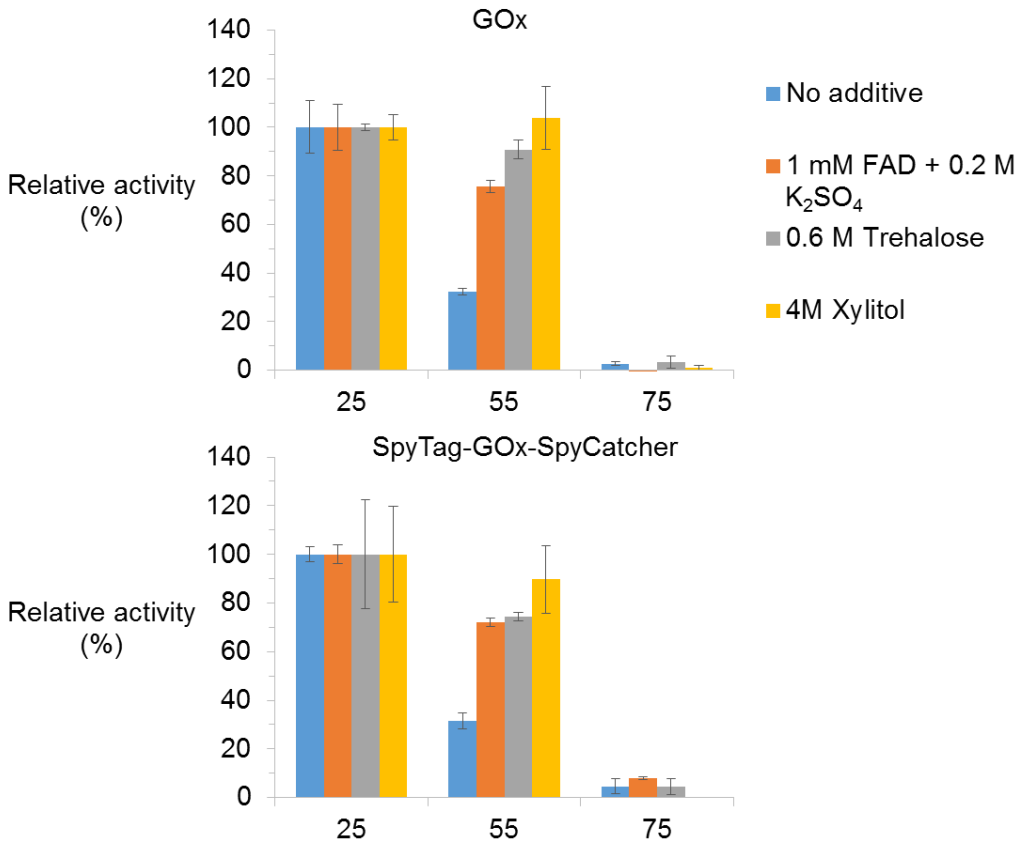


Figure 4.19: Additives did not preferably stabilize SpyTag-GOx-SpyCatcher. Samples were incubated in PBS pH 7.4 with different additives for 10 min at the indicated temperature and then cooled to 10 °C, before returning to RT and then measuring the residual initial activity. 100% was defined as the initial activity of each enzyme after incubating at 25 °C in each buffer (mean of triplicate ± 1 standard deviation).

4.5 Discussion

SpyTag/SpyCatcher-mediated cyclization was capable of enhancing the resilience of three distinct enzymes, BLA, DHFR and PhyC. However, I was not able to increase the resilience of GOx and FLuc. Therefore, while isopeptide-mediated cyclization can be a powerful tool to enhance a currently unknown number of enzymes, the current method has room for improvement. The most important concepts to study in the future include obtaining predictive power to know when isopeptide-mediated cyclization will be effective as well as improving the current SpyTag/SpyCatcher system. For both of these areas of study, a better understanding of the mechanism of thermal resilience and a better understanding of protein folding will be crucial.

Between BLA, DHFR and PhyC there is little similarity in terms of tertiary structure (Figure 3.1, 4.1, and 4.8). Moreover, the distance between the termini range from 7 to 29 Å providing evidence that there is a degree of flexibility in the system. The sizes of the three proteins are 29 (BLA), 18 (DHFR) and 39 (PhyC) kDa while the size of the SpyTag/SpyCatcher domain is 14 kDa. Therefore, SpyTag/SpyCatcher is able to stabilize enzymes that are 2.8-fold larger which is an improvement over using a formylmethanofuran:tetrahydromethanopterin formyltransferase (Ftr) fusion which is only able to keep proteins soluble if they are 0.6-fold the size (~20 kDa) of the Ftr domain (de Marco et al., 2004).

Coenzymes/cofactors also do not appear to be an issue as DHFR requires NADPH for activity (Morrison and Stone, 1988) and PhyC requires calcium for activity and stability (Kerovuo et al., 1998). Therefore, I have been able to stabilize monomeric enzymes of bacterial origin where the maximum length between the termini is 29 Å, maximum size has been 39 kDa and the enzyme is or is not cofactor-dependent.

DSC and CD data presented in this chapter indicate that the stabilization I have achieved with SpyTag/SpyCatcher uses a similar mechanism in all three cases. SpyTag/SpyCatcher does not increase the Tm of the target domain but instead increases the reversibility of the denaturation.

FLuc is 61 kDa and GOx is 63 kDa in size thereby far exceeding the maximum size where I have shown that cyclization would be effective. Moreover, the distance between the termini of FLuc is 38 Å and GOx is 36 Å. Therefore, one possibility why cyclization was ineffective in the case of FLuc and GOx might be that there is a maximum length of amino acids where cyclization can still introduce rigidity to the protein backbone. If the chain of amino acids is too long, it might be necessary to introduce internal isopeptide bonds to rigidify the protein backbone. Therefore, it might be necessary to test whether GOx and FLuc can be stabilized by inserting SpyTag in different loops within the protein sequence.

A good understanding of the irreversibly denatured states of a protein or the process of irreversible denaturation is also crucial in determining whether the protein is a good target for cyclization. For example, FLuc's misfolded conformation was found to have the N and C-terminus close to each other (Scholl et al., 2014) which might explain the ineffectiveness of cyclization. FLuc also requires chaperones to refold after heat-denaturation (Schröder et al., 1993) and therefore, enzymes that require chaperones to fold correctly might not benefit from cyclization.

GOx is a homodimeric glycoprotein and since I only tested one example of whether cyclization could stabilize a multimeric protein or whether glycosylated proteins benefit from cyclization, it is difficult to say whether glycosylated or multimeric proteins should be avoided. However, GOx's denaturation has been linked to the irreversible dissociation of $FADH_2$. Also, the denatured state of GOx is able to form dimers, trimers and tetramers (Zoldák et al., 2004). Moreover, GOx retains ~70% of its native secondary structure ellipticity after thermal stress (Zoldák et al., 2004). Therefore, GOx's irreversible denaturation might be dependent on a small localized change near the $FADH_2$ binding pocket which might not be influenced by the tethering of the N- and C-termini.

Isopeptide-mediated cyclization was effective at increasing the thermal resilience of PhyC. This result is encouraging as phytases are valuable enzymes in the agricultural industry and there is a need for a more heat-resilient phytase (Lei et al., 2013). An advantage of PhyC is that it is active in the presence of high concentrations of calcium, whereas histidine acid phytases are affected negatively by the calcium (Lei et al., 2013). Therefore, PhyC might be useful for preparing the calcium-rich feed that egg-laying hens need to produce a lot of eggs. However, PhyC is only able to hydrolyse 3 of the 6 phosphates in a phytic acid molecule (Kerovuo, 2000) whereas most fungal phytases are able to hydrolyse 5 of the 6 phosphates (Wyss et al., 1999). Therefore, stabilization of a fungal phytase or other bacterial phytases might be a worthwhile future endeavour.

Chapter 5: Phage display as an approach to improve the SpyTag/SpyCatcher interaction.

Chapter 5: Phage display as an approach to improve the SpyTag/SpyCatcher interaction

As discussed in chapter 3, the Tm of the isopeptide-containing domain might be linked with the thermal resilience of a cyclized enzyme. A faster reaction rate might increase the yield of the monomeric species in comparison to the multimeric species thereby increasing the monodispersity and yield of the monomeric product. Moreover, increasing the solubility of the Tag/Catcher domain further could also be beneficial for increasing the yield of cyclized enzyme. Phage display and directed evolution might lead us to discover novel tag/catcher pairs which are more stable, soluble and have a faster reaction rate.

The SpyTag/SpyCatcher interaction has a second order rate constant of 1.4 x 10^3 M^{-1} s^{-1} (Zakeri et al., 2012). A diffusion-limited association of two folded proteins has been hypothesized to be between 10^5 and 10^6 M^{-1} s^{-1} (Rogers et al., 2013). Therefore, there might be a possibility to increase the reaction rate constant of SpyCatcher/SpyTag by a factor of 10^2 to 10^3. A faster reacting SpyCatcher/SpyTag pair could have multiple uses which are not currently possible with the SpyCatcher/SpyTag system. The uses could include: capturing molecules at very low concentrations for proteomics studies or trapping molecular motors like myosin (Hartman and Spudich, 2012) in certain conformational states. Moreover, the discovery of novel Tags and Catchers might lead to the development of even more orthogonal pairs (Veggiani et al., 2016).

To address the aforementioned potential, I decided to try and set-up a phage display (Smith, 1985) approach that would allow us to screen and select for new and improved SpyTag/SpyCatcher pairs. Phage display would allow us to display the Tags or Catchers on the surface of a phage particle which would contain the sequence information within the phage's genome, thereby directly linking the phenotype to the genotype (Smith, 1985). Combining phage display technology with directed evolution

can be a powerful tool for the selection of strong interactions (McCafferty et al., 1990; Zhou et al., 2007).

In this chapter, I describe how I set up a phage display screen for the selection of novel Tags capable of forming an isopeptide bond with SpyCatcher. Then, I developed a selection method which allowed me to extract peptides able to form a covalent interaction with SpyCatcher from a commercially available phage display peptide library. I developed a method using qPCR to quantify the amount of phage DNA in a sample (Anderson et al., 2011). Using the qPCR screen I found that none of the clones that I tested appeared to be faster than the SpyTag/SpyCatcher reaction. Moreover, the library became enriched with a single clone that did not react faster than SpyTag/SpyCatcher. However, after a 1 amino acid mutation, the peptide exhibited interesting thermal properties when substituted for the SpyTag in the SpyTag-BLA-SpyCatcher construct.

5.1 Design of the phage display selection approach

5.1.1 Introduction to the phage display approach

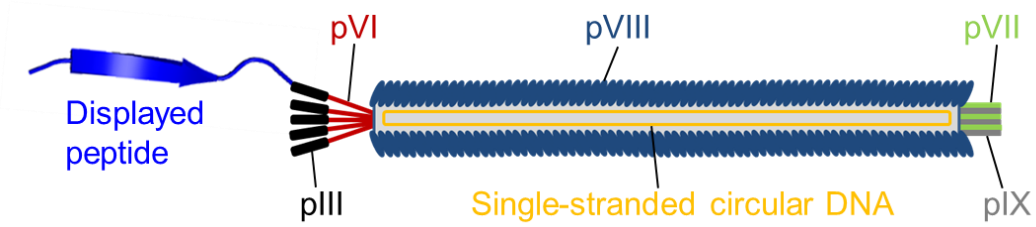


Figure 5.1: Diagram of a peptide displaying M13 filamentous phage. The coat proteins have been labelled and the peptide has been fused to the N-terminus of the pIII coat protein (parts not drawn to scale).

I chose M13 filamentous phage for our phage display approach. By fusing the protein to be displayed on the N-terminus of the pIII protein it is relatively easy to control the amount of displayed protein through the use of phagemids (Qi et al., 2012), helper phages (Russel et al., 1986) and/or hyperphages (Rondot et al., 2001). pIII is usually present as 5 copies on a phage but, through a combination of a helper phage (eg, R408) and phagemid (eg, pFab5cHis), the displayed protein or peptide can be made monovalent (Clackson and Lowman, 2004) as in Figure 5.1. Initial selection optimization was carried out with SpyTag and SpyTag DA phagemids because both of these peptides had already been cloned into the pFab5cHis phagemid. Since we still had aliquots of a commercially available phage display library (PhD-12) library, we tested out the selection and screening protocols on this library. It is worth noting that the PhD-12 library is M13KE-derived and therefore each phage displayed ~5 copies of the peptides. Multivalency could complicate the selection process due to an increase in the avidity of the multivalent phages.

A disadvantage of using the M13 filamentous phage is that the proteins need to be secreted to the periplasm for phage assembly and this might be problematic for the display of proteins that fold well in the cytosol (Velappan et al., 2010).

5.1.2 Design of panning to select for better Tags

When selecting for covalent bonds it is important to be able to use harsh washing conditions to remove any non-covalently bound phage. To achieve this, we used a Sulfo-LC-SPDP biconjugate linker to covalently bind a SpyCatcher containing an N-terminal cysteine residue to amine-coated beads. The selection strategy is depicted in Figure 5.2A.

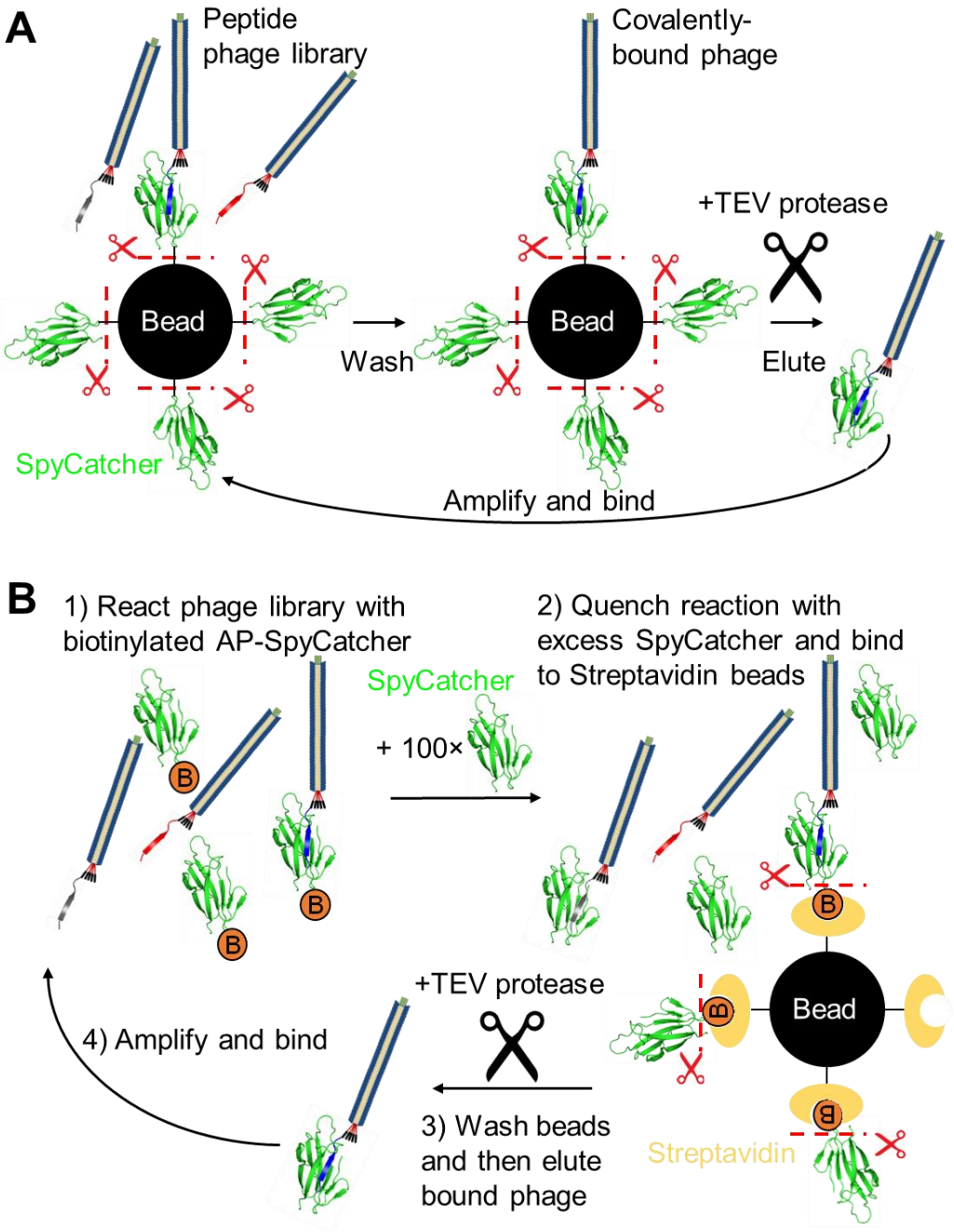


Figure 5.2: Selection strategies for peptide selection. A) Strategy to select for phages that form a covalent bond with SpyCatcher. B) Strategy for the selection of phages that form fast interactions with SpyCatcher.

While applying the selection strategy outline in Figure 5.2A I found it difficult to titre the amount of SpyCatcher I was using in the reaction. Therefore, to ensure that I would be pulling out the peptides with the fastest on-rate (Dueñas et al., 1996), I developed a second selection approach which used biotinylated AP-SpyCatcher as the bait molecule. The AP-tag (acceptor peptide), a 15 amino acid peptide tag with the sequence GLNDIFEAQKIEWHE, is a substrate for *E.* coli biotin ligase (BirA) which biotinylates the lysine residue (Fairhead and Howarth, 2015). The biotinylated AP-SpyCatcher could be titered down to lower concentrations and could then be captured using streptavidin coated superparamagnetic beads (Figure 5.2B).

During initital optimization of the two selection procedures outlined in Figure 5.2, I tested the recovery after TEV cleavage by first cleaving with TEV as outlined in Table 5.1 and then testing the titre of the eluate using a phage infectivity assay. Then, I incubated the samples with PBS containing 100 mM DTT at 25 °C 1,000 RPM for 1 hour and tested the titre of the second eluate. TEV protease was able to cleave ~90 % of the phage and the rest was recovered after eluting with DTT. TEV protease was aliquoted into single use vials and frozen at -80 °C to ensure that TEV cleavage would be effective throughout all my selections.

5.2 Verification of phage display as a potential approach for the selection of improved peptides

5.2.1 Western Blot verified the formation of a covalent bond between the SpyTag-displaying phages and SpyCatcher

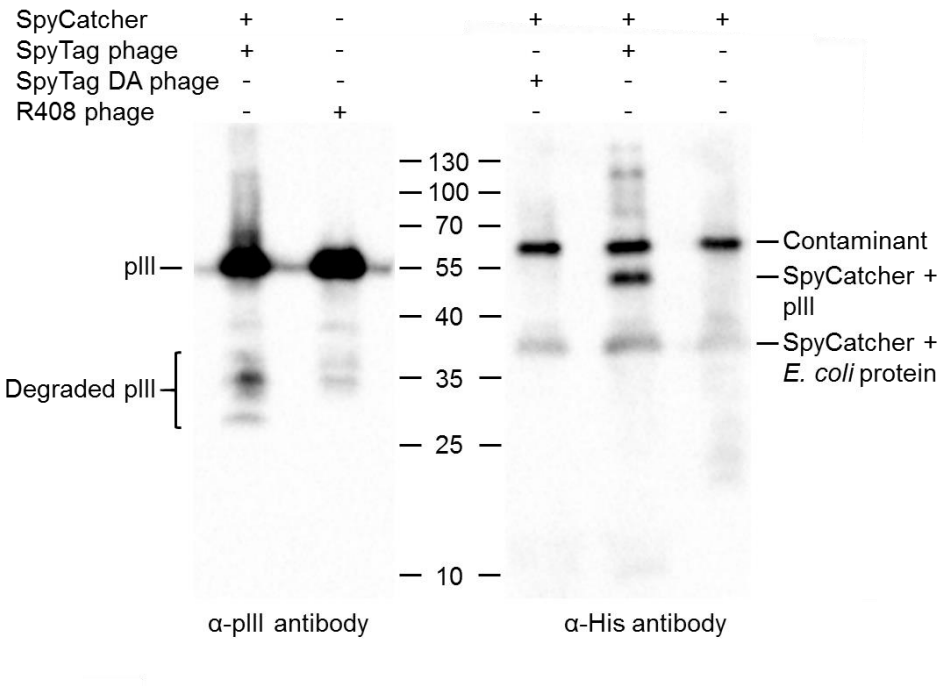


Figure 5.3: α-pIII and α-His Western Blots of SpyTag phage provide some evidence for the formation of a second product.

As described previously (Zakeri, 2011), one cannot see the emergence of a new band when mixing SpyTag phage with SpyCatcher protein when using an α-pIII antibody (Figure 5.3). Instead, one can only see a band corresponding to the pIII protein at around 60 kDa in size (Tesar et al., 1995).

When I used the α-His antibody as my primary antibody, I observed a band at around 37 kDa consistent with previous results (Zakeri et al., 2012) where it was hypothesized that this band might be the SpyCatcher protein reacting with an endogenous *E. coli* protein (Zakeri et al., 2012). The expected size of the covalent adduct was 39 kDa for SpyCatcher + SpyTag-ΔpIII. However, I observed a band at around 50 kDa (Figure 5.3). The difference in size between the expected and observed size for the adduct might be linked with why pIII is only ~42 kDa in size but travels the same distance on an SDS-PAGE as a protein of a molecular weight of 60 kDa (Tesar et al., 1995). I also observed a band at ~65 kDa which could be the antibody binding to an *E. coli* protein with a section containing multiple histidines. The absence of a SpyCatcher band at ~15 kDa in the α-His blot could be related to the SpyCatcher not

transferring well onto the membrane (Personal communication with Dr. Jacob O. Fierer).

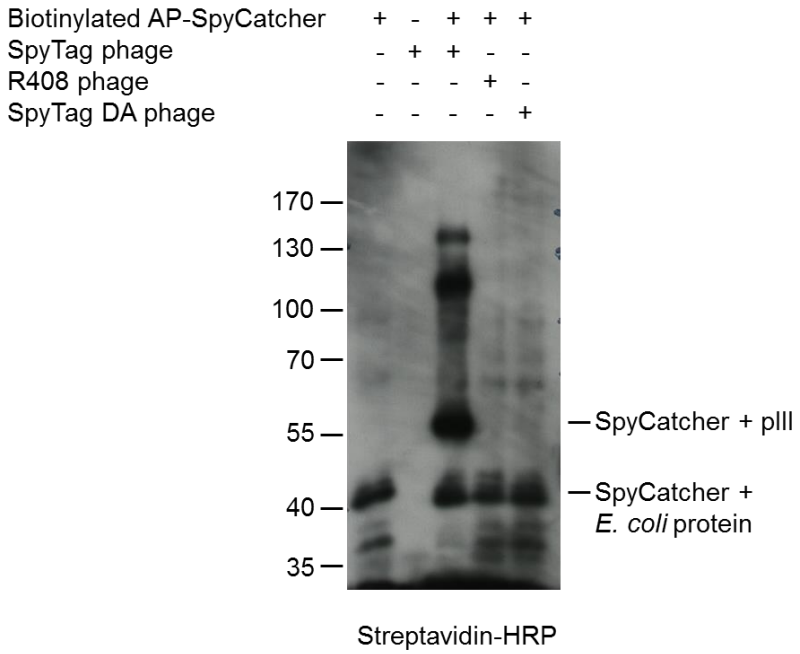


Figure 5.4: Streptavidin-HRP Western Blot provides further evidence that the SpyTag phage is reacting with SpyCatcher.

However, using biotinylated AP-SpyCatcher, I was able to use Streptavidin-HRP detection in my Western Blot. I observed multiple bands appear when I reacted the biotinylated AP-SpyCatcher with the SpyTag-displaying phage (Figure 5.4). The expected size of the covalent adduct was 41 kDa for AP-SpyCatcher + SpyTag-ΔpIII but I observed a major band of ~55 kDa in size (Figure 5.4). The discrepancy in expected size compared to observed size was similar to that observed in the α-His blot (Figure 5.3). The additional bands might be due to read-through of the AP-SpyCatcher gene since the construct only has a single TAA stop codon (Major et al., 2002). Read-through would also explain the bands between the 40 and 35 kDa lane markers which might be AP-SpyCatcher with additional amino acids added on the C-terminus due to read-through and therefore fade in the presence of the SpyTag-displaying phage because these longer AP-SpyCatcher's could still be reactive (Zhang et al., 2013).

The additional bands might only be visible in the Streptavidin-HRP blot and not in the α-His blot due to a difference in sensitivity between the two Western Blot approaches. No band appeared when mixing R408 helper phage or SpyTag DA phage (a phage displaying a peptide where there catalytic aspartic acid residue has been mutated to a non-reactive alanine residue) with the biotinylated AP-SpyCatcher (Figure 5.4). Therefore, I concluded that the peptide was properly displayed and that SpyCatcher would still be able to react with the exposed peptide.

5.2.2 The wash protocol was optimized to improve phage viability

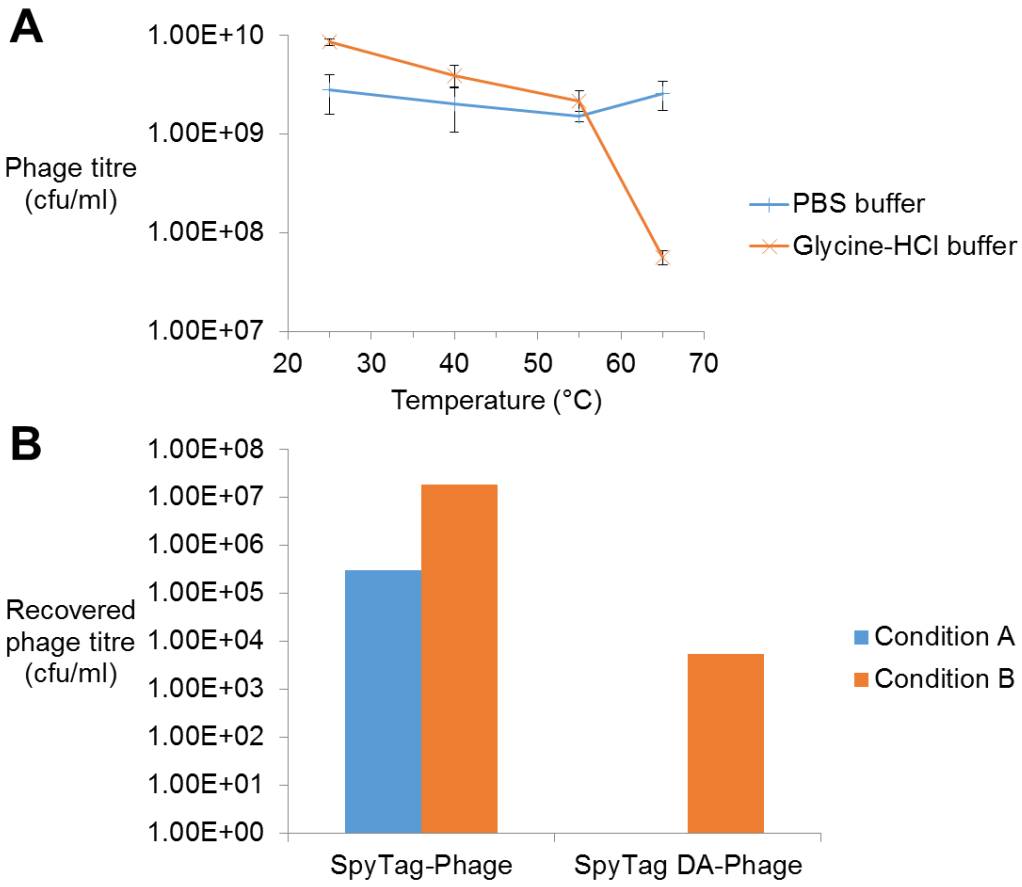


Figure 5.5: The wash protocol was optimized to increase phage survival. A) Phage was incubated in PBS pH 7.4 or 0.2 M Glycine-HCl pH 2.2 for 10 min at the indicated temperatures. A phage titre was then determined using an infectivity assay (mean of triplicate ± 1 standard deviation). B) SpyTag phage and SpyTag DA phage were panned using the approach in Figure 5.2A. Washing conditions were used from Table 5.1.

Previous work in our laboratory showed that the washing and panning conditions used previously did not allow for the selection of peptides that could form covalent bonds (Zakeri, 2011). My first assumption was that the washing conditions might be too harsh and might be killing a lot of the phage in the PhD-12 library in the first round of panning. We had decided to use phage infectivity to propagate the library because the approach is fast, cheap and is widely used in phage display selections (Clackson and Lowman, 2004). However, the approach can lead to replication bias (Derda et al., 2010) and is reliant on the phage remaining viable. To make sure that the phage remains viable, I tested what effects a low pH wash at higher temperatures could have on the viability of phage. I found that incubating the phage in 0.2 M glycine-HCl pH 2.2 at 65 °C for 10 min led to a loss of viability of the phage (Figure 5.5A).

Condition A	Condition B
1. Wash beads 3 times with 100 µl of TBS +0.5%.	1. Wash beads 3 times with 100 µl of TBS +0.5%.
2. Wash beads 1 time with TBS +1% Tween-20. Incubate for 10 minutes at 65 °C at 1,200 RPM.	2. Wash beads 1 time with TBS +1% Tween-20.
3. Wash 1 time with PBS +0.9M NaCl. Incubate for 10 minutes at 65 °C at 1,200 RPM.	3. Wash 1 time with PBS +0.9M NaCl.
4. Wash 1 time with 0.2 M glycine-HCl pH 2.2. Incubate for 10 minutes at 65 °C at 1,200 RPM.	4. Wash 1 time with 0.2 M glycine-HCl pH 2.2.
5. Wash 3 times with 100 µl TBS + 1% Tween-20.	5. Wash 3 times with 100 µl TBS + 1% Tween-20.
6. Wash once with PBS.	6. Wash once with PBS.
7. Elute with 100 µl 38 µM TEV protease in 50 mM Tris-HCl buffer pH 8.0 + 0.5 mM EDTA at 34 °C for 2 h at 1,200 RPM.	7. Elute with 100 µl 38 µM TEV protease in 50 mM Tris-HCl buffer pH 8.0 + 0.5 mM EDTA at 34 °C for 2 h at 1,200 RPM.

Table 5.1: Washing conditions used for experiment highlighted in Figure 5.5B.

I discovered that by removing all 10 min 65 °C incubations (Table 5.1) I could increase the amount of viable phage I recovered from a round of panning (Figure 5.5B). However, this increase in phage viability came with a decrease in stringency as a low amount of non-covalent interacting phages were also recovered (Figure 5.5B). The K_d of the interaction between SpyTag DA and SpyCatcher was measured to be

0.2 µM (Zakeri et al., 2012). Thankfully, the stringency of the new conditions was enough to allow for a recovery of covalent interactions that was 10^3 times higher than that of the non-covalent interactions.

I decided to use the condition B for the first panning round (Table 5.1) with increasing stringency until I used condition A in the 4th step of panning to remove non-covalent interacting phages. Then, I used the strategy highlighted in Figure 5.2B to select for the fastest reacting peptides in panning rounds 6 and 7 (Chapter 2).

5.2.3 A qPCR approach was developed for the rapid screening of selected phages

After selection, clones then need to be screened in order to be able to isolate promising peptide sequences. Previous work has shown that qPCR can be used to quantify the amount of phage in a sample, a faster and higher throughput alternative to the standard plaque assay (Anderson et al., 2011). I designed primers to amplify a portion of the pIII gene in the pFab5cHis phagemid (used to make the SpyTag phage and SpyTag DA phage, as well as all subsequent SpyCatcher-displaying phages). I also designed the primers to be compatible with the M13KE genome used in the PhD-12 library. The primers worked well for both the M13KE derived phages and the pFab5cHis derived phagemids (Figure 5.6).

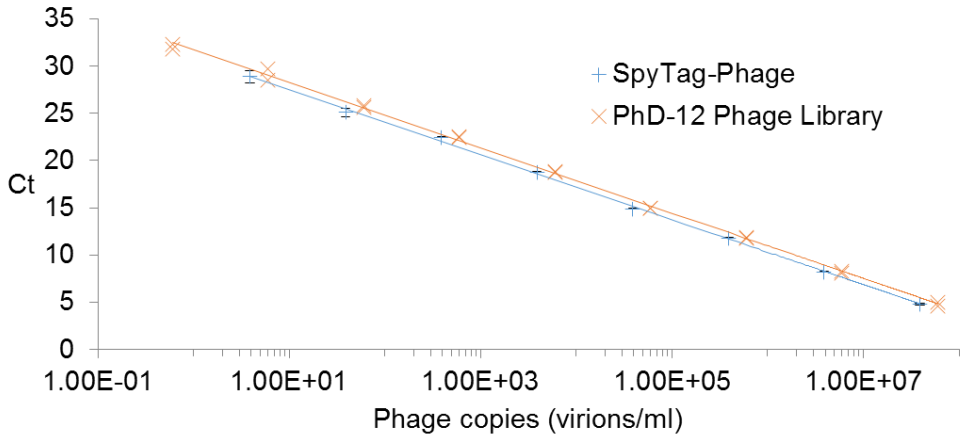


Figure 5.6: qPCR can be used to quantify the amount of phage in a sample. Standard curve relating the threshold fluorescence (Ct, value at which fluorescence passes a threshold set over the background noise) to the amount of phage particles in a sample (calibrated by phage infectivity assay). SpyTag phage samples are shown in triplicate (mean ± 1 standard deviation) and PhD-12 samples were analysed in duplicates (all data points shown), to save library samples for future panning work.

The screening assay was performed by adapting the scheme highlighted in Figure 5.2B for use in 96 well format. Also, the phage was not eluted from the Streptavidin beads using TEV protease but was instead boiled at 95 °C to break open the phages and release the DNA. The supernatant was then analysed using qPCR with the primers used to make the standard curve in Figure 5.6.

5.3 Panning the NEB PhD-12 library

5.3.1 Panning the PhD-12 library led to an enrichment of covalent-bond forming clones

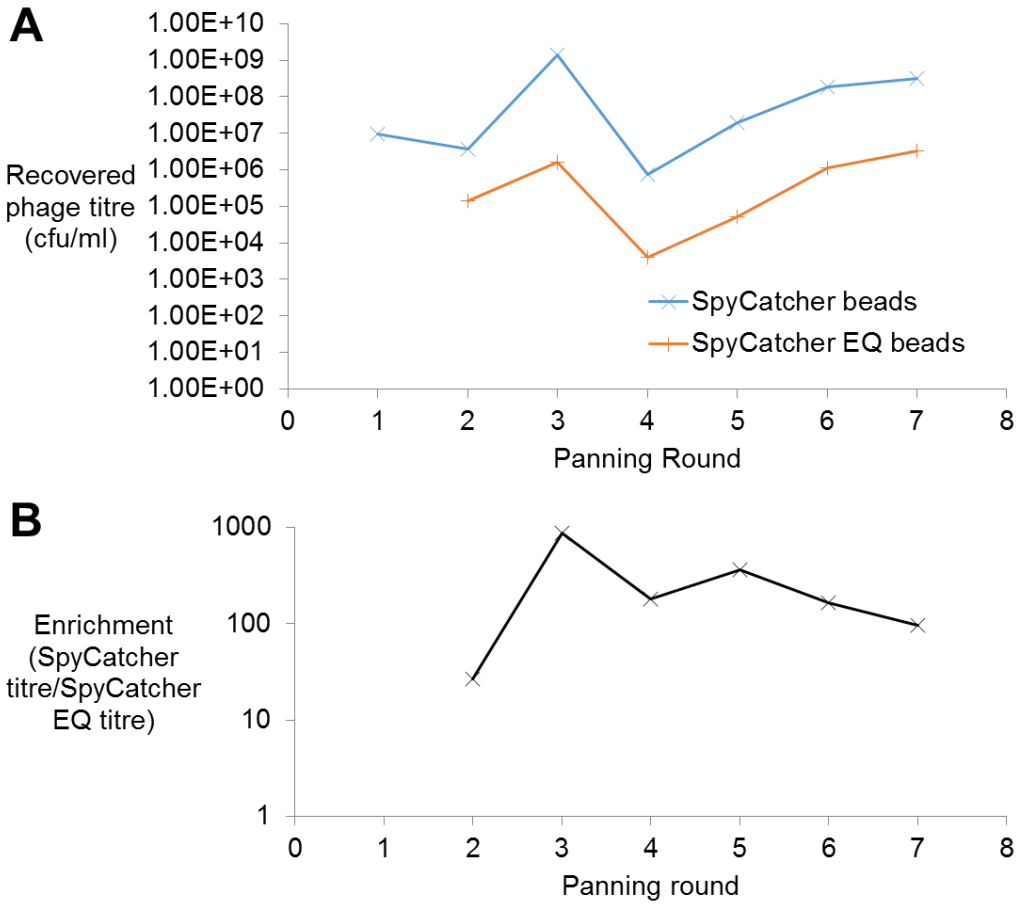


Figure 5.7: Selection conditions led to an increase in enrichment of covalent bond-forming peptide. A) Recovered phage titer after panning of the PhD-12 library against beads coated with SpyCatcher or SpyCatcher EQ (to determine the enrichment of covalent interactions). B) Enrichment of covalent interactions at each panning round.

The third round of panning was enriched for covalent bond forming peptides by nearly 10^3-fold (Figure 5.7B). At the fourth round of panning I used the most stringent conditions to select for only covalent bond forming peptides which decreased the recovered titre considerably (Figure 5.7A). Then, in round 6-7 I used the strategy highlighted in Figure 5.2B and titred down the amount of bait protein in each subsequent round of selection (Chapter 2) to select for the peptides with the fastest on-rate (Dueñas et al., 1996).

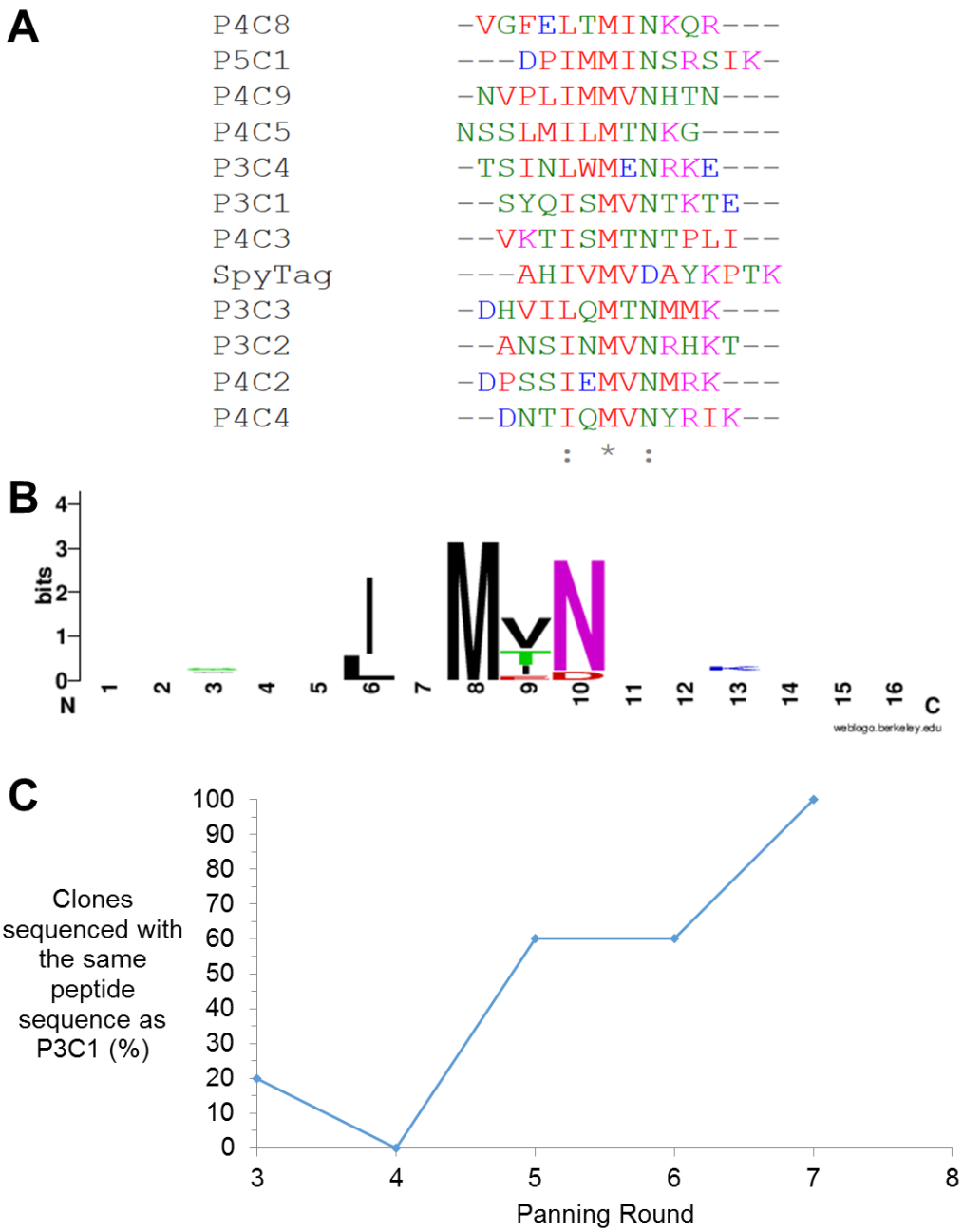


Figure 5.8: Peptide diversity was large but the population boiled down to a single peptide sequence. A) Clustal Omega sequence alignment of unique clones extracted from the PhD-12 library. P3C1 denotes the first clone isolated from the eluate of the third round of panning. : represents a residue where the properties have been conserved. * represents a residue which has been completely conserved. RED: small and/or hydrophobic, BLUE: acidic, MAGENTA: basic, GREEN: contains hydroxyl and/or amine, or sulfhydryl or is glycine. B) WebLogo (Schneider and Stephens, 1990) generated using the alignment generated by Clustal Omega in A). C) Enrichment of the peptide library for a single clone over multiple rounds of selection. Percentage value was based on the amount of P3C1 sequences in each round when I sequenced at least 5 clones.

5.3.2 Selected clones provided insight into the peptide sequence needed to be able to react with SpyCatcher

The PhD-12 library contains 10^9 unique sequences and I started by panning ~10^{11} phage. The diversity of the library was monitored by sequencing 5 clones per round. While 5 clones is not an adequate sample size, limited resources prevented me from sequencing more clones. Therefore, any loss of diversity due to too high or too low stringency could not be monitored. However, isolated clones from each round of selection were quite diverse as seen in Figure 5.8A. All the clones that I selected had a conserved methionine residue at position 8 as well as a conserved leucine/isoleucine at position 6 (Figure 5.8B). Instead of an aspartic acid residue, all the clones had an asparagine residue that could also react with SpyCatcher. The CnaB2 domain also was shown to be able to form an isopeptide bond when the aspartic acid was mutated to an asparagine. However, the isopeptide bond formed slower (Hagan et al., 2010). Asparagine is the reactive residue in both Spy0128 (Kang et al., 2007) and RrgA (Izoré et al., 2010). Therefore, a reactive asparagine in the peptide sequences was not surprising. During the panning rounds the library became enriched by a single peptide motif, “SYQISMVNTKTE” (Figure 5.8C).

5.3.3 Initial screens showed that none of the clones were faster-reacting than SpyTag

Initial screens with the qPCR method did not yield any results where the clones were clearly better than the M13KE-derived SpyTag phage (Figure 5.9A). However, the results obtained from a single time point were sometimes so close to the value obtained for SpyTag that it was hard to tell if the phage reacted faster. I showed that multiple time points can paint a clearer picture of which peptide reacts faster (Figure 5.9B). Therefore, I suggest that in future the screens should be done in two rounds. The first round should be done in single or duplicate with a single time point and will weed out any clones that are not promising. The second round will then use at least three time points and should be done in triplicate.

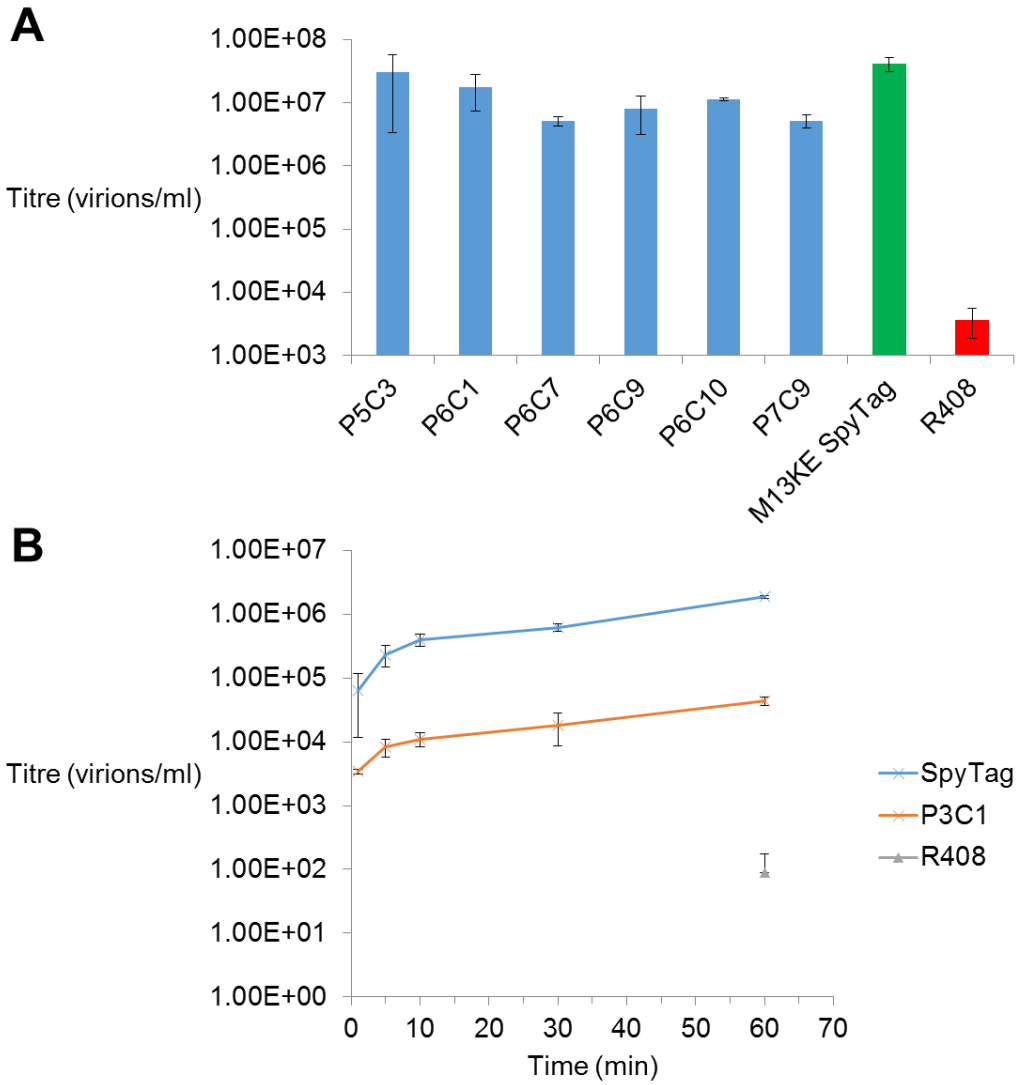


Figure 5.9: Selected clones did not react faster than SpyTag. A) Typical results of a screening experiment. Phages were allowed to react with 630 nM biotinylated AP-SpyCatcher for 1 min before the reaction was quenched with excess SpyCatcher. R408 helper phage was used as a negative control (mean of duplicate ± 1 standard deviation). B) SpyTag-phage reacted faster with biotin-AP-SpyCatcher than the P3C1 clone. qPCR-derived kinetics of SpyTag and P3C1 (SYQISMVNTKTE) obtained by quenching the reaction at different time points. R408 was used as a negative control (mean of triplicate ± 1 standard deviation).

5.3.4 Selected clones were fused to MBP and shown to react slower than SpyTag

Next, I fused the peptide sequence to the N-terminus of MBP. I devised a cloning strategy based on CPEC (Quan and Tian, 2009) which could insert the peptides at the N-terminus of MBP without the need of ordering new primers for each clone. To take into account the additional amino acids I had to add to the MBP for the

annealing of the peptide DNA sequences in the cloning step, I cloned PhD12C SpyTag-MBP which is exactly the same as the isolated peptide-MBP construct except for the SpyTag in the place of the isolated peptide (Figure 5.10). While MBP was chosen for testing due to MBP's solubility, the peptide might still react differently when attached to the N-terminus of MBP than when attached to the N-terminus of pIII.



Figure 5.10: Alignment of the Tag-MBP constructs. In order to make a cloning strategy that would not require the sequencing of the clones and the designing of new primers for every isolated peptide sequence I devised a restriction-less cloning strategy. The nucleotides encoding PLVVPFYSHS and GGGSAE were used to anneal the primers to the M13KE phage genome.

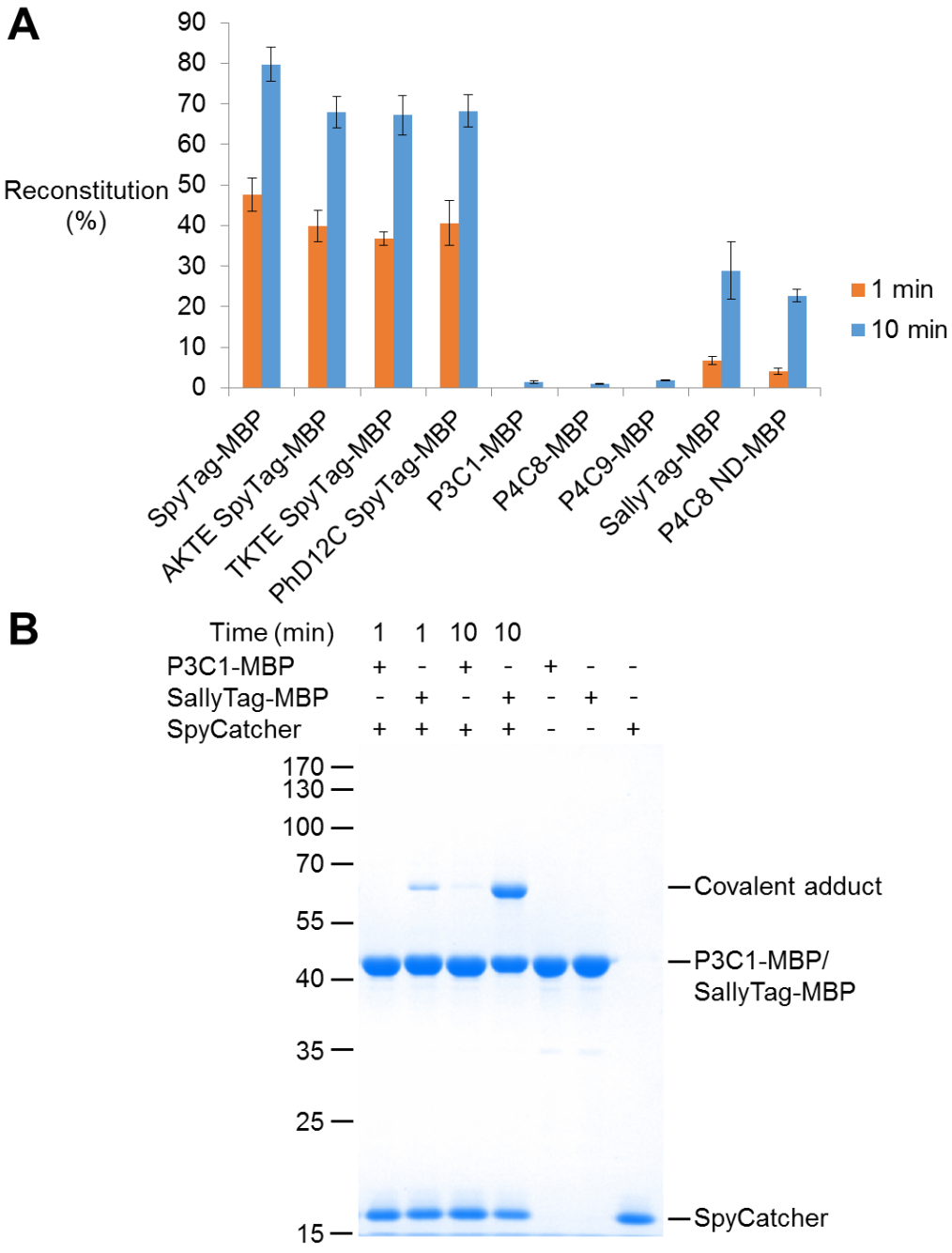


Figure 5.11: SpyTag is still the fastest reacting peptide. A) Reconstitution percentage of MBP constructs (Figure 5.10) at 1 min and 10 min. Reactions were quenched with 6x SDS loading buffer and boiled at 95 °C for 7 min. Samples were loaded on a gel and the band corresponding to the covalent adduct was quantified via densitometry (mean of triplicate ± 1 standard deviation). B) N→D mutation is beneficial. Demonstrative gel showing the SallyTag-MBP reacted faster than P3C1-MBP.

I found that the isolated clones reacted very slowly with SpyCatcher (Figure 5.11A) in comparison with the PhD12C SpyTag-MBP construct. I thought that one of the possible reasons for the slow reactivity of the peptides was the asparagine.

Therefore, I tried mutating two of the peptides. P4C8-MBP became far more reactive after the N→D mutation (Figure 5.10A). P3C1 also was a lot more reactive when the asparagine was substituted with an aspartic acid (P3C1 ND was renamed SallyTag) (Figure 5.11). However, the N→D mutation was not enough to make the peptides react faster than SpyTag.

Furthermore, I tried to see whether mutating the "YKPTK" component to be more similar to the "TKTE" C-terminus of SallyTag would improve the reactivity of SpyTag. The "YKPTK" sequence was chosen because I believed that since the YKPTK sequence was not part of the original CnaB2 sequence (Zakeri et al., 2012) the sequence might not form many interactions with SpyCatcher and therefore could be improved upon. Mutating AYKPTK to AKTE or TKTE led to a decrease in reactivity when compared to SpyTag-MBP (Figure 5.10A). Shortly after this experiment had been performed, a crystal structure of SpyTag/SpyCatcher was elucidated (Li et al., 2013). The crystal structure provided evidence that the YKPTK portion does form interactions with SpyCatcher which would explain the loss in reaction speed of the AKTE or TKTE tags.

5.3.5 SallyTag was able to cyclize BLA and confer an increase in thermal resilience

To test whether SallyTag would have any different effects on the cyclized BLA construct, I substituted SpyTag for SallyTag. I decided to focus on SallyTag because P3C1 was the dominant clone in the population of the library at the late stages of selection (Figure 5.8B).

SallyTag-BLA-SpyCatcher could still cyclize BLA as shown with MS (Figure 5.12A-B). SallyTag-BLA-SpyCatcher and SpyTag-BLA-SpyCatcher were similar in terms of their thermal resilience (Figure 3.7A and Figure 5.12C). Surprisingly, SallyTag DA-BLA-SpyCatcher was more resilient than SpyTag DA-BLA-SpyCatcher (Figure 3.11 and Figure 5.12C). Moreover, the thermal resilience of SallyTag DA-BLA-

SpyCatcher only appears to be slightly worse than that of BLA-CnaB2 (Figure 3.11 and Figure 5.12C).

Therefore, I hypothesize that there is room for improving the thermal resilience of SpyTag-BLA-SpyCatcher and phage display might provide interesting Tags and Catchers that will prove to be useful for future cyclization approaches.

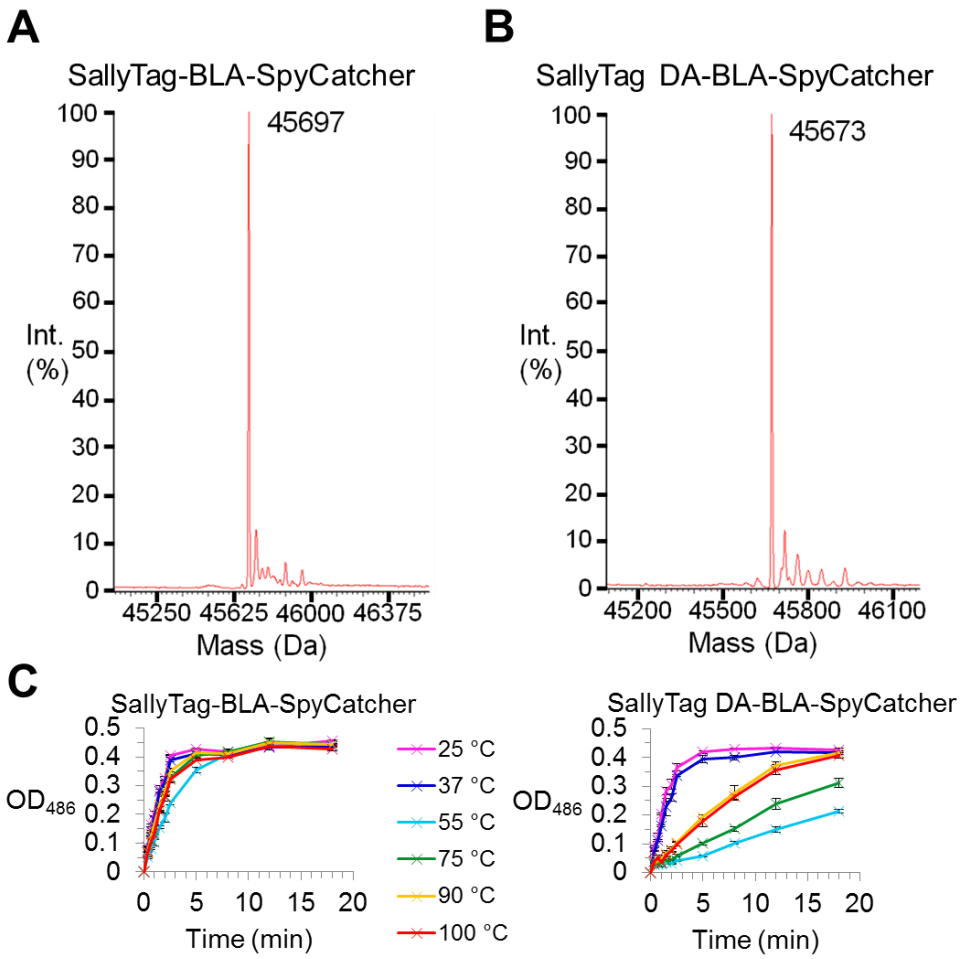


Figure 5.12: SallyTag/SpyCatcher can cyclize BLA and the construct has high thermal resilience. A) MS of SallyTag-BLA-SpyCatcher. Predicted molecular weight was 45,715.2 – 18 = 45,697.2 Da. B) MS of SallyTag DA-BLA-SpyCatcher. Predicted molecular weight was 45,671.2 Da. C) SallyTag-BLA-SpyCatcher and SallyTag DA-BLA-SpyCatcher were incubated for 10 min at the indicated temperature and then cooled to 10 °C, before returning to RT and running a nitrocefin colorimetric activity assay (mean of triplicate ± 1 standard deviation).

5.4 Development of SpyCatcher-displaying phages

5.4.1 Strategy devised to make a SpyCatcher library

Due to the novelty of the SpyTag/SpyCatcher system, little was known about the structural components and mutation hotspots which might be exploited to make a focused library. Therefore, the first SpyCatcher library should be made with the aim of understanding which residues are important for the SpyTag/SpyCatcher reaction. Hence, I decided to plan for a random mutagenesis based library.

By analyzing the sequence of SpyCatcher using a mutagenesis assistant program (MAP) (Wong et al., 2006), I decided to use error-prone PCR (Wilson and Keefe, 2001). MAP is a browser-based statistical analysis tool which utilizes the information obtained from mutating four different proteins using 19 different mutagenesis approaches to calculate the mutation bias of each method for an input sequence (Wong et al., 2006). The approach recommended by the software utilized Taq polymerase, unbalanced dNTP's and $MnCl_2$ to insert mutations with the least amount of bias throughout the SpyCatcher sequence.

Therefore, the cloning strategy to make the library consisted of amplifying the insert using error-prone PCR conditions followed by insertion of the SpyCatcher at the N-terminus of the pIII protein in the pFab5cHis phagemid by using Gibson assembly (Gibson et al., 2009). I also inserted a TEV protease cleavage site between the SpyCatcher and pIII protein for elution in future selections.

The cloning conditions were efficient and I could make an error-prone PCR library consisting of 10^9 clones.

5.4.2 Western blot showed that SpyCatcher was not being displayed properly on the phage

I cloned SpyCatcher on the N-terminus of the pIII protein in the pFab5cHis phagemid and produced SpyCatcher phagemid using the R408 helper phage. As I had done previously with the SpyTag phage, I tested whether a covalent adduct had formed by using a Western Blot. I used α-pIII primary antibody to try to ascertain whether I could see any SpyCatcher-pIII fusions and whether I could see the fusion

forming covalent adducts with SpyTag-MBP. Unfortunately, I did not observe any SpyCatcher-pIII fusions or covalent adducts (Figure 5.13). Then, I tried using biotinylated AP-SpyTag-MBP in combination with Streptavidin-HRP to see if I could observe any covalent adducts in the sample. I could not observe any covalent adducts (Figure 5.13B). Therefore, I concluded that the current construct was not able to display SpyCatcher on the phage particles.

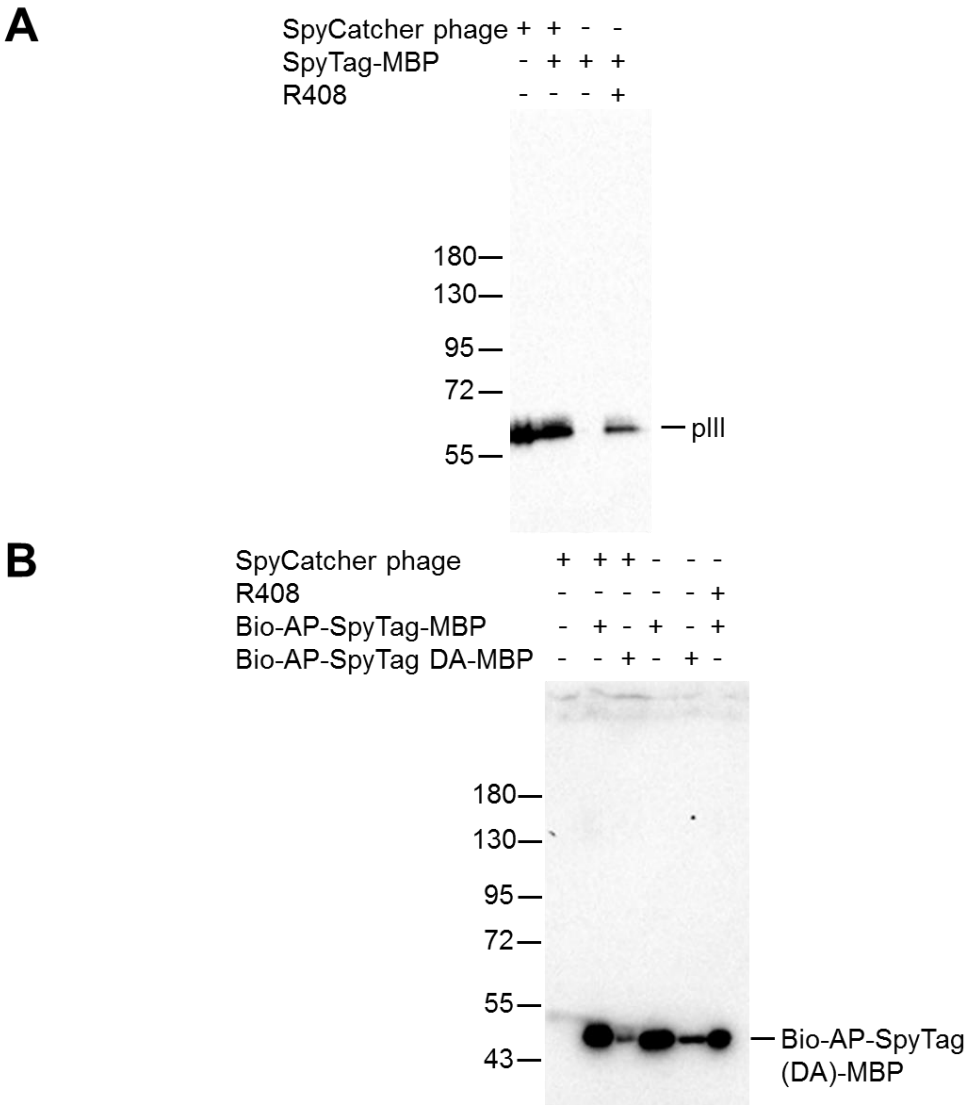


Figure 5.13: Western Blots showed that SpyCatcher was not displayed on the phage. A) Western Blot using α-pIII antibody as the primary antibody. B) Western Blot using Streptavidin-HRP as the detection probe.

5.5 Discussion

In this chapter, I have successfully demonstrated that phage display can be used to select for peptides which are able to form a covalent interaction with SpyCatcher. I have determined that the α-pIII antibody cannot be used to determine whether a peptide-displaying phage is able to react with SpyCatcher (Figure 5.3). However, both α-His and Streptavidin-HRP can be used to determine the formation of a covalent bond (Figure 5.3 and Figure 5.4).

The greatest enrichment of covalent forming peptides appeared to have occurred in the 3rd round of panning (Figure 5.7). However, since every panning step was performed differently (Chapter 2), one needs to be cautious when analyzing the enrichment results. Therefore, a more interesting study would have been to monitor the population of the peptides in each panning step via next generation sequencing techniques (Dijk et al., 2014).

Little biochemical data was available on the sequence space of the SpyTag/SpyCatcher interaction. Therefore, it was difficult to design a library for SpyTag without having access to more information. The PhD-12 library allowed me to gain an insight into the necessary motif for a SpyCatcher reactive peptide. The motif could be the foundation for a new library design which could be more successful.

All of the clones isolated contained a catalytic asparagine instead of a catalytic aspartic acid (Figure 5.8A). I found that, in at least two of the clones tested, the asparagine did not improve the reaction rate of the peptide (Figure 5.11A). So, the fact that all the peptides have a reactive asparagine is probably indicative of the limitations of the PhD-12 peptide library's sequence bias (Krumpe et al., 2006). Moreover, a peptide library where every single peptide would be present would need to have 20^{12} (~4.1 x 10^{15}) unique clones. The PhD-12 library contains ~10^9 unique sequences. Hence, the PhD-12 library only covers a tiny portion of all possible sequences and it is not inconceivable that the library might not include clones that have a catalytic aspartic acid. Therefore, a focused library based on the SpyTag sequence and using

the sequence information I have obtained (Figure 5.8A-B) via phage display might yield faster-reacting Tags.

The concentration of bait protein and reaction time with the phage before adding the competitor in panning rounds 6 and 7 were calculated by taking into account the reaction rate of SpyTag/SpyCatcher. Reaction time and bait concentration were calculated to select for faster-reacting peptides than SpyTag. However, a peptide was selected which reacts significantly slower than SpyTag. A cause for this discrepancy could be that the PhD-12 phages are multivalent and therefore might exhibit an avidity effect for SpyCatcher. Another reason that the P3C1 peptide was selected could be that the peptide has a very fast on-rate and a very slow off-rate but is not able to form an isopeptide bond as quickly as SpyCatcher. Therefore, it would be beneficial to determine the on- and off-rate of P3C1. Spiking the PhD-12 library with a small amount of M13KE-derived SpyTag could also provide evidence on whether the P3C1 peptide has a fast on-rate or whether the selection strategy needs to be improved upon.

The production of SpyCatcher-displaying phages proved to be problematic (Figure 5.13). The leader sequence I used was the PelB sequence which utilizes the Sec-dependent translocation mechanism (Huber et al., 2005) to move pIII protein into the periplasm where phage assembly occurs. The Sec-dependent pathway usually requires that the protein being translocated is unfolded and therefore is a problematic system to use when trying to export a protein that folds well in the cytosol (Steiner et al., 2006). Therefore, it might be possible to produce SpyCatcher-displaying phage by using the signal recognition particle (SRP)-dependent pathway (Steiner et al., 2006) or the twin-arginine translocation (Tat) pathway (Speck et al., 2011).

Work on phage display has been carried on by Dr. Samuel C. Reddington who discovered that the DsbA signal sequence, which uses the SRP-dependent pathway, led to the production of SpyCatcher-displaying phages (Personal communication with

Dr. Samuel C. Reddington). The project is currently being continued using the techniques and approaches I have developed in this chapter.

Chapter 6: Summary and general discussion

Chapter 6: Summary and general discussion

Enzyme stability is an important topic in the field of Biochemistry which has garnered a lot of attention from the scientific community (Packer and Liu, 2015; Yang et al., 2015). The growing field of white biotechnology reflects the importance of enzymes in our current society (Bevan and Franssen, 2006). Directed evolution (Packer and Liu, 2015) and rational design (Yang et al., 2015) have been the most common methods of improving the stability of enzymes, closely followed by bioprospecting of thermophilic organisms (Turner et al., 2007). However, enzyme cyclization is an alternative which has the potential of being more broadly applicable and easy to perform (Iwai and Plückthun, 1999). To further study enzyme cyclization, I genetically encoded BLA in-between different Tag/Catcher pairs (Veggiani et al., 2016; Zakeri and Howarth, 2010; Zakeri et al., 2012) and studied the resultant cyclized enzyme.

In Chapter 3, I observed that the cyclization process is easy, efficient and occurs spontaneously upon expression of the genetic construct. The resultant protein was able to withstand higher temperatures than the wild-type control, even withstanding 100 °C. I studied the importance of the isopeptide bond and postulated a possible stabilization mechanism based on data obtained from controls and biophysical characterization with DSC.

In Chapter 4, I tested the generality of this approach by cyclizing PhyC, DHFR, FLuc and GOx. I provided evidence that the SpyTag/SpyCatcher cyclization approach increased the thermal resilience of both PhyC and DHFR. However, the cyclization approach was not able to confer a significant increase in thermal resilience to GOx and FLuc, providing evidence that there is a limitation to the system.

In Chapter 5, I developed and validated a Phage Display-based approach for the selection and screening of covalent-bond forming peptides which can react with SpyCatcher. This directed evolution work might hold the key to better understanding

the SpyTag/SpyCatcher interaction and possibly discovering novel Tag/Catcher pairs which might confer a higher degree of thermal resilience to enzymes.

The work presented in chapters 3 and 4 provides an insight into the field of protein engineering via cyclization and opened up a plethora of questions. Some of the most intriguing questions that need further attention revolve around the stabilization mechanism of the cyclized enzymes. Previous work using nuclear magnetic resonance (NMR) on an intein-cyclized SH3 domain provided evidence that cyclization increased the rigidity of the protein backbone. (Schumann et al., 2015). It would be interesting to carry out further structural and dynamics studies on the cyclized proteins described in this thesis.

A combination of NMR and/or X-ray crystallography, small-angle x-ray scattering (SAXS) and molecular dynamics (Wen et al., 2014) could provide a lot of interesting information. NMR could be used to study both the structure and dynamics of the cyclized enzymes in solution at ambient temperature (Schumann et al., 2015). However, if the cyclized enzymes crystallize using the same conditions used to obtain the structure of the wild-type enzymes (often not the case), X-ray crystallography might be the simpler option. Either approach could lead to the elucidation of an atomic-resolution structure that could then be used as a model for the SAXS and molecular dynamics experiments (Weeks et al., 2014). Moreover, SAXS could be used to study the tertiary structure and unfolding of the BLA domain of SpyTag-BLA-SpyCatcher at different temperatures (Bianconi et al., 2012). All the structural and biophysical characterization could then be used in a simulation to understand the unfolding process at an atomic level (Toofanny and Daggett, 2012).

The limitations of the SpyTag/SpyCatcher cyclization approach need to be better understood to be able to predict whether SpyTag/SpyCatcher cyclization will increase the thermal resilience of an enzyme. The data I have gathered so far is insufficient to obtain a prediction on whether cyclization will be effective in increasing enzyme resilience. The aforementioned biophysical characterizations might provide

further information helpful for prediction. However, with the decreasing costs of DNA synthesis (Carlson, 2009) and construction of the UK's first DNA foundry (Bawden, 2016) it might be possible to construct an automated system for the cloning of cyclized constructs (SpyTag-Enzyme-SpyCatcher).

By creating an automated process for cloning and expression of cyclized enzymes, it might be possible to determine more information on what are the limitations of the SpyTag/SpyCatcher system. However, a bottleneck still exists in figuring out how to test the thermal resilience of the enzymes in a high-throughput manner. In order to screen for enzymes with high thermal resilience, an activity assay would have to be developed for each type of enzyme, which would slow down the process. Therefore, it might be worth initially focusing on enzymes where there is a known need for thermally resilient variants. I identified phytase as a good target for cyclization because of how phytase is applied in the agricultural industry (Lei et al., 2013). However, no single phytase is ideal for all applications (Lei et al., 2013) and therefore, it might be wise to focus on this one class of enzyme first.

Other enzymes that might be worth investigating include agricultural additives such as xylanase, subtilisin and α-amylase (Li et al., 2012). Increasing the thermal resilience of enzymes used in the production of biofuels also appears to be a promising application (Wang et al., 2016).

In some cases, even if the thermal resilience of an enzyme is improved, there are still some issues which might hinder a real world application. For the most part, industrial biotechnology has not expanded in the commercial world as quickly as expected (Chen, 2012). The issue at hand is that industrial biotechnology has to compete with extremely cheap and efficient processes. For example, the biotechnological production of polyhydroxyalkanoates (PHA) competing with the chemical production of polyethylene (Chen, 2012). There are issues at hand that need to be tackled to move biotechnology forward including improving the fermentation process and recovery methods (Li et al., 2012).

The role of the linkers also deserves further study. Previous research has described the effects of different linker lengths and linker types on protein fusions (Chen et al., 2013). However, it will probably not be possible to find one linker that will work ideally with all cyclized enzymes. Linkers might interact differently with one enzyme compared to another. Therefore, linker optimization is something that should be studied individually for each enzyme in an attempt to further improve the resilience.

In Chapter 3, I studied cyclization mediated by three distinct Tag/Catcher pairs and found that there might be a link between the Tm of the Tag/Catcher domain and the resilience conferred upon the cyclized enzymes. Therefore, it might be possible to improve upon the technology by either improving the Tm of the CnaB2 domain via rational design or directed evolution, or to look for other isopeptide-bond containing proteins which might be present in extremophiles. Other possibilities of increasing the Tm of CnaB2 might include using software such as FireProt (Bednar et al., 2015) or determining ancestral mutants which might be more thermostable (Watanabe et al., 2006). Considering that both Pilin-C/Isopeptag and SpyTag/SpyCatcher were able to cyclize BLA efficiently, it might not even be necessary to split the protein and optimize the Tag/Catcher interaction. Simply introducing the enzyme ORF between the two β-sheets involved in the isopeptide bond formation of the novel thermostable isopeptide-bond-containing protein might be enough.

An enzyme's optimum temperature can be seen as the temperature at which the reaction progresses quickest, while not being affected by the denaturation of said enzyme (Barton, 1979). I observed that BLA had a Tm of 41 °C by DSC but already started aggregating at 37 °C. Therefore, I postulate that SpyTag-BLA-SpyCatcher's optimum temperature would be higher, because it did not aggregate at 37 °C and therefore might not start losing activity until the sample reaches 41 °C. The increase in optimum temperature has been shown to occur in the SpyTag/SpyCatcher cyclized lichenase (Wang et al., 2016). Other effects that should be studied are the effects of isopeptide-mediated cyclization on the pH optimum. If the structure of the enzyme

remains unperturbed by the cyclization, the effects on the optimum pH of the enzyme should be minimal. In the case of the SpyTag/SpyCatcher-cyclized lichenase, the pH optimum only changed from 7.0 to 6.4 when compared to the linear construct (Wang et al., 2016).

In Chapter 4 I described that some enzymes will not be stabilized via the cyclization approach developed in this thesis. In the case of FLuc and GOx, other approaches should be considered. I suggest that it might be interesting to try inserting the SpyTag into the different loops of GOx and FLuc instead of fusing the peptide to the N-terminus of the enzyme. Instead of using the full 13 amino acid sequence, one could simply try the core AHIVMVDA motif (Zakeri et al., 2012) or the SallyTag peptide, which might cause less perturbation if inserted into certain loops of the enzyme. Through this approach, one might be able to introduce more rigidity to a section of the enzyme instead of a small amount of rigidity to the whole protein backbone. If it were crucial for a section of the enzyme close to the catalytic site to be able to move freely to go through the catalytic cycle, promoting localized rigidity might be more effective than global rigidity of the protein backbone.

Other approaches to stabilize GOx could try to mimic the positive effects that adsorption and immobilization can have on the stability of GOx (He et al., 2009; Pekel et al., 2005). Cross-linked enzyme aggregates (CLEAs) have also been shown to increase the stability of enzymes (Cerdobbel et al., 2010; Sheldon, 2007). However, in order to form CLEAs, one needs first to precipitate the enzyme and then cross-link the enzyme using non-specific cross-linking agents such as glutaraldehyde. Non-specific cross-linking can be problematic for some enzymes (Mateo et al., 2004). Another issue with CLEAs are their small particle size, which make them difficult to recover (Cui and Jia, 2015). Moreover, CLEAs tend to cluster and form clumps when centrifuged or undergoing filtration treatments, which can further decrease activity (Cui and Jia, 2015).

Recently, it has been shown that the spider silk protein can be shortened and produced recombinantly in *E. coli* (Stark et al., 2007). The engineered spider silk protein (4RepCT) was still able to form silk-like materials (Stark et al., 2007). Also, 4RepCT can be functionalized by fusing domains to the N-terminus (Jansson et al., 2014). Therefore, it might be possible to fuse different Tags and Catchers to the N-terminus of 4RepCT to create an easily functionalizable protein matrix.

Combining the Tag- or Catcher-functionalized 4RepCT fibers with BiCatcher (SpyCatcher-SnoopCatcher) (Veggiani et al., 2016) and N- and C-terminally tag-fused GOx, it might be possible to increase the thermal tolerance of GOx via a similar mechanism achieved with using amyloid fibers and glutaraldehyde (Kim et al., 2012). However, by making the cross-linking isopeptide-mediated, the enzyme structure might be less perturbed and therefore more active. Moreover, the fibers that are formed by the 4RepCT are 1-10 cm long and 15 mm in diameter (Stark et al., 2007) so they should be easy to recover from a reaction medium. An alternative approach to using spider silk might be through the use of filamentous phage displaying the isopeptide bond-forming Tags and Catchers and creating a large matrix by cross-linking the phage (Hess et al., 2012).

Isopeptide-mediated cyclization could also be used in directed evolution campaigns. There is evidence that there is a positive correlation between evolvability and stability of a protein. In the case of P450 propane monooxygenase, it was necessary to insert stabilizing mutations to find further beneficial mutations (Romero and Arnold, 2009). Therefore, isopeptide-mediated cyclization might allow for the expression of less stable but more active mutants in a directed evolution campaign. Isopeptide-mediated cyclization could lead to the discovery of new pathways in the energy landscape which were not accessible before.

Currently, I have elucidated the following guidelines for the application of isopeptide-mediated cyclization. A $(GSG)_2$ linker between the N-terminal SpyTag and protein of interest and a $(GSG)_2$ linker between the protein of interest and the C-

terminal SpyCatcher have been shown to be effective. These linkers have been proven to be adequate for the cyclization of termini up to 29 Å apart and potentially has also managed to cyclize enzymes where the termini are 38 Å apart. Isopeptide-mediated cyclization has effectively stabilized an enzyme of 357 amino acids in length which might be reaching the upper limit for the technique. For larger enzymes, N- to C-terminal cyclization might not be adequate. Choosing the correct Tag/Catcher pair will depend on the application. For applications where the enzyme needs to be incubated at and around 55 °C, Isopeptag/Pilin-C might be best. For higher temperatues, SpyTag/SpyCatcher is the obvious choice.

In Chapter 5, I was able to isolate peptides able to form a covalent bond with SpyCatcher from a commercially available library. However, none of the peptides reacted faster than SpyTag. Therefore I suggested creating a more focused library to select for peptides that might react faster than SpyTag. The main factor limiting the size of a library is the transformation efficiency of *E. coli* (Zhu and Dimitrov, 2009). Therefore, a library with 10^9 clones is realistically the largest that we could construct. However, looking at the peptides extracted from the PhD-12 (NEB) library, only 4 residues appear to be highly conserved among all the constructs. Therefore, the peptide library should be built based on XXIXMVDXXXXX, where X could be any amino acid. However, that would mean that there are theoretically 5.1 x 10^{11} unique clones possible. Moreover, if we take into account the degeneracy of the genetic code and were to use NNK (where N is any nucleotide and K is either guanine or thymine) for every variable amino acid position, there would be a theoretical maximum of 3.5 x 10^{13} unique clones. Therefore, it would make more sense to use NDT (where D is adenine, thymine or guanine), which would only require 12 different codons but codes for 12 different amino acids and has been shown to be an effective way of tackling the number problem (Reetz et al., 2008). Therefore, for the next library I suggest removing one of the C-terminal amino acids to shorten the tag and to use the XXIXMVDXXXXX

motif where X is encoded for by a random NDT codon. This library would have a theoretical size of 4.3 x 10^8 unique sequences.

Other tests that could be done with the PhD-12 library would be to select for peptides that can form a covalent interaction with SnoopCatcher. The data obtained from this experiment could then be used to tailor a library to select for better Tags that can react with SnoopCatcher. Furthermore, Phage Display might be a good approach for the evolution of orthogonal Tag/Catcher systems (Zhou et al., 2007).

In Chapter 5 I also did preliminary work on displaying SpyCatcher on the surface of the phage, which was ineffective. Dr. Samuel C. Reddington changed the secretion signal sequence and managed to get the phage to express the SpyCatcher on the surface. However, even with a new secretion signal sequence, the yield of SpyCatcher-displaying phage was low (Personal communication with Dr. Samuel C. Reddington). Therefore, an alternative approach to using M13 phage would be using T7 phage, which is assembled in the cytosol and does not require any secretion because T7 is a lytic phage (Slootweg et al., 2006). Proteins of up to 1000 amino acids can be displayed at low copy number on the surface of the T7 phage (Slootweg et al., 2006) and commercial kits like T7Select (Merck Millipore) are available. Alternative approaches such as ribosome display (Plückthun, 2012), protein fragment complementation assays (Remy et al., 2007) and yeast surface display (Gera et al., 2013) might also be worth considering. However, the main advantage of phage is its resistance to stringent washes, to select for covalent interactions instead of possibly also selecting for non-covalent interactions.

Bibliography

Abe, H., Wakabayashi, R., Yonemura, H., Yamada, S., Goto, M., and Kamiya, N. (2013). Split Spy0128 as a Potent Scaffold for Protein Cross-Linking and Immobilization. Bioconjug. Chem. *24*, 242–250.

Amid, M., Manap, Y., and Zohdi, N.K. (2014). Microencapsulation of purified amylase enzyme from pitaya (Hylocereus polyrhizus) peel in Arabic gum-chitosan using freeze drying. Mol. Basel Switz. *19*, 3731–3743.

Anderson, B., Rashid, M.H., Carter, C., Pasternack, G., Rajanna, C., Revazishvili, T., Dean, T., Senecal, A., and Sulakvelidze, A. (2011). Enumeration of bacteriophage particles: Comparative analysis of the traditional plaque assay and real-time QPCR- and nanosight-based assays. Bacteriophage *1*, 86–93.

Antos, J.M., Popp, M.W.-L., Ernst, R., Chew, G.-L., Spooner, E., and Ploegh, H.L. (2009). A straight path to circular proteins. J. Biol. Chem. *284*, 16028–16036.

Atwell, S., and Wells, J.A. (1999). Selection for improved subtiligases by phage display. Proc. Natl. Acad. Sci. *96*, 9497–9502.

Baase, W.A., Liu, L., Tronrud, D.E., and Matthews, B.W. (2010). Lessons from the lysozyme of phage T4. Protein Sci. Publ. Protein Soc. *19*, 631–641.

Barton, J.S. (1979). Denaturation at the optimum temperature. Biochem. Educ. *7*, 13–14.

Bastioli, C. (2005). Handbook of Biodegradable Polymers (iSmithers Rapra Publishing).

Bawden, T. (2016). “DNA factory” opens at Imperial College London.

Bazan, J., Całkosiński, I., and Gamian, A. (2012). Phage display—A powerful technique for immunotherapy. Hum. Vaccines Immunother. *8*, 1817–1828.

Bednar, D., Beerens, K., Sebestova, E., Bendl, J., Khare, S., Chaloupkova, R., Prokop, Z., Brezovsky, J., Baker, D., and Damborsky, J. (2015). FireProt: Energy- and Evolution-Based Computational Design of Thermostable Multiple-Point Mutants. PLOS Comput Biol *11*, e1004556.

Bessler, C., Schmitt, J., Maurer, K.-H., and Schmid, R.D. (2003). Directed evolution of a bacterial α-amylase: Toward enhanced pH-performance and higher specific activity. Protein Sci. Publ. Protein Soc. *12*, 2141–2149.

Bevan, M.W., and Franssen, M.C.R. (2006). Investing in green and white biotech. Nat. Biotechnol. *24*, 765–767.

Bianconi, A., Ciasca, G., Tenenbaum, A., Battisti, A., and Campi, G. (2012). Temperature and solvent dependence of the dynamical landscape of tau protein conformations. J. Biol. Phys. *38*, 169–179.

Bisht, D., Yadav, S.K., and Darmwal, N.S. (2014). An oxidant and organic solvent tolerant alkaline lipase by P. aeruginosa mutant: Downstream processing and biochemical characterization. Braz. J. Microbiol. *44*, 1305–1314.

Bloom, J.D., Labthavikul, S.T., Otey, C.R., and Arnold, F.H. (2006). Protein stability promotes evolvability. Proc. Natl. Acad. Sci. *103*, 5869–5874.

Bohlke, R.A., Thaler, R.C., and Stein, H.H. (2005). Calcium, phosphorus, and amino acid digestibility in low-phytate corn, normal corn, and soybean meal by growing pigs. J. Anim. Sci. *83*, 2396–2403.

Botyanszki, Z., Tay, P.K.R., Nguyen, P.Q., Nussbaumer, M.G., and Joshi, N.S. (2015). Engineered catalytic biofilms: Site-specific enzyme immobilization onto E. coli curli nanofibers. Biotechnol. Bioeng. *112*, 2016–2024.

Brady, D., Steenkamp, L., Skein, E., Chaplin, J.A., and Reddy, S. (2004). Optimisation of the enantioselective biocatalytic hydrolysis of naproxen ethyl ester using ChiroCLEC-CR. Enzyme Microb. Technol. *34*, 283–291.

Brune, K.D., Leneghan, D.B., Brian, I.J., Ishizuka, A.S., Bachmann, M.F., Draper, S.J., Biswas, S., and Howarth, M. (2016). Plug-and-Display: decoration of Virus-Like Particles via isopeptide bonds for modular immunization. Sci. Rep. *6*, 19234.

Bryksin, A.V., and Matsumura, I. (2010). Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. BioTechniques *48*, 463–465.

Camarero, J.A., and Muir, T.W. (1999). Biosynthesis of a head-to-tail cyclized protein with improved biological activity. J. Am. Chem. Soc. *121*.

Cao, L., van Rantwijk, F., and Sheldon, R.A. (2000). Cross-linked enzyme aggregates: a simple and effective method for the immobilization of penicillin acylase. Org. Lett. *2*, 1361–1364.

Carlson, R. (2009). The changing economics of DNA synthesis. Nat. Biotechnol. *27*, 1091–1094.

Carrea, G., and Riva, S. (2000). Properties and Synthetic Applications of Enzymes in Organic Solvents. Angew. Chem. Int. Ed Engl. *39*, 2226–2254.

Cerdobbel, A., De Winter, K., Desmet, T., and Soetaert, W. (2010). Sucrose phosphorylase as cross-linked enzyme aggregate: improved thermal stability for industrial applications. Biotechnol. J. *5*, 1192–1197.

Chambers, P.J., and Pretorius, I.S. (2010). Fermenting knowledge: the history of winemaking, science and yeast research. EMBO Rep. *11*, 914–920.

Chen, G.-Q. (2012). New challenges and opportunities for industrial biotechnology. Microb. Cell Factories *11*, 111.

Chen, C.C.H., and Herzberg, O. (1999). Relocation of the catalytic carboxylate group in class A β-lactamase: the structure and function of the mutant enzyme Glu166→Gln:Asn170→Asp. Protein Eng. *12*, 573–579.

Chen, K., and Arnold, F.H. (1993). Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. Proc. Natl. Acad. Sci. U. S. A. *90*, 5618–5622.

Chen, A.Y., Deng, Z., Billings, A.N., Seker, U.O.S., Lu, M.Y., Citorik, R.J., Zakeri, B., and Lu, T.K. (2014). Synthesis and patterning of tunable multiscale materials with engineered cells. Nat. Mater. *13*, 515–523.

Chen, X., Zaro, J., and Shen, W.-C. (2013). Fusion Protein Linkers: Property, Design and Functionality. Adv. Drug Deliv. Rev. *65*, 1357–1369.

Chiu, J., March, P.E., Lee, R., and Tillett, D. (2004). Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. Nucleic Acids Res. *32*, e174–e174.

Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F.B., and Xu, M.Q. (1996). Protein splicing involving the Saccharomyces cerevisiae VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. J. Biol. Chem. *271*, 22159–22168.

Cioci, F., and Lavecchia, R. (1994). Effect of polyols and sugars on heat-induced flavin dissociation in glucose oxidase. Biochem. Mol. Biol. Int. *34*, 705–712.

Clackson, T., and Lowman, H.B. (2004). Phage Display : A Practical Approach (Oxford University Press).

Clark, R.J., and Craik, D.J. (2010). Native chemical ligation applied to the synthesis and bioengineering of circular peptides and proteins. Biopolymers *94*, 414–422.

Colgrave, M.L., and Craik, D.J. (2004). Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot. Biochemistry (Mosc.) *43*, 5965–5975.

Conrado, R.J., Wu, G.C., Boock, J.T., Xu, H., Chen, S.Y., Lebar, T., Turnšek, J., Tomšič, N., Avbelj, M., Gaber, R., et al. (2012). DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. Nucleic Acids Res. *40*, 1879–1889.

Craig, D.B., and Dombkowski, A.A. (2013). Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinformatics *14*, 346.

Cui, J.D., and Jia, S.R. (2015). Optimization protocols and improved strategies of cross-linked enzyme aggregates technology: current development and future challenges. Crit. Rev. Biotechnol. *35*, 15–28.

Cui, J.D., Sun, L.M., and Li, L.L. (2013). A simple technique of preparing stable CLEAs of phenylalanine ammonia lyase using co-aggregation with starch and bovine serum albumin. Appl. Biochem. Biotechnol. *170*, 1827–1837.

Damborsky, J., and Brezovsky, J. (2014). Computational tools for designing and engineering enzymes. Curr. Opin. Chem. Biol. *19*, 8–16.

Daniel, R.M., and Danson, M.J. (2013). Temperature and the catalytic activity of enzymes: A fresh understanding. FEBS Lett. *587*, 2738–2743.

Davis, J.H., Rubin, A.J., and Sauer, R.T. (2011). Design, construction and characterization of a set of insulated bacterial promoters. Nucleic Acids Res. *39*, 1131–1141.

Dawson, P.E., Muir, T.W., Clark-Lewis, I., and Kent, S.B. (1994). Synthesis of proteins by native chemical ligation. Science *266*, 776–779.

Derda, R., Tang, S.K.Y., and Whitesides, G.M. (2010). Uniform Amplification of Phage with Different Growth Characteristics in Individual Compartments Consisting of Monodisperse Droplets. Angew. Chem. Int. Ed. *49*, 5301–5304.

Dijk, E.L. van, Auger, H., Jaszczyszyn, Y., and Thermes, C. (2014). Ten years of next-generation sequencing technology. Trends Genet. *30*, 418–426.

Dombkowski, A.A., Sultana, K.Z., and Craig, D.B. (2014). Protein disulfide engineering. FEBS Lett. *588*, 206–212.

Dueñas, M., Malmborg, A.-C., Casalvilla, R., Ohlin, M., and Borrebaeck, C.A.K. (1996). Selection of phage displayed antibodies based on kinetic constants. Mol. Immunol. *33*, 279–285.

Ebrahimi, M., Hosseinkhani, S., Heydari, A., Khavari-Nejad, R.A., and Akbari, J. (2012). Improvement of thermostability and activity of firefly luciferase through [TMG][Ac] ionic liquid mediator. Appl. Biochem. Biotechnol. *168*, 604–615.

Eijsink, V.G.H., Gåseidnes, S., Borchert, T.V., and van den Burg, B. (2005). Directed evolution of enzyme stability. Biomol. Eng. *22*, 21–30.

Engberg, J., Johansen, L.K., Westengaard-Hildinge, M., Riise, E.S., and Albrechtsen, B. (2000). Phage-Display Libraries of Murine and Human Antibody Fab Fragments. In The Nucleic Acid Protocols Handbook, R. Rapley, ed. (Humana Press), pp. 449–478.

Evans, T.C., Benner, J., and Xu, M.Q. (1999a). The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins. J. Biol. Chem. *274*, 18359–18363.

Evans, T.C., Jr, Benner, J., and Xu, M.Q. (1999b). The in vitro ligation of bacterially expressed proteins using an intein from Methanobacterium thermoautotrophicum. J. Biol. Chem. *274*, 3923–3926.

Fairhead, M., and Howarth, M. (2015). Site-specific biotinylation of purified proteins using BirA. Methods Mol. Biol. Clifton NJ *1266*, 171–184.

Fairhead, M., Veggiani, G., Lever, M., Yan, J., Mesner, D., Robinson, C.V., Dushek, O., van der Merwe, P.A., and Howarth, M. (2014). SpyAvidin hubs enable precise and ultrastable orthogonal nanoassembly. J. Am. Chem. Soc. *136*, 12355–12363.

Fan, F., Binkowski, B.F., Butler, B.L., Stecha, P.F., Lewis, M.K., and Wood, K.V. (2008). Novel Genetically Encoded Biosensors Using Firefly Luciferase. ACS Chem. Biol. *3*, 346–351.

Ferri, S., Kojima, K., and Sode, K. (2011). Review of Glucose Oxidases and Glucose Dehydrogenases: A Bird's Eye View of Glucose Sensing Enzymes. J. Diabetes Sci. Technol. *5*, 1068–1076.

Fierer, J.O. Personal communication.

Flory, P.J. (1956). Theory of Elastic Mechanisms in Fibrous Proteins. J. Am. Chem. Soc. *78*, 5222–5235.

Franks, N.P., Jenkins, A., Conti, E., Lieb, W.R., and Brick, P. (1998). Structural basis for the inhibition of firefly luciferase by a general anesthetic. Biophys. J. *75*, 2205–2211.

Fuglsang, A. (2003). Codon optimizer: a freeware tool for codon optimization. Protein Expr. Purif. *31*, 247–249.

Fujinami, S., and Fujisawa, M. (2010). Industrial applications of alkaliphiles and their enzymes--past, present and future. Environ. Technol. *31*, 845–856.

Fukumori, F., Kudo, T., Sashihara, N., Nagata, Y., Ito, K., and Horikoshi, K. (1988). The third cellulase of alkalophilic Bacillus sp. strain N-4: evolutionary relationships within the cel gene family. Gene *76*, 289–298.

Galarneau, A., Primeau, M., Trudeau, L.-E., and Michnick, S.W. (2002). Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. Nat. Biotechnol. *20*, 619–622.

Garbe, D., Thiel, I.V., and Mootz, H.D. (2010). Protein trans-splicing on an M13 bacteriophage: towards directed evolution of a semisynthetic split intein by phage display. J. Pept. Sci. Off. Publ. Eur. Pept. Soc. *16*, 575–581.

Garrett, J.B., Kretz, K.A., O'Donoghue, E., Kerovuo, J., Kim, W., Barton, N.R., Hazlewood, G.P., Short, J.M., Robertson, D.E., and Gray, K.A. (2004). Enhancing the Thermal Tolerance and Gastric Performance of a Microbial Phytase for Use as a Phosphate-Mobilizing Monogastric-Feed Supplement. Appl. Environ. Microbiol. *70*, 3041–3046.

Gasteiger, E., Hoogland, C., Gatticker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocols Handbook, (Humana Press), pp. 571–607.

Gaucher, E.A., Govindarajan, S., and Ganesh, O.K. (2008). Palaeotemperature trend for Precambrian life inferred from resurrected proteins. Nature *451*, 704–707.

Geoghegan, K.F., Dixon, H.B., Rosner, P.J., Hoth, L.R., Lanzetti, A.J., Borzilleri, K.A., Marr, E.S., Pezzullo, L.H., Martin, L.B., LeMotte, P.K., et al. (1999). Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. Anal. Biochem. *267*, 169–184.

Gera, N., Hussain, M., and Rao, B.M. (2013). Protein selection using yeast surface display. Methods San Diego Calif *60*, 15–26.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Goldenberg, D.P., and Creighton, T.E. (1983). Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. J. Mol. Biol. *165*, 407–413.

Goldsmith, M., and Tawfik, D.S. (2013). Enzyme engineering by targeted libraries. Methods Enzymol. *523*, 257–283.

Gong, M., Gong, F., and Yanofsky, C. (2006). Overexpression of tnaC of Escherichia coli inhibits growth by depleting tRNA2Pro availability. J. Bacteriol. *188*, 1892–1898.

Goodman, D.B., Church, G.M., and Kosuri, S. (2013). Causes and effects of N-terminal codon bias in bacterial genes. Science *342*, 475–479.

Gouda, M.D., Singh, S.A., Rao, A.G.A., Thakur, M.S., and Karanth, N.G. (2003). Thermal inactivation of glucose oxidase. Mechanism and stabilization using additives. J. Biol. Chem. *278*, 24324–24333.

Gould, S.J., and Subramani, S. (1988). Firefly luciferase as a tool in molecular and cell biology. Anal. Biochem. *175*, 5–13.

Guerois, R., Nielsen, J.E., and Serrano, L. (2002). Predicting Changes in the Stability of Proteins and Protein Complexes: A Study of More Than 1000 Mutations. J. Mol. Biol. *320*, 369–387.

Guo, H.H., Choe, J., and Loeb, L.A. (2004). Protein tolerance to random amino acid change. Proc. Natl. Acad. Sci. U. S. A. *101*, 9205–9210.

Gustafsson, C. (2009). Tools Designed to Regulate Translational Efficiency. In The Metabolic Pathway Engineering Handbook, (CRC Press), pp. 9-1-9–14.

Haering, C.H., Farcas, A.-M., Arumugam, P., Metson, J., and Nasmyth, K. (2008). The cohesin ring concatenates sister DNA molecules. Nature *454*, 297–301.

Hagan, R.M., Bjornsson, R., McMahon, S.A., Schomburg, B., Braithwaite, V., Buhl, M., Naismith, J.H., and Schwarz-Linek, U. (2010). NMR spectroscopic and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond. Angew. Chem. *49*, 8421–8425.

Hartman, M.A., and Spudich, J.A. (2012). The myosin superfamily at a glance. J Cell Sci *125*, 1627–1632.

He, C., Liu, J., Xie, L., Zhang, Q., Li, C., Gui, D., Zhang, G., and Wu, C. (2009). Activity and Thermal Stability Improvements of Glucose Oxidase upon Adsorption on Core−Shell PMMA−BSA Nanoparticles. Langmuir *25*, 13456–13460.

Heller, A., and Feldman, B. (2008). Electrochemical Glucose Sensors and Their Applications in Diabetes Management. Chem. Rev. *108*, 2482–2505.

Hess, G.T., Cragnolini, J.J., Popp, M.W., Allen, M.A., Dougan, S.K., Spooner, E., Ploegh, H.L., Belcher, A.M., and Guimaraes, C.P. (2012). An M13 bacteriophage display framework that allows sortase-mediated modification of surface-accessible phage proteins. Bioconjug. Chem. *23*, 1478–1487.

van ’t Hof, W., Hansenová Maňásková, S., Veerman, E.C.I., and Bolscher, J.G.M. (2015). Sortase-mediated backbone cyclization of proteins and peptides. Biol. Chem. *396*, 283–293.

Hoover, D.M., and Lubkowski, J. (2002). DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. Nucleic Acids Res. *30*, e43.

Houde, D.J., and Berkowitz, S.A. (2014). Biophysical Characterization of Proteins in Developing Biopharmaceuticals (Newnes).

Howarth, M., and Ting, A.Y. (2008). Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin. Nat. Protoc. *3*, 534–545.

Hu, C.-D., Chinenov, Y., and Kerppola, T.K. (2002). Visualization of Interactions among bZIP and Rel Family Proteins in Living Cells Using Bimolecular Fluorescence Complementation. Mol. Cell *9*, 789–798.

Hu, X., Hu, H., Melvin, J.A., Clancy, K.W., McCafferty, D.G., and Yang, W. (2011). Autocatalytic Intramolecular Isopeptide Bond Formation in Gram-Positive Bacterial Pili: A QM/MM Simulation. J. Am. Chem. Soc. *133*, 478–485.

Huber, D., Boyd, D., Xia, Y., Olma, M.H., Gerstein, M., and Beckwith, J. (2005). Use of thioredoxin as a reporter to identify a subset of Escherichia coli signal sequences that promote signal recognition particle-dependent translocation. J. Bacteriol. *187*, 2983–2991.

Hugonnet, J.-E., and Blanchard, J.S. (2007). Irreversible inhibition of the Mycobacterium tuberculosis beta-lactamase by clavulanate. Biochemistry (Mosc.) *46*, 11998–12004.

Hui, Y.H., and Sherkat, F. (2005). Handbook of Food Science, Technology, and Engineering - 4 Volume Set (CRC Press).

Huse, W.D., Sastry, L., Iverson, S.A., Kang, A.S., Alting-Mees, M., Burton, D.R., Benkovic, S.J., and Lerner, R.A. (1989). Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. Science *246*, 1275–1281.

Ito, S., Shikata, S., Ozaki, K., Kawai, S., Okamoto, K., Inoue, S., Takei, A., Ohta, Y.-I., Satoh, T., and Ito, S. (1989). Alkaline cellulase for laundry detergents: Production by bacillus sp. ksm-635 and enzymatic properties. Agric. Biol. Chem. *53*, 1275–1281.

Iwai, H., and Plückthun, A. (1999). Circular β-lactamase: stability enhancement by cyclizing the backbone. FEBS Lett. *459*, 166–172.

Iwai, H., Lingel, A., and Plückthun, A. (2001). Cyclic Green Fluorescent Protein Produced in Vivo Using an Artificially Split PI-PfuI Intein from Pyrococcus furiosus. J. Biol. Chem. *276*, 16548–16554.

Izoré, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., Di Guilmi, A.M., and Dessen, A. (2010). Structural basis of host cell recognition by the pilus adhesin from Streptococcus pneumoniae. Structure *18*, 106–115.

Jacobson, G.R., Schaffer, M.H., Stark, G.R., and Vanaman, T.C. (1973). Specific chemical cleavage in high yield at the amino peptide bonds of cysteine and cystine residues. J. Biol. Chem. *248*, 6583–6591.

Jansson, R., Thatikonda, N., Lindberg, D., Rising, A., Johansson, J., Nygren, P.-Å., and Hedhammar, M. (2014). Recombinant Spider Silk Genetically Functionalized with Affinity Domains. Biomacromolecules *15*, 1696–1706.

Jayaraj, S., Reid, R., and Santi, D.V. (2005). GeMS: an advanced software package for designing synthetic genes. Nucleic Acids Res. *33*, 3011–3016.

Jia, X., Kwon, S., Wang, C.-I.A., Huang, Y.-H., Chan, L.Y., Tan, C.C., Rosengren, K.J., Mulvenna, J.P., Schroeder, C.I., and Craik, D.J. (2014). Semienzymatic Cyclization of Disulfide-rich Peptides Using Sortase A. J. Biol. Chem. *289*, 6627–6638.

Jiang, Y., Sun, W., Zhou, L., Ma, L., He, Y., and Gao, J. (2016). Improved Performance of Lipase Immobilized on Tannic Acid-Templated Mesoporous Silica Nanoparticles. Appl. Biochem. Biotechnol.

Jochens, H., and Bornscheuer, U.T. (2010). Natural diversity to guide focused directed evolution. Chembiochem Eur. J. Chem. Biol. *11*, 1861–1866.

Johannes, T.W., Woodyer, R.D., and Zhao, H. (2005). Directed Evolution of a Thermostable Phosphite Dehydrogenase for NAD(P)H Regeneration. Appl. Environ. Microbiol. *71*, 5728–5734.

Joo, H.-S., Kumar, C.G., Park, G.-C., Paik, S.R., and Chang, C.-S. (2003). Oxidant and SDS-stable alkaline protease from Bacillus clausii I-52: production and some properties. J. Appl. Microbiol. *95*, 267–272.

Juillerat, A., Gronemeyer, T., Keppler, A., Gendreizig, S., Pick, H., Vogel, H., and Johnsson, K. (2003). Directed evolution of O6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. Chem. Biol. *10*, 313–317.

Kaluz, S., Kaluzová, M., and Pastorek, J. (1999). Inverse PCR-generated internally deleted constructs for direct characterization of promoter regulatory regions. BioTechniques *26*, 446–450.

Kang, H.J., and Baker, E.N. (2009). Intramolecular isopeptide bonds give thermodynamic and proteolytic stability to the major pilin protein of Streptococcus pyogenes. J. Biol. Chem. *284*, 20729–20737.

Kang, H.J., and Baker, E.N. (2011). Intramolecular isopeptide bonds: protein crosslinks built for stress? Trends Biochem. Sci. *36*, 229–237.

Kang, H.J., Coulibaly, F., Clow, F., Proft, T., and Baker, E.N. (2007). Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. Science *318*, 1625–1628.

Kellogg, E.H., Leaver-Fay, A., and Baker, D. (2011). Role of conformational sampling in computing mutation-induced changes in protein structure and stability. Proteins *79*, 830–838.

Keresztessy, Z., Csosz, E., Hársfalvi, J., Csomós, K., Gray, J., Lightowlers, R.N., Lakey, J.H., Balajthy, Z., and Fésüs, L. (2006). Phage display selection of efficient glutamine-donor substrate peptides for transglutaminase 2. Protein Sci. Publ. Protein Soc. *15*, 2466–2480.

Kerovuo, J. (2000). A Novel Phytase from Bacillus. Characterization and Production of the Enzyme. University of Helsinki.

Kerovuo, J., Lauraeus, M., Nurminen, P., Kalkkinen, N., and Apajalahti, J. (1998). Isolation, Characterization, Molecular Gene Cloning, and Sequencing of a Novel Phytase from Bacillus subtilis. Appl. Environ. Microbiol. *64*, 2079–2085.

Kim, C.-S., Pierre, B., Ostermeier, M., Looger, L.L., and Kim, J.R. (2009). Enzyme stabilization by domain insertion into a thermophilic protein. Protein Eng. Des. Sel. *22*, 615–623.

Kim, S., Bae, S.Y., Lee, B.Y., and Kim, T.D. (2012). Coaggregation of amyloid fibrils for the preparation of stable and immobilized enzymes. Anal. Biochem. *421*, 776–778.

Kintses, B., Simon, Z., Gyimesi, M., Tóth, J., Jelinek, B., Niedetzky, C., Kovács, M., and Málnási-Csizmadia, A. (2006). Enzyme Kinetics above Denaturation Temperature: A Temperature-Jump/Stopped-Flow Apparatus. Biophys. J. *91*, 4605–4610.

Kintses, B., Hein, C., Mohamed, M.F., Fischlechner, M., Courtois, F., Lainé, C., and Hollfelder, F. (2012). Picoliter Cell Lysate Assays in Microfluidic Droplet Compartments for Directed Enzyme Evolution. Chem. Biol. *19*, 1001–1009.

Kommoju, P.-R., Chen, Z., Bruckner, R.C., Mathews, F.S., and Jorns, M.S. (2011). Probing oxygen activation sites in two flavoprotein oxidases using chloride as an oxygen surrogate. Biochemistry (Mosc.) *50*, 5521–5534.

Krumpe, L.R.H., Atkinson, A.J., Smythers, G.W., Kandel, A., Schumacher, K.M., McMahon, J.B., Makowski, L., and Mori, T. (2006). T7 lytic phage-displayed peptide libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide libraries. Proteomics *6*, 4210–4222.

Kuipers, R.K.P., Joosten, H.-J., Verwiel, E., Paans, S., Akerboom, J., van der Oost, J., Leferink, N.G.H., van Berkel, W.J.H., Vriend, G., and Schaap, P.J. (2009). Correlated mutation analyses on super-family alignments reveal functionally important residues. Proteins *76*, 608–616.

Le, Q.A.T., Joo, J.C., Yoo, Y.J., and Kim, Y.H. (2012). Development of thermostable Candida antarctica lipase B through novel in silico design of disulfide bridge. Biotechnol. Bioeng. *109*, 867–876.

Lehmann, M., Loch, C., Middendorf, A., Studer, D., Lassen, S.F., Pasamontes, L., van Loon, A.P.G.M., and Wyss, M. (2002). The consensus concept for thermostability engineering of proteins: further proof of concept. Protein Eng. *15*, 403–411.

Lei, X.G., Weaver, J.D., Mullaney, E., Ullah, A.H., and Azain, M.J. (2013). Phytase, a New Life for an “Old” Enzyme. Annu. Rev. Anim. Biosci. *1*, 283–309.

Levin, I., and Aharoni, A. (2012). Evolution in Microfluidic Droplet. Chem. Biol. *19*, 929–931.

Li, L., Fierer, J.O., Rapoport, T.A., and Howarth, M. (2013). Structural Analysis and Optimization of the Covalent Association between SpyCatcher and a Peptide Tag. J. Mol. Biol.

Li, S., Yang, X., Yang, S., Zhu, M., and Wang, X. (2012). Technology Prospecting on Enzymes: Application, Marketing and Engineering. Comput. Struct. Biotechnol. J. *2*.

Liao, Y.-D., Jeng, J.-C., Wang, C.-F., Wang, S.-C., and Chang, S.-T. (2004). Removal of N-terminal methionine from recombinant proteins by engineered E. coli methionine aminopeptidase. Protein Sci. Publ. Protein Soc. *13*, 1802–1810.

van Lieshout, J.F.T., Pérez Gutiérrez, O.N., Vroom, W., Planas, A., de Vos, W.M., van der Oost, J., and Koutsopoulos, S. (2012). Thermal stabilization of an endoglucanase by cyclization. Appl. Biochem. Biotechnol. *167*, 2039–2053.

Lin-Cereghino, G.P., Stark, C.M., Kim, D., Chang, J., Shaheen, N., Poerwanto, H., Agari, K., Moua, P., Low, L.K., Tran, N., et al. (2013). The Effect of α-Mating Factor Secretion Signal Mutations on Recombinant Protein Expression in Pichia pastoris. Gene *519*, 311–317.

Linder, M. (2012). COMPUTATIONAL ENZYME DESIGN: ADVANCES, HURDLES AND POSSIBLE WAYS FORWARD. Comput. Struct. Biotechnol. J. *2*, 1–8.

Liu, L., Yang, H., Shin, H., Chen, R.R., Li, J., Du, G., and Chen, J. (2013). How to achieve high-level expression of microbial enzymes: strategies and perspectives. Bioengineered *4*, 212–223.

Liu, Z., Zhou, H., Wang, W., Tan, W., Fu, Y.-X., and Zhu, M. (2014). A novel method for synthetic vaccine construction based on protein assembly. Sci. Rep. *4*, 7266.

Lundin, A., Rickardsson, A., and Thore, A. (1976). Continuous monitoring of ATP-converting reactions by purified firefly luciferase. Anal. Biochem. *75*, 611–620.

Major, L.L., Edgar, T.D., Yee Yip, P., Isaksson, L.A., and Tate, W.P. (2002). Tandem termination signals: myth or reality? FEBS Lett. *514*, 84–89.

de Marco, A., Casatta, E., Savaresi, S., and Geerlof, A. (2004). Recombinant proteins fused to thermostable partners can be purified by heat incubation. J. Biotechnol. *107*, 125–133.

Marquart, M., Walter, J., Deisenhofer, J., Bode, W., and Huber, R. (1983). The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. Acta Crystallogr. B *39*, 480–490.

Marques, S.M., and Esteves da Silva, J.C.G. (2009). Firefly bioluminescence: a mechanistic approach of luciferase catalyzed reactions. IUBMB Life *61*, 6–17.

Marraffini, L.A., Dedent, A.C., and Schneewind, O. (2006). Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. Microbiol. Mol. Biol. Rev. MMBR *70*, 192–221.

Masakapalli, S.K., Ritala, A., Dong, L., van der Krol, A.R., Oksman-Caldentey, K.-M., Ratcliffe, R.G., and Sweetlove, L.J. (2014). Metabolic flux phenotype of tobacco hairy roots engineered for increased geraniol production. Phytochemistry *99*, 73–85.

Mateo, C., Palomo, J.M., van Langen, L.M., van Rantwijk, F., and Sheldon, R.A. (2004). A new, mild cross-linking methodology to prepare cross-linked enzyme aggregates. Biotechnol. Bioeng. *86*, 273–276.

Matsumura, M., Becktel, W.J., Levitt, M., and Matthews, B.W. (1989). Stabilization of phage T4 lysozyme by engineered disulfide bonds. Proc. Natl. Acad. Sci. *86*, 6562–6566.

Matsunaga, R., Yanaka, S., Nagatoishi, S., and Tsumoto, K. (2013). Hyperthin nanochains composed of self-polymerizing protein shackles. Nat. Commun. *4*, 2211.

McCafferty, J., Griffiths, A.D., Winter, G., and Chiswell, D.J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. Nature *348*, 552–554.

Meng, Y., Zhao, M., Yang, M., Zhang, Q., Hao, J., and Meng, Y. (2014). Production and characterization of recombinant glucose oxidase from Aspergillus niger expressed in Pichia pastoris. Lett. Appl. Microbiol. *58*, 393–400.

Miyazaki, K., and Arnold, F.H. (1999). Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. J. Mol. Evol. *49*, 716–720.

Mohamad, N.R., Marzuki, N.H.C., Buang, N.A., Huyop, F., and Wahab, R.A. (2015). An overview of technologies for immobilization of enzymes and surface analysis techniques for immobilized enzymes. Biotechnol. Biotechnol. Equip. *29*, 205–220.

Moore, B.D., Deere, J., Edrada-Ebel, R., Ingram, A., and van der Walle, C.F. (2010). Isolation of recombinant proteins from culture broth by co-precipitation with an amino acid carrier to form stable dry powders. Biotechnol. Bioeng. *106*, 764–773.

Morra, R., Shankar, J., Robinson, C.J., Halliwell, S., Butler, L., Upton, M., Hay, S., Micklefield, J., and Dixon, N. (2016). Dual transcriptional-translational cascade permits cellular level tuneable expression control. Nucleic Acids Res. *44*, e21.

Morrison, J.F., and Stone, S.R. (1988). Mechanism of the reaction catalyzed by dihydrofolate reductase from Escherichia coli: pH and deuterium isotope effects with NADPH as the variable substrate. Biochemistry (Mosc.) *27*, 5499–5506.

Nguyen, G.K.T., Wang, S., Qiu, Y., Hemu, X., Lian, Y., and Tam, J.P. (2014). Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. Nat. Chem. Biol. *10*, 732–738.

Nguyen, G.K.T., Kam, A., Loo, S., Jansson, A.E., Pan, L.X., and Tam, J.P. (2015). Butelase 1: A Versatile Ligase for Peptide and Protein Macrocyclization. J. Am. Chem. Soc. *137*, 15398–15401.

Niesen, F.H., Berglund, H., and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat. Protoc. *2*, 2212–2221.

Ogino, H., Watanabe, F., Yamada, M., Nakagawa, S., Hirose, T., Noguchi, A., Yasuda, M., and Ishikawa, H. (1999). Purification and characterization of organic solvent-stable protease from organic solvent-tolerant Pseudomonas aeruginosa PST-01. J. Biosci. Bioeng. *87*, 61–68.

Oke, M., Carter, L.G., Johnson, K.A., Liu, H., McMahon, S.A., Yan, X., Kerou, M., Weikart, N.D., Kadi, N., Sheikh, M.A., et al. (2010). The Scottish Structural Proteomics Facility: targets, methods and outputs. J. Struct. Funct. Genomics *11*, 167–180.

O'Loughlin, T.L., Patrick, W.M., and Matsumura, I. (2006). Natural history as a predictor of protein evolvability. Protein Eng. Des. Sel. *19*, 439–442.

Ortega, A., Amorós, D., and García de la Torre, J. (2011). Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. Biophys. J. *101*, 892–898.

Pace, C.N., Grimsley, G.R., Thomson, J.A., and Barnett, B.J. (1988). Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. J. Biol. Chem. *263*, 11820–11825.

Packer, M.S., and Liu, D.R. (2015). Methods for the directed evolution of proteins. Nat. Rev. Genet. *16*, 379–394.

Parthasarathy, R., Subramanian, S., and Boder, E.T. (2007). Sortase A as a novel molecular "stapler" for sequence-specific protein conjugation. Bioconjug. Chem. *18*, 469–476.

Parthiban, V., Gromiha, M.M., and Schomburg, D. (2006). CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Res. *34*, W239-242.

Paulmurugan, R., Umezawa, Y., and Gambhir, S.S. (2002). Noninvasive imaging of protein–protein interactions in living subjects by using reporter protein complementation and reconstitution strategies. Proc. Natl. Acad. Sci. *99*, 15608–15613.

Paz-Alfaro, K.J., Ruiz-Granados, Y.G., Uribe-Carvajal, S., and Sampedro, J.G. (2009). Trehalose-mediated thermal stabilization of glucose oxidase from Aspergillus niger. J. Biotechnol. *141*, 130–136.

Pekel, N., Salih, B., and Güven, O. (2005). Enhancement of stability of glucose oxidase by immobilization onto metal ion-chelated poly(N-vinyl imidazole) hydrogels. J. Biomater. Sci. Polym. Ed. *16*, 253–266.

Pelletier, J.N., Arndt, K.M., Plückthun, A., and Michnick, S.W. (1999). An in vivo library-versus-library selection of optimized protein-protein interactions. Nat. Biotechnol. *17*, 683–690.

Perry, L.J., and Wetzel, R. (1984). Disulfide bond engineered into T4 lysozyme: stabilization of the protein toward thermal inactivation. Science *226*, 555–557.

Pertzev, A.V., and Nicholson, A.W. (2006). Characterization of RNA sequence determinants and antideterminants of processing reactivity for a minimal substrate of Escherichia coli ribonuclease III. Nucleic Acids Res. *34*, 3708–3721.

Pickering, G.J., Heatherbell, D.A., and Barnes, M.F. (1999). The Production of Reduced-Alcohol Wine Using Glucose Oxidase-Treated Juice. Part II. Stability and SO2-Binding. Am. J. Enol. Vitic. *50*, 299–306.

Pierre, B., Xiong, T., Hayles, L., Guntaka, V.R., and Kim, J.R. (2011). Stability of a guest protein depends on stability of a host protein in insertional fusion. Biotechnol. Bioeng. *108*, 1011–1020.

Pierre, B., Labonte, J.W., Xiong, T., Aoraha, E., Williams, A., Shah, V., Chau, E., Helal, K.Y., Gray, J.J., and Kim, J.R. (2015). Molecular Determinants for Protein Stabilization by Insertional Fusion to a Thermophilic Host Protein. ChemBioChem *16*, 2392–2402.

Ping, S.-Z., Zhang, W., Lin, M., Yan, Y.-L., Zhao, Z.-L., Lu, W., and Chen, M. (2010). Cyclized phytase with improved heat stability and proteinase resistance.

Piotukh, K., Geltinger, B., Heinrich, N., Gerth, F., Beyermann, M., Freund, C., and Schwarzer, D. (2011). Directed evolution of sortase A mutants with altered substrate selectivity profiles. J. Am. Chem. Soc. *133*, 17536–17539.

Plückthun, A. (2012). Ribosome display: a perspective. Methods Mol. Biol. Clifton NJ *805*, 3–28.

Polyak, S.W., Chapman-Smith, A., Mulhern, T.D., Cronan, J.E., Jr, and Wallace, J.C. (2001). Mutational analysis of protein substrate presentation in the post-translational attachment of biotin to biotin domains. J. Biol. Chem. *276*, 3037–3045.

Popat, A., Hartono, S.B., Stahr, F., Liu, J., Qiao, S.Z., and Lu, G.Q. (Max) (2011). Mesoporous silica nanoparticles for bioadsorption, enzyme immobilisation, and delivery carriers. Nanoscale *3*, 2801–2818.

Popp, M.W., Dougan, S.K., Chuang, T.-Y., Spooner, E., and Ploegh, H.L. (2011). Sortase-catalyzed transformations that improve the properties of cytokines. Proc. Natl. Acad. Sci. *108*, 3169–3174.

Qi, H., Lu, H., Qiu, H.-J., Petrenko, V., and Liu, A. (2012). Phagemid Vectors for Phage Display: Properties, Characteristics and Construction. J. Mol. Biol. *417*, 129–143.

Quan, J., and Tian, J. (2009). Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways. PLoS ONE *4*, e6441.

Rajan, A., and Abraham, T.E. (2007). Studies on crystallization and cross-linking of lipase for biocatalysis. Bioprocess Biosyst. Eng. *31*, 87–94.

Ramakrishnan, V., Srinivasan, S.P., Salem, S.M., Matthews, S.J., Colón, W., Zaki, M., and Bystroff, C. (2012). Geofold: topology-based protein unfolding pathways capture the effects of engineered disulfides on kinetic stability. Proteins *80*, 920–934.

Reddington, S.C. Personal communication.

Reetz, M.T., and Carballeira, J.D. (2007). Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. Nat. Protoc. *2*, 891–903.

Reetz, M.T., Carballeira, J.D., and Vogel, A. (2006). Iterative Saturation Mutagenesis on the Basis of B Factors as a Strategy for Increasing Protein Thermostability. Angew. Chem. Int. Ed. *45*, 7745–7751.

Reetz, M.T., Kahakeaw, D., and Lohmer, R. (2008). Addressing the numbers problem in directed evolution. Chembiochem Eur. J. Chem. Biol. *9*, 1797–1804.

de Regil, R., and Sandoval, G. (2013). Biocatalysis for Biobased Chemicals. Biomolecules *3*, 812–847.

Remy, I., Campbell-Valois, F.X., and Michnick, S.W. (2007). Detection of protein–protein interactions using a simple survival protein-fragment complementation assay based on the enzyme dihydrofolate reductase. Nat. Protoc. *2*, 2120–2125.

Robinson, P.K. (2015). Enzymes: principles and biotechnological applications. Essays Biochem. *59*, 1–41.

Rogers, J.M., Steward, A., and Clarke, J. (2013). Folding and Binding of an Intrinsically Disordered Protein: Fast, but Not “Diffusion-Limited.” J. Am. Chem. Soc. *135*, 1415–1422.

Romero, P.A., and Arnold, F.H. (2009). Exploring protein fitness landscapes by directed evolution. Nat. Rev. Mol. Cell Biol. *10*, 866–876.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlén, M., and Nyrén, P. (1996). Real-Time DNA Sequencing Using Detection of Pyrophosphate Release. Anal. Biochem. *242*, 84–89.

Rondot, S., Koch, J., Breitling, F., and Dübel, S. (2001). A helper phage to improve single-chain antibody presentation in phage display. Nat. Biotechnol. *19*, 75–78.

Rossi, F., Charlton, C.A., and Blau, H.M. (1997). Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation. Proc. Natl. Acad. Sci. U. S. A. *94*, 8405–8410.

Russel, M., Kidd, S., and Kelley, M.R. (1986). An improved filamentous helper phage for generating single-stranded plasmid DNA. Gene *45*, 333–338.

Sagar, D.M., Aoudjane, S., Gaudet, M., Aeppli, G., and Dalby, P.A. (2013). Optically induced thermal gradients for protein characterization in nanolitre-scale samples in microfluidic devices. Sci. Rep. *3*, 2130.

Sahdev, S., Khattar, S.K., and Saini, K.S. (2008). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol. Cell. Biochem. *307*, 249–264.

Sakamoto, R., Nishikori, S., and Shiraki, K. (2004). High temperature increases the refolding yield of reduced lysozyme: implication for the productive process for folding. Biotechnol. Prog. *20*, 1128–1133.

Salis, H.M., Mirsky, E.A., and Voigt, C.A. (2009). Automated Design of Synthetic Ribosome Binding Sites to Precisely Control Protein Expression. Nat. Biotechnol. *27*, 946–950.

Sanchez-Romero, I., Ariza, A., Wilson, K.S., Skjøt, M., Vind, J., Maria, L.D., Skov, L.K., and Sanchez-Ruiz, J.M. (2013). Mechanism of Protein Kinetic Stabilization by Engineered Disulfide Crosslinks. PLOS ONE *8*, e70013.

Sangeetha, K., and Emilia Abraham, T. (2008). Preparation and characterization of cross-linked enzyme aggregates (CLEA) of Subtilisin for controlled release applications. Int. J. Biol. Macromol. *43*, 314–319.

Savitzky, A., and Golay, M.J.E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem. *36*, 1627–1639.

Sawaya, M.R., and Kraut, J. (1997). Loop and subdomain movements in the mechanism of Escherichia coli dihydrofolate reductase: crystallographic evidence. Biochemistry (Mosc.) *36*, 586–603.

Scharf, M.E., Karl, Z.J., Sethi, A., and Boucias, D.G. (2011). Multiple Levels of Synergistic Collaboration in Termite Lignocellulose Digestion. PLOS ONE *6*, e21709.

Schneider, T.D., and Stephens, R.M. (1990). Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. *18*, 6097–6100.

Schnell, J.R., Dyson, H.J., and Wright, P.E. (2004). Structure, Dynamics, and Catalytic Function of Dihydrofolate Reductase. Annu. Rev. Biophys. Biomol. Struct. *33*, 119–140.

Schoene, C., Fierer, J.O., Bennett, S.P., and Howarth, M. (2014). SpyTag/SpyCatcher Cyclization Confers Resilience to Boiling on a Mesophilic Enzyme. Angew. Chem. *53*, 6101–6104.

Schoene, C., Bennett, S.P., and Howarth, M. (2016a). SpyRing interrogation: analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries. Sci. Rep. *6*, 21151.

Schoene, C., Bennett, S.P., and Howarth, M. (2016b). SpyRings Declassified: A Blueprint for Using Isopeptide-Mediated Cyclization to Enhance Enzyme Thermal Resilience. B.-M. in Enzymology, ed. (Academic Press), p.

Scholl, Z.N., Yang, W., and Marszalek, P.E. (2014). Chaperones Rescue Luciferase Folding by Separating Its Domains. J. Biol. Chem. *289*, 28607–28618.

Schröder, H., Langer, T., Hartl, F.U., and Bukau, B. (1993). DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. EMBO J. *12*, 4137–4144.

Schumann, F.H., Varadan, R., Tayakuniyil, P.P., Grossman, J.H., Camarero, J.A., and Fushman, D. (2015). Changing the topology of protein backbone: the effect of backbone cyclization on the structure and dynamics of a SH3 domain. Chem. Biol. *3*, 26.

Scott, C.P., Abel-Santos, E., Wall, M., Wahnon, D.C., and Benkovic, S.J. (1999). Production of cyclic peptides and proteins in vivo. Proc. Natl. Acad. Sci. *96*, 13638–13643.

Selimova, L.M., Zaĭdes, V.M., and Zhdanov, V.M. (1981). Effect of the inter- and intramolecular disulfide bonds on the electrophoretic properties of influenza virus proteins. Vopr. Virusol. 358–365.

Selle, P.H., Ravindran, V., Bryden, W.L., and Scott, T. (2006). Influence of Dietary Phytate and Exogenous Phytase on Amino Acid Digestibility in Poultry: A Review. J. Poult. Sci. *43*, 89–103.

Shah, N.H., and Muir, T.W. (2014). Inteins: Nature's Gift to Protein Chemists. Chem. Sci. R. Soc. Chem. 2010 *5*, 446–461.

Sharma, A., Kawarabayasi, Y., and Satyanarayana, T. (2012). Acidophilic bacteria and archaea: acid stable biocatalysts and their potential applications. Extrem. Life Extreme Cond. *16*, 1–19.

Sheldon, R.A. (2007). Cross-linked enzyme aggregates (CLEAs): stable and recyclable biocatalysts. Biochem. Soc. Trans. *35*, 1583–1587.

Sheldon, R.A. (2011). Characteristic features and biotechnological applications of cross-linked enzyme aggregates (CLEAs). Appl. Microbiol. Biotechnol. *92*, 467–477.

Sheldon, R.A., and van Pelt, S. (2013). Enzyme immobilisation in biocatalysis: why, what and how. Chem. Soc. Rev. *42*, 6223–6235.

Shi, S., Semple, A., Cheung, J., and Shameem, M. (2013). DSF method optimization and its application in predicting protein thermal aggregation kinetics. J. Pharm. Sci. *102*, 2471–2483.

Siebold, C., and Erni, B. (2002). Intein-mediated cyclization of a soluble and a membrane protein in vivo: function and stability. Biophys. Chem. *96*, 163–171.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Sinha, J., Plantz, B.A., Inan, M., and Meagher, M.M. (2005). Causes of proteolytic degradation of secreted recombinant proteins produced in methylotrophic yeast Pichia pastoris: case study with recombinant ovine interferon-tau. Biotechnol. Bioeng. *89*, 102–112.

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., and Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. Protein Sci. Publ. Protein Soc. *18*, 936–948.

Slabinski, L., Jaroszewski, L., Rodrigues, A.P.C., Rychlewski, L., Wilson, I.A., Lesley, S.A., and Godzik, A. (2007). The challenge of protein structure determination—lessons from structural genomics. Protein Sci. Publ. Protein Soc. *16*, 2472–2482.

Slootweg, E.J., Keller, H.J.H.G., Hink, M.A., Borst, J.W., Bakker, J., and Schots, A. (2006). Fluorescent T7 display phages obtained by translational frameshift. Nucleic Acids Res. *34*, e137–e137.

Smith, G.P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science *228*, 1315–1317.

Speck, J., Arndt, K.M., and Müller, K.M. (2011). Efficient phage display of intracellularly folded proteins mediated by the TAT pathway. Protein Eng. Des. Sel. *24*, 473–484.

St. Clair, N.L., and Navia, M.A. (1992). Cross-linked enzyme crystals as robust biocatalysts. J. Am. Chem. Soc. *114*, 7314–7316.

Stark, M., Grip, S., Rising, A., Hedhammar, M., Engström, W., Hjälm, G., and Johansson, J. (2007). Macroscopic fibers self-assembled from recombinant miniature spider silk proteins. Biomacromolecules *8*, 1695–1701.

Stec, B., Holtz, K.M., Wojciechowski, C.L., and Kantrowitz, E.R. (2005). Structure of the wild-type TEM-1 beta-lactamase at 1.55 A and the mutant enzyme Ser70Ala at 2.1 A suggest the mode of noncovalent catalysis for the mutant enzyme. Acta Crystallogr. D Biol. Crystallogr. *61*, 1072–1079.

Steiner, K., and Schwab, H. (2012). Recent advances in rational approaches for enzyme engineering. Comput. Struct. Biotechnol. J. *2*.

Steiner, D., Forrer, P., Stumpp, M.T., and Plückthun, A. (2006). Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. Nat. Biotechnol. *24*, 823–831.

Stemmer, W.P. (1994). Rapid evolution of a protein in vitro by DNA shuffling. Nature *370*, 389–391.

Stepankova, V., Bidmanova, S., Koudelakova, T., Prokop, Z., Chaloupkova, R., and Damborsky, J. (2013). Strategies for Stabilization of Enzymes in Organic Solvents. ACS Catal. *3*, 2823–2836.

Sumner, J.B. (1944). A Method for the Colorimetric Determination of Phosphorus. Science *100*, 413–414.

Sun, F., Zhang, W.-B., Mahdavi, A., Arnold, F.H., and Tirrell, D.A. (2014). Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. Proc. Natl. Acad. Sci. *111*, 11269–11274.

Sweetlove, L.J., Obata, T., and Fernie, A.R. (2014). Systems analysis of metabolic phenotypes: what have we learnt? Trends Plant Sci. *19*, 222–230.

Takahashi, H., Arai, M., Takenawa, T., Sota, H., Xie, Q.H., and Iwakura, M. (2007). Stabilization of Hyperactive Dihydrofolate Reductase by Cyanocysteine-mediated Backbone Cyclization. J. Biol. Chem. *282*, 9420–9429.

Tesar, M., Beckmann, C., Röttgen, P., Haase, B., Faude, U., and Timmis, K.N. (1995). Monoclonal antibody against pIII of filamentous phage: an immunological tool to study pIII fusion protein expression in phage display systems. Immunotechnology Int. J. Immunol. Eng. *1*, 53–64.

Thrane, S., Janitzek, C.M., Matondo, S., Resende, M., Gustavsson, T., de Jongh, W.A., Clemmensen, S., Roeffen, W., van de Vegte-Bolmer, M., van Gemert, G.J., et al. (2016). Bacterial superglue enables easy development of efficient virus-like particle based vaccines. J. Nanobiotechnology *14*, 30.

Toofanny, R.D., and Daggett, V. (2012). Understanding protein unfolding from molecular simulations. Wiley Interdiscip. Rev. Comput. Mol. Sci. *2*, 405–423.

Torres-Salas, P., Mate, D.M., Ghazi, I., Plou, F.J., Ballesteros, A.O., and Alcalde, M. (2013). Widening the pH Activity Profile of a Fungal Laccase by Directed Evolution. ChemBioChem *14*, 934–937.

Turner, P., Mamo, G., and Karlsson, E.N. (2007). Potential and utilization of thermophiles and thermostable enzymes in biorefining. Microb. Cell Factories *6*, 9.

Uwajima, T., Oshiro, T., Eguchi, T., Kuge, Y., Horiguchi, A., Igarashi, A., Mochida, K., and Iwakura, M. (1990). Chemo-enzymatic synthesis of optically pure l-leucovorin, an augmentor of 5-fluorouracil cytotoxicity against cancer. Biochem. Biophys. Res. Commun. *171*, 684–689.

Veggiani, G., Nakamura, T., Brenner, M.D., Gayet, R.V., Yan, J., Robinson, C.V., and Howarth, M. (2016). Programmable polyproteams built using twin peptide superglues. Proc. Natl. Acad. Sci. *113*, 1202–1207.

Velappan, N., Fisher, H.E., Pesavento, E., Chasteen, L., D'Angelo, S., Kiss, C., Longmire, M., Pavlik, P., and Bradbury, A.R.M. (2010). A comprehensive analysis of filamentous phage display vectors for cytoplasmic proteins: an analysis with different fluorescent proteins. Nucleic Acids Res. *38*, e22.

Velasco-Lozano, S., López-Gallego, F., Vázquez-Duhalt, R., Mateos-Díaz, J.C., Guisán, J.M., and Favela-Torres, E. (2014). Carrier-Free Immobilization of Lipase from Candida rugosa with Polyethyleneimines by Carboxyl-Activated Cross-Linking. Biomacromolecules *15*, 1896–1903.

Vellard, M. (2003). The enzyme as drug: application of enzymes as pharmaceuticals. Curr. Opin. Biotechnol. *14*, 444–450.

Vigneshvar, S., Sudhakumari, C.C., Senthilkumaran, B., and Prakash, H. (2016). Recent Advances in Biosensor Technology for Potential Applications – An Overview. Front. Bioeng. Biotechnol. *4*.

Villalobos, A., Ness, J.E., Gustafsson, C., Minshull, J., and Govindarajan, S. (2006). Gene Designer: a synthetic biology tool for constructing artificial DNA segments. BMC Bioinformatics *7*, 285.

Vyas, V.K., Ukawala, R.D., Ghate, M., and Chintha, C. (2012). Homology Modeling a Fast Tool for Drug Discovery: Current Perspectives. Indian J. Pharm. Sci. *74*, 1–17.

Waldhauer, M.C., Schmitz, S.N., Ahlmann-Eltze, C., Gleixner, J.G., Schmelas, C.C., Huhn, A.G., Bunne, C., Büscher, M., Horn, M., Klughammer, N., et al. (2015). Backbone circularization of Bacillus subtilis family 11 xylanase increases its thermostability and its resistance against aggregation. Mol. Biosyst. *11*, 3231–3243.

Wang, W., and Roberts, C.J. (2013). Non-Arrhenius Protein Aggregation. AAPS J. *15*, 840–851.

Wang, C.K.L., Kaas, Q., Chiche, L., and Craik, D.J. (2008). CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering. Nucleic Acids Res. *36*, D206-210.

Wang, J., Wang, Y., Wang, X., Zhang, D., Wu, S., and Zhang, G. (2016). Enhanced thermal stability of lichenase from Bacillus subtilis 168 by SpyTag/SpyCatcher-mediated spontaneous cyclization. Biotechnol. Biofuels *9*, 79.

Wang, T., Liu, X., Yu, Q., Zhang, X., Qu, Y., Gao, P., and Wang, T. (2005). Directed evolution for engineering pH profile of endoglucanase III from Trichoderma reesei. Biomol. Eng. *22*, 89–94.

Watanabe, K., Ohkuri, T., Yokobori, S., and Yamagishi, A. (2006). Designing thermostable proteins: ancestral mutants of 3-isopropylmalate dehydrogenase designed by using a phylogenetic tree. J. Mol. Biol. *355*, 664–674.

Weeks, S.D., Baranova, E.V., Heirbaut, M., Beelen, S., Shkumatov, A.V., Gusev, N.B., and Strelkov, S.V. (2014). Molecular structure and dynamics of the dimeric human small heat shock protein HSPB6. J. Struct. Biol. *185*, 342–354.

Wen, B., Peng, J., Zuo, X., Gong, Q., and Zhang, Z. (2014). Characterization of protein flexibility using small-angle x-ray scattering and amplified collective motion simulations. Biophys. J. *107*, 956–964.

Williams, N.K., Prosselkov, P., Liepinsh, E., Line, I., Sharipo, A., Littler, D.R., Curmi, P.M.G., Otting, G., and Dixon, N.E. (2002). In Vivo Protein Cyclization Promoted by a Circularly Permuted Synechocystis sp. PCC6803 DnaB Mini-intein. J. Biol. Chem. *277*, 7790–7798.

Williamson, M. (2011). How Proteins Work (Garland Science).

Wilson, D.S., and Keefe, A.D. (2001). Random mutagenesis by PCR. Curr. Protoc. Mol. Biol. Ed. Frederick M Ausubel Al *Chapter 8*, Unit8.3.

Wong, T.S., Roccatano, D., Zacharias, M., and Schwaneberg, U. (2006). A Statistical Analysis of Random Mutagenesis Methods Used for Directed Protein Evolution. J. Mol. Biol. *355*, 858–871.

Wood, D.W., and Camarero, J.A. (2014). Intein Applications: From Protein Purification and Labeling to Metabolic Control Methods. J. Biol. Chem. *289*, 14512–14519.

Wu, C.-H., Liu, I.-J., Lu, R.-M., and Wu, H.-C. (2016). Advancement and applications of peptide phage display technology in biomedical science. J. Biomed. Sci. *23*, 8.

Wyss, M., Brugger, R., Kronenberger, A., Rémy, R., Fimbel, R., Oesterhelt, G., Lehmann, M., and van Loon, A.P.G.M. (1999). Biochemical Characterization of Fungal Phytases (myo-Inositol Hexakisphosphate Phosphohydrolases): Catalytic Properties. Appl. Environ. Microbiol. *65*, 367–373.

Xia, X.-X., Qian, Z.-G., Ki, C.S., Park, Y.H., Kaplan, D.L., and Lee, S.Y. (2010). Native-sized recombinant spider silk protein produced in metabolically engineered Escherichia coli results in a strong fiber. Proc. Natl. Acad. Sci. *107*, 14059–14063.

Xu, J., Jiao, F., and Yu, L. (2008). Protein structure prediction using threading. Methods Mol. Biol. Clifton NJ *413*, 91–121.

Xu, Y., Lu, Y., Xing, Y., and Xi, T. (2015). Characterization of solvent, detergent and oxidizing agent stable protease from isolated Antarctic marine Streptomyces sp. XE-1. Wei Sheng Wu Xue Bao *55*, 311–320.

Yang, G., and Withers, S.G. (2009). Ultrahigh-throughput FACS-based screening for directed enzyme evolution. Chembiochem Eur. J. Chem. Biol. *10*, 2704–2715.

Yang, H., Liu, L., Li, J., Chen, J., and Du, G. (2015). Rational Design to Improve Protein Thermostability: Recent Advances and Prospects. ChemBioEng Rev. *2*, 87–94.

Yennamalli, R.M., Rader, A.J., Kenny, A.J., Wolt, J.D., and Sen, T.Z. (2013). Endoglucanases: insights into thermostability for biofuel applications. Biotechnol. Biofuels *6*, 136.

Yin, S., Ding, F., and Dokholyan, N.V. (2007). Eris: an automated estimator of protein stability. Nat. Methods *4*, 466–467.

Yoo, E.-H., and Lee, S.-Y. (2010). Glucose Biosensors: An Overview of Use in Clinical Practice. Sensors *10*, 4558–4576.

Zakeri, B. (2011). Peptide targeting by spontaneous isopeptide bond formation. University of Oxford.

Zakeri, B., and Howarth, M. (2010). Spontaneous Intermolecular Amide Bond Formation between Side Chains for Irreversible Peptide Targeting. J. Am. Chem. Soc. *132*, 4526–4527.

Zakeri, B., Fierer, J.O., Celik, E., Chittock, E.C., Schwarz-Linek, U., Moy, V.T., and Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. Proc. Natl. Acad. Sci. *109*, E690–E697.

Zeng, Y.-F., Ko, T.-P., Lai, H.-L., Cheng, Y.-S., Wu, T.-H., Ma, Y., Chen, C.-C., Yang, C.-S., Cheng, K.-J., Huang, C.-H., et al. (2011). Crystal Structures of Bacillus Alkaline Phytase in Complex with Divalent Metal ions and Inositol Hexasulfate. J. Mol. Biol. *409*, 214–224.

Zhang, W.-B., Sun, F., Tirrell, D.A., and Arnold, F.H. (2013). Controlling Macromolecular Topology with Genetically Encoded SpyTag–SpyCatcher Chemistry. J. Am. Chem. Soc. *135*, 13988–13997.

Zhao, Z., Ma, X., Li, L., Zhang, W., Ping, S., Xu, M.-Q., and Lin, M. (2010). Protein cyclization enhanced thermostability and exopeptidase-resistance of green fluorescent protein. J. Microbiol. Biotechnol. *20*, 460–466.

Zhou, Z., Cironi, P., Lin, A.J., Xu, Y., Hrvatin, S., Golan, D.E., Silver, P.A., Walsh, C.T., and Yin, J. (2007). Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases. ACS Chem. Biol. *2*, 337–346.

Zhu, Z., and Dimitrov, D.S. (2009). Construction of a Large Naïve Human Phage-Displayed Fab Library Through One-Step Cloning. Methods Mol. Biol. Clifton NJ *525*, 129–xv.

Zinchenko, A., Devenish, S.R.A., Kintses, B., Colin, P.-Y., Fischlechner, M., and Hollfelder, F. (2014). One in a Million: Flow Cytometric Sorting of Single Cell-Lysate Assays in Monodisperse Picolitre Double Emulsion Droplets for Directed Evolution. Anal. Chem. *86*, 2526–2533.

Zoldák, G., Zubrik, A., Musatov, A., Stupák, M., and Sedlák, E. (2004). Irreversible Thermal Denaturation of Glucose Oxidase from Aspergillus niger Is the Transition to the Denatured State with Residual Structure. J. Biol. Chem. *279*, 47601–47609.

Appendix

Appendix

A.1 Migration of cyclized construct is acrylamide percentage-dependent

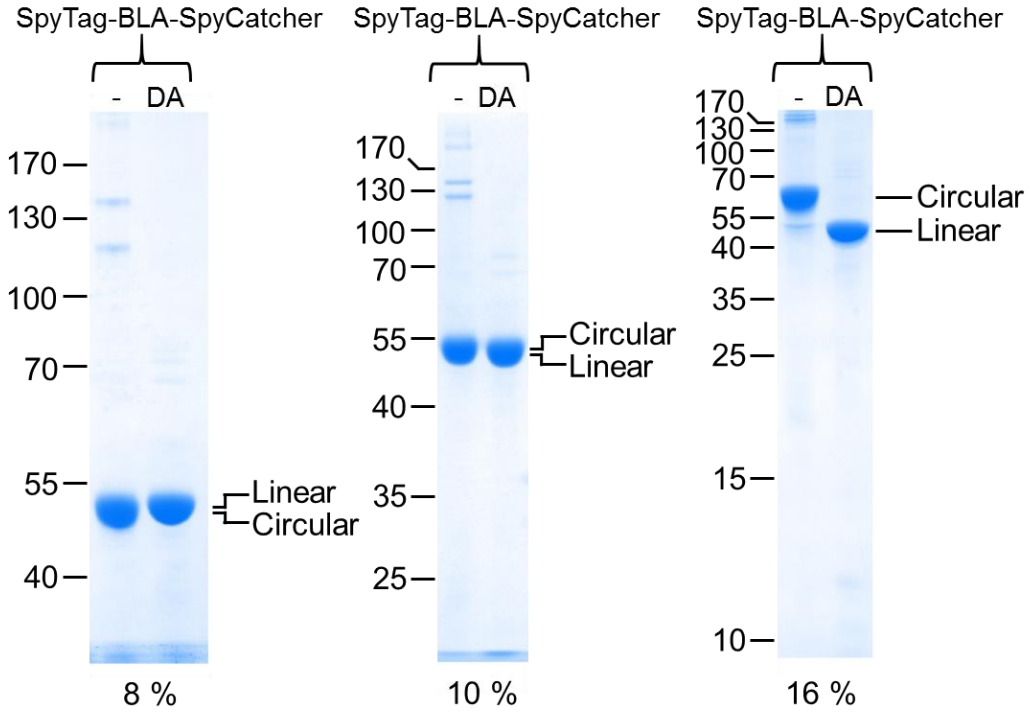
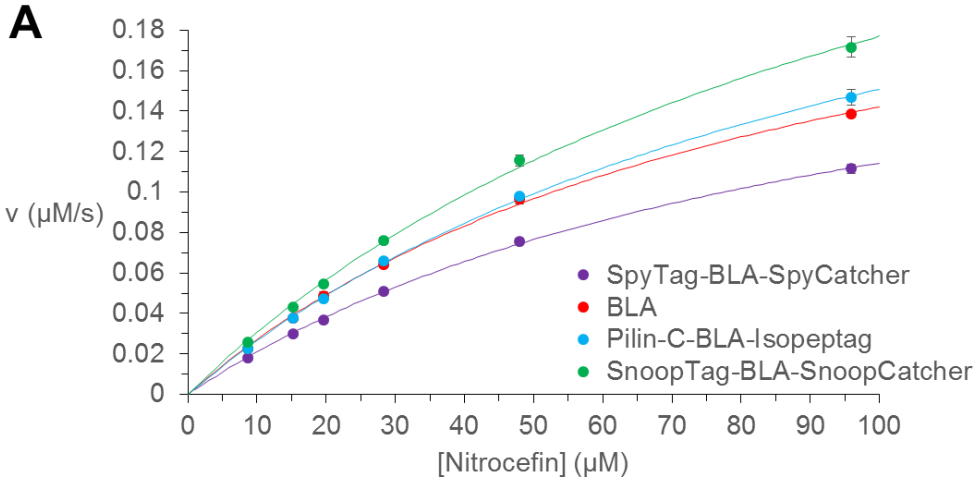


Figure A.1: Effects of polyacrylamide gel percentage on protein migration. Reducing 8 %, 10 % and 16 % SDS-PAGE gels of SpyTag (DA)-BLA-SpyCatcher. The SpyTag D→A mutation removes one of the three amino acids needed to form the isopeptide bond and therefore was not cyclized.

A.2 Cyclization did not substantially alter the activity of BLA.

A



B

Sample	Vmax (μM/s)	K_M (μM)
SpyTag-BLA-SpyCatcher	0.2265 ± 0.0109	98.1 ± 5.6
BLA	0.2691 ± 0.0065	89.3 ± 2.5
Pilin-C-BLA-Isopeptag	0.3149 ± 0.0291	108.8 ± 14.1
SnoopTag-BLA-SnoopCatcher	0.3769 ± 0.0358	113.1 ± 15.1

Figure A.2: Kinetic parameters of cyclized BLA construct and WT BLA. A) Plot of initial velocity vs. concentration of nitrocefin of BLA, SpyTag-BLA-SpyCatcher, SnoopTag-BLA-SnoopCatcher and Pilin-C-BLA-Isopeptag. 0.25 nM of enzyme was allowed to react with different concentrations of nitrocefin in 50 mM NaH_2PO_4 pH 7.0 at 30 °C (mean ± 1 standard deviation). A line of best fit has been drawn using kinetic parameters obtained from non-linear curve-fitting. B) Kinetic parameters inferred from A) (mean ± 1 standard deviation based on triplicate measurements). Data were previously published in (Schoene et al., 2016a).

A.3 Cyclization did not substantially alter the specific activity of PhyC.

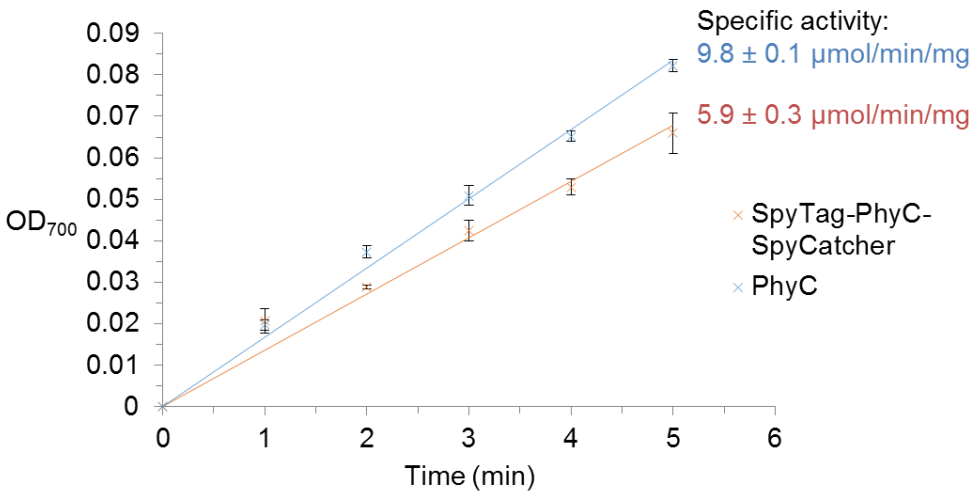


Figure A.3: Specific activity of PhyC and SpyTag-PhyC-SpyCatcher. 40 nM of PhyC construct was allowed to react with 1.6 mM phytic acid for 1, 2, 3, 4, or 5 min before the reaction was quenched. The amount of phosphate released was quantified through the molybdenum blue method. A standard curve using known quantities of phosphate was used to determine the relationship between phosphate concentration and OD_{700} (mean ± 1 standard deviation). Data were previously published in (Schoene et al., 2016a).

Protein Engineering Hot Paper

DOI: 10.1002/anie.201402519

SpyTag/SpyCatcher Cyclization Confers Resilience to Boiling on a Mesophilic Enzyme**

*Christopher Schoene, Jacob O. Fierer, S. Paul Bennett, and Mark Howarth**

Abstract: *SpyTag is a peptide that spontaneously forms an amide bond with its protein partner SpyCatcher. SpyTag was fused at the N terminus of β-lactamase and SpyCatcher at the C terminus so that the partners could react to lock together the termini of the enzyme. The wild-type enzyme aggregates above 37°C, with irreversible loss of activity. Cyclized β-lactamase was soluble even after heating at 100°C; after cooling, the catalytic activity was restored. SpyTag/SpyCatcher cyclization led to a much larger increase in stability than that achieved through point mutation or alternative approaches to cyclization. Cyclized dihydrofolate reductase was similarly resilient. Analyzing unfolding through calorimetry indicated that cyclization did not increase the unfolding temperature but rather facilitated refolding after thermal stress. SpyTag/SpyCatcher sandwiching represents a simple and efficient route to enzyme cyclization, with potential to greatly enhance the robustness of biocatalysts.*

Biological catalysts frequently show superlative regioselectivity and stereoselectivity compared to chemical catalysts but they suffer from instability.[1] The stabilization of enzymes can be achieved without the use of chemical modifications by looking for homologues in thermophiles,[2] inferring a consensus or ancestral sequence,[3–6] selection from libraries,[7,8] or structure-based design.[9,10] Proteins obtained from thermophiles usually achieve their optimum catalytic efficiency at a high temperature,[11] a costly trait for application to biotransformations. Therefore, it is desirable to obtain a protein with a high catalytic efficiency at ambient temperature, whilst maintaining the thermal tolerance of thermophile-derived proteins.

Directed evolution and rational design of enzymes have led to important successes, but these approaches require time-consuming individual optimization and the gains are often marginal. For example, after more than 700 mutations on T4 lysozyme, the best thermostabilization from combining individual stabilizing point mutations was 8°C, at a cost of losing most of the catalytic activity.[12] Therefore, generic approaches that require neither extensive structural knowledge nor library selection are needed to provide a faster and more broadly applicable route to enzyme stabilization.

To overcome the limited stability of peptide interactions, we previously developed a peptide tag that effects spontaneous intermolecular amide bond formation. This tag was developed through the dissection and modification of a protein domain from Gram-positive bacteria.[13–16] We engineered SpyTag, a 13 amino acid peptide, to rapidly form an irreversible amide bond to the 15 kDa protein SpyCatcher.[14] The system is genetically encodable and the tag and protein are functional on either terminus.[14] Since the termini are one of the most flexible regions of proteins and fluctuations here may initiate unfolding,[17] we hypothesized that connecting the N and C termini of a protein through SpyTag/SpyCatcher cyclization could enhance stability (Figure 1A).

Protein cyclization has been previously achieved through the use of carbodiimide cross-linking,[18] sortase,[19] or (most often) inteins.[20–23] TEM-1 β-lactamase (BLA) is an important model system for enzyme evolution, as well as having clinical relevance from the emergence of bacteria with broad spectrum antibiotic resistance.[6,24] Previous cyclization of BLA with a split intein achieved a 5°C increase in thermal tolerance.[21] We genetically fused SpyTag to the N terminus of BLA and SpyCatcher to the C terminus (Figure 1A). This SpyTag-BLA-SpyCatcher construct expressed efficiently in *Escherichia coli*. To test whether we had successfully cyclized BLA, we generated negative controls unable to react in both SpyTag and SpyCatcher: we mutated the reactive Asp in SpyTag to Ala (SpyTagDA) or the catalytic Glu residue in SpyCatcher to Gln (SpyCatcherEQ).[14] When we analyzed the linear mutants alongside the cyclized construct by SDS-PAGE, we found that the cyclized form had a lower mobility, a result consistent with efficient cyclization (75%; Figure 1B and Figure S1a in the Supporting Information).

To further confirm that we had successfully cyclized BLA, we introduced a TEV protease cleavage site between the BLA and SpyCatcher. Following TEV protease cleavage, the cyclized SpyTag-BLA-TEV-SpyCatcher migrated to the same apparent molecular weight as the linear construct (Figure 1B). We also witnessed the disappearance of polymeric forms (generated from a low level of intermolecular reaction of SpyTag-BLA-SpyCatcher) and the appearance of new bands from polymer cleavage (Figure 1B).

Furthermore, we verified the molecular weight of SpyTag-BLA-SpyCatcher, as well as of the DA and EQ controls, by

[*] C. Schoene, J. O. Fierer, Dr. S. P. Bennett,[+] Prof. M. Howarth
Department of Biochemistry, University of Oxford
South Parks Road, Oxford OX1 3QU (UK)
E-mail: mark.howarth@bioch.ox.ac.uk

[+] Current address: Sekisui Diagnostics (UK) Ltd.
50 Gibson Drive, Kings Hill, West Malling
Kent ME19 4AF (UK)

[**] Funding was provided by BBSRC (C.S.), Sekisui Diagnostics (C.S., S.P.B.), the Clarendon Fund (J.O.F.), New College Oxford (J.O.F.), Worcester College Oxford (M.H.), and Oxford University Department of Biochemistry (M.H.). We thank Nick Major of Sekisui Diagnostics for helpful advice and the Molecular Biophysics Facility at Oxford University Department of Biochemistry.

Supporting information for this article (including experimental details) is available on the WWW under http://dx.doi.org/10.1002/anie.201402519.

 © 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

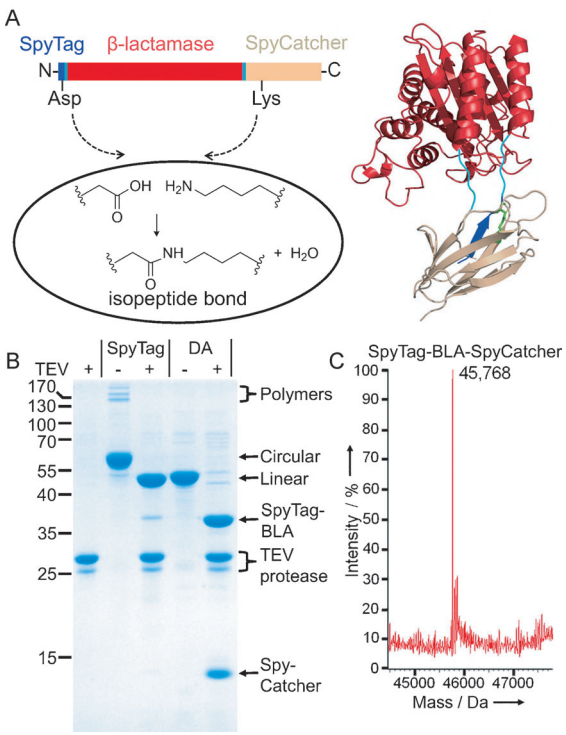


Figure 1. SpyTag/SpyCatcher-mediated cyclization of β-lactamase. A) A cartoon of the cyclization of BLA. B) SDS-PAGE with Coomassie staining for TEV protease or SpyTag-BLA-TEV-SpyCatcher and SpyTagDA-BLA-TEV-SpyCatcher with or without TEV protease. C) Mass spectrometry of SpyTag-BLA-SpyCatcher.

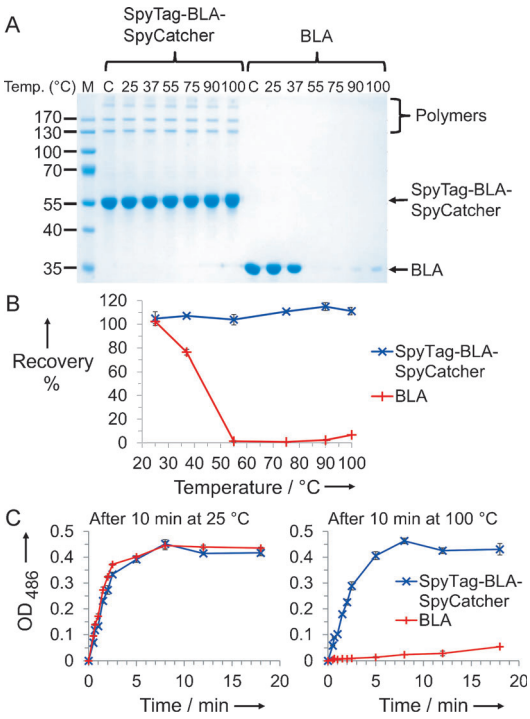


Figure 2. Increased thermal tolerance of the cyclized enzyme. A) Cyclized or original enzyme was heated at the indicated temperature for 10 min, centrifuged, and the supernatant analyzed by SDS-PAGE with Coomassie staining. M = molecular weight markers. C = control without any incubation. B) Quantification of soluble fractions from (A). The data points show the mean of triplicate samples ± 1 standard deviation (SD). C) Nitrocefin assay of enzyme activity after 10 min incubation at 25 or 100 °C (mean of triplicate ± 1 SD). Some error bars are too small to be seen. OD_{486} = absorbance at 486 nm.

using electrospray ionization mass spectrometry (calculated M_w of SpyTag-BLA-SpyCatcher before reaction 45,786.4 Da; calculated M_w of SpyTag-BLA-SpyCatcher after loss of H_2O from spontaneous amide bond formation 45,768.4 Da; observed M_w 45,768 Da; Figure 1C and Figure S1 in the Supporting Information).

To test the effect of cyclization on protein thermal stability, we incubated SpyTag-BLA-SpyCatcher and unmodified BLA in phosphate-buffered saline over a wide range of temperatures. Surprisingly, we found that SpyTag-BLA-SpyCatcher was not lost from solution at any of the temperatures tested, even up to 100 °C. On the other hand, BLA started aggregating at 37 °C and had completely aggregated at 55 °C (Figure 2A,B). Therefore, SpyTag/SpyCatcher cyclization conferred an increase in aggregation temperature of more than 60 °C. It was conceivable that the different stability resulted from different protease contamination of the preparations. However, after mixing BLA and SpyTag-BLA-SpyCatcher, there was still a much greater loss of BLA than SpyTag-BLA-SpyCatcher upon heating (Figure S2).

To test whether SpyTag-BLA-SpyCatcher retained activity as well as solubility, we measured enzymatic activity with the colorimetric substrate nitrocefin after incubation at 25 or 100 °C. Nearly all BLA activity was lost after high-temperature incubation, whereas nearly all the activity of SpyTag-BLA-SpyCatcher was retained after 100 °C exposure (Figure 2C).

To understand to what extent the fused components conferred resistance to aggregation, we tested the noncyclized SpyTagDA-BLA-SpyCatcher in equivalent assays. Interestingly, SpyTagDA-BLA-SpyCatcher did show enhanced resistance to aggregation compared to BLA. However, covalent cyclization was important, because the recovery of activity following heating of SpyTagDA-BLA-SpyCatcher was inferior to SpyTag-BLA-SpyCatcher (Figure S3). SpyTag-BLA-SpyCatcherEQ (with the E77Q mutation blocking the SpyCatcher reaction)[14] showed low resistance to aggregation and weak recovery of catalytic activity (Figure S4). Fusing the whole CnaB2 domain (to give BLA-CnaB2, which undergoes isopeptide formation within the CnaB2 domain but without connecting the termini of BLA)[14] gave excellent resistance to aggregation but only moderate recovery of catalytic activity (Figure S4).

To see whether the effect of SpyTag/SpyCatcher could be extended to another class of enzyme, we tested our approach on dihydrofolate reductase (DHFR). SpyTag-DHFR-SpyCatcher cyclized in high yield, as shown by SDS-PAGE and mass spectrometry (Figure S5). SpyTag-DHFR-SpyCatcher, like the SpyTagDA-DHFR-SpyCatcher control, resisted aggregation at 100 °C (Figure S5).[25] However, SpyTag-DHFR-SpyCatcher retained its catalytic activity following the 100 °C heating, unlike SpyTagDA-DHFR-SpyCatcher (Figure S5).

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

To dissect the mechanism of SpyTag/SpyCatcher stabilization, we used differential scanning calorimetry (DSC) to test the thermal unfolding profiles of SpyTag-BLA-SpyCatcher, SpyTagDA-BLA-SpyCatcher, and BLA. SpyTag-BLA-SpyCatcher and SpyTagDA-BLA-SpyCatcher each generated two peaks (Figure 3A). The common peak for all

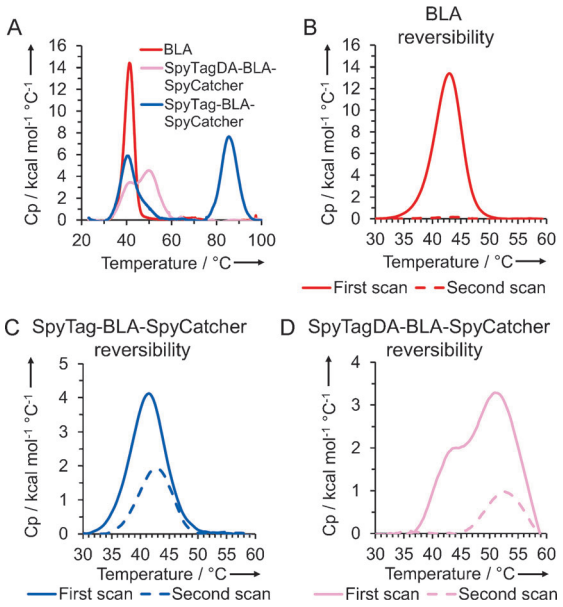


Figure 3. Thermal unfolding transitions of BLA, SpyTag-BLA-SpyCatcher, and SpyTagDA-BLA-SpyCatcher. A) DSC giving the specific heat capacity (Cp) of each enzyme construct scanned from 20–110 °C at 1 °C min^{-1}. B) DSC of BLA scanned from 20–60 °C at 2 °C min^{-1} before cooling and then repeating the scan. C) DSC as in (B) for SpyTag-BLA-SpyCatcher. D) DSC as in (B) for SpyTagDA-BLA-SpyCatcher.

constructs, corresponding to BLA domain unfolding, showed a melting temperature (T_m) of 39.6 °C for SpyTag-BLA-SpyCatcher, 39.6 °C for SpyTagDA-BLA-SpyCatcher, and 41.3 °C for BLA (Figure 3A). The second peak for SpyTag-BLA-SpyCatcher, with a T_m value of 85.4 °C, likely corresponds to unfolding of the reconstituted SpyTag/SpyCatcher domain. The second peak for SpyTagDA-BLA-SpyCatcher, related to the noncovalent SpyTagDA/SpyCatcher complex, showed a T_m value of 54.0 °C (Figure 3A).

To test the capability of the BLA domain to refold, we used DSC to scan from 20 to 60 °C (where the signature from BLA unfolding is observed), allowed the sample to cool, and then scanned over this temperature range again. BLA gave no peak on the second scan, a result consistent with an inability to refold (Figure 3B). By contrast, SpyTag-BLA-SpyCatcher produced a signal of similar shape but reduced intensity on the second scan (Figure 3C), thus suggesting that the BLA domain is able to refold successfully in this cyclized context. SpyTagDA-BLA-SpyCatcher gave no peak in the range 40–45 °C, thus indicating an absence of BLA refolding; the peak from 45–60 °C corresponds to the SpyTagDA/SpyCatcher moiety (Figure 3D). Therefore, cyclization through SpyTag-SpyCatcher chemistry did not appear to increase resistance to unfolding, but instead increased the ability to refold.

In conclusion, we have developed a method to dramatically increase the ability of a protein to recover from denaturation by using cyclization through spontaneous isopeptide bond formation. SpyTag/SpyCatcher cyclization conferred resistance to aggregation and enabled the recovery of catalytic activity following heating. DSC showed that the T_m value for the BLA domain was largely unaffected by cyclization, thus the enhanced resilience is likely due to an increase in the ability of the cyclized protein to refold. The greatly enhanced stability to aggregation conferred by SpyTag/SpyCatcher cyclization compared to other cyclization strategies may relate to the shielding of interprotein associations by the presence of a stable domain,[26] but intriguingly, DSC indicates that the stabilization effect extends beyond temperatures at which the SpyCatcher/SpyTag complex can unfold.

Disulfide bonds may also be engineered to lock protein termini but can interfere with existing disulfides, will not form in reducing environments, can break through β-elimination at elevated temperatures,[27] and often only give modest stability enhancements.[12] While this work was in progress, elastin-like peptides containing terminal SpyTag and SpyCatcher were shown to cyclize, but no functional effect of cyclization was established.[28] For some protein architectures, it will be hard to achieve SpyTag/SpyCatcher cyclization, but approximately 50% of single-domain proteins in the PDB have N- and C-terminal structural elements within 5 Å,[29] and future work may be able to harness the reactivity of SpyTag and SpyCatcher in internal regions of proteins.[14,28] Since protein stability can confer tolerance to mutations advantageous for function but deleterious for folding to the native structure,[30] it will also be interesting to explore whether SpyTag/SpyCatcher stabilization could facilitate library-based evolution of novel protein function.

Received: February 17, 2014
Published online: May 9, 2014

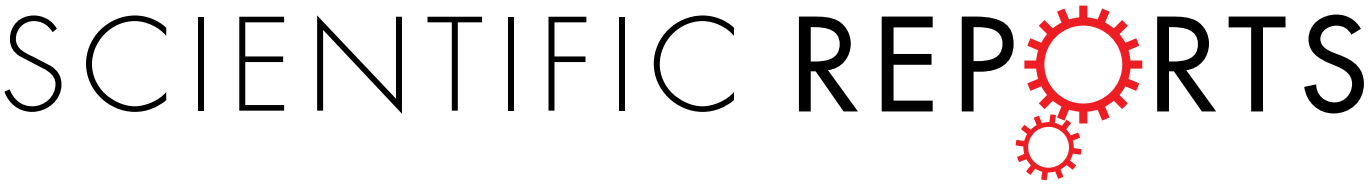
Keywords: enzymes · peptides · protein engineering · protein modifications · protein stability

[1] K. M. Polizzi, A. S. Bommarius, J. M. Broering, J. F. Chaparro-Riggers, *Curr. Opin. Chem. Biol.* **2007**, *11*, 220–225.
[2] A. Chien, D. B. Edgar, J. M. Trela, *J. Bacteriol.* **1976**, *127*, 1550–1557.
[3] M. Lehmann, D. Kostrewa, M. Wyss, R. Brugger, A. D'Arcy, L. Pasamontes, A. P. van Loon, *Protein Eng.* **2000**, *13*, 49–57.
[4] E. A. Gaucher, S. Govindarajan, O. K. Ganesh, *Nature* **2008**, *451*, 704–707.
[5] K. Watanabe, T. Ohkuri, S. Yokobori, A. Yamagishi, *J. Mol. Biol.* **2006**, *355*, 664–674.
[6] V. A. Risso, J. A. Gavira, D. F. Mejia-Carmona, E. A. Gaucher, J. M. Sanchez-Ruiz, *J. Am. Chem. Soc.* **2013**, *135*, 2899–2902.
[7] M. T. Reetz, *J. Am. Chem. Soc.* **2013**, *135*, 12480–12496.
[8] M. Goldsmith, D. S. Tawfik, *Curr. Opin. Struct. Biol.* **2012**, *22*, 406–412.
[9] P. V. Iyer, L. Ananthanarayan, *Process Biochem.* **2008**, *43*, 1019–1032.

 © 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[10] A. Korkegian, M. E. Black, D. Baker, B. L. Stoddard, *Science* **2005**, *308*, 857 – 860.
[11] C. Vieille, G. J. Zeikus, *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1 – 43.
[12] W. A. Baase, L. Liu, D. E. Tronrud, B. W. Matthews, *Protein Sci.* **2010**, *19*, 631 – 641.
[13] B. Zakeri, M. Howarth, *J. Am. Chem. Soc.* **2010**, *132*, 4526 – 4527.
[14] B. Zakeri, J. O. Fierer, E. Celik, E. C. Chittock, U. Schwarz-Linek, V. T. Moy, M. Howarth, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E690 – E697.
[15] H. J. Kang, E. N. Baker, *Trends Biochem. Sci.* **2011**, *36*, 229 – 237.
[16] R. W. Hendrix, J. E. Johnson, *Adv. Exp. Med. Biol.* **2012**, *726*, 351 – 363.
[17] E. Jacob, R. Unger, *Bioinforma. Oxf. Engl.* **2007**, *23*, e225 – 230.
[18] D. P. Goldenberg, T. E. Creighton, *J. Mol. Biol.* **1983**, *165*, 407 – 413.
[19] R. Parthasarathy, S. Subramanian, E. T. Boder, *Bioconjugate Chem.* **2007**, *18*, 469 – 476.
[20] T. C. Evans, Jr., J. Benner, M. Q. Xu, *J. Biol. Chem.* **1999**, *274*, 3923 – 3926.
[21] H. Iwai, A. Plückthun, *FEBS Lett.* **1999**, *459*, 166 – 172.
[22] J. A. Camarero, T. W. Muir, *J. Am. Chem. Soc.* **1999**, *121*, 5597 – 5598.
[23] C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnon, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13638 – 13643.
[24] S. Bershtein, K. Goldin, D. S. Tawfik, *J. Mol. Biol.* **2008**, *379*, 1029 – 1044.
[25] S. Bershtein, W. Wu, E. I. Shakhnovich, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4857 – 4862.
[26] L. Li, J. O. Fierer, T. A. Rapoport, M. Howarth, *J. Mol. Biol.* **2014**, *426*, 309 – 317.
[27] D. B. Volkin, A. M. Klibanov, *J. Biol. Chem.* **1987**, *262*, 2945 – 2950.
[28] W.-B. Zhang, F. Sun, D. A. Tirrell, F. H. Arnold, *J. Am. Chem. Soc.* **2013**, *135*, 13988 – 13997.
[29] M. M. G. Krishna, S. W. Englander, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 1053 – 1058.
[30] J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5869 – 5874.

 © 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



OPEN

SpyRing interrogation: analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries

Received: 12 October 2015
Accepted: 18 January 2016
Published: 10 February 2016

Christopher Schoene[1], S. Paul Bennett[2] & Mark Howarth[1]

Enzymes catalyze reactions with exceptional selectivity and rate acceleration but are often limited by instability. Towards a generic route to thermo-resilience, we established the SpyRing approach, cyclizing enzymes by sandwiching between SpyTag and SpyCatcher (peptide and protein partners which lock together via a spontaneous isopeptide bond). Here we first investigated the basis for this resilience, comparing alternative reactive peptide/protein pairs we engineered from Gram-positive bacteria. Both SnoopRing and PilinRing cyclization gave dramatic enzyme resilience, but SpyRing cyclization was the best. Differential scanning calorimetry for each ring showed that cyclization did not inhibit unfolding of the inserted β-lactamase. Cyclization conferred resilience even at 100 °C, where the cyclizing domains themselves were unfolded. Phytases hydrolyze phytic acid and improve dietary absorption of phosphate and essential metal ions, important for agriculture and with potential against human malnutrition. SpyRing phytase (PhyC) resisted aggregation and retained catalytic activity even following heating at 100 °C. In addition, SpyRing cyclization made it possible to purify phytase simply by heating the cell lysate, to drive aggregation of non-cyclized proteins. Cyclization via domains forming spontaneous isopeptide bonds is a general strategy to generate resilient enzymes and may extend the range of conditions for isolation and application of enzymes.

Purified enzymes are finding a range of important uses, from high-cost low-scale applications such as GMP-grade asparaginase in treatment of leukemia[1], to low-cost high-scale applications such as the millions of tons per year of high fructose corn syrup biotransformed by glucose isomerase[2]. When enzymes are being developed for application in new areas, a major challenge is the stability of the enzyme to inactivation during isolation or use in its new environment. Stability can come from sourcing enzymes from harsh environments or from enzyme-specific optimization campaigns using iterative mutation[3]. However, it is also important to seek generic approaches capable of stabilizing a wide range of enzymes. One such approach is cyclization, connecting the enzyme's N-terminus to the C-terminus. Cyclization can be accomplished via split inteins[4–7], sortases[8], carbodiimide cross-linking[9] and, most recently, via the design of SpyRings[10]. SpyTag is a peptide tag that we engineered from *Streptococcus pyogenes* FbaB adhesion protein to form a spontaneous isopeptide bond upon binding to its protein partner SpyCatcher[11,12]. SpyRing cyclization utilizes SpyTag on the N-terminus and SpyCatcher on the C-terminus to cyclize an enzyme of interest. We showed this cyclization increased the thermal tolerance of TEM-1 β-lactamase (BLA) by >60 °C, both in terms of aggregation resistance and retaining activity after heating[10]. This approach to enzyme resilience is not known in nature and so it is important to further explore this phenomenon, as well as to test whether SpyRing cyclization could be extended to an industrially important class of enzymes.

At least 2 billion people are thought to suffer from deficiency of micronutrients, including vitamin A, iodine, iron, niacin, zinc and calcium[13]. Such deficiency, termed the "silent hunger", has global consequences, including decreased bone mineralization, anemia, impaired learning, and decreased resistance to infectious diseases such as AIDS and tuberculosis[13,14]. Major effort has gone into increasing the nutritional value of major food crops around the world, notably introducing the β-carotene pathway into rice[15]. Most crops store phosphate as phytic acid (inositol hexakisphosphate), making the phosphate unavailable for humans as well as other agriculturally-important

[1]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK. [2]Sekisui Diagnostics UK Ltd., Operations Building, Liphook Way, Allington, Maidstone, Kent, ME16 0LQ, UK. Correspondence and requests for materials should be addressed to M.H. (email: mark.howarth@bioch.ox.ac.uk)

non-ruminants (e.g. poultry, fish, and pigs). Moreover, phytic acid is an anti-nutrient which chelates essential mineral cations (iron, zinc, calcium)[16,17]. Adding phytase to the diet breaks down phytic acid and has been widely used in agriculture to improve animal health[18]. The use of phytase also reduces the danger, if phosphate is directly added to feedstock, of excess phosphate polluting waterways and leading to eutrophication[18]. The thermal tolerance of phytase has great significance because feed-stock must be subjected to temperatures of 75–90 °C to kill *Salmonella* and other dietary pathogens[18].

Here we have explored the principle of SpyRing resilience by evaluating enzyme cyclization with alternative recently-established isopeptide-forming protein chemistries. We then extended the scope of SpyRing cyclization by enhancing the thermo-resilience of a phytase. We studied the stability effects of cyclization by differential scanning calorimetry (DSC) and circular dichroism (CD). In addition we opened up a new approach for purification of SpyRing-cyclized enzymes based simply upon heating.

Results

Enzyme cyclization using Isopeptag/Pilin-C or SnoopTag/SnoopCatcher spontaneous isopeptide bond formation. To understand the dramatic increase in thermal resilience caused by SpyRing-mediated cyclization, we compared the effects of cyclizing with two different Tag/Catcher pairs, each able to form a spontaneous isopeptide bond (Fig. 1A). First we sandwiched BLA between Isopeptag and Pilin-C, a Tag/Catcher pair we had previously engineered by splitting the *Streptococcus pyogenes* major pilin Spy0128[19,20]. We termed this approach PilinRing cyclization. We also tested sandwiching BLA between SnoopTag and SnoopCatcher, a peptide/protein pair we recently engineered from the RrgA adhesin of *Streptococcus pneumoniae*[21,22] (SnoopRing cyclization, Fig. 1A). SpyRing cyclization occurs through reaction of an amine with a carboxylic acid with release of water, but both SnoopRing and PilinRing cyclization occur through reaction of an amine with a carboxamide with release of ammonia (Fig. 1A)[20,22]. Cyclization was evaluated by the change in SDS-PAGE mobility upon cyclization, compared to negative controls having the key reactive lysine mutated to alanine (SnoopTag KA-BLA-SnoopCatcher and Pilin-C KA-BLA-Isopeptag) (Fig. 1B). We found that the PilinRing system was efficient at cyclizing BLA when Pilin-C was fused to the N-terminus and Isopeptag was fused to the C-terminus (Fig. 1B). SnoopRing cyclization was efficient when fusing SnoopTag on the N-terminus and SnoopCatcher on the C-terminus (Fig. 1B). Electrospray ionization mass spectrometry (MS) confirmed the loss of ammonia for both the PilinRing and SnoopRing upon isopeptide bond formation (Fig. S1). Small amounts of high molecular weight forms were produced by both PilinRing and SnoopRing systems (but not the point mutant controls), consistent with a proportion of intermolecular reaction, as seen with SpyRing cyclization (Fig. 1B). The k_{cat} and K_M values of the SpyRing, SnoopRing and PilinRing cyclized BLA constructs were comparable to those of the wild-type BLA construct (Fig. S2).

Comparison of different cyclization domains for β-lactamase resilience. We evaluated SnoopRing and PilinRing aggregation-resistance by heating the enzymes for 10 min at various temperatures, centrifuging at high speed to remove aggregates, and analyzing the remaining soluble protein by SDS-PAGE with Coomassie staining. SpyRing was previously found to provide full aggregation-resistance to temperatures up to 100 °C[10]. For BLA itself, little protein remained in solution above 37 °C (Fig. 2A). SnoopRing and PilinRing cyclization both led to major increases in aggregation resistance (Fig. 2A,B). For Pilin-C-BLA-Isopeptag, there was little loss of soluble protein even after heating to 100 °C (Fig. 2A,B). For SnoopTag-BLA-SnoopCatcher, solubility was well preserved after heating to 100 °C, although there was more aggregation at 75 °C than at 100 °C (Fig. 2A,B).

We then tested recovery of catalytic activity after this heating step followed by cooling to room temperature. Similar to the aggregation data, nearly all catalytic activity of BLA itself was lost after 55 °C heating (Fig. 2C). SnoopRing cyclization and PilinRing cyclization greatly improved the recovery of catalytic activity, although the retention of activity was less than with SpyRing cyclization, particularly after 75 °C heating (Fig. 2C). However, Pilin-C-BLA-Isopeptag exhibited the best thermal resilience after incubation at 55 °C (Fig. 2C).

The non-cyclized control SnoopTag KA-BLA-SnoopCatcher showed little difference in aggregation-resistance or catalytic recovery compared to BLA (Fig. S3A,C). The non-cyclized control Pilin-C KA-BLA-Isopeptag did have enhanced resilience to aggregation and more retention of catalytic activity, when compared to BLA, but was clearly inferior to the cyclized PilinRing (Fig. S3B,C).

Calorimetry of β-lactamase unfolding with alternative cyclization domains. To explore the difference in resilience for BLA cyclized by different domains, we performed DSC on SpyRing, PilinRing and SnoopRing systems. First we investigated the unfolding of the individual domains. SpyCatcher after reconstitution with SpyTag peptide unfolded at 88 °C, higher than SnoopCatcher reconstituted with SnoopTag peptide (70 °C) or Spy0128 (77 °C), the parent domain split to generate Pilin-C/Isopeptag (Fig. 3A).

We previously showed that SpyTag-BLA-SpyCatcher did not change the Tm for BLA unfolding, giving a similar peak at ~41 °C for BLA (Fig. 3A) and SpyTag-BLA-SpyCatcher[10] (Fig. 3B,C). For SpyTag-BLA-SpyCatcher we also observed a peak from SpyTag/SpyCatcher unfolding at 85.0 °C (Fig. 3B,C)[10,23]. For the SnoopRing and its KA control, surprisingly there was a small increase in the BLA Tm to 47.0 °/47.6 °C (Fig. 3B,C). We suggest that the 64.1 °C peak is likely to correspond to unfolding of the SnoopTag/SnoopCatcher domain and the 52.2 °C peak to unfolding of the SnoopTag KA/SnoopCatcher domain (Fig. 3B,C).

For the PilinRing and its non-cyclized control, the Tm for BLA unfolding was ~41 °C, the same as SpyTag-BLA-SpyCatcher (Fig. 3B,C). We suggest that the 72.9 °C peak is likely to correspond to unfolding of the Pilin-C/Isopeptag domain and the 55.9 and 62.4 °C peaks to unfolding of Pilin-C KA/Isopeptag (Fig. 3B,C). Therefore, as for SpyTag/SpyCatcher, cyclization via an isopeptide bond-forming domain had no effect (PilinRing) or a minor effect (SnoopRing) on the transition temperature of the sandwiched BLA domain.

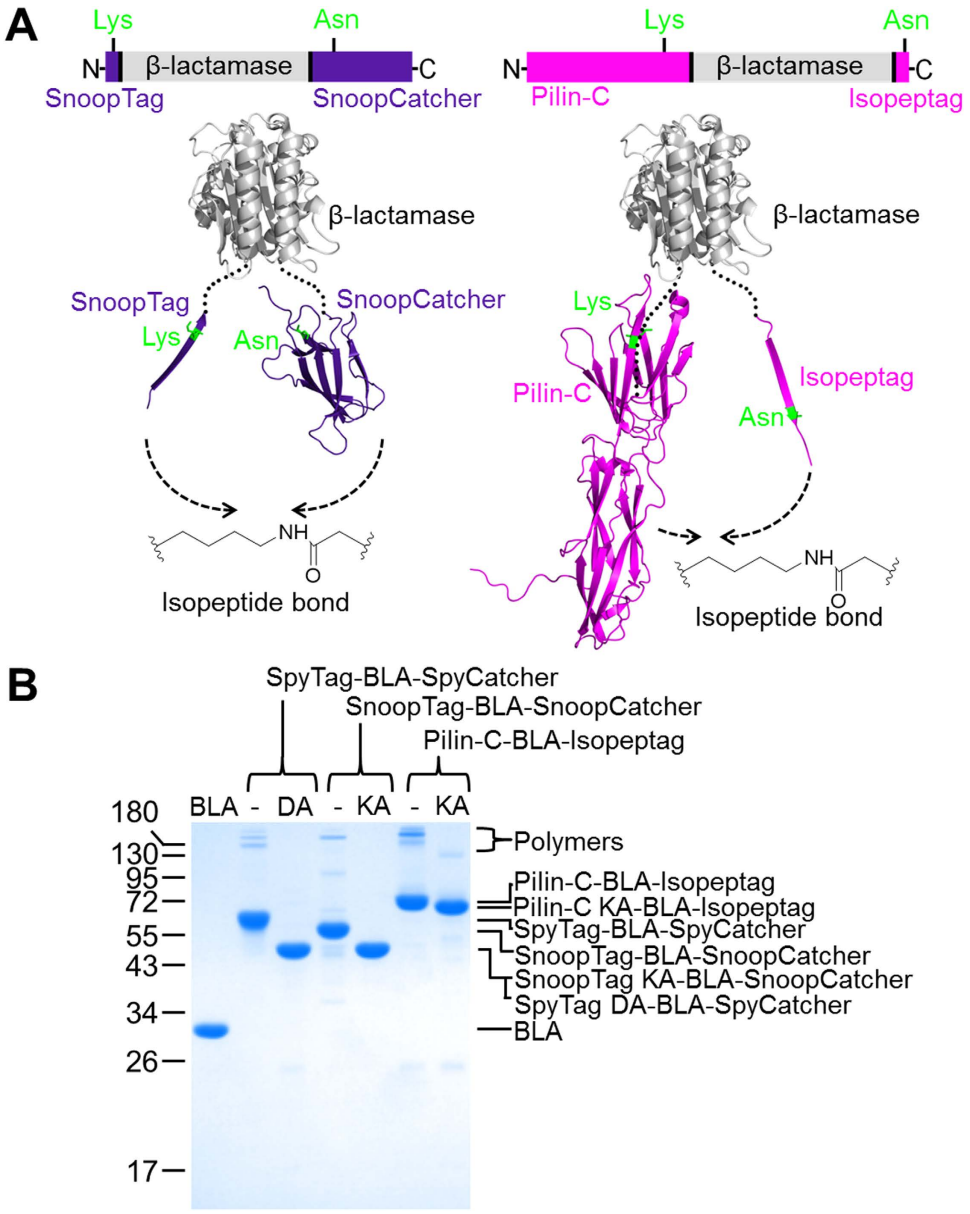


Figure 1. Enzyme cyclization using alternative isopeptide-forming domains. (**A**) Schematic of SnoopRing and PilinRing cyclization of β-lactamase (BLA). Proteins are shown in cartoon format, with reacting residues shown in green in stick format. (**B**) SDS-PAGE analysis of SpyRing, SnoopRing and PilinRing cyclization. Each purified protein is shown alongside the non-reactive control (DA or KA), after staining with Coomassie.

Overall, PilinRing cyclization improved enzyme resilience substantially, but gave less catalytic recovery than SpyRing cyclization. Since Pilin-C is also much larger than SpyCatcher (31 kDa rather than 12 kDa), subsequent investigation focused on advancing and understanding SpyRing cyclization.

SpyRing cyclization made phytase thermo-resilient. We previously showed how SpyRing cyclization conferred resilience on BLA (7 Å between termini) and dihydrofolate reductase (15 Å between termini)[10]. To explore an industrially-important enzymatic activity with a much larger distance between termini, we tested cyclization of a phytase. The crystal structure of PhyC phytase from *Bacillus subtilis* has been previously determined[24]: the direct distance between the resolved N- and C-termini of PhyC is 29 Å (Fig. 4A). SpyTag-PhyC-SpyCatcher cyclized efficiently, based on the loss of H_2O from isopeptide bond formation in MS (Fig. S4A) and the reduction in gel mobility compared to the non-cyclized SpyTag DA-PhyC-SpyCatcher control (Fig. 4B). SpyTag-PhyC-SpyCatcher remained active, exhibiting comparable specific activity to PhyC (Fig. S5). Heat-inactivation of phytases is a major challenge limiting their application for improving nutrition[18]. PhyC remained soluble after heating at 55 °C but aggregated at 75 °C (Fig. 4C). SpyRing cyclization led to a major increase in PhyC thermal resilience, with most of the enzyme remaining in solution even after heating at 100 °C

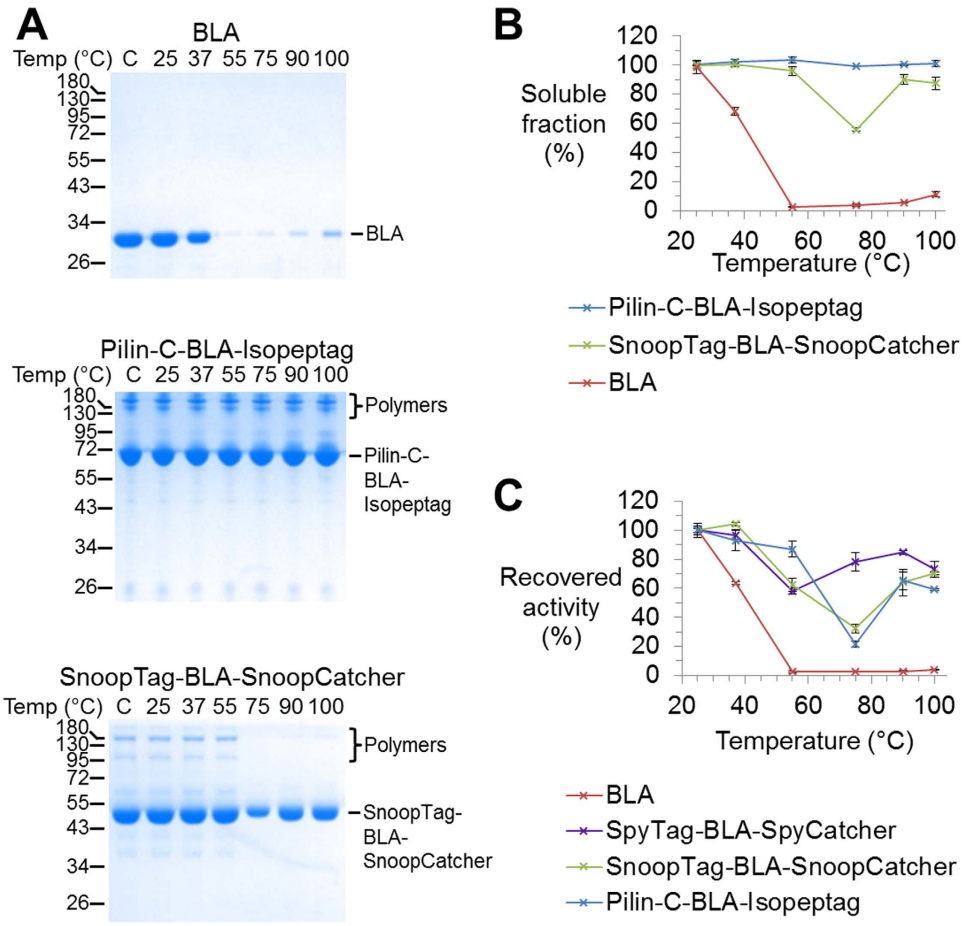


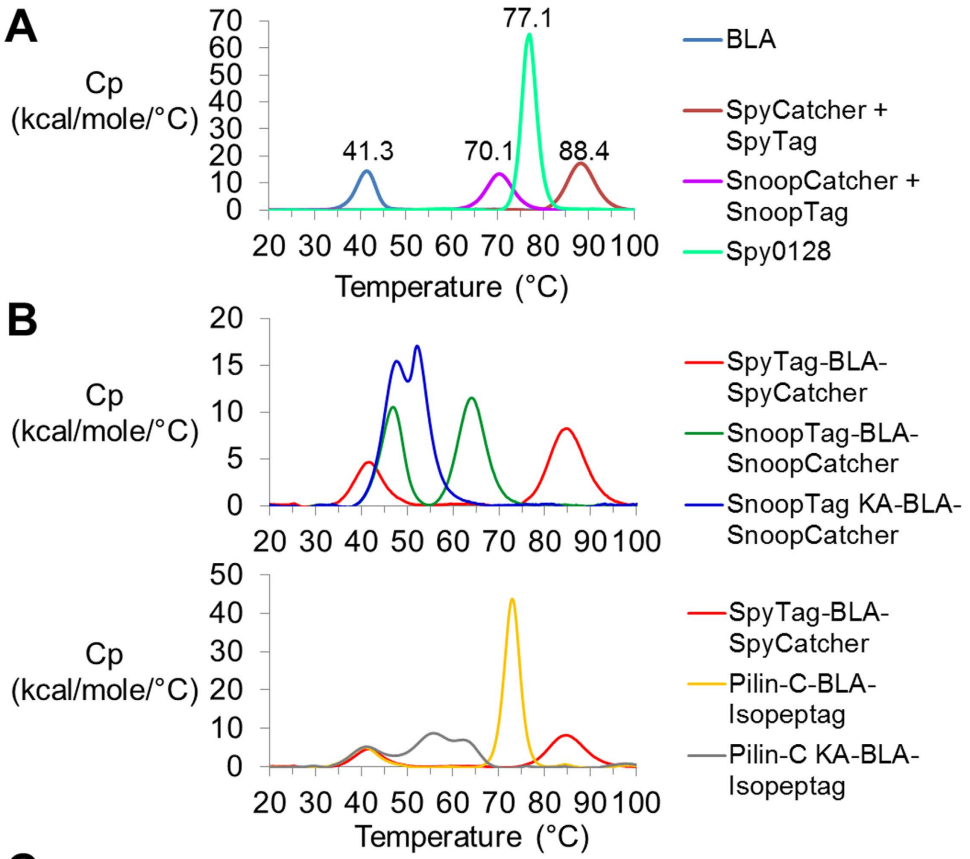
Figure 2. SnoopRing and PilinRing cyclization conferred thermal resilience. (**A**) Cyclization decreased heat-induced aggregation. BLA, Pilin-C-BLA-Isopeptag and SnoopTag-BLA-SnoopCatcher were heated at the indicated temperature for 10 min, centrifuged, and the supernatant analyzed by SDS-PAGE with Coomassie staining. C is control without incubation. (**B**) Plot of heat-induced aggregation, from quantifying SDS-PAGE. (**C**) Cyclization enhanced the catalytic activity following heating. BLA, SnoopTag-BLA-SnoopCatcher, Pilin-C-BLA-Isopeptag and SpyTag-BLA-SpyCatcher were incubated for 10 min at the indicated temperature, before returning to RT and then running a nitrocefin colorimetric activity assay. The initial rates of reaction were calculated using the linear portion of the enzyme curves. Data were normalized using the initial rate of reaction at 25 °C as 100% recovered activity. (All are mean of triplicate ± 1 s.d.; some error bars are too small to be visible.)

(Fig. 4C). Most of PhyC's catalytic activity was lost following 75 °C heating, but the SpyRing PhyC retained most of its activity even after 90 °C heating (Fig. 4D).

The non-cyclized control, SpyTag DA-PhyC-SpyCatcher behaved similarly to PhyC in both aggregation-resistance and recovery of catalytic activity following heating, so covalent cyclization was essential for SpyRing-mediated resilience of PhyC (Fig. S4).

Secondary structure and calorimetric investigation of SpyRing phytase resilience. To understand better the basis for SpyRing phytase resilience, we looked for signatures of unfolding transitions upon heating. The Tm for unfolding of PhyC by DSC was not substantially shifted upon cyclization (PhyC 69.8 °C, SpyTag-PhyC-SpyCatcher 70.0 °C) (Fig. 5). For SpyTag-PhyC-SpyCatcher (but not PhyC or SpyTag DA-PhyC-SpyCatcher) a 2nd transition was seen at 91.5 °C, which is likely to correspond to unfolding of the SpyTag/SpyCatcher domain (Fig. 5). This temperature is slightly higher than the 85.0 °C SpyTag/SpyCatcher transition from SpyTag-BLA-SpyCatcher (Fig. 3A). The difference in Tm might be due to a difference in buffer used for DSC (Tm of 85.0 °C in PBS pH 7.4 and Tm of 91.5 °C in 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$) or might be an effect of PhyC stabilizing the SpyTag/SpyCatcher domain.

We then used circular dichroism (CD) to explore how heating affected secondary structure. We measured the ellipticity of the sample from 185–260 nm. For PhyC and SpyTag DA-PhyC-SpyCatcher there was little change in the spectrum at 55 °C, but at 75 °C and higher there was a general loss of signal, which did not recover upon cooling (Fig. S6). We observed aggregation for these proteins in the cuvette. However, for SpyRing phytase there was less perturbation at 75 and 90 °C (Fig. S6). The CD signal for SpyRing phytase did recover after cooling from



C

Sample	Peak	Temperature (°C)
SpyTag-BLA-SpyCatcher	Tm1	41.6
	Tm2	85.0
SnoopTag-BLA-SnoopCatcher	Tm1	47.0
	Tm2	64.1
SnoopTag KA-BLA-SnoopCatcher	Tm1	47.6
	Tm2	52.2
Pilin-C-BLA-Isopeptag	Tm1	40.9
	Tm2	72.9
Pilin-C KA-BLA-Isopeptag	Tm1	41.0
	Tm2	55.9
	Tm3	62.4

Figure 3. DSC comparison of SpyRing, SnoopRing and PilinRing BLA. (**A**) DSC showing the melting profiles of BLA overlaid with SpyCatcher + SpyTag peptide, SnoopCatcher + SnoopTag peptide, and Spy0128, giving the specific heat capacity (Cp) of each protein scanned from 20–100 °C at 1 °C/min. (**B**) DSC showing the melting profile of SpyTag-BLA-SpyCatcher overlaid with SnoopTag (KA)-BLA-SnoopCatcher or Pilin-C (KA)-BLA-Isopeptag, analyzed as in (A). (**C**) Tm values for peaks observed in (B).

75 and 90 °C (Fig. S6), although interestingly the recovery was less complete than anticipated from the recovery of catalytic activity (Fig. 4D).

SpyRing enzymes could be purified simply by heating. Enzyme purification often uses multiple chromatography steps, which can be time-consuming and expensive. We hypothesized that the ability of SpyRing cyclization to give extreme resilience to heating could permit a simple route to purification, since most of the endogenous cellular proteins might aggregate with heating[25]. Rather than starting with purified enzymes as previously, we explored heating the lysate obtained from enzyme over-expression in *E. coli*. PhyC was the major band on the Coomassie-stained gel of *E. coli* total cell lysate. While the heating at 75–90 °C removed most of the background proteins, heating also aggregated most of the PhyC (Fig. 6A). However, the SpyTag-PhyC-SpyCatcher band in the lysate showed minimal change in intensity after heating from 75–90 °C, while most endogenous proteins were depleted (Fig. 6A). Since 75 °C heating gave similar depletion of endogenous proteins as 90 °C

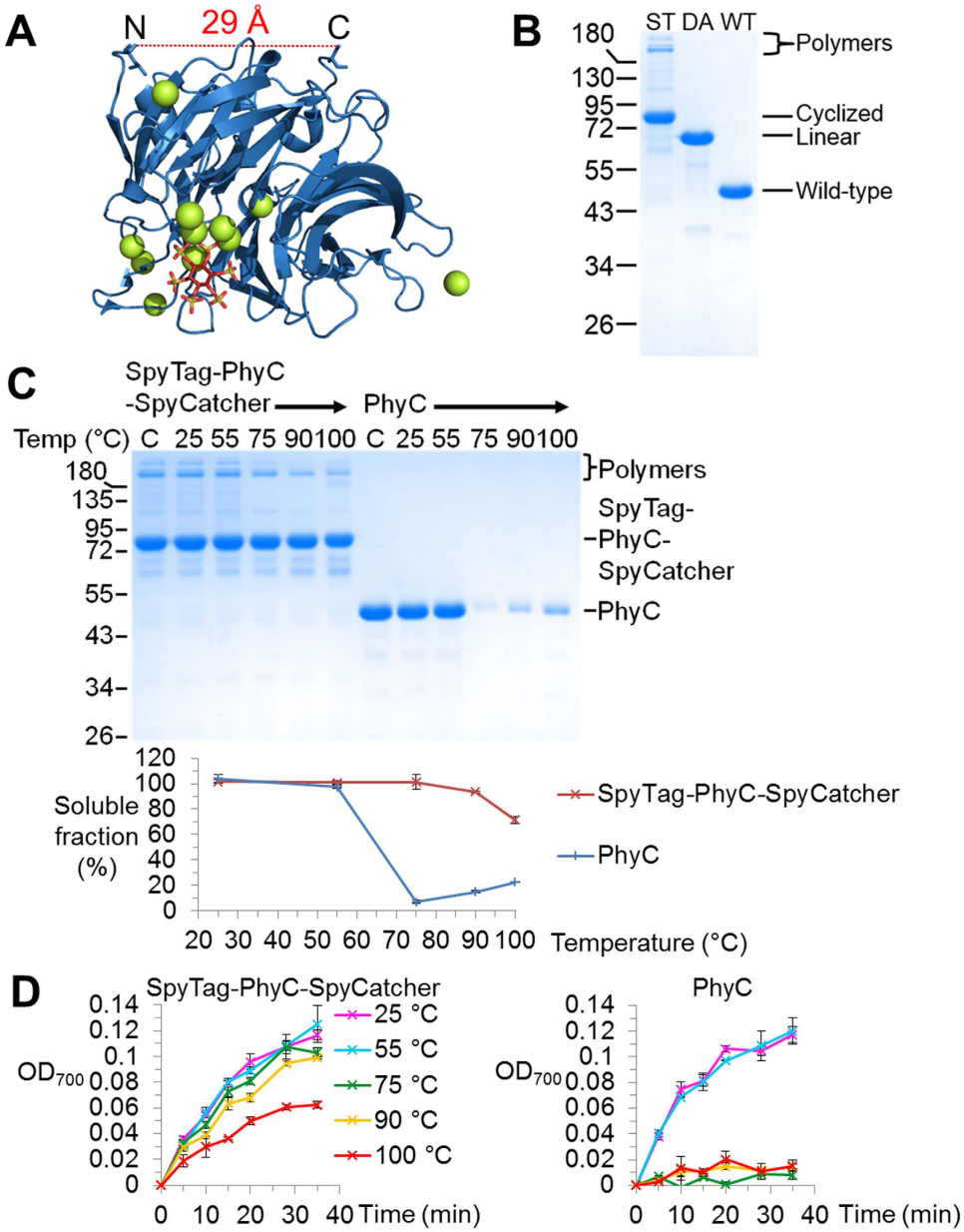


Figure 4. SpyRing cyclization of phytase increased the thermal resilience. (**A**) Crystal structure of PhyC phytase from *Bacillus subtilis* in cartoon format, with the substrate analog myo-inositol hexasulfate in stick format (based on PDB 3AMR). Spheres represent calcium ions and the distance between termini is marked. (**B**) Cyclization caused a gel shift. SpyTag-PhyC-SpyCatcher (ST), SpyTag DA-PhyC-SpyCatcher (DA) and PhyC (WT) were boiled in SDS-loading buffer and analyzed by SDS-PAGE with Coomassie staining. (**C**) SpyRing cyclization improved aggregation-resistance. SpyTag-PhyC-SpyCatcher or PhyC were heated at the indicated temperature for 10 min, centrifuged and the supernatant analyzed by SDS-PAGE with Coomassie staining. C is control without incubation. Data from triplicate measurements were then plotted. (**D**) SpyRing cyclization improved the recovered activity after heating. SpyTag-PhyC-SpyCatcher or PhyC were incubated for 10 min at the indicated temperature, before returning to RT and then running a colorimetric activity assay for phosphate release from phytic acid. (All are mean of triplicate ± 1 s.d.; some error bars are too small to be visible.)

(Fig. 6A), 75 °C heating was used subsequently. Typically the first step in purification does not purify to homogeneity, but SpyTag-PhyC-SpyCatcher heat purification gave a good purity in comparison to Ni-NTA purification (Fig. 6B). There were two major impurities from heat purification of SpyTag-PhyC-SpyCatcher at ~110 kDa and >180 kDa, neither of which were present in the heated PhyC sample (Fig. 6B). Therefore we hypothesize that the >180 kDa band is a product of intermolecular SpyTag-PhyC-SpyCatcher reaction.

Performing a similar approach for heat purification of SpyRing BLA also showed good recovery, with some impurities present and also more recovery of heat-stable SpyTag-BLA-SpyCatcher polymers than with Ni-NTA

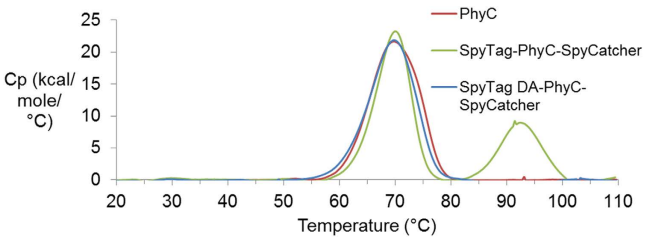


Figure 5. Analysis of SpyRing phytase by DSC. DSC of SpyTag (DA)-PhyC-SpyCatcher and PhyC. DSC gave the specific heat capacity (Cp) of each enzyme, scanning from 20–110 °C at 1 °C/min.

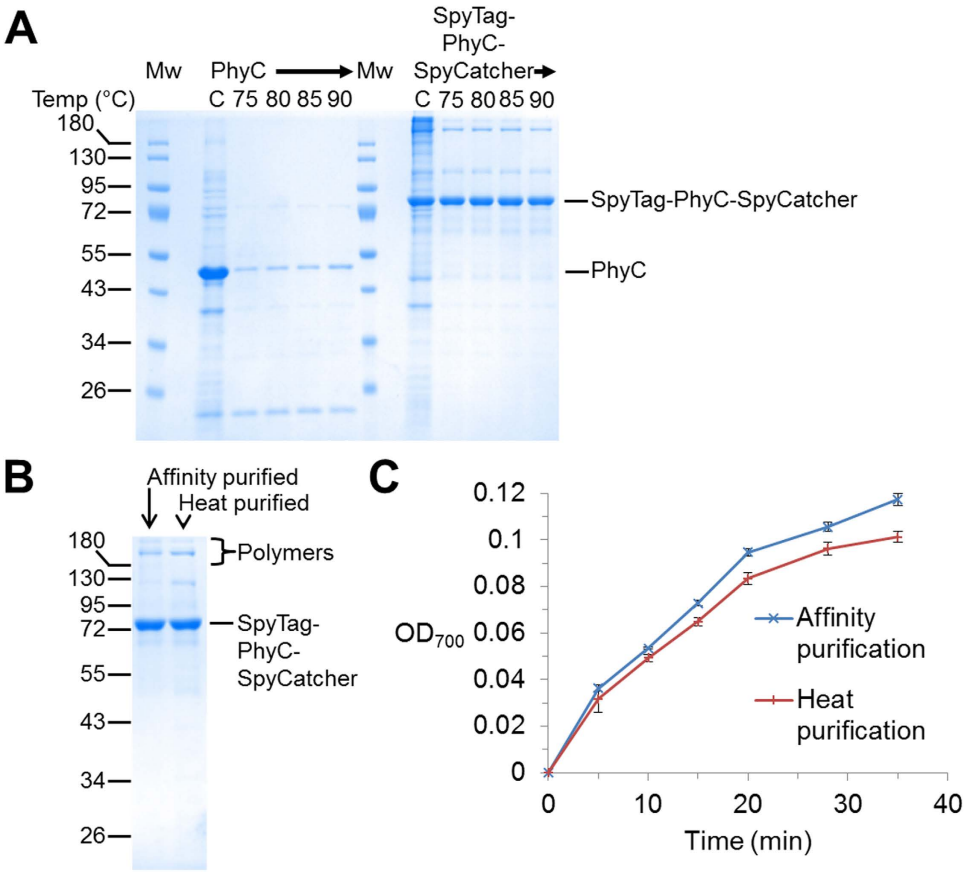


Figure 6. SpyRing cyclization allowed enzyme purification just by heating. (**A**) Heat-based purification of phytase. *E. coli* lysate expressing PhyC or SpyTag-PhyC-SpyCatcher was heated for 10 min at the indicated temperature, centrifuged, and the supernatant analyzed by SDS-PAGE with Coomassie staining. C is control without incubation. Mw are the molecular weight markers. (**B**) Comparing purity between heat-purified and affinity-purified phytase. SpyTag-PhyC-SpyCatcher either affinity-purified by Ni-NTA or heat-purified (normalized to give an equal intensity of the SpyTag-PhyC-SpyCatcher band) was boiled in SDS-loading buffer and analyzed by SDS-PAGE with Coomassie staining. (**C**) Heat purified phytase was still active. Catalytic activity of SpyTag-PhyC-SpyCatcher purified by either affinity chromatography or by heat purification (normalized by densitometry for an equal intensity of the SpyTag-PhyC-SpyCatcher band) was tested by a colorimetric activity assay for phosphate release from phytic acid. (All are mean of triplicate ± 1 s.d.; some error bars are too small to be visible.)

purification (Fig. S7A,B). Importantly for both SpyRing phytase (Fig. 6C) and SpyRing BLA (Fig. S7C), the recovered protein had comparable catalytic activity from heat purification when compared to affinity purification.

Discussion

We have shown that cyclization with SpyTag/SpyCatcher gave superior resilience to other spontaneous isopeptide-forming domains of distinct sizes and reactive chemistries. SpyRing cyclization conferred resilience on phytase, an industrially-important class of enzyme with termini 29 Å apart. The mechanism of SpyRing

resilience was studied through DSC and CD, indicating that cyclization by an isopeptide bond-forming domain did not change the temperature for initial unfolding but is likely to facilitate refolding. The resilience to thermal inactivation of SpyRings allowed normally heat-labile enzymes to be purified simply by heating of cell lysate.

The alternative isopeptide-forming partners, Pilin-C/Isopeptag and SnoopCatcher/SnoopTag, allowed efficient enzyme cyclization, even though their reaction rate (when tested intermolecularly) is slower than SpyCatcher/SpyTag[11,19,21]. Both PilinRing and SnoopRing cyclization gave resilience to heat-induced aggregation/inactivation, even at 100 °C, so enzyme resilience from cyclization by isopeptide bond-forming domains is a general phenomenon. Non-reactive point mutants confirmed that isopeptide bond formation was important to the effect on enzyme resilience. Inability to form the isopeptide bond had two effects: lowering the Tm of the fused partner and untethering the termini, both likely to decrease the enzyme's resilience. The degree of resilience from PilinRing or SnoopRing cyclization for BLA was not as complete as with SpyRing cyclization. For PilinRing and SnoopRing cyclization, recovery of activity after 75 °C heating was less efficient than after 100 °C heating. Previous studies have suggested that temperature may influence the partition between different unfolded states, some of which favor refolding and some of which favor aggregation[26,27]. The importance of the cyclization framework for determining the extent of resilience is consistent with the large difference in BLA thermo-resilience reported through split intein cyclization (~5 °C increase)[5] compared to SpyRing cyclization (>60 °C increase)[10]. It is also consistent with our previous work which showed that the cyclized SpyTag-BLA-SpyCatcher construct exhibited a greater thermal resilience than a BLA-CnaB2 fusion construct[10]. PilinRing cyclization did give excellent resistance to aggregation, but after heating not all the soluble enzyme was functional. Resilience did not correlate with the size of the cyclizing domain, since Pilin-C is much bigger than SpyCatcher, but resilience did correlate with the unfolding temperature of the cyclizing domain. Interestingly all three cyclization systems stabilized the sandwiched enzyme from aggregation at 100 °C where DSC indicates that the Tag/Catcher domain would have unfolded.

SpyRing cyclization had a major effect in increasing the thermal resilience of PhyC phytase. Extensive efforts have been made to engineer different classes of phytases for high stability but phytase inactivation when food pellets are heat-sterilized is still a significant issue and so our SpyRing phytase provides a new direction for this challenge[18,28]. Phytase supplementation has already shown success in human trials for improving dietary uptake of iron and zinc[29]. Alternatives to phytase supplementation include the engineering of crops expressing phytase[18] or "Enviropig", a genetically-modified pig producing phytase in its saliva[30]. However, eating transgenic organisms encounters a lot of resistance from the general public[31].

It is commonly known that enzymes from hyperthermophiles, such as Taq DNA polymerase from *Thermus aquaticus*, can be expressed in *E. coli* and purified simply by heating[25,32]. Establishing such a simple purification route for enzymes that are not intrinsically thermostable could have academic and industrial importance. We showed here that SpyRing cyclization allowed substantial heat-purification of BLA and phytase with high retention of activity. Purification was not dissimilar to that achieved by affinity purification, although the protein was not purified to homogeneity[33]. Nevertheless orthogonal polishing steps may then be used to purify further. Also, for many industrial processes, such as nutritional supplementation or biotransformations, moderate purity is often sufficient. Fusing proteins to thermostable domains was previously shown to allow stabilization[34] and heat-purification of proteins of interest, but this required the protein of interest to be substantially smaller than the thermostable domain[35]. In contrast, here SpyRing cyclization was effective in conferring resilience on fused enzymes substantially larger than SpyTag/SpyCatcher.

SpyRings represent a simple and versatile approach to increase the thermal resilience of enzymes. Future work will apply this technology to other globally-important enzymes and endeavor to improve resilience even further, by increasing SpyTag/SpyCatcher's Tm or by harnessing the bacterial diversity of isopeptide-forming domains[36].

Methods

Cloning. KOD Hot Start DNA Polymerase (Roche) was used as the primary polymerase for PCR. DNA was transformed into competent *E. coli* DH5α (Life Technologies). All constructs contained an N-terminal His_6 tag encoded in the pET28a plasmid (Novagen). The *Bacillus subtilis* PhyC insert (GenBank: JQ437256.1) was assembled using the DNAWorks method[37]. The PhyC insert was amplified using primers 5′- CCGACGAAGGGTTCAGGGGGTTCCGGTAAACATAAACTGAGCGATCCGTATCATTTCAC and 5′- GCGCCTCCGCTGCCACCACTCCCTTTGCCCGAACGATCGGTCAATTTACG. The insert was cloned into the SpyTag/SpyCatcher cassette using the method previously described[10] to make pET28a SpyTag-PhyC-SpyCatcher (GenBank accession number KU608330) and pET28a SpyTag DA-PhyC-SpyCatcher. pET28a PhyC without the SpyTag/SpyCatcher fusions was cloned by amplifying the PhyC gene from pET28a SpyTag-PhyC-SpyCatcher using primers 5′- TATACATATGGGAAAACATAAACTGAGCGATCCGTATCATTTCAC and 5′- TTTTAAGCTTTCATTATTTGCCCGAACGATCGGTCAATTTACG. The amplified product was inserted into pET28a using restriction digestion with NdeI (NEB) and HindIII (NEB), followed by ligation using T4 DNA ligase (NEB). pET28a SpyTag-BLA-SpyCatcher (GenBank KJ645919, Addgene ID 70943), pET28a SpyTag DA-BLA-SpyCatcher and pET28a BLA were described previously[10]. pDEST14 SpyCatcher[11], pET28a Spy0128[19] and pET28a RrgACatcher G842T (SnoopCatcher G848D)[21] were as described. pET28a SnoopTag-BLA-SnoopCatcher (GenBank accession number KU608331) was cloned using overlap extension PCR[38] followed by digestion with NdeI and HindIII and ligation into pET28a. SnoopTag is based on the N-terminal β-strand of RrgA's D4 domain (residues 734–748, numbering from PDB ID 2WW8)[22]. SnoopCatcher is based on residues 749–860 of *Streptococcus pneumoniae* adhesin RrgA's D4 domain (numbering from PDB ID 2WW8). The first fragment containing the SnoopTag-BLA component was amplified from SpyTag-BLA-SpyCatcher using primers 5′- ATTACATATGGGAAAACTGGGGGACATCGAATTCATCAAAGTAAACAAAGGTTCAGGGGGTTCCGGTCACC and 5′- CTAAACACGGCACCACGCAGCGGCTTTCCACTGCCACCACTCCCCCAATGC. The second fragment containing the SnoopCatcher component was amplified from the RrgACatcher G842T construct[21] using primers

5′- AAGCCGCTGCGTGGTGCCG and 5′- ATTAAAGCTTTCATTATTTCGGCGGTATCGGTTCATTGGTG. pET28a SnoopTag KA-BLA-SnoopCatcher (K742 to A) was generated using pET28a SnoopTag-BLA-SnoopCatcher as a template and primer 5′- CCCTGAACCTTTGTTTACTGCGATGAATTCGATGTCCCC and its reverse complement by QuikChange. pET28a Pilin-C-BLA-Isopeptag (GenBank accession number KU608332) was cloned using Gibson assembly[39] with Gibson Assembly Master Mix (NEB). The vector was amplified using primers 5′- TCCCATATGGCTGCCGCGCGG and 5′- AAGCTTGCGGCCGCACTCGAGC using pET28a SpyTag-BLA-SpyCatcher as a template. The Pilin-C fragment was generated using primers 5′- CCGCGCGGCAGCCATATGGGAGCTACAACAGTTCACGGGGAGAC and 5′- GGTGACCGGAACCCCCTGAACCAATGGTCATATCTTTATCAGTAGATGTCTCTTG using pET28a Pilin-C[19] as a template. The BLA-Isopeptag fragment was amplified using primers 5′- GGTTCAGGGGGTTCCGGTCAC and 5′- GCTCGAGTGCGGCCGCAAGCTTTCATTATTCCGCATCTT TTTTGTTGGTAAAGGTAATGGTCATATCTTTATCGGTTCCACTGCCACCACTCCCCC using pET28a SpyTag-BLA-SpyCatcher as a template. pET28a Pilin-C KA-BLA-Isopeptag was generated using pET28a Pilin-C-BLA-Isopeptag as a template and primer 5′- TCTACTACATTAACGGTGAAGGCAAAAGTTTCAGGTA CCGGTGG and its reverse complement by QuikChange. All constructs were verified by Sanger sequencing.

Recombinant protein expression and purification. Samples for MS were expressed in *E. coli* B834 DE3 cells (Novagen) in the absence of glucose to avoid gluconylation[40]. Saturated cultures were diluted 1 in 100 and grown at 37 °C in LB with 0.5 mg/mL kanamycin to an OD_{600} of 0.5 and induced using 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Melford Labs) at 30 °C for 3 hours. For all other purposes, samples were expressed in *E. coli* BL21 DE3 RIPL (Stratagene), except Pilin-C (KA)-BLA-Isopeptag was expressed in *E. coli* B834 DE3. Samples were grown in LB with 0.8% glucose and 0.5 mg/mL kanamycin overnight. Saturated cultures were diluted 1 in 100 and grown at 37 °C until cultures reached an OD_{600} of 0.5. The cultures were induced with 0.4 mM IPTG and grown for 16 hours at 18 °C. Samples were spun down at 4,000 g for 20 minutes and the pellet was re-suspended in 50 mM Tris-HCl pH 7.8 with 300 mM NaCl (Ni-NTA buffer) containing 1 mM phenylmethylsulfonyl fluoride (PMSF, Roche) and 1× EDTA-free mixed protease inhibitors (Roche). Protein purification was performed using Ni-NTA resin (Qiagen). Samples were allowed to bind to the resin for 30 minutes at 4 °C. The sample was added onto a poly-prep column (Bio-Rad), washed with 40 resin volumes of Ni-NTA buffer with 10 mM imidazole, then washed with 20 resin volumes of Ni-NTA buffer with 30 mM imidazole and eluted with 4 resin volumes of Ni-NTA buffer with 75 mM imidazole. SpyTag (DA)-BLA-SpyCatcher, BLA, SnoopTag (KA)-BLA-SnoopCatcher, and Pilin-C (KA)-BLA-Isopeptag samples used in solubility assays, enzymatic assays and SDS-PAGE assays were dialyzed 3 times in PBS. SpyTag (DA)-PhyC-SpyCatcher and PhyC samples used in solubility assays, enzymatic assays and SDS-PAGE assays were dialyzed 3 times in 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$. SnoopTag (KA)-BLA-SnoopCatcher, Pilin-C (KA)-BLA-Isopeptag, and SpyTag (DA)-PhyC-SpyCatcher samples for MS were dialyzed 3 times in 10 mM ammonium acetate. SpyTag (DA)-PhyC-SpyCatcher and PhyC samples used for CD were dialyzed 3 times into 10 mM Tris-HCl pH 8.0 with 2 mM $CaCl_2$ and then purified again using anion-exchange to polish the samples and remove polymeric species. Anion-exchange chromatography was performed using a 1 mL HiTrap HP Q column (GE Healthcare) on an ÄKTA purifier (GE Healthcare). Fractions were collected and concentrated using a Vivaspin 6 centrifugal concentrator (Sartorius) with a 5 kDa cut-off and dialyzed 3 times in 10 mM Tris-SO_4 (Tris base with pH adjusted using H_2SO_4) pH 7.0 with 2 mM $CaSO_4$. Samples used to obtain values for kinetic parameters or units of activity were further purified by gel filtration chromatography to remove multimers. Gel filtration was performed using a Superdex 200 10/300 GL column (GE Healthcare) on an ÄKTA purifier. Fractions containing only enzyme monomer were collected and concentrated using a Vivaspin 6 centrifugal concentrator with a 5 kDa cut-off. Concentrations were determined from OD_{280} and the calculated molar extinction coefficients from ProtParam. For the comparative kinetic and activity studies, protein concentrations were determined using a Pierce bicinchoninic acid (BCA) Protein Assay Kit (Thermo Fisher Scientific).

Reaction of SpyCatcher with SpyTag peptide and SnoopCatcher with SnoopTag peptide. SpyTag peptide (GAHIVMVDAYKPTK) and SnoopTag peptide (GKLGDIEFIKVNKGY) were synthesized by Insight Biotechnology at >95% purity, validated by HPLC and mass spectrometry. SpyCatcher was allowed to react with a 4-fold molar excess of SpyTag at 20 °C for 2 hours in PBS. SnoopCatcher was allowed to react with a 3-fold molar excess of SnoopTag at 20 °C for 20 hours in PBS. Each sample was then dialyzed 3 times in PBS. The reaction was confirmed by the gel shift on 20% SDS-PAGE[19].

SDS-PAGE. SDS-PAGE was run on 12% (for phytase constructs and heat purification gels) or 14% (for β-lactamase constructs) polyacrylamide gels in an XCell SureLock module (Life Technologies). Gels were run at 200 V for 1 hour and then stained with InstantBlue Coomassie (Expedeon). For β-lactamase constructs, a final concentration of 100 mM dithiothreitol (DTT, Sigma) was added. For phytase constructs, a final concentration of 100 mM ethylenediaminetetraacetic acid (EDTA) was added to remove tightly-bound divalent cations affecting phytase mobility. Samples were mixed with 6× SDS-PAGE loading buffer (0.23 M Tris-HCl, 0.24% glycerol, 6.7% SDS, and 12 mM bromophenol blue) and heated at 95 °C for 7 minutes in a Bio-Rad C1000 Thermal cycler. Gels were imaged on a Bio-Rad ChemiDoc XRS+ and analyzed using Image Lab 3.0 software (Bio-Rad).

Temperature-dependent solubility assay. 20 μL 25 μM BLA construct in PBS with 100 mM DTT was incubated at 10, 25, 37, 55, 75, 90 or 100 °C for 10 minutes and then cooled to 10 °C on a Bio-Rad C1000 Thermal Cycler. 20 μL 10 μM phytase construct in 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$ was incubated at 10, 25, 55, 75, 90 or 100 °C for 10 minutes and then cooled to 10 °C on a Bio-Rad C1000 Thermal Cycler. The ramp rate was 3 °C/s. Samples were spun at 17,000 g at 4 °C for 30 minutes and the supernatant was removed for SDS-PAGE.

Every gel had a triplicate loading control to normalize the samples. Loading samples consisted of the protein sample that had been kept at 10 °C. The loading control was defined as 100% soluble fraction. Bands were quantified using densitometry values obtained from the Image Lab 3.0 software.

β-lactamase thermal resilience enzymatic assay. A 96 well clear flat-bottom polystyrene plate (Greiner) was pre-blocked with 200 μL PBS with 3% bovine serum albumin (BSA, ≥ 98%, Sigma) at 37 °C for 2 hours. The plate was then rinsed twice using 0.1 M NaH_2PO_4 pH 7.0 with 1 mM EDTA. 20 μL 25 μM BLA construct in PBS with 100 mM DTT was incubated at 10, 25, 37, 55, 75, 90 or 100 °C for 10 minutes and then cooled to 10 °C on a Bio-Rad C1000 Thermal Cycler (3 °C/s). The protein was diluted 1 in 400 using 0.1 M NaH_2PO_4 pH 7.0 with 1 mM EDTA and 1% BSA. 3 nM of BLA construct was allowed to react with 100 μM nitrocefin (Merck) in 0.1 M NaH_2PO_4 pH 7.0 with 1 mM EDTA at 20 °C. The reaction was quenched at different time-points by adding an equal volume of 500 μM potassium clavulanate (Sigma) dissolved in 0.1 M NaH_2PO_4 pH 7.0 with 1 mM EDTA. Care was taken to protect the plates from direct light. Half of the volume was transferred into a clean 96 well clear flat-bottom polystyrene plate to remove bubbles that could interfere with absorbance measurements. OD_{486} was measured using a SpectraMax M3 microplate reader (Molecular Devices). Absorbance was corrected for the dilution caused by the addition of potassium clavulanate. Samples were blanked against a control lacking enzyme.

β-lactamase kinetics assay. BLA activity was measured as previously[41]. 2.0 mL sample cups (Kartell Labware) were pre-blocked using 50 mM NaH_2PO_4 pH 7.0 with 3% BSA at 37 °C for 2 hours. The samples cups were then rinsed twice using 50 mM NaH_2PO_4 pH 7.0. Enzymes were diluted in the sample cup to a concentration of 0.25 nM using 50 mM NaH_2PO_4 pH 7.0 with 0.1% BSA. Samples were then allowed to react in 50 mM NaH_2PO_4 pH 7.0 with 100, 50, 30, 20, 15, or 10 μM Nitrocefin (Oxoid) in a Gallery Plus Automated Photometric Analyzer (Thermo Scientific) at 30 °C. Readings were taken every 9 seconds for a period of 99 seconds at a wavelength of 500 nm ($\Delta\varepsilon_{500} = 15{,}900\ M^{-1}cm^{-1}$)[42]. The first 6 data points were linear and were used to obtain the initial rates of reaction. A Lineweaver-Burk plot was used to extrapolate the kinetic parameters. Errors for kinetic values were obtained by calculating the kinetics for 3 separate sets of data and then calculating the average and standard deviation of the 3 kinetic values.

Phytase thermal resilience enzymatic assay. A 96 well clear flat-bottom polystyrene plate (Greiner) was pre-blocked with 200 μL 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$ and 3% BSA at 37 °C for 2 hours. The plate was then rinsed twice using 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$. 20 μl of 10 μM phytase construct in 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$ was incubated at 10, 55, 75, 90 or 100 °C for 10 minutes and then cooled to 10 °C on a Bio-Rad C1000 Thermal Cycler (3 °C/s). The protein was diluted 1 in 26 using 50 mM Tris-HCl pH 7.0, 2 mM $CaCl_2$, 10% glycerol, 1% BSA. 77 nM phytase construct was allowed to react with 2 mM phytic acid (Sigma) in 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$ and 10% glycerol at 50 °C. The reaction was quenched at different time points by adding an equal volume of 20% trichloroacetic acid. A solution containing 12 mM ammonium heptamolybdate (Sigma) and 4.3% H_2SO_4 was mixed in a 4 : 1 ratio with a solution containing 178 mM $FeSO_4$ (Scientific Laboratory Supplies) to make the color mix solution. 1 volume of color mix solution was added for every 2 volumes of quenched reaction solution. The plates were spun at 3,500 g at 20 °C in a Multifuge X3R (Thermo Scientific) for 10 minutes. Within 8 minutes, 60% of the solution (90 μL of 150 μL) was transferred into a clean 96 well clear flat-bottom polystyrene plate and OD_{700} was measured using a SpectraMax M3 microplate reader. Blank consisted of all the components minus phytase.

Phytase specific activity assay. Phytase activity was measured as previously[43]. 40 nM phytase was incubated with 1.6 mM phytic acid at 37 °C for 1, 2, 3, 4, or 5 minutes in 100 mM Tris-HCl pH 7.5 with 1 mM $CaCl_2$. The reaction was quenched by adding an equal volume of 20% trichloroacetic acid. Samples were spun at 17,000 g at 20 °C for 10 minutes and 90% of the supernatant was moved into a clean tube. A solution containing 12 mM ammonium heptamolybdate (Sigma) and 4.3% H_2SO_4 was mixed in a 4:1 ratio with a solution containing 178 mM $FeSO_4$ (Scientific Laboratory Supplies) to make the color mix solution. 1 volume of color mix solution was added for every 2 volumes of quenched reaction solution. The samples were then incubated at 25 °C for 20 minutes and then OD_{700} was measured using a Varian Cary 50 UV-Vis Spectrophotometer. For the blank, the enzyme was first added to 20% trichloroacetic acid to inactivate the enzyme before mixing with the substrate. In order to determine the specific activity, OD_{700} was converted to μg of phosphate using a set of standards containing 20 to 100 μM NaH_2PO_4. A straight line of best fit was used to obtain the initial rate. The specific activity for each repeat was individually calculated and averaged. The error refers to 1 standard deviation from the mean.

Heat purification. Bacterial pellets obtained from an 800 mL bacterial culture were frozen at −80 °C and then thawed. Cells were resuspended in 8 mL 50 mM Tris-HCl pH 7.8 with 300 mM NaCl containing 1 mM phenylmethylsulfonyl fluoride (PMSF, Roche) and 1× EDTA-free mixed protease inhibitors (Roche, 1 tablet resuspended in 1.5 ml H_2O was 100× stock). The OD_{600} of the solution was 110–150 (as determined on a diluted sample of the resuspended cells). Cells were sonicated and centrifuged at 17,000 g for 20 minutes at 4 °C. The soluble fraction was removed. The OD_{280} of the solution was 60 (as determined on a diluted sample). Half of the cell lysate was purified using Ni-NTA affinity chromatography as described above and dialyzed 3 times into 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$ (PhyC constructs) or PBS pH 7.4 (BLA constructs) for use in the activity comparison later on. The other half was diluted 1 in 10 into 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$ (PhyC constructs) or PBS pH 7.4 (BLA constructs) The sample was split into 50 μL aliquots and heated at 75, 80, 85 or 90 °C for 10 minutes in a Bio-Rad C1000 Thermal Cycler and then cooled to 10 °C. Samples were then centrifuged at 17,000 g at 4 °C for 30 minutes and the supernatant was removed and further analyzed. Samples heated to 75, 80, 85 or 90 °C were analyzed via SDS-PAGE to check their purity. For testing purity and activity, the concentration of SpyTag-PhyC-SpyCatcher from affinity purification was matched to the concentration of

SpyTag-PhyC-SpyCatcher from heat purification using densitometry of the SpyTag-PhyC-SpyCatcher band based on SDS-PAGE with Coomassie staining.

Mass spectrometry. The samples were analyzed using a Micromass LCT time-of-flight electrospray ionization mass spectrometer (Micromass UK). MassLynx V 4.00.00 (Waters Corporation) was used to convert the m/z spectrum to a mass spectrum using a Maximum Entropy algorithm. Predicted masses were obtained using ProtParam and by taking into account the change in mass that occurs when an isopeptide bond is formed (−18 for SpyCatcher, −17 for SnoopCatcher and Pilin-C)[10,20,22].

DSC. DSC profiles of PhyC, SpyTag (DA)-PhyC-SpyCatcher, SpyTag-BLA-SpyCatcher, SnoopTag (KA)-BLA-SnoopCatcher, Pilin-C (KA)-BLA-Isopeptag, BLA, Spy0128, SpyCatcher + SpyTag and SnoopCatcher + SnoopTag were measured on the VP Cap DSC (GE Healthcare). The DSC conditions to determine the Tm were 20 μM protein in PBS pH 7.4 (BLA constructs, Spy0128, SpyCatcher + SpyTag and SnoopCatcher +SnoopTag) or 50 mM Tris-HCl pH 7.0 with 2 mM $CaCl_2$ (PhyC constructs) from 20 to 110 °C at a scan rate of 1 °C/min at a pressure of 3 atm. The baseline traces were subtracted from the sample traces and then normalized for concentration and volume using the MicroCal DSC Origin Pro 7.0 software (GE Healthcare). Samples containing only buffer were used to obtain the baseline traces.

CD. Circular Dichroism experiments were performed in 10 mM Tris-SO_4 pH 7.0 with 2 mM $CaSO_4$ buffer. Samples were analyzed on a J-815 Circular Dichroism Spectrometer (Jasco) using a 1 mm path-length Spectrosil quartz cuvette (Starna Scientific). The temperature was controlled using a CDF-426S Peltier unit (Jasco). Far-UV CD spectra were obtained by incubating 0.15 mg/mL protein at 10 °C, before ramping at 12 °C/min to 55, 75 or 90 °C. Samples were held at that temperature for 3 minutes before starting a read (the "During heating" trace). The scanning speed of the instrument was set to 50 nm/min and scans were performed from 260 to 185 nm. After the read, samples were cooled to 10 °C. Samples were kept at 10 °C for 3 minutes before taking another read (the "After cooling" trace). Four scans were recorded from each independent sample, the recorded spectra were averaged, and then subtracted by the averaged baseline spectrum (buffer alone). Spectra were smoothened using the Savitzky-Golay method[44] on the SpectraManager 2 software (Jasco). Ellipticity was converted to molar ellipticity using the SpectraManager 2 software.

Molecular visualization. Protein structures were rendered in PyMOL (DeLano Scientific), based on Protein Data Bank files 1ZG4[45], 3B2M[20], 2WW8[22], and 3AMR[24].

References

1. Kawedia, J. D. & Rytting, M. E. Asparaginase in Acute Lymphoblastic Leukemia. *Clin. Lymphoma Myeloma Leuk.* **14,** Supplement, S14–S17 (2014).
2. Bhosale, S. H., Rao, M. B. & Deshpande, V. V. Molecular and industrial aspects of glucose isomerase. *Microbiol. Rev.* **60,** 280–300 (1996).
3. Reetz, M. T. & Carballeira, J. D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* **2,** 891–903 (2007).
4. Evans, T. C., Jr, Benner, J. & Xu, M. Q. The *in vitro* ligation of bacterially expressed proteins using an intein from Methanobacterium thermoautotrophicum. *J. Biol. Chem.* **274,** 3923–3926 (1999).
5. Iwai, H. & Plückthun, A. Circular β-lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett.* **459,** 166–172 (1999).
6. Camarero, J. A. & Muir, T. W. Biosynthesis of a Head-to-Tail Cyclized Protein with Improved Biological Activity. *J. Am. Chem. Soc.* **121,** 5597–5598 (1999).
7. Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C. & Benkovic, S. J. Production of cyclic peptides and proteins *in vivo*. *Proc. Natl. Acad. Sci.* **96,** 13638–13643 (1999).
8. Parthasarathy, R., Subramanian, S. & Boder, E. T. Sortase A as a novel molecular 'stapler' for sequence-specific protein conjugation. *Bioconjug. Chem.* **18,** 469–476 (2007).
9. Goldenberg, D. P. & Creighton, T. E. Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **165,** 407–413 (1983).
10. Schoene, C., Fierer, J. O., Bennett, S. P. & Howarth, M. SpyTag/SpyCatcher Cyclization Confers Resilience to Boiling on a Mesophilic Enzyme. *Angew. Chem. Int. Ed.* **53,** 6101–6104 (2014).
11. Zakeri, B. *et al.* Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci.* **109,** E690–E697 (2012).
12. Veggiani, G., Zakeri, B. & Howarth, M. Superglue from bacteria: unbreakable bridges for protein nanotechnology. *Trends Biotechnol.* **32,** 506–512 (2014).
13. Bailey, R. L., West, K. P. & Black, R. E. The epidemiology of global micronutrient deficiencies. *Ann. Nutr. Metab.* **66** Suppl 2, 22–33 (2015).
14. Allen, L., de Benoist, B., Dary, O. & Hurrell, R. *Guidelines on Food Fortification with Micronutrients*. (World Health Organization, 2006).
15. Ye, X. *et al.* Engineering the Provitamin A (β-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science* **287,** 303–305 (2000).
16. Bohlke, R. A., Thaler, R. C. & Stein, H. H. Calcium, phosphorus, and amino acid digestibility in low-phytate corn, normal corn, and soybean meal by growing pigs. *J. Anim. Sci.* **83,** 2396–2403 (2005).
17. Selle, P. H., Ravindran, V., Bryden, W. L. & Scott, T. Influence of Dietary Phytate and Exogenous Phytase on Amino Acid Digestibility in Poultry: A Review. *J. Poult. Sci.* **43,** 89–103 (2006).
18. Lei, X. G., Weaver, J. D., Mullaney, E., Ullah, A. H. & Azain, M. J. Phytase, a New Life for an 'Old' Enzyme. *Annu. Rev. Anim. Biosci.* **1,** 283–309 (2013).
19. Zakeri, B. & Howarth, M. Spontaneous Intermolecular Amide Bond Formation between Side Chains for Irreversible Peptide Targeting. *J. Am. Chem. Soc.* **132,** 4526–4527 (2010).
20. Kang, H. J., Coulibaly, F., Clow, F., Proft, T. & Baker, E. N. Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. *Science* **318,** 1625–1628 (2007).
21. Veggiani, G. *et al.* Programmable polyproteams built using twin peptide superglue. *Proc. Natl. Acad. Sci.* doi:10.1073/pnas.1519214113 2016.
22. Izoré, T. *et al.* Structural basis of host cell recognition by the pilus adhesin from Streptococcus pneumoniae. *Structure* **18,** 106–115 (2010).

23. Hagan, R. M. *et al.* NMR Spectroscopic and Theoretical Analysis of a Spontaneously Formed Lys-Asp Isopeptide Bond. *Angew. Chem. Int. Ed.* **49,** 8421–8425 (2010).
24. Zeng, Y.-F. *et al.* Crystal Structures of Bacillus Alkaline Phytase in Complex with Divalent Metal ions and Inositol Hexasulfate. *J. Mol. Biol.* **409,** 214–224 (2011).
25. Kirk, N. & Cowan, D. Optimising the recovery of recombinant thermostable proteins expressed in mesophilic hosts. *J. Biotechnol.* **42,** 177–184 (1995).
26. Wang, W. & Roberts, C. J. Non-Arrhenius Protein Aggregation. *AAPS J.* **15,** 840–851 (2013).
27. Sakamoto, R., Nishikori, S. & Shiraki, K. High temperature increases the refolding yield of reduced lysozyme: implication for the productive process for folding. *Biotechnol. Prog.* **20,** 1128–1133 (2004).
28. Ping, S.-Z., Zhang, W., Lin, M., Yan, Y.-L., Zhao, Z.-L., Lu, W. & Chen, M. Inventors; Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, assignee. Cyclized phytase with improved heat stability and proteinase resistance. People's Republic of China patent CN101638642 A. 2010 Feb 3.
29. Troesch, B., Jing, H., Laillou, A. & Fowler, A. Absorption studies show that phytase from Aspergillus niger significantly increases iron and zinc bioavailability from phytate-rich foods. *Food Nutr. Bull.* **34,** S90–101 (2013).
30. Golovan, S. P. *et al.* Pigs expressing salivary phytase produce low-phosphorus manure. *Nat. Biotechnol.* **19,** 741–745 (2001).
31. Wu, F. Explaining public resistance to genetically modified corn: an analysis of the distribution of benefits and risks. *Risk Anal. Off. Publ. Soc. Risk Anal.* **24,** 715–726 (2004).
32. Tanaka, T., Kawano, N. & Oshima, T. Cloning of 3-Isopropylmalate Dehydrogenase Gene of an Extreme Thermophile and Partial Purification of the Gene Product. *J. Biochem. (Tokyo)* **89,** 677–682 (1981).
33. Kwon, S., Jung, Y. & Lim, D. Proteomic analysis of heat-stable proteins in Escherichia coli. *BMB Rep.* **41,** 108–111 (2008).
34. Pierre, B., Xiong, T., Hayles, L., Guntaka, V. R. & Kim, J. R. Stability of a guest protein depends on stability of a host protein in insertional fusion. *Biotechnol. Bioeng.* **108,** 1011–1020 (2011).
35. de Marco, A., Casatta, E., Savaresi, S. & Geerlof, A. Recombinant proteins fused to thermostable partners can be purified by heat incubation. *J. Biotechnol.* **107,** 125–133 (2004).
36. Kang, H. J. & Baker, E. N. Intramolecular isopeptide bonds: protein crosslinks built for stress? *Trends Biochem. Sci.* **36,** 229–237 (2011).
37. Hoover, D. M. & Lubkowski, J. DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res.* **30,** e43 (2002).
38. Bryksin, A. V. & Matsumura, I. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *BioTechniques* **48,** 463–465 (2010).
39. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).
40. Geoghegan, K. F. *et al.* Spontaneous alpha-N-6-phosphogluconoylation of a 'His tag' in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **267,** 169–184 (1999).
41. Raquet, X. *et al.* TEM β-Lactamase Mutants Hydrolysing Third-generation Cephalosporins: A Kinetic and Molecular Modelling Analysis. *J. Mol. Biol.* **244,** 625–639 (1994).
42. Chen, C. C. H. & Herzberg, O. Relocation of the catalytic carboxylate group in class A β-lactamase: the structure and function of the mutant enzyme Glu166→Gln: Asn170→Asp. *Protein Eng.* **12,** 573–579 (1999).
43. Guerrero-Olazarán, M., Rodríguez-Blanco, L., Carreon-Treviño, J. G., Gallegos-López, J. A. & Viader-Salvadó, J. M. Expression of a Bacillus Phytase C Gene in Pichia pastoris and Properties of the Recombinant Enzyme. *Appl. Environ. Microbiol.* **76,** 5601–5608 (2010).
44. Savitzky, A. & Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **36,** 1627–1639 (1964).
45. Stec, B., Holtz, K. M., Wojciechowski, C. L. & Kantrowitz, E. R. Structure of the wild-type TEM-1 beta-lactamase at 1.55 A and the mutant enzyme Ser70Ala at 2.1 A suggest the mode of noncovalent catalysis for the mutant enzyme. *Acta Crystallogr. D Biol. Crystallogr.* **61,** 1072–1079 (2005).

Acknowledgements

Funding was provided by BBSRC (C.S.), Sekisui Diagnostics (C.S., S.P.B.), and Oxford University Department of Biochemistry (M.H.). We thank Nick Major of Sekisui Diagnostics for helpful advice and the Molecular Biophysics Facility at Oxford University Department of Biochemistry.

Author Contributions

C.S. performed the experiments. C.S. and M.H. wrote the paper. C.S., S.P.B. and M.H. designed the research and analyzed data. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: M.H. is an inventor on a patent concerning peptide reaction via spontaneous amide bond formation (EP2534484).

How to cite this article: Schoene, C. *et al.* SpyRing interrogation: analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries. *Sci. Rep.* **6**, 21151; doi: 10.1038/srep21151 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

SpyRings Declassified: A Blueprint for Using Isopeptide-Mediated Cyclization to Enhance Enzyme Thermal Resilience

C. Schoene*, S.P. Bennett†, M. Howarth*,1
*University of Oxford, Oxford, United Kingdom
†Sekisui Diagnostics UK Ltd., Maidstone, Kent, United Kingdom
1Corresponding author: e-mail address: mark.howarth@bioch.ox.ac.uk

Contents

1. Introduction 2
2. Isopeptide-Mediated Enzyme Cyclization 3
   2.1 Partners for Spontaneous Isopeptide Bond Formation 3
   2.2 Cyclization with Isopeptide Bonds 5
   2.3 Selection of Enzymes for Cyclization 5
   2.4 Cloning Genes of Interest into the SpyRing Cassette 6
   2.5 Confirming Successful Cyclization 7
3. Characterization of Enzyme Resilience After SpyRing Cyclization 11
   3.1 Industrial Relevance of Phytase 11
   3.2 Measuring Aggregation Can Provide Evidence for an Increase in Thermal Resilience 11
   3.3 Recovered Activity Is the Most Important Measurement When Testing Thermal Resilience 12
   3.4 FLuc Is an Example of a Poor Target for SpyRing Cyclization 14
   3.5 Dynamic Scanning Calorimetry Is a Useful Biophysical Tool to Study SpyRing Cyclization 15
4. Concluding Remarks 16
Acknowledgments 16
References 16

Abstract

Enzymes often have marginal stability, with unfolding typically leading to irreversible denaturation. This sensitivity is a major barrier, both for de novo enzyme development and for expanding enzyme impact beyond the laboratory. Seeking an approach to enhance resilience to denaturation that could be applied to a range of different enzymes, we developed SpyRing cyclization. SpyRings contain genetically encoded SpyTag (13 amino acids) on the N-terminus and SpyCatcher (12 kDa) on the

Methods in Enzymology
ISSN 0076-6879
http://dx.doi.org/10.1016/bs.mie.2016.05.004

© 2016 Elsevier Inc.
All rights reserved.

ARTICLE IN PRESS

C-terminus of the enzyme, so that the Spy partners spontaneously react together through an irreversible isopeptide bond. SpyRing cyclization gave major increases in thermal resilience, including on a model for enzyme evolution, β-lactamase, and an industrially important enzyme in agriculture and nutrition, phytase. We outline the SpyRing rationale, including comparison of SpyRing cyclization to other cyclization strategies. The cloning strategy is presented for the simple insertion of enzyme genes for recombinant expression. We discuss structure-based approaches to select suitable enzyme cyclization targets. Approaches to evaluate the cyclization reaction and its effect on enzyme resilience are described. We also highlight the use of differential scanning calorimetry to understand how SpyRing cyclization promotes enzyme refolding. Efficiently searching sequence space will continue to be important for enzyme improvement, but the SpyRing platform may be a valuable rational adjunct for conferring resilience.

1. INTRODUCTION

Understanding how to extend the scope of enzyme use is a central challenge in protein engineering. The potential of enzymes is illustrated by the fact that there are more than 100 enzymes already in large-scale industrial use. These applications range from laundry detergents and food processing to the cosmetics industry (Li, Yang, Yang, Zhu, & Wang, 2012). Major challenges for taking promising enzymes from the laboratory into large-scale application include the high cost of production (Howard, Abotsi, Jansen van Rensburg, & Howard, 2003) and poor stability at high temperatures (Lei, Weaver, Mullaney, Ullah, & Azain, 2013). The most common strategies to enhance the thermal resilience of enzymes are structure based or library based, but these approaches are usually performed on a case-by-case basis (Packer & Liu, 2015; Yang, Liu, Li, Chen, & Du, 2015).

To identify generic approaches to enhance stability, protein cyclization has been explored. Protein cyclization consists of covalently tethering the N- and C-termini of a protein, thereby reducing flexibility in two of the most mobile parts of the protein, while reducing the entropy gain on unfolding (Schumann et al., 2015). Previous work on protein cyclization has utilized carbodiimide cross-linking (Goldenberg & Creighton, 1983), split inteins (Camarero & Muir, 1999; Evans, Benner, & Xu, 1999; Iwai & Plückthun, 1999; Scott, Abel-Santos, Wall, Wahnon, & Benkovic, 1999), and sortase (Parthasarathy, Subramanian, & Boder, 2007). In some cases, an increase in thermal resilience has been reported

ARTICLE IN PRESS

(Antos et al., 2009; Iwai & Plückthun, 1999; Popp, Dougan, Chuang, Spooner, & Ploegh, 2011; van Lieshout et al., 2012). We have developed an approach to cyclize proteins based on spontaneous isopeptide bond formation, which has been able to achieve >60°C increases in thermal resilience (Reddington & Howarth, 2015; Schoene, Bennett, & Howarth, 2016; Schoene, Fierer, Bennett, & Howarth, 2014).

2. ISOPEPTIDE-MEDIATED ENZYME CYCLIZATION

2.1 Partners for Spontaneous Isopeptide Bond Formation

We have developed a range of different peptide/protein pairs able to react irreversibly. The story starts when Ted Baker's group solved the crystal structure of Spy0128, a pilin subunit originating from *Streptococcus pyogenes*, discovering the formation of an isopeptide bond in each of two domains between lysine and asparagine side chains (Kang, Coulibaly, Clow, Proft, & Baker, 2007). Mass spectrometry of Spy0128 validated isopeptide bond formation, while mutational analysis confirmed the three residues necessary for the formation of this bond (Kang et al., 2007). We were able to split the C-terminal β-strand of Spy0128 into a 16 amino acid tag, termed Isopeptag. Residues 18–299 of Spy0128 were separately expressed and called Pilin-C. By genetically fusing Isopeptag to maltose-binding protein (MBP), we found that Isopeptag-MBP could reconstitute and form an isopeptide bond with Pilin-C (Zakeri & Howarth, 2010).

Further work was then carried out to find a faster and smaller reactive pair than Pilin-C/Isopeptag. The CnaB2 domain of the fibronectin-binding protein (FbaB) of *S. pyogenes* was found to form a spontaneous isopeptide bond between aspartic acid and lysine (Hagan et al., 2010; Oke et al., 2010). SpyTag, a 13 amino acid peptide tag containing the reactive aspartic acid residue, consisted of the C-terminal β-strand of the CnaB2 domain as well as the succeeding five residues of FbaB not included in the crystal structure. SpyCatcher was generated from the remainder of CnaB2, which was further optimized by mutating two exposed hydrophobic residues (Fig. 1A). SpyTag-MBP was able to react efficiently with SpyCatcher (second-order rate constant $1.4 \times 10^3\ M^{-1}s^{-1}$). Reaction occurred in a wide range of buffers, including reducing or oxidizing conditions (Zakeri et al., 2012).

Next, we developed an orthogonal pair to SpyTag/SpyCatcher. The D4 domain of RrgA adhesin from *Streptococcus pneumoniae* was found to have

ARTICLE IN PRESS

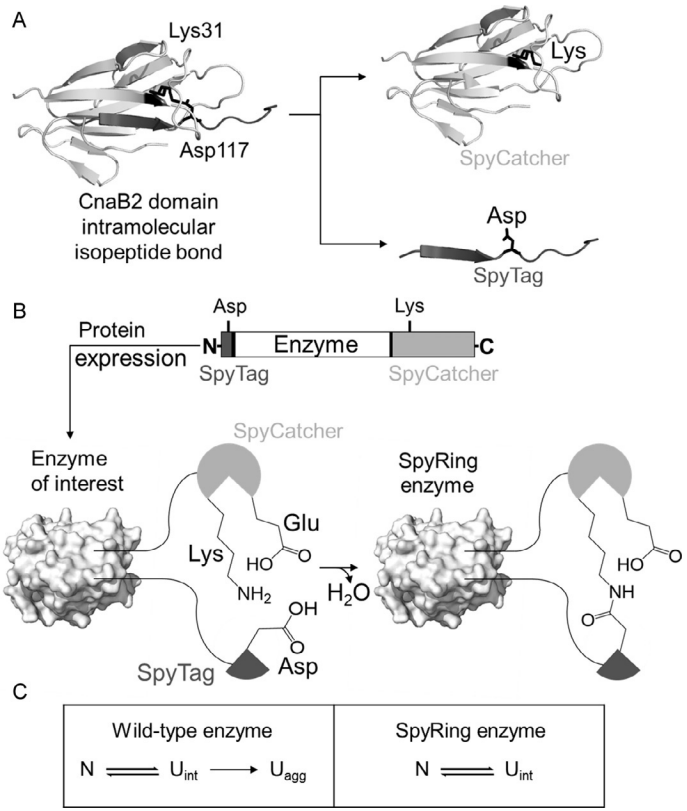


Fig. 1 SpyTag and SpyCatcher for enzyme cyclization. (A) The CnaB2 domain of FbaB was split into SpyTag and SpyCatcher. Cartoon based on PDB 2X5P (Oke et al., 2010). (B) Expression of the SpyRing genetic construct leads to in vivo cyclization. The formation of the isopeptide bond is dependent on the three illustrated residues: reaction between Lys and Asp promoted by an apposed Glu. (C) Proposed principle of SpyRing-mediated thermal resilience, by favoring reversible transition between the native (N) and unfolded intermediate states (U_{int}), avoiding irreversible transition to the unfolded aggregate state (U_{agg}).

an isopeptide bond between lysine and asparagine side chains (Izoré et al., 2010). The N-terminal β-strand containing the reactive lysine residue (residues 734–745) was split and named SnoopTag. Residues 749–860 provided SnoopCatcher, which was further optimized by mutating two residues to stabilize secondary structure. SnoopTag/SnoopCatcher and SpyTag/SpyCatcher were found to be mutually unreactive. The combined use of both Tag/Catcher systems made it possible to build sequence-programmed protein chains (Veggiani et al., 2016).

ARTICLE IN PRESS

2.2 Cyclization with Isopeptide Bonds

Our initial studies with these peptide/protein ligation tools all focused on opportunities for locking together different protein units (Fairhead et al., 2014; Fierer, Veggiani, & Howarth, 2014; Veggiani et al., 2016). However, the efficient reaction and flexible tag location suggested to us the opportunity to apply these pairs to drive intramolecular reaction. SpyTag/SpyCatcher (SpyRing, Fig. 1B), Pilin-C/Isopeptag (PilinRing), and SnoopTag/SnoopCatcher (SnoopRing) were all shown to cyclize β-lactamase (BLA) efficiently (Schoene et al., 2016). All these cyclized constructs exhibited majorly enhanced thermal resilience, in comparison to the untagged control. We placed Pilin-C, SpyTag, and SnoopTag on the N-terminus and Isopeptag, SpyCatcher, and SnoopCatcher on the C-terminus of BLA but we have not verified whether flipping the components around might work better. Therefore, further work could be carried out on testing whether the Tag should be on the N or C-terminus and vice versa for the Catcher. However, we found SpyRing cyclization to give the greatest enhancement in resilience (Schoene et al., 2016), so this approach will be the focus of this chapter. Such comparisons, along with testing linear controls, show how SpyRing resilience exceeds that simply from cyclization or fusion to a thermostable domain (Schoene et al., 2016, 2014). We proposed a mechanism whereby SpyRing cyclization confers resilience through favoring enzyme refolding after denaturation, blocking irreversible transitions to an unfolded aggregate (Fig. 1C) (Schoene et al., 2016, 2014; Schumann et al., 2015).

We describe the selection of appropriate enzyme targets for cyclization, their insertion into the SpyRing construct, and the methods used to verify successful cyclization. Moreover, we focus on the cyclization of a β-propeller phytase (PhyC) and the enzymatic, solubility, and biophysical characterization approaches one might use to study SpyRing-cyclized proteins.

2.3 Selection of Enzymes for Cyclization

SpyRing cyclization was envisioned to provide a simple method to stabilize different enzymes. In 2 weeks, one can go from obtaining the DNA for an enzyme of interest to testing the SpyRing enzyme's thermal resilience. However, we postulate that the following criteria are helpful when prioritizing enzyme candidates for SpyRing cyclization.

If a structure of the enzyme (or a close ortholog) is available, it is helpful to consider:

- the distance between the N- and C-termini,
- whether the active site is on the same face as the termini.

ARTICLE IN PRESS

We suggest that enzymes with a short distance (ie, <15 Å) between the N- and C-termini and where the active site is not closely apposed to the termini would be ideal candidates for SpyRing cyclization. In such cases, our experience is that SpyRing cyclization gave little or no change in enzyme activity. For BLA (7 Å from N- to C-terminus), we showed minimal difference in k_{cat} or K_m (Schoene et al., 2016) between wild-type and SpyRing forms. The distribution of termini in the Protein Data Bank has been surveyed (Krishna & Englander, 2005) and a large fraction of single-domain proteins should be amenable based on this distance criterion. However, even where the termini are 29 Å apart as in PhyC, we obtained efficient SpyRing cyclization and increased thermal resilience (Schoene et al., 2016).

Previous successes with other cyclization approaches, as well as with circular permutations or disulfide bond engineering, might also give confidence that SpyRing cyclization would work. Intein-mediated cyclization of BLA led to an increase in aggregation resistance of ~5°C (Iwai & Plückthun, 1999). Dihydrofolate reductase (DHFR) had previously been stabilized via cyanocysteine-mediated cyclization (Takahashi et al., 2007). We were able to show that both DHFR and BLA were rendered resilient by SpyRing cyclization (Schoene et al., 2014). However, if no data on cyclization are available, it is worth searching the literature for whether the enzyme remains active when fused to another protein, or whether a fusion to a thermostable protein increases the thermal resilience (Pierre, Xiong, Hayles, Guntaka, & Kim, 2011).

SpyRing cyclization is unlikely to substitute for the activity of chaperones. If a protein requires chaperones to fold or refold, like RuBisCO (Liu et al., 2010) or firefly luciferase (FLuc) (Svetlov, Kommer, Kolb, & Spirin, 2006), SpyRing cyclization might not be the right approach. Enzyme cofactors might also cause complications, especially if an enzyme cannot refold correctly following the loss of a cofactor, eg, glucose oxidase (Zoldák, Zubrik, Musatov, Stupák, & Sedlák, 2004).

So far we have worked primarily with monomeric enzymes. Generic strategies for stabilizing multimeric proteins will require further study.

2.4 Cloning Genes of Interest into the SpyRing Cassette

SpyRing cyclization is achieved by genetically encoding the SpyTag on the N-terminus and the SpyCatcher on the C-terminus of an enzyme of interest. The expressed construct is then able to spontaneously form an isopeptide bond inside the expressing cell (Fig. 1B). Therefore, SpyRing cyclization

ARTICLE IN PRESS

requires the insertion of the open-reading frame of the enzyme within the SpyRing expression cassette.

To achieve a quick, reliable, and modular approach to insert different enzymes into the SpyRing cassette, we designed a ligation-independent cloning strategy (Fig. 2). The cloning method requires two custom primers which can anneal to the coding sequence of the enzyme. The custom primers have tail-regions which can hybridize to the vector. The 5′ tail-region is 27 bp long and the 3′ tail-region is 23 bp long and both regions overlap with the vector sequence. This approach produces a glycine/serine spacer (GSGGSG) on both sides of the enzyme, to facilitate efficient cyclization by SpyTag/SpyCatcher and to minimize interference with enzyme folding. The template from which to amplify the vector sequence is readily available from the Addgene plasmid repository (ID: 52656). Once the insert and vector have been amplified and purified, the fragments can be linked together by circular polymerase extension cloning (CPEC) (Quan & Tian, 2009). We use KOD Hot Start DNA Polymerase (Roche) in the CPEC reaction. 500 ng of vector are mixed with an equimolar amount of insert in a 50 μL reaction. Cycling conditions are: 95°C for 3 min followed by 10 cycles of 95°C for 30 s, 64°C for 30 s, and 68°C for 4 min. 1 μL of the CPEC reaction can be used to transform competent *Escherichia coli* cells (a $RecA^-$ strain such as XL-1 or DH5α).

We recommend CPEC because the method does not require any additional materials not already used to amplify the vector and insert. However, methods such as Gibson assembly (Gibson et al., 2009) should also be suitable.

2.5 Confirming Successful Cyclization

The cloned construct can then be expressed in *E. coli* protein production strains. We have obtained good yields using the BL21 DE3 RIPL strains from Agilent, which enables efficient translation even if the gene contains rare codons (Schoene et al., 2016, 2014). Two important parameters to consider are the temperature at which the constructs are expressed and the concentration of isopropyl β-D-1-thiogalactopyranoside (IPTG). There are two types of covalent interactions that can occur once the SpyRing protein has been expressed: intramolecular covalent interactions leading to enzyme cyclization and intermolecular covalent interactions leading to enzyme oligomerization (Fig. 3A). By decreasing the induction temperature and IPTG concentration, one can increase the proportion of monomeric-cyclized

ARTICLE IN PRESS

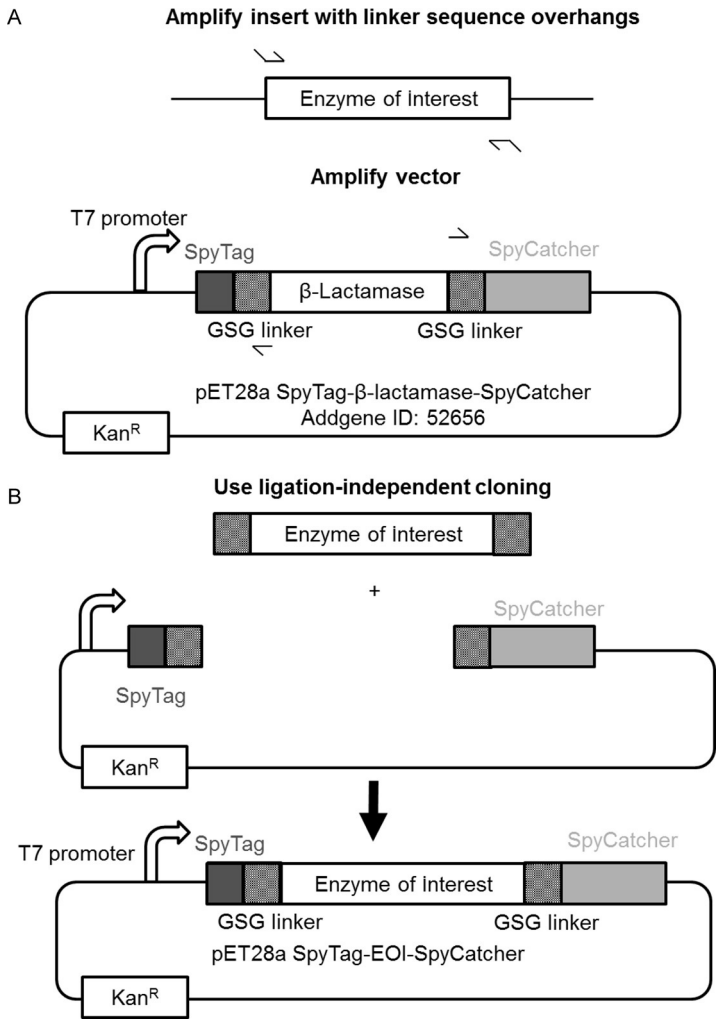


Fig. 2 Cloning strategy to insert an enzyme into the SpyRing expression cassette. (A) The enzyme of interest is amplified by PCR using primers containing adaptor sequences, to generate an overlap at each end with the vector. The vector is amplified using pET28a SpyTag-BLA-SpyCatcher as a template, with the primers hybridizing to linker regions of the construct. (B) The vector and insert are then combined through a ligation-independent cloning method, such as CPEC, to generate the desired SpyRing expression plasmid.

protein (Zhang, Sun, Tirrell, & Arnold, 2013). We found that growing the culture to an OD_{600} of ~0.5 and then inducing protein expression at 18°C for 16 h using 0.4 m*M* IPTG worked well to obtain a good yield of monomeric-cyclized product.

ARTICLE IN PRESS

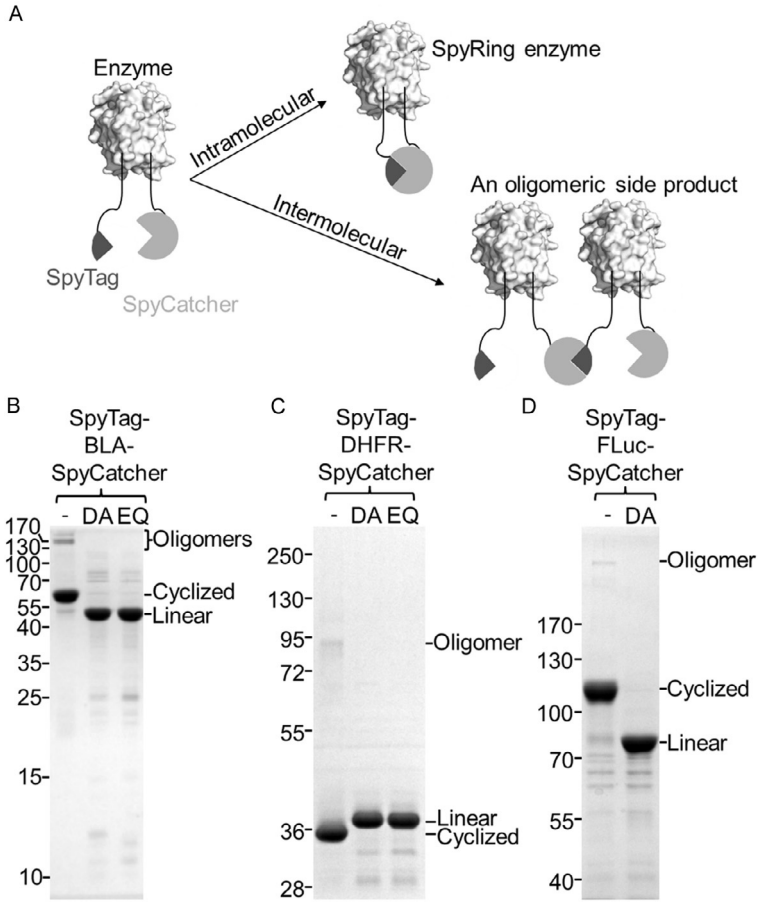


Fig. 3 Gel analysis of SpyRing enzyme cyclization. (A) Cartoon illustrating how fusing SpyTag on the N-terminus and SpyCatcher at the C-terminus can lead to intermolecular or intramolecular covalent bond formation. (B) SpyRing cyclization decreased the mobility of BLA. Reducing SDS–PAGE with Coomassie staining, showing SpyTag-BLA-SpyCatcher compared to the noncyclized controls of SpyTag DA-BLA-SpyCatcher and SpyTag-BLA-SpyCatcher EQ. "Oligomers" indicates the side products from intermolecular SpyTag/SpyCatcher reaction. (C) SpyRing cyclization increased the mobility of DHFR. Reducing SDS–PAGE with Coomassie staining for SpyTag-DHFR-SpyCatcher and DA or EQ linear controls. (D) SpyRing cyclization decreased the mobility of FLuc. Reducing SDS–PAGE with Coomassie staining for SpyTag-FLuc-SpyCatcher and the DA control. *Panels (B) and (C): Data previously published in Schoene, C., Fierer, J. O., Bennett, S. P., & Howarth, M. (2014). SpyTag/SpyCatcher cyclization confers resilience to boiling on a mesophilic enzyme.* Angewandte Chemie*, 53, 6101–6104.*

ARTICLE IN PRESS

The SpyRing cassette has an N-terminal His_6 tag, which can be used to purify the cyclized proteins via immobilized metal ion affinity chromatography (IMAC). To avoid eluting most of the oligomeric protein, we recommend using low concentrations of imidazole (ie, 75 m*M*) to favor elution of proteins that have only a single His_6 tag (Howarth & Ting, 2008). If desired, complete removal of oligomeric side products can be achieved by a subsequent size-exclusion chromatography step (Schoene et al., 2016).

An alternative method for the purification of the SpyRing enzyme is to use heat to precipitate *E. coli*'s native proteins, while the SpyRing enzyme largely remains soluble. The SpyRing enzyme in the soluble fraction can be separated from the precipitate using centrifugation (Schoene et al., 2016). Heat-mediated purification is an attractive option because it is rapid and cheap. An important consideration is the total concentration of protein in the cell lysate. If the concentration is too high, the cyclized protein may be lost to the precipitate as well, through interaction with other unfolded proteins. A cell lysate OD_{280} of ~6 is recommended. Furthermore, we found that the presence of NaCl can help improve the solubility of the cyclized protein (Schoene et al., 2016). Heat-mediated purification has also been achieved previously by using thermostable proteins as fusion partners (de Marco, Casatta, Savaresi, & Geerlof, 2004).

A first quick test for cyclization is to perform SDS–PAGE with Coomassie staining on the SpyRing enzyme, looking for a mobility shift compared to a linear point mutant unable to form an isopeptide bond (SpyTag DA or SpyCatcher EQ) (Zakeri et al., 2012). Complicated effects on mobility are seen from cyclization, depending on enzyme size and charge (Fig. 3B–D). Moreover, the acrylamide percentage of the SDS–PAGE can alter the difference in mobility between the cyclized enzyme and the non-cyclized point mutant. The change in mobility resulting from cyclization is much larger than expected simply from the effect of the mutation on protein charge, reflecting the complex interaction between the conformation of the unfolded protein coated in SDS and the polyacrylamide network (Rath, Cunningham, & Deber, 2013).

For a more definitive method to confirm isopeptide bond formation, electrospray ionization mass spectrometry (ESI-MS) can be employed to determine the exact mass of the protein constructs (Schoene et al., 2016, 2014). Since the formation of the isopeptide bond between SpyTag and SpyCatcher leads to the loss of a water molecule (Fig. 1B), the mass determined by ESI-MS should be 18 Da lighter than the predicted molecular weight of the linear protein from the ProtParam online tool

ARTICLE IN PRESS

(Gasteiger et al., 2005). A possible complication in the spectrum from *E. coli* expression is gluconoylation of the His_6 tag (Geoghegan et al., 1999). For ESI-MS, samples should be expressed in *E. coli* B834 DE3, grown in the absence of any supplementary glucose in the LB media to an OD_{600} of ~0.5, and induced at 30°C for 3 h using 0.4 m*M* IPTG. Samples should then be purified using the previously described IMAC approach, before dialyzing into 10 m*M* ammonium acetate buffer (a low ionic strength buffer with volatile components is important to obtain clean spectra).

3. CHARACTERIZATION OF ENZYME RESILIENCE AFTER SpyRing CYCLIZATION

3.1 Industrial Relevance of Phytase

We illustrate the process of characterizing enzyme resilience with particular reference to SpyRing cyclization of phytase. Most plants store phosphates as phytic acid (inositol hexakisphosphate). Nonruminant animals such as chicken, pigs, and salmon are unable to degrade phytic acid, and therefore find it difficult to absorb the necessary dietary phosphate for a healthy diet. Moreover, phytic acid is an antinutrient, which is able to chelate and thereby block absorption of micronutrients like zinc and iron (Bohlke, Thaler, & Stein, 2005; Selle, Ravindran, Bryden, & Scott, 2006). Therefore, the enzyme phytase is widely used to release phosphate from phytic acid in animal feed. However, during the pelleting process, feedstock is heated to 75–90°C to kill pathogenic bacteria, notably *Salmonella* (Lei et al., 2013). A phytase variant able to regain full activity after this pelleting process would be a significant advance.

Humans are also unable to degrade phytic acid. This antinutrient may contribute to the global health problem of micronutrient deficiency (Bailey, West, & Black, 2015), the "silent hunger" with a range of short-term and long-term health consequences (Allen, de Benoist, Dary, & Hurrell, 2006). Food security is a major global issue; increasing the nutritional value of food that is already available will be an important part of the solution.

3.2 Measuring Aggregation Can Provide Evidence for an Increase in Thermal Resilience

The first simple and quick test we use to evaluate a SpyRing enzyme is to measure aggregation at elevated temperature. Heat-induced aggregation correlates with a loss of activity and irreversible inactivation of enzymes.

ARTICLE IN PRESS

To test whether SpyRing cyclization has improved an enzyme's thermal resilience, we recommend that a SpyTag D → A linear point mutant as well as a construct lacking both SpyTag and SpyCatcher are used as controls. The test should be run in the wild-type enzyme's preferred buffer. Enzyme concentration can be critical to the extent of aggregation. Therefore, we recommend using a concentration relevant to the application one has in mind for the enzyme of interest. 10 μ*M* enzyme is a good starting point.

We recommend the use of a PCR machine to heat and then cool the protein for a precise amount of time with a reproducible ramp rate. For example, heat the proteins at 25°C, 37°C, 55°C, 75°C, 90°C, and 100°C for 10 min and then cool to 10°C for 1 h. After cooling for 1 h, centrifuge the samples for 30 min at 17,000 × *g*, before analyzing the soluble fraction by SDS–PAGE. Coomassie-stained bands can then be quantified using densitometry.

In the case of PhyC, SpyRing cyclization improved the solubility of the enzyme at temperatures at and above 75°C (Fig. 4A). However, an aggregation assay alone should not be used as conclusive evidence on whether SpyRing cyclization is helpful.

3.3 Recovered Activity Is the Most Important Measurement When Testing Thermal Resilience

Aggregation can be used as an initial indicator to determine whether SpyRing cyclization increases the thermal resilience of an enzyme. However, the aggregation assay does not take into account the formation of soluble aggregates and the possibility that an enzyme does not aggregate but misfolds and becomes irreversibly inactivated. For example, the noncyclized control SpyTag DA-DHFR-SpyCatcher did not aggregate at 100°C but the activity assay showed that the enzyme was irreversibly inactivated at this temperature (Schoene et al., 2014). Therefore, the best assay to use is one that measures the residual enzyme activity after heating.

The enzyme's temperature treatment should be the same as described earlier, so a correlation can be drawn between the enzyme's recovered activity and solubility. However, since the enzyme will be at a high concentration (∼10 μ*M*) for the solubility assay, one might need to dilute the enzyme down to n*M* concentrations for the activity assay. Finding the right concentration of enzyme for your assay is important, so an enzyme titration to where the initial activity can be measured using at least five time points is a good starting place. When diluting the enzyme, reagent stability will need to be taken into account. Enzymes at low concentrations will readily stick to

ARTICLE IN PRESS

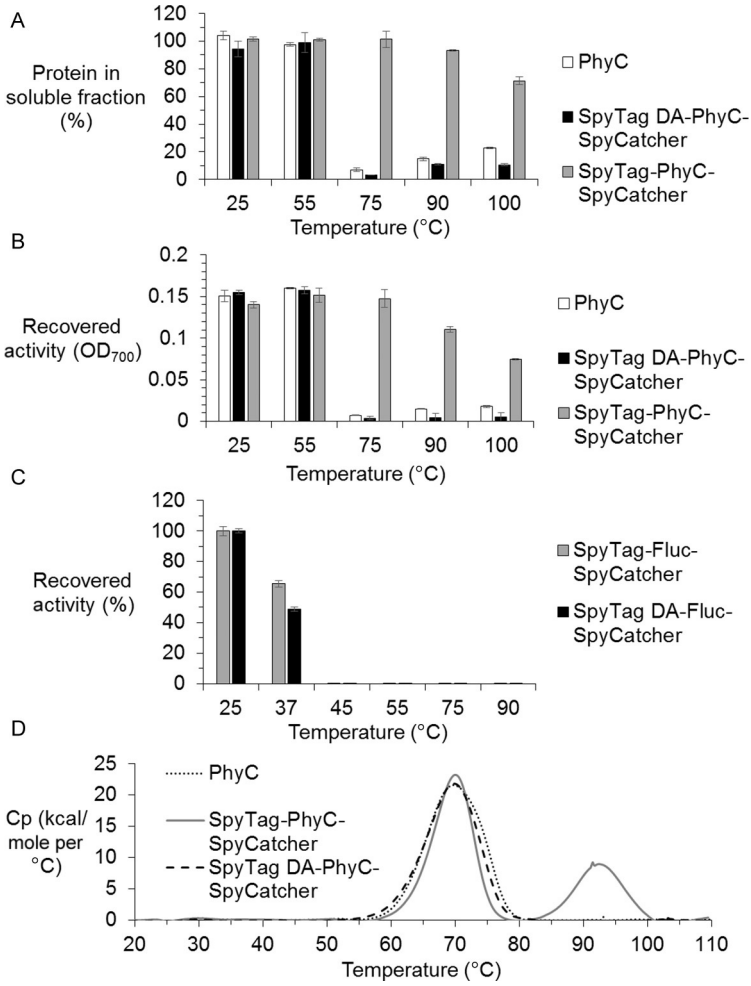


Fig. 4 SpyRing cyclization effect on thermal resilience. (A) SpyRing cyclization improved the solubility of PhyC at high temperatures. Samples were heated to the indicated temperature for 10 min, cooled to 10°C, and the supernatant was analyzed by SDS–PAGE and Coomassie staining (mean of triplicates ± 1 s.d.). (B) Cyclization improved the recovered activity of PhyC. Samples were heated to the indicated temperature for 10 min and cooled to 10°C. Phosphate release from phytase was quantified colorimetrically (mean of triplicate ± 1 s.d.). (C) Cyclization had a minor effect on FLuc activity recovered after heating. Samples at 5 μ*M* in PBS, pH 7.4, containing 100 m*M* DTT were heated to the indicated temperature for 10 min and cooled to 10°C. 12.5 n*M* enzyme was reacted with 300 μ*M* luciferin in 50 m*M* Tris-acetate, pH 7.8, 10 m*M* $MgSO_4$, 4 m*M* ATP, and 2 m*M* EDTA. Luminescence was measured after a 15 s delay (mean of triplicates ± 1 s.d.). (D) DSC of SpyTag-PhyC-SpyCatcher overlaid with the untagged PhyC construct and the DA point mutant. Scanning from 20°C to110°C at 1°C/min. *Panels (A) and (D): Data previously published in Schoene, C., Bennett, S. P., & Howarth, M. (2016). SpyRing interrogation: Analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries.* Scientific Reports*, 6, 21151.*

ARTICLE IN PRESS

plastic surfaces and be removed from solution. Therefore, plastic containers should be preblocked with 1–3% bovine serum albumin (BSA) at 37°C for 2 h and the dilution buffer should include BSA at a concentration of 0.1–1%.

We recommend that the activity recovery assay is set up in 96-well plate format to allow for high throughput. Ideally, an enzyme substrate should be used with an absorbance or fluorescence readout. For plate-reader measurement, the delay time between each read should be considered and how that might affect the reading of the initial catalytic activity.

In the case of PhyC, there are no fluorescent or colorimetric analogs of phytic acid currently available, so reactions were quenched with trichloroacetic acid and the released phosphate was measured colorimetrically using the ammonium molybdate method (Schoene et al., 2016). Cyclization provided a large improvement in recovered activity of PhyC after heating to the industrially relevant temperatures of 75°C and 90°C (Fig. 4B).

3.4 FLuc Is an Example of a Poor Target for SpyRing Cyclization

FLuc is an important reporter enzyme used to study gene expression, monitor protein–protein interactions, and sequence DNA. However, FLuc's instability over 30°C is a problem for various applications (Ebrahimi, Hosseinkhani, Heydari, Khavari-Nejad, & Akbari, 2012). SpyRing cyclization of FLuc decreased the mobility on SDS–PAGE, compared to the linear DA control (Fig. 3D), and therefore we postulated that the cyclization reaction had been successful. However, when we tested the recovered activity of the cyclized FLuc, improvements were only marginal when compared to the linear point mutant (Fig. 4C).

The following tips might help to avoid testing cyclization on poor targets. A circular permutation of FLuc in combination with intein-mediated cyclization has been shown to inactivate the enzyme (Kanno, Yamanaka, Hirano, Umezawa, & Ozawa, 2007). Cyclization of a circularly permuted FLuc was used to create a reporter for caspase-3, by placing a caspase cleavage site between the N- and C-termini. Upon cleavage the termini are able to move and the enzyme becomes active (Kanno et al., 2007). While our cyclized FLuc was still active (Fig. 4C), enzymes which have been shown to be less active when the termini are restrained will probably be hard to work with.

FLuc has also been shown to require the help of chaperones when refolding from a heat-denatured state (Schröder, Langer, Hartl, & Bukau, 1993). Moreover, misfolding involves the formation of a stable nonnative

ARTICLE IN PRESS

conformation where the N- and C-terminal residues are interacting (Scholl, Yang, & Marszalek, 2014). Therefore, FLuc SpyRing cyclization might not be able to stop this interaction from occurring unless chaperones are present.

3.5 Dynamic Scanning Calorimetry Is a Useful Biophysical Tool to Study SpyRing Cyclization

Dynamic scanning calorimetry (DSC) has provided an effective approach to follow the thermal unfolding and refolding of SpyRing-cyclized proteins (Schoene et al., 2014). DSC measures the difference in heat required to increase the temperature of a protein sample in comparison to a reference cell. When proteins unfold, the hydration of the core causes an increase in heat capacity (Privalov & Makhatadze, 1992). Preparation of samples for DSC is simple: no dyes are necessary since the signal monitored is the change in specific heat capacity. We have used the VP Cap DSC system from GE Healthcare, which requires 400 μL of 20 μ*M* protein and ~4 mL of buffer for a single run.

The ramp rate for a single scan from 20°C to 110°C was set to 1°C/min, which produced high-resolution data (Fig. 4D). DSC data indicated that the SpyRing did not increase the transition temperature of PhyC (PhyC: 69.8°C, SpyTag DA-PhyC-SpyCatcher: 69.7°C, SpyTag-PhyC-SpyCatcher: 70.0°C). SpyTag-PhyC-SpyCatcher unfolded as two transitions. The second transition had a T_m of 91.5°C, which is close to the T_m we measured for SpyCatcher + SpyTag peptide melting (Schoene et al., 2016).

Such data are consistent with the mechanism proposed for the increase in thermal resilience observed in PhyC, DHFR, and BLA, where SpyRing cyclization does not block unfolding but promotes efficient refolding (Fig. 1C). The increase in thermal resilience of the cyclized enzyme also extended higher than the T_m of the SpyTag/SpyCatcher domain (Schoene et al., 2016, 2014).

A downside of using DSC is that it is low throughput. Even automated DSC systems are unable to process more than 50 samples/day (Plotnikov et al., 2002). Dynamic scanning fluorimetry (DSF) is able to provide a much higher throughput (Niesen, Berglund, & Vedadi, 2007). However, DSF requires the use of a dye which can interfere with the transition temperature of a protein (Shi, Semple, Cheung, & Shameem, 2013). Alternatively, microscale thermophoresis is a recently developed high-throughput method, which can monitor unfolding from the intrinsic protein fluorescence (Alexander et al., 2014).

ARTICLE IN PRESS

4. CONCLUDING REMARKS

We have developed three distinct peptide–protein pairs (Veggiani et al., 2016; Zakeri et al., 2012; Zakeri & Howarth, 2010) which are able to cyclize inside cells and increase enzyme resilience (Schoene et al., 2016). SpyRing cyclization requires the genetic encoding of a 13 amino acid tag, SpyTag, at the N-terminus of a protein and a 12 kDa protein, SpyCatcher, on the C-terminus. Cloning of an enzyme of interest into the SpyRing cassette is simple and modular. SpyRing enzymes can easily be purified through IMAC or heat purification. SpyRing cyclization may provide a quick and generic route toward enzyme stabilization. We have provided guidelines for the selection of appropriate targets for SpyRing cyclization but exploring this strategy on a wider diversity of enzymes (expression hosts, protein folds, multimerization state, and enzyme classes) will help to advance the scope of this approach.

ACKNOWLEDGMENTS

Funding was provided by BBSRC (C.S.), Sekisui Diagnostics (C.S., S.P.B.), and Oxford University Department of Biochemistry (M.H.).

REFERENCES

Alexander, C. G., Wanner, R., Johnson, C. M., Breitsprecher, D., Winter, G., Duhr, S., et al. (2014). Novel microscale approaches for easy, rapid determination of protein stability in academic and commercial settings. *Biochimica et Biophysica Acta (BBA)—Proteins and Proteomics*, *1844*, 2241–2250.

Allen, L., de Benoist, B., Dary, O., & Hurrell, R. (2006). *Guidelines on food fortification with micronutrients*. Geneva, Switzerland: World Health Organization.

Antos, J. M., Popp, M. W.-L., Ernst, R., Chew, G.-L., Spooner, E., & Ploegh, H. L. (2009). A straight path to circular proteins. *The Journal of Biological Chemistry*, *284*, 16028–16036.

Bailey, R. L., West, K. P., & Black, R. E. (2015). The epidemiology of global micronutrient deficiencies. *Annals of Nutrition & Metabolism*, *66*(Suppl. 2), 22–33.

Bohlke, R. A., Thaler, R. C., & Stein, H. H. (2005). Calcium, phosphorus, and amino acid digestibility in low-phytate corn, normal corn, and soybean meal by growing pigs. *Journal of Animal Science*, *83*, 2396–2403.

Camarero, J. A., & Muir, T. W. (1999). Biosynthesis of a head-to-tail cyclized protein with improved biological activity. *Journal of the American Chemical Society*, *121*, 5597–5598.

de Marco, A., Casatta, E., Savaresi, S., & Geerlof, A. (2004). Recombinant proteins fused to thermostable partners can be purified by heat incubation. *Journal of Biotechnology*, *107*, 125–133.

Ebrahimi, M., Hosseinkhani, S., Heydari, A., Khavari-Nejad, R. A., & Akbari, J. (2012). Improvement of thermostability and activity of firefly luciferase through [TMG][Ac] ionic liquid mediator. *Applied Biochemistry and Biotechnology*, *168*, 604–615.

ARTICLE IN PRESS

Evans, T. C., Jr., Benner, J., & Xu, M. Q. (1999). The in vitro ligation of bacterially expressed proteins using an intein from *Methanobacterium thermoautotrophicum*. *The Journal of Biological Chemistry*, *274*, 3923–3926.

Fairhead, M., Veggiani, G., Lever, M., Yan, J., Mesner, D., Robinson, C. V., et al. (2014). SpyAvidin hubs enable precise and ultrastable orthogonal nanoassembly. *Journal of the American Chemical Society*, *136*, 12355–12363.

Fierer, J. O., Veggiani, G., & Howarth, M. (2014). SpyLigase peptide-peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. *Proceedings of the National Academy of Sciences*, *111*, E1176–E1181.

Gasteiger, E., Hoogland, C., Gatticker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., et al. (2005). Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook* (pp. 571–607). New York City, NY: Humana Press.

Geoghegan, K. F., Dixon, H. B., Rosner, P. J., Hoth, L. R., Lanzetti, A. J., Borzilleri, K. A., et al. (1999). Spontaneous alpha-N-6-phosphogluconoylation of a 'His tag' in *Escherichia coli*: The cause of extra mass of 258 or 178 Da in fusion proteins. *Analytical Biochemistry*, *267*, 169–184.

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*, 343–345.

Goldenberg, D. P., & Creighton, T. E. (1983). Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *Journal of Molecular Biology*, *165*, 407–413.

Hagan, R. M., Bjornsson, R., McMahon, S. A., Schomburg, B., Braithwaite, V., Buhl, M., et al. (2010). NMR spectroscopic and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond. *Angewandte Chemie*, *49*, 8421–8425.

Howard, R. I., Abotsi, E., Jansen van Rensburg, E. I., & Howard, S. (2003). Lignocellulose biotechnology: Issues of bioconversion and enzyme production. *African Journal of Biotechnology*, *2*, 602–619.

Howarth, M., & Ting, A. Y. (2008). Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin. *Nature Protocols*, *3*, 534–545.

Iwai, H., & Plückthun, A. (1999). Circular β-lactamase: Stability enhancement by cyclizing the backbone. *FEBS Letters*, *459*, 166–172.

Izoré, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., et al. (2010). Structural basis of host cell recognition by the pilus adhesin from *Streptococcus pneumoniae*. *Structure*, *18*, 106–115.

Kang, H. J., Coulibaly, F., Clow, F., Proft, T., & Baker, E. N. (2007). Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. *Science*, *318*, 1625–1628.

Kanno, A., Yamanaka, Y., Hirano, H., Umezawa, Y., & Ozawa, T. (2007). Cyclic luciferase for real-time sensing of caspase-3 activities in living mammals. *Angewandte Chemie*, *46*, 7595–7599.

Krishna, M. M. G., & Englander, S. W. (2005). The N-terminal to C-terminal motif in protein folding and function. *Proceedings of the National Academy of Sciences*, *102*, 1053–1058.

Lei, X. G., Weaver, J. D., Mullaney, E., Ullah, A. H., & Azain, M. J. (2013). Phytase, a new life for an 'old' enzyme. *Annual Review of Animal Biosciences*, *1*, 283–309.

Li, S., Yang, X., Yang, S., Zhu, M., & Wang, X. (2012). Technology prospecting on enzymes: Application, marketing and engineering. *Computational and Structural Biotechnology Journal*, *2*, 1–11.

Liu, C., Young, A. L., Starling-Windhof, A., Bracher, A., Saschenbrecker, S., Rao, B. V., et al. (2010). Coupled chaperone action in folding and assembly of hexadecameric Rubisco. *Nature*, *463*, 197–202.

Niesen, F. H., Berglund, H., & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*, *2*, 2212–2221.

ARTICLE IN PRESS

Oke, M., Carter, L. G., Johnson, K. A., Liu, H., McMahon, S. A., Yan, X., et al. (2010). The Scottish structural proteomics facility: Targets, methods and outputs. *Journal of Structural and Functional Genomics*, *11*, 167–180.

Packer, M. S., & Liu, D. R. (2015). Methods for the directed evolution of proteins. *Nature Reviews. Genetics*, *16*, 379–394.

Parthasarathy, R., Subramanian, S., & Boder, E. T. (2007). Sortase A as a novel molecular 'stapler' for sequence-specific protein conjugation. *Bioconjugate Chemistry*, *18*, 469–476.

Pierre, B., Xiong, T., Hayles, L., Guntaka, V. R., & Kim, J. R. (2011). Stability of a guest protein depends on stability of a host protein in insertional fusion. *Biotechnology and Bioengineering*, *108*, 1011–1020.

Plotnikov, V., Rochalski, A., Brandts, M., Brandts, J. F., Williston, S., Frasca, V., et al. (2002). An autosampling differential scanning calorimeter instrument for studying molecular interactions. *Assay and Drug Development Technologies*, *1*, 83–90.

Popp, M. W., Dougan, S. K., Chuang, T.-Y., Spooner, E., & Ploegh, H. L. (2011). Sortase-catalyzed transformations that improve the properties of cytokines. *Proceedings of the National Academy of Sciences*, *108*, 3169–3174.

Privalov, P. L., & Makhatadze, G. I. (1992). Contribution of hydration and non-covalent interactions to the heat capacity effect on protein unfolding. *Journal of Molecular Biology*, *224*, 715–723.

Quan, J., & Tian, J. (2009). Circular polymerase extension cloning of complex gene libraries and pathways. *PloS One*, *4*, e6441.

Rath, A., Cunningham, F., & Deber, C. M. (2013). Acrylamide concentration determines the direction and magnitude of helical membrane protein gel shifts. *Proceedings of the National Academy of Sciences*, *110*, 15668–15673.

Reddington, S. C., & Howarth, M. (2015). Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher. *Current Opinion in Chemical Biology*, *29*, 94–99.

Schoene, C., Bennett, S. P., & Howarth, M. (2016). SpyRing interrogation: Analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries. *Scientific Reports*, *6*, 21151.

Schoene, C., Fierer, J. O., Bennett, S. P., & Howarth, M. (2014). SpyTag/SpyCatcher cyclization confers resilience to boiling on a mesophilic enzyme. *Angewandte Chemie*, *53*, 6101–6104.

Scholl, Z. N., Yang, W., & Marszalek, P. E. (2014). Chaperones rescue luciferase folding by separating its domains. *Journal of Biological Chemistry*, *289*, 28607–28618.

Schröder, H., Langer, T., Hartl, F. U., & Bukau, B. (1993). DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *The EMBO Journal*, *12*, 4137–4144.

Schumann, F. H., Varadan, R., Tayakuniyil, P. P., Grossman, J. H., Camarero, J. A., & Fushman, D. (2015). Changing the topology of protein backbone: The effect of backbone cyclization on the structure and dynamics of a SH3 domain. *Frontiers in Chemistry*, *3*, 26.

Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C., & Benkovic, S. J. (1999). Production of cyclic peptides and proteins in vivo. *Proceedings of the National Academy of Sciences*, *96*, 13638–13643.

Selle, P. H., Ravindran, V., Bryden, W. L., & Scott, T. (2006). Influence of dietary phytate and exogenous phytase on amino acid digestibility in poultry: A review. *The Journal of Poultry Science*, *43*, 89–103.

Shi, S., Semple, A., Cheung, J., & Shameem, M. (2013). DSF method optimization and its application in predicting protein thermal aggregation kinetics. *Journal of Pharmaceutical Sciences*, *102*, 2471–2483.

ARTICLE IN PRESS

Svetlov, M. S., Kommer, A., Kolb, V. A., & Spirin, A. S. (2006). Effective cotranslational folding of firefly luciferase without chaperones of the Hsp70 family. *Protein Science*, *15*, 242–247.

Takahashi, H., Arai, M., Takenawa, T., Sota, H., Xie, Q. H., & Iwakura, M. (2007). Stabilization of hyperactive dihydrofolate reductase by cyanocysteine-mediated backbone cyclization. *Journal of Biological Chemistry*, *282*, 9420–9429.

van Lieshout, J. F. T., Pérez Gutiérrez, O. N., Vroom, W., Planas, A., de Vos, W. M., van der Oost, J., et al. (2012). Thermal stabilization of an endoglucanase by cyclization. *Applied Biochemistry and Biotechnology*, *167*, 2039–2053.

Veggiani, G., Nakamura, T., Brenner, M. D., Gayet, R. V., Yan, J., Robinson, C. V., et al. (2016). Programmable polyproteams built using twin peptide superglues. *Proceedings of the National Academy of Sciences*, *113*, 1202–1207.

Yang, H., Liu, L., Li, J., Chen, J., & Du, G. (2015). Rational design to improve protein thermostability: Recent advances and prospects. *ChemBioEng Reviews*, *2*, 87–94.

Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., et al. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proceedings of the National Academy of Sciences*, *109*, E690–E697.

Zakeri, B., & Howarth, M. (2010). Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting. *Journal of the American Chemical Society*, *132*, 4526–4527.

Zhang, W.-B., Sun, F., Tirrell, D. A., & Arnold, F. H. (2013). Controlling macromolecular topology with genetically encoded SpyTag–SpyCatcher chemistry. *Journal of the American Chemical Society*, *135*, 13988–13997.

Zoldák, G., Zubrik, A., Musatov, A., Stupák, M., & Sedlák, E. (2004). Irreversible thermal denaturation of glucose oxidase from *Aspergillus niger* is the transition to the denatured state with residual structure. *Journal of Biological Chemistry*, *279*, 47601–47609.