Synthetic biology biosensor design for medical diagnostics

Nicolas Kylilis

Imperial College London

Department of Medicine

Doctor of Philosophy (PhD)

Abstract

Biosensors are compact, integrated analytical devices that use a biologically derived sensor element to specifically detect their target analyte. Their design makes them convenient for the performance of diagnostic tests at the site where the sample is procured. Synthetic biology has been extensively applied to biosensor design to produce genetically encoded devices that detect metal and organic contaminants, drug molecules and metabolites among others. However, there is only a very limited number of synthetic biology biosensor tools available for the detection of analytes that do not readily transport across the cell membrane. This is especially the case for protein analytes despite the potential of such tools for use in diagnostic applications at the point of care.

This research investigated synthetic biology biosensor designs for the detection of protein analytes and developed such technologies featuring high specificity sensor elements. One of the technologies developed features a single domain antibody (nanobody) which is surface displayed on *E. coli* cells. By the use of a GFP binding nanobody, this biosensor technology was successfully demonstrated to detect dimeric GFP in a plate agglutination assay format. A second technology developed features a RNA-aptamer sensor element incorporated into a RNA-switch device. For prototyping this technology, a thrombin binding RNA-aptamer incorporated into the RNA-switch device was demonstrated to detect thrombin by controlling GFP expression in *E. coli* cell-extract assay. Additionally, the design space for both prototyped technologies was successfully investigated for strategies that would enable the modulation of their analytical characteristics. Lastly, the modularity of the biosensor designs was investigated by the incorporation of sensors for alternative protein analytes.

The pool of known biomarkers for medical conditions includes a large number of protein molecules. Since the prototyped designs feature sensor elements that can be evolved *in vitro* for the specific recognition of any protein target at will, the technologies developed here can potentially be utilised as synthetic biology platforms for the development of low cost biosensor devices for use in medical diagnostic applications.

Declaration of Originality

I hereby declare that this document submitted for the degree of Doctor of Philosophy is the product of my own research work, and that all external contributions to this work have been appropriately referenced and acknowledged.

Copyright Declaration

'The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work'

Acknowledgements

I would like to express my special appreciation and gratitude to my advisors Prof. Paul Freemont and Dr. Karen Polizzi for their valuable advice on my research and their continuous support during this PhD. Your enthusiasm for this work and generally towards the field of Synthetic Biology provided the motivation, endurance and inspiration needed to become the scientist I am today.

I would also like to thank my friends and colleagues in the Freemont Lab and at the Centre for Synthetic Biology and Innovation for their assistance throughout the years, and for providing a friendly and supportive work environment. You know who you are and I thank you for everything.

I would also like to thank all of my friends outside of Imperial College who their support during this time was invaluable to me.

A special thanks to my family. Words cannot express how grateful I am to my mother and father for all the sacrifices they have made for me all these years, and for never stopping to encourage and support me towards achieving my goals in life no matter what these are. I dedicate this thesis to you.

Finally to SC. I would like to express my deepest appreciation to you who has always stand next to me through the bests and the worsts and whose unconditional support and trust makes everything worthwhile.

I. Table of Contents

I. Table of Contents 5

II. List of Figures 9

III. List of Tables 12

IV. Abbreviations 13

V. Thesis Outline 14

1 Introduction 16

1.1 Biosensors 16

1.2 Synthetic Biology 20

1.3 Synthetic Biology applied to biosensor design 25

1.4 Research motivation and objectives 32

2 Materials and Methods 35

2.1 General molecular biology materials and methods 35

2.1.1 Bacterial strains 35

2.1.2 Antibiotic selection 35

2.1.3 Test digest of DNA plasmid constructs 36

2.1.4 DNA agarose gel electrophoresis 36

2.1.5 Plasmid construction procedures 36

2.1.6 Plasmid DNA isolation from bacterial cell cultures 39

2.1.7 DNA Sequencing Reactions 39

2.1.8 Glycerol Stock preparation of *E. coli* cell lines 39

2.2 Materials and methods for Chapter 3 40

2.2.1 Plasmid constructs for Chapter 3 40

2.2.2 Transformation of electrocompetent *E. coli* cells 42

2.2.3 SDS polyacrylamide gel electrophoresis (SDS-PAGE) 42

2.2.4 GFP and GFP_2 overexpression and protein purification 42

2.2.5 Fluorescence of purified GFP and GFP_2 samples....44
2.2.6 Cell toxicity assay....44
2.2.7 NVgfp protein expression, cell localisation and nanobody functionality....45
2.2.8 Quantification of displayed anti-GFP nanobodies....46
2.2.9 Levels of anti-GFP nanobody display in biosensor device variants....47
2.2.10 Plate agglutination assay....48
2.2.11 Simulation of the agglutination reaction....49
2.2.12 Fibrinogen stock solutions....49
2.3 Materials and methods for Chapter 4....50
2.3.1 Plasmid constructs for Chapter 4....50
2.3.2 Chemically synthesized RNA-aptamers....51
2.3.3 In vitro transcription....51
2.3.4 RNA gel electrophoresis....52
2.3.5 Proteins stock solution preparation....53
2.3.6 Binding reactions....53
2.3.7 Electrophoretic Mobility Shift Assay (EMSA)....53
2.3.8 Cell-extract expression assay using a commercial S30 kit....54
2.3.9 TX-TL *E. coli* cell extract production....54
2.3.10 Cell-extract expression assay using homemade cell-extract....55
2.3.11 Data analysis for cell-extract assays....55
3 Biosensor design: Nanobody & whole cell chassis....57
3.1 Background Information....57
3.1.1 Antibodies and antibody fragments....57
3.1.2 Engineering surface display in *E. coli*....62
3.1.3 The agglutination reaction as an immunologic diagnostic technique....66
3.2 Experimental Results....69
3.2.1 Biosensor design....69

3.2.2 Prototype design and device construction 71
3.2.3 Characterisation of the pNVgfp construct in *E. coli* (DH10B) cell chassis 74
3.2.4 Development of the model biomarker protein: GFP_2 81
3.2.5 Simulation of the bacterial agglutination reaction 83
3.2.6 Testing prototype biosensor device for agglutination 88
3.2.7 Engineering the analytical characteristics of the agglutination assay 92
3.2.8 Testing the biosensor design for sensor element modularity 102
3.3 Technical Discussion 106
4 Biosensor design: RNA-aptamer & *E. coli* cell-extract 111
4.1 Background Information 111
4.1.1 RNA-aptamers 111
4.1.2 Bacterial translation initiation and translation initiation rate prediction 116
4.1.3 Ligand inducible control of bacterial translation initiation 120
4.1.4 RNA secondary structure prediction 124
4.1.5 Transcription/translation cell-free technologies 126
4.2 Experimental Results 128
4.2.1 Biosensor design 128
4.2.2 Prototype design and construction 130
4.2.3 Investigating RNA-switch functionality of the prototype device 134
4.2.4 Investigating thrombin binding to prototype device by EMSA 141
4.2.5 Testing a low-cost cell-extract system for the biosensor device 153
4.2.6 Strategies for modulating the biosensing assay characteristics 160
4.2.7 Investigating biosensor design modularity 172
4.2.8 *In silico* folding analysis of RNA-switch devices 191
4.3 Technical Discussion 198
5 Conclusions 208
5.1 Discussion 208

5.2 Future Work 213

6 References 217

7 Appendix A – Extended materials and methods 235

7.1 Materials 235

7.2 Plasmid construct sequences 242

8 Appendix B – Agglutination reaction mathematical model 250

9 Appendix C – RNA sequences folding analysis 263

10 Appendix D – Reprint permissions 279

II. List of Figures

Figure 1.1: Biosensor device components. 19

Figure 1.2: Synthetic biology biosensor design. 28

Figure 3.1: Structure of antibody molecules. 58

Figure 3.2: Classical (type Va) autotrasporter protein biogenesis pathway. 64

Figure 3.3: Agglutination assay. 67

Figure 3.4 Mechanism and test assay work flow of nanobody/whole-cell biosensor design. 70

Figure 3.5: NVgfp display vector and the pNVgfp construct plasmid map. 72

Figure 3.6: Test digest of pNVgfp DNA construct. 73

Figure 3.7: Cell toxicity assay for NVgfp protein expression in *E. coli* DH10B cell strain. 75

Figure 3.8: Flow cytometry analysis of *E. coli* DH10B cells harbouring pNVgfp. 77

Figure 3.9: Quantification of anti-GFP nanobodies displayed on prototype device. 79

Figure 3.10: GFP fluorescence versus GFP_2 fluorescence. 82

Figure 3.11: Bacterial cell pellet as a visual output for bacterial agglutination reactions. 85

Figure 3.12: *In silico* simulation of bacterial agglutination reaction. 86

Figure 3.13: Plate agglutination assay using the designed prototype biosensor device. 89

Figure 3.14: Biosensor device variants mean cell fluorescence and agglutination assays. 96

Figure 3.15: Bacterial cell agglutination of anti-fibrinogen biosensor devices. 104

Figure 4.1: Systematic Evolution of Ligands by Exponential Enrichment (SELEX). 113

Figure 4.2: Secondary structure of RNA-aptamers. 115

Figure 4.3: Translation initiation in bacteria. 117

Figure 4.4: Mechanisms of ligand inducible control of translation initiation. 121

Figure 4.5: RNA-switch biosensor design. 129

Figure 4.6: Plasmid maps of DNA constructs pNK401 and pNK402. 131

Figure 4.7: Test digest of plasmid constructs pNK401 and pNK402. 133

Figure 4.8: RNA electrophoresis of control device and prototype thrombin switch device. 133

Figure 4.9: Cell-extract repression assay for prototype device. 136

Figure 4.10: Specificity assay for prototype device. 138

Figure 4.11: Input / output characteristics of prototype device. 140

Figure 4.12: EMSA of anti-thrombin RNA-aptamer. 143

Figure 4.13: EMSA of prototype device. 145

Figure 4.14: EMSA of the prototype device partial sequence (~150 nucleotides). 149

Figure 4.15: EMSA of the prototype device partial sequence (~250 nucleotides). 151

Figure 4.16: GFP expression in *E. coli* BL21(DE3) strain cell-extract system. 155

Figure 4.17: Buffer salts concentration optimization for cell-extract reactions. 157

Figure 4.18: TXTL cell-extract repression assay for prototype thrombin device. 159

Figure 4.19: GFP expression of thrombin switch device variants in cell-extract reactions...163

Figure 4.20: TIR prediction for thrombin switch device variants by RBScalculator. 165

Figure 4.21: RNA electrophoresis of thrombin switch device variants partial sequences. ...167

Figure 4.22: EMSA of the thrombin switch device variants. 168

Figure 4.23: Input / output characteristics of thrombin switch device v04. 170

Figure 4.24: EMSA of the anti-VEGF RNA-aptamer. 174

Figure 4.25: RNA-electrophoresis of designed VEGF switch devices (partial sequences). ...176

Figure 4.26: EMSA of designed VEGF switch devices. 178

Figure 4.27: GFP expresion of designed VEGF switch devices. 180

Figure 4.28: EMSA of toggle-25t RNA-aptamer. 183

Figure 4.29: RNA electrophoresis of designed toggle switch devices (partial sequences). ..185

Figure 4.30: EMSA of designed toggle switch devices. 187

Figure 4.31: GFP expresion of designed toggle RNA-switch devices. 189

Figure 4.32: Minimum free energy structure of VEGF switch device v02. 194

Figure 4.33: Minimum free energy structure of toggle switch device v02. 195

III. List of Tables

Table 2.1: Bacterial strains.....35

Table 2.2: Plasmid constructs for Chapter 3.....40

Table 2.3: Plasmid constructs for Chapter 4.....50

Table 2.4: Chemically synthesised RNA-aptamers.....51

Table 2.5: Primers used for *in vitro* transcription procedures.....51

Table 3.1: Biosensor device variants for GFP_2 analyte detection.....93

Table 3.2: Comparison of reported relative promoter activity versus measured activity. 98

Table 3.3: Comparison of designed RBS TIR activity versus measured activity. 99

Table 4.1: Thrombin swicth device variants predicted deltaG values for incorporated RNA-aptamer.....161

Table 4.2: Designed VEGF RNA-switch devices.....176

Table 4.3: Designed toggle RNA-switch devices. 185

Table 4.4: Folding analysis of designed RNA-switch devices. 193

Table 7.1: Buffer recipes 235

Table 7.2: Restriction Enzymes 236

Table 7.3: Proteins 236

Table 7.4: Chemicals 237

Table 7.5: Molecular biology kits 239

Table 7.6: Instrumentation 239

Table 7.7: Software tools 240

Table 7.8: Biological parts DNA sequences.....240

IV. Abbreviations

Ab	antibody
APS	ammonium persulphate
au	arbitrary units
bp	base pair
ddH_2O	double distilled water
dsDNA	double stranded DNA
IF	Initiation Factor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
L	Litre
LB	Lysogeny Broth
ml	millilitres
MWCO	Molecular weight cut off
ng	nanograms
nm	nanometers
$O.D._{600}$	optical density value at 600 nanometers
PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
PNK	Polynucleotide Kinase
rcf	relative centrifugal force
rpm	Rotations per minute
SDS	Sodium dodecyl sulfate
TAE	Tris-Acetate-EDTA
TEMED	N,N,N',N'-Tetramethylethylenediamine
TIR	Translation initiation Rate
UTR	Untranslated region
v/v	Volume to volume
w/v	Weight to volume
µl	microlitres
°C	degrees Celsius

V. Thesis Outline

This document describes the development of two synthetic biology biosensor technologies for the detection of protein analytes, that incorporate design features favourable for medical diagnostic applications.

Chapter 1 is an introductory chapter that places this research project in a scientific context, and aims to give the reader a review of the general research concepts. First, the chapter introduces the concept of biosensor technology as this applies for the development of analytical devices for the detection of molecules of interest. Then, the chapter introduces the field of Synthetic Biology as a scientific discipline and presents its history, principles and areas of application to date. Following, the chapter converges biosensor technology and synthetic biology discipline by reviewing the scientific literature of synthetic biology design biosensors. Finally, it identifies an unaddressed need in this research theme and sets the objectives for this research.

Chapter 2 describes the materials and methods for this research. It contains detailed accounts of the experimental protocols and the data analysis for the results presented in chapters 3 & 4 of this document.

Chapter 3 describes the development of a synthetic biology whole-cell biosensor technology that features a single domain antibody as its sensor element. Initially, the chapter provides background information on technologies and procedures used for the biosensor development, and these include antibody engineering technology, *E. coli* cell surface display technology and the agglutination reaction as a diagnostic assay. Following, the experimental results are presented that include the description of the proposed biosensor design, the development and testing of a prototype device and further experiments that investigated the design space and the modularity of the technology. Finally, a technical discussion is provided that critically examines the experimental procedures and results in this chapters.

Chapter 4 describes the development of a synthetic biology biosensor technology implemented in *E. coli* cell-extract that features a RNA-aptamer domain as its sensor element. Initially, the chapter provides background information of technologies and concepts used for the biosensor development, and these include RNA-aptamer technology,

the process of bacterial translation initiation and the mechanisms of control of this process by ligand molecules, nucleic acid folding prediction technology and cell-free expression systems. Following, the experimental results are presented that include the description of the proposed biosensor design, the development and testing of a prototype device and further experiments that investigated the design space and the modularity of the technology. Finally, a technical discussion is provided that critically examines the experimental procedures and results in this chapters.

Chapter 5 concludes this document by placing its finding in the context of relevant scientific literature, as well as proposing direction for future research work on these technologies and similar synthetic biology biosensor technologies.

Chapter 1

1 Introduction

1.1 Biosensors

Biosensors are compact, integrated analytical devices with a design that makes them ideal for the performance of diagnostic tests at the site where the sample is procured. Such sites, often called point-of-care or point-of-concern, are primary health centres, the bedside of patients at hospitals, at home and where infrastructure is lacking such as accident sites and field sites. In contrast to analytical techniques performed in central laboratories, the use of biosensor devices eliminates or reduces the need for highly qualified personnel to perform the test, avoids time wasted by transferring samples to a central laboratory and allows for immediate action following diagnosis (Turner, 2013).

Most of the commercially available biosensor devices are employed for medical diagnostic applications (J. Wang, 2008). Amongst them, glucose biosensors dominate the biosensor market (Mascini & Tombelli, 2008). Following the initial demonstration of enzyme electrodes (Clark & Lyons, 1962), first generation glucose biosensors used the glucose oxidase enzyme as the sensor element for the measurement of glucose concentration (Guilbault & Lubrano, 1973). The enzyme uses glucose as a substrate in a reaction that results in the reduction of a flavin group in the enzyme. The enzyme is then regenerated by the re-oxidation of the flavin group by molecular oxygen that results in the production of hydrogen peroxide. Then, an electrochemical transduction mechanism was used to transfer electrons between oxygen or hydrogen peroxide to an electrode and thus relate glucose concentrations in the sample to measured electric current in the electrode. Subsequently, second generation glucose biosensors used redox mediators instead of oxygen and hydrogen peroxide to shuttle electrons to the electrode (Cass et al., 1984) and eventually led, in combination with other innovations (J. Wang, 2008), to the commercialisation of self-monitoring blood glucose devices for use at home (Matthews et al., 1987). Since then, extensive research has expanded the use of biosensors for use in medical diagnostics with newer generations of biosensor making use of high affinity and specificity sensor elements,

such as antibodies and nucleic acids, for the detection of cancer biomarkers, biomarkers of cardiac diseases and hormones among others (Mascini & Tombelli, 2008).

In biosensor devices the recognition/sensor element for the target analyte is of biological origin or design. When operated, a biosensor relies on the specificity of its sensor element for the particular target molecule of choice to provide analytical information about the test sample in question. The basic components of a biosensor device include the recognition/sensor element and the transducer. The sensor element of the device is able to interact with the analyte molecules selectively, as compared to interactions with other molecules in the sample, and this interaction changes some chemical or physical property of the sensor element as a function of the analyte concentration. The transducer element of the biosensor device translates the information of the recognition event to a different kind of signal that is presented to the user of the device as the analysis results (Bănică, 2012; C. R. Lowe, 2008; Thevenot, Toth, Durst, & Wilson, 2001).

The analyte recognition process proceeds though the following general reaction scheme:

$$\mathrm{A} + S \underset{k\,off}{\overset{k\,on}{\rightleftharpoons}} P \qquad \text{(Eq. 1)}$$

where (A) is the analyte molecule recognised by the sensor element (S) to form the analyte/sensor interaction product (P). The molecular recognition arises from a number of non-covalent chemical bonds between the two species and is usually a reversible process. When the reaction is at equilibrium, the recognition process is characterised by an equilibrium association constant K_a that is defined as:

$$Ka = \frac{[\,P]}{[\,A\,] \times [\,S\,]} = \frac{k\,on}{k\,off} \qquad \text{(Eq. 2)}$$

which translates to K_a being equal to the concentration of product species [P], divided by the product of the concentrations of the analyte species [A] and the sensor element species [S]. This equilibrium association constant indicates the affinity of the sensor element for the analyte. In addition, the terms: i) k_{on} (association rate constant) and ii) k_{off} (dissociation rate constant) determine the kinetics by which the forward and the reverse reactions take place.

Another chemical species in the test sample *i.e.* species B, will also have an equilibrium association constant for the sensor element, but the selectivity of the sensor element for the target analyte (A) arises from the fact that K_a for (A) is much greater than the K_a for (B) (Bănică, 2012).

Biosensor devices make use of various kinds of recognition elements (Figure 1.1) that allow selective molecular recognition of the target analyte. Biological sensor elements such antibodies and nucleic acid aptamers recognise their target analyte molecules through affinity interactions. The affinity interactions include a large number of non-covalent chemical bonds between the interacting species, such as ionic bonds, hydrogen bonds and van der Waals forces. These result in the formation of a complex in which the interacting species are complementary in shape and chemical reactivity in the three-dimensional space. In addition to affinity interactions, nucleic acid molecules can recognise their target analyte through base pairing hybridisation interactions. These interactions involve the recognition of a target analyte through hydrogen bonding between well-defined pairs of nitrogenous bases of DNA or RNA. Enzymes are another class of sensor elements that are employed for the construction of biosensor devices. In enzyme based biosensors the recognition process of the analyte of interest (called substrate or inhibitor depending on application- for enzyme biosensors) is a 3-step process that is initiated by the substrate/enzyme binding, followed by a chemical conversion catalysis reaction and then product release (Bănică, 2012; Thevenot et al., 2001).

The transducer component of the biosensor is sensitive to the change in the physical or chemical properties of the recognition element resulting from the interaction of the recognition element with the analyte molecules. Upon interaction, the transducer component translates the above change into a quantity that can be measured and assessed by the user of the device. Signal transduction can be achieved by electrochemical, optical and mass sensitive mechanisms among others (Vo-Dinh & Cullum, 2000). In electrochemical transducers, the recognition event is converted into an electrical signal though the generation of electrical current, voltage difference or change in the resistive properties of the medium (Grieshaber, MacKenzie, Vörös, & Reimhult, 2008). In the case of optical transducers, the information of the binding event is transduced by a change in the number

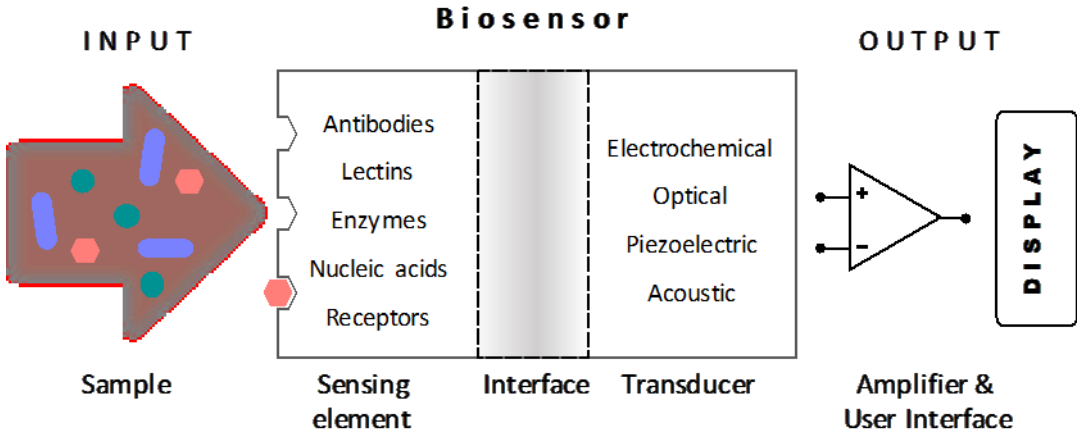


Figure 1.1: Biosensor device components. Biosensors are analytical devices made of a sensing module and a transducer module. The sensing module incorporates a biologically derived sensor that is able to specifically interact with the target analyte of interest. The transducer module is sensitive to the change upon analyte and sensor interaction and translates this information to a signal that can be read by measured and assessed by the user of the device.

of photons as a result of optical phenomena such as fluorescence, absorbance and refractive index effects among others (Borisov & Wolfbeis, 2008). Lastly, in mass sensitive mechanisms the binding event information is transduced through the use of piezoelectric crystals (Bănică, 2012).

1.2 Synthetic Biology

Synthetic Biology is an interdisciplinary research and technology field which strives to develop tools and methodologies that aid the design of biological systems and living organisms for useful purposes (Khalil & Collins, 2010). As opposed to traditional genetic engineering, synthetic biology places strong emphasis in engineering principles and methodology in the design and construction of biological systems (Church, Elowitz, Smolke, Voigt, & Weiss, 2014). In the past, biotechnology project designs tended to be *ad hoc* and this meant that there was little transfer of knowledge between projects. As a result, project timelines for the development of biotechnological applications were lengthy, time consuming and very costly. With synthetic biology, traditional molecular biologists, microbiologists and cell biologists come together with engineers, such as electrical and software engineers among others, to develop concepts and procedures relevant to the engineering of living matter. These concepts and procedures enable the realisation of project designs that build upon previous work as well as designs that allow for future work to build upon them (Endy, 2005).

Historically, the landmark work that established synthetic biology as a discipline was carried out at the beginning of the 21st century (Elowitz & Leibler, 2000; Gardner, Cantor, & Collins, 2000), but there were some earlier key discoveries and inventions that allowed for the emergence of the field. Such a key discovery was the elucidation of the regulatory mechanisms that allow a cell to respond to the perturbations and changes in its environment, and this was achieved by the study of the *lac* operon in *E. coli* (Jacob & Monod, 1961). Also, in the 1970s and 1980s the discovery of restriction enzymes and the development of molecular cloning techniques were central to the ability of scientists to produce recombinant/synthetic DNA for design purposes (Danna & Nathans, 1971; Roberts, 2005; Smith & Wilcox, 1970), and this was complemented by the invention of the polymerase chain reaction (PCR) that allowed the easy amplification and isolation of DNA sequences in molecular biology research (Mullis et al., 1987) and the development of methods that allow the chemical synthesis of DNA oligonucleotides (Caruthers, 1985). Later, automated DNA sequencing and computational tools made whole microbial and eukaryotic genomes available for the scientist to use as a pool of available biological parts, diverse enough in functionality to carry out any purposeful task imaginable (Blattner et al., 1997;

Fleischmann et al., 1995; Goffeau et al., 1996). At the same time, studies in the field of systems biology, that used computational tools to analyse myriads of interactions between cellular components, revealed hierarchical structures within cellular networks of biomolecules that although vast and "messy" resembled man-made engineered systems (Barabasi & Oltvai, 2004; Hartwell, Hopfield, Leibler, & Murray, 1999). As previously mentioned, in early 2000 two scientific papers were published in which the authors used engineering approaches to build genetic circuits that carried out a designed, purposeful function and now are widely regarded as marking the beginning of the field of synthetic biology. Both papers used mathematical modelling to describe and tune an interacting system of biomolecules, not normally encountered together in natural systems, which they viewed as a biological circuit to produce a toggle switch device (Gardner et al., 2000) and an oscillator device (Elowitz & Leibler, 2000). Following these initial publications, synthetic biology principles were widely adopted to produce biological circuits of genetically encoded, synthetic logic gates (Guet, Elowitz, Hsing, & Leibler, 2002), cell-to-cell communication systems (Tamsir, Tabor, & Voigt, 2011) and event counters (Friedland et al., 2009) among others. Notable achievements in the field include the heterologous production of antimalarial drug artemisinin from a metabolic pathway engineered in yeast instead of its low yield natural production from *Artemisia annua* plant (Ro et al., 2006) and the recreation of a viable *Mycoplasma mycoides* bacterial cell controlled entirely by synthesised DNA created chemically from digitized genome sequence information (Gibson et al., 2010). More recently, synthetic biologists have constructed the first designed eukaryotic chromosome in yeast (Annaluru et al., 2014) and efforts are on the way for constructing a designer yeast genome, while there has been great successes in the development of genome engineering and editing tools such as MAGE (Gallagher, Li, Lewis, & Isaacs, 2014) and CRISPR (Mali et al., 2013).

Although synthetic biology is mostly concerned with the design of biological systems and living organisms, the field is strongly influenced by engineering concepts such as standardisation and abstraction, and practises such as computer aided design (CAD). The reason for this influence is that such concepts and practises enable the scale-up of project design complexity and transfer of knowledge between projects. Standardisation is the process of developing and implementing technical standards or technical norms which

allows for increased compatibility between projects developed separately. The iGEM Registry or Registry of Standard Biological parts is an example of the concept of standardisation as this applies to the field of synthetic biology. This particular registry promotes the use of a technical standard for the process of DNA assembly in which particular restriction digestion sites are found on each site of each biological part DNA sequence. These sites make use of a set of particular restriction enzymes to enable the assembly of biological parts into larger constructs. The adoption of this technical standard allows for cost reduction and efficient use of time. These benefits arise due to the fact that all registered biological parts can be assembled together using a small set of enzymes, and the circuit does not need to spend time to check or refactor DNA sequences of biological parts against assembly procedure incompatibility. Nevertheless, the popularity of iGEM standard is declining due to newer DNA assembly technologies -such as Gibson assembly (Gibson et al., 2009) and Golden Gate assembly (Engler, Kandzia, & Marillonnet, 2008)- that offer other procedure advantages and due to the affordable prices of large DNA fragments chemical synthesis. Abstraction is another concept adopted from engineering that allows managing complexity of very complicated systems, by dividing the system into levels of hierarchies. (Endy, 2005). A usual abstraction hierarchy used in synthetic biology includes the abstraction levels of “DNA sequence”, “Part”, “Device” and “System”. For example, a synthetic biologist with the task of designing an oscillator device will be working on the “Part” level of the hierarchy. The relevant information that the designer will have to consider are transcription factors (repressor proteins) with appropriate repression function characteristics; but the complexity of DNA sequence for the coding of the protein product or of the operator site functionality is hidden (or abstracted) from him, as this problem is dealt at the DNA sequence hierarchy level. Lastly, the field places a strong emphasis to the development and use of CAD tools. These tools can help reduce time and costs towards a functional prototype device by running test simulations *in silico* rather than the lengthier wet lab experiments. RBScalculator is one CAD tool that is widely used in the synthetic biology community and it allows the *in silico* design of ribosome binding sites to tune protein expression at the desired levels (Salis, Mirsky, & Voigt, 2009).

Synthetic biology is considered an emergent and enabling technology with applications in many sectors such as the healthcare, environmental protection and

bioenergy (Synthetic_Biology_Leadership_Council, 2016). For the healthcare sector, designed biological systems could be used for drug discovery applications. In one such example, a chemical compound library was screened with the use of a drug responsive transcription factor that aided in the identification of an anti-tuberculosis drug (Weber et al., 2008). Additionally, designed biological organisms can be used to fight diseases, and one such application was demonstrated with the use of engineered therapeutic *E. coli* cells that invaded cancer cell lines when they encountered hypoxic environment or high cell density (Anderson, Clarke, Arkin, & Voigt, 2006). Another use of designed organisms in the healthcare sector is for the production of pharmaceutically active chemicals in a cost-efficient manner. This concept was demonstrated by the production of the anti-cancer drug taxol in engineered *E. coli* (Ajikumar et al., 2010) and by the production of the anti-malarial drug artemisinin in engineered yeast (Ro et al., 2006).

Environmental protection has attracted a lot of interest from synthetic biologists and a large number of applications have been demonstrated targeted for this sector. Engineered biological systems have been used to monitor the environment for pollutants, as in the case of engineered *P.putida* cells that utilised non-natural, directed evolution derived transcription factors that recognised nitrotoluenes for the purpose of tracing explosives in the soil (Garmendia, de las Heras, Galvao, & de Lorenzo, 2008). Furthermore, in addition to monitoring and measuring chemicals in the environment, designed organisms were engineered to degrade pollutants in order to clean up a polluted site, as in the case of an engineered *E. coli* that was designed to seek and destroy the atrazine herbicide (Sinha, Reyes, & Gallivan, 2010). Moreover, engineered organisms have been designed to relieve the strain put on the environment by human activity by producing commodity chemicals from renewable resources. For example, plastics produced from mineral oils (distillate of petroleum) have a multifaceted negative impact on the environment, and in order to try to address those problems a study designed the biomass crop switchgrass to produce a biodegradable chemical with plastic like characteristics (Somleva et al., 2008).

Synthetic biologists have also shown a strong interest in developing solutions for the bioenergy sector. High energy density alcohols, such as butanol and derivatives, have been produced using engineered cells that make use of feedstock from renewable carbon sources

such as glucose (Atsumi, Hanai, & Liao, 2008), from non-edible plant parts (Higashide, Li, Yang, & Liao, 2011), or even from carbon dioxide (Atsumi, Higashide, & Liao, 2009).

In addition to providing practical solutions, the engineering of biological systems can be used to enable basic research by helping to elucidate the mechanisms of various biological phenomena (Cagatay, Turcotte, Elowitz, Garcia-Ojalvo, & Suel, 2009).

1.3 Synthetic Biology applied to biosensor design

One of the first applications explored by synthetic biologists was the design and construction of biosensor devices. In these synthetic biology biosensor devices, rationally designed genetic circuits incorporate cellular sensing elements that are able to recognize a target analyte of interest. Then, this chemical detection event is translated/transduced into a quantifiable signal through various mechanisms that result in the expression of a reporter protein.

Sensor elements that were used for the construction of synthetic biology biosensor devices to date were mostly extracted from nature. Biological organisms naturally evolved sensor elements to monitor their environment for the availability of nutrients and respond to adverse conditions (Khalil & Collins, 2010). Biological organisms have also developed sensor elements to monitor substrates, intermediates and products of processes of their metabolism. One class of naturally derived sensor element is transcription factor proteins and these have been used extensively in biosensing applications (van der Meer & Belkin, 2010). For example, a biosensor device was constructed in which the ArsR transcription factor protein from the *E. coli ars* operon was used as the sensing element to measure arsenic levels in water (Ramanathan, Shi, Rosen, & Daunert, 1997). A plethora of ligand responsive transcription factor sensor elements can be found in the literature, in online databases of metabolic pathways and in the Registry of Standard Biological Parts (Goers et al., 2013). Nevertheless, for some analytes of interest there is no known natural sensor that can be utilised or the natural system that recognizes the particular analyte of interest is poorly characterised. For such a case, a study presented an approach in which the DmpR transcription factor protein derived from *P. putida* that recognises phenols was altered by mutagenic PCR to recognise a number of substituted phenol chemicals pollutants (Wise & Kuske, 2000). Alternatively, biosensing devices have been constructed using non-protein, RNA-based sensing elements that are not naturally occurring but have been evolved in the lab to recognise target analyte molecules. For example, the RNA-aptamer sensing element that specifically recognises the small molecule Theophylline (Jenison, Gill, Pardi, & Polisky, 1994) was incorporated into a reporter mRNA expression platform for the construction of a riboswitch biosensor device (Lynch, Desai, Sajja, & Gallivan, 2007).

Synthetic biology biosensor designs usually couple the sensing of the analyte of interest to the production of a spectroscopically active reporter protein. Green fluorescent protein (GFP) from *A. victoria* and various other fluorescent variants (Shaner, Steinbach, & Tsien, 2005) are some of the most widely used reporter proteins. GFP is an auto-fluorescent monomeric protein that emits light following light excitation at specific wavelengths and its detection does not require the addition of any co-factor or exogenously added substrate (Daunert et al., 2000). Many biosensing applications use reporter systems that utilise members of the luciferase family of enzymes to catalyse light emitting reactions. Bacterial luciferase genes were derived from the lux operon of various bacteria such as *V. fisheri*, *V. harveyi*, *P. luminescens* etc (Meighen, 1991). The bacterial luciferase enzyme is a heterodimeric protein resulting from the gene products of the lux operon genes luxA and luxB. The bacterial luciferase enzyme catalyses the oxidation of a reduced flavin mononucleotide ($FMNH_2$) and a long chain fatty aldehyde to FMN and the corresponding fatty acid in the presence of oxygen with a concomitant emission of light. This reporter system requires the addition of a fatty aldehyde substrate (for example decanal) and $FMNH_2$, but co-expressing the luxA and luxB genes with the operon genes luxCDE (with gene products Acyl-CoA reductase, acyl transferase and acyl-protein synthetase respectively) gives the advantage of producing a bioluminescence signal without requiring the addition of the aldehyde substrate (Daunert et al., 2000; van der Meer & Belkin, 2010). In contrast, the Firefly luciferase from lucFF gene is a monomeric protein and it catalyses the oxidation of its substrate luciferin to oxyluciferin in the presence of ATP, O_2, and Mg^{2+}, to produce CO_2 and visible light (de Wet, Wood, Helinski, & DeLuca, 1985). This reporter system requires the addition of Mg^{2+}, ATP and luciferin substrate. Another enzymatic reporter protein is β-galactosidase, the product of lacZ gene from *E. coli*, that catalyses the hydrolysis of the glycosidic bond of β-D-galactopyranosides (Juers, Matthews, & Huber, 2012). β-galactosidase enzyme is a homotetramer and many detection methods can be used to monitor its activity including colorimetry, chemilluminescence, electrochemistry and fluorescence depending on the substrate chemical used in the assay (Daunert et al., 2000).

Synthetic biology biosensors have been engineered using a wide variety of transduction mechanisms in order to relay the information of analyte binding to the user of the device (Figure 1.2). These transduction module mechanisms usually act through the

control of some stage of the protein expression process. A large number of constructed biosensing devices rely on the control of the transcription process as the transduction mechanism. In most cases, the transduction mechanism involves a set of transcription factor and promoter elements, where analyte recognition by the transcription factor results in the transcription factor binding to an operator sequence of the promoter element to activate or repress the process of transcription initiation (Anderson, Voigt, & Arkin, 2007). Alternatively, a number of transduction mechanisms have been demonstrated that act through the control of the translation process. For example in riboswitch devices, upon ligand recognition the transduction mechanism involves RNA-secondary structure rearrangements that expose or shield the ribosome binding site from ribosome binding (Lynch et al., 2007), while in ribozyme devices the transduction mechanism involves destabilization of the mRNA transcript (Kennedy, Vowles, d'Espaux, & Smolke, 2014). Additionally, post-translational transduction mechanisms have been demonstrated in which light sensing by an extracellular sensor domain activated an engineered bacterial two component system (Levskaya et al., 2005) and in another example redesigned periplamic proteins recognising small ligand molecules activated naturally occurring two-component systems (Looger, Dwyer, Smith, & Hellinga, 2003).

In some instances, the output signal from the biosensing devices needs to be amplified in order to match amplitude specifications of downstream processes. To address this need, synthetic biologists have constructed genetic amplifiers with a function equivalent to their electronic counterparts (Coughlin & Driscoll, 2001). One genetic amplifier design managed signal amplification and digitisation by connecting the output signal of a transcriptional transducer module to the expression of integrase enzyme (Courbet, Endy, Renard, Molina, & Bonnet, 2015). In this case the signal amplification and digitisation was necessary to tackle low sensitivity of detection of the device. In another study a heterodimeric transcription factor was used to build a transcriptional amplifier module that allowed tuneable signal gain (B. Wang, Barahona, & Buck, 2014). The tuning of the amplifier was achieved with the use of another protein that was able to repress the formation of the heterodimeric transcription factor complex.

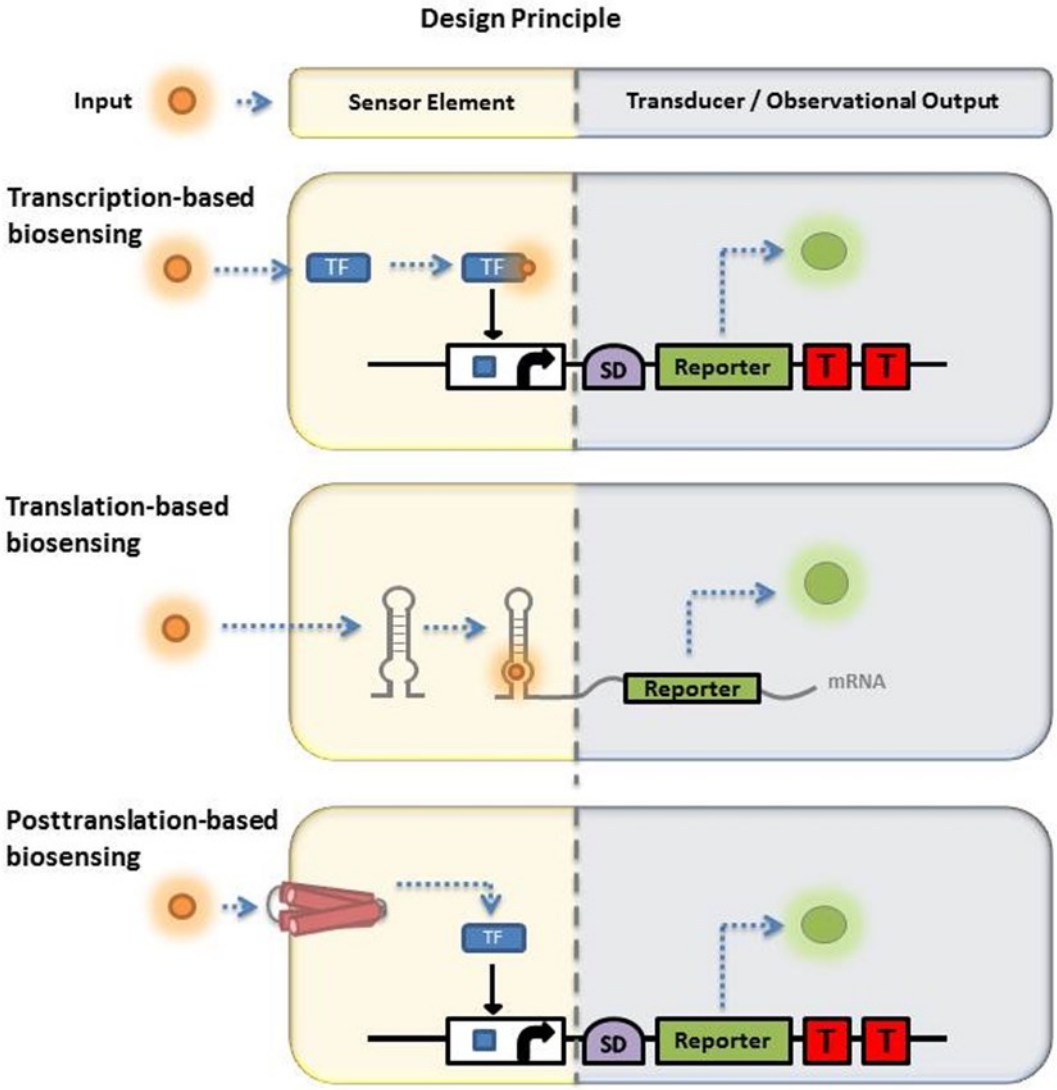


Figure 1.2: Synthetic biology biosensor design. Transduction mechanisms commonly employed by synthetic biology biosensor designs act through the control of some stage of the protein expression process to relay the information of analyte sensing. The analyte of interest (glowing orange) is recognised by a biological sensor molecule and this signal is processed to transduce the information which is reported in the form of protein expression (glowing green). At the transcriptional level the input signal is bound by a signal specific transcription factor (TF) that binds an operator sequence to control transcription rate. At the translational level the input signal is bound by a signal specific RNA aptamer that modulates the rate of translation initiation. At the post-translational level the input signal is recognised by a surface receptor extracellularly, that then activates an intracellular TF which regulates the rate of transcription initiation

Whole cell biosensors

The use of whole cells as biosensing platforms can confer a number of favourable characteristics to biosensor design: i) they are a self-replicating manufacturing platform and their costs of culturing can be very cheap, ii) following prototype development, production of devices can be readily scaled up using existing industrial infrastructure, iii) they avoid expensive purification procedures, iv) they have high biocatalytic activity, v) they are robust and tend to be stable to environmental changes and vi) can provide physiologically relevant data and evaluate bioavailability (Courbet et al., 2015; Raut, O'Connor, Pasini, & Daunert, 2012; van der Meer & Belkin, 2010). Early implementations of whole cell biosensors designs relied on general effects affecting the host cell physiology. For example, a *S. typhimurium* based whole cell bioassay that detected carcinogenic chemicals relied on DNA damage inducing the umu operon that was genetically fused to beta-galactosidase reporter protein (Oda, Nakamura, Oki, Kato, & Shinagawa, 1985). Another example is the Microtox test that is based on the principle that the inherent luminescence of *Vibrio fisheri* bacterial cells would diminish in the presence of toxicants (Bulich, 1979).

Advances in molecular biology and genetic engineering allowed the construction of whole cell biosensors with synthetic DNA circuits that used purpose built sensor elements to specifically detect target analytes at the molecular level. For example, whole cell biosensors have been constructed for the detection of arsenite in groundwater in which the toxic chemical was specifically recognised by the ArsR transcription factor (Trang, Berg, Viet, Van Mui, & Van Der Meer, 2005), as well as whole cell biosensors that detected the theophylline chemical that was specifically recognised by a RNA-aptamer sensor element specific for this small molecule (Lynch et al., 2007). For both biosensors, in order for the sensor element to specifically recognise the molecular analyte, the analyte molecules must first transport past the cell membrane barrier. In the case of the arsenic biosensor, the hydrophilic analyte enters the cell through *E. coli* transmembrane transporter proteins (Zangi & Filella, 2012), while hydrophobic molecules are able to diffuse across lipid bilayers (Gutknecht & Walter, 1981). An alternative strategy for the construction of whole cell biosensors is the detection of analytes extracellularly. This strategy was employed for the specific detection of ribose in the periplasmic space by a periplasmic binding protein and then followed by signal transduction through a transmembrane two component system (Reimer, Yagur-Kroll, Belkin,

Roy, & van der Meer, 2014). Another example of the concept of extracellular detection was the use of cell surface displayed peptide tags enabling the detection of elastase protease activity of the parasitic worm *S. mansoni* (Webb et al., 2016). Finally, researchers developed a modular transmembrane receptor and signal transduction platform -named MESA- for the detection of extracellular ligands in mammalian cell host (Daringer, Dudek, Schwarz, & Leonard, 2014). This platform was demonstrated to detect the small molecule rapamycin extracellularly and induce reporter expression intracellularly.

Cell-extract based biosensors

The development of protocols that allow a straightforward and low cost production of enzymatically active *E. coli* cell-extract material (Kwon & Jewett, 2015; Sun et al., 2013) sparked interest in developing synthetic biology biosensor designs in this cell-free platform. This technology platform shares some of the favourable characteristics of whole cell biosensors platforms such as being a low cost manufacturing platform for biosensing devices. In addition, the cell-extract platform I) eliminates the need and/or complications of analyte transport across the cell membrane, ii) allows for rapid prototyping of designed genetic circuits (Chappell, Jensen, & Freemont, 2013; Sun, Yeung, Hayes, Noireaux, & Murray, 2014), iii) allows for easily modifiable reaction conditions and iv) provides a mode for safe and sterile deployment of applications outside the lab environment.

A number of studies have demonstrated the suitability of the *E. coli* cell-extract technology for biosensor applications. In proof of concept studies, the platform was used for the detection of mRNAs of antibiotic resistance genes and for distinguishing between different strains of the Ebola virus (Pardee et al., 2014). In this biosensor design, the absence of cell membrane allowed the recognition of mRNA sequences -the analyte of interest- by RNA sensor elements through a base pair hybridisation mechanism (Green, Silver, Collins, & Yin, 2014). Further work on this technology allowed the development of a biosensor device for the detection of Zika virus RNA at sample concentrations down to femtomolar concentrations (Pardee, Green, et al., 2016). This was achieved by the incorporation of a RNA sample amplification step into the procedure workflow. Also, biosensor designs were demonstrated with the use of eukaryotic cell-extract systems. A study used a cell-extract system derived from germ wheat for the detection of theophylline and other small molecules by the use of riboswitch technology (Ogawa, 2011), and another

study used a cell-extract system derived from HeLa cells for the detection of glucose using genetically encoded FRET sensor (Pardee et al., 2014).

Synthetic circuits as intracellular biosensing devices

In some instances, the 'analyte of interest' is the intracellular state of the cell, and so the challenges of analyte transport across the cell membrane are irrelevant to the biosensor design. In these cases, the reporting of the cell state or a combination of sensing/correcting of the disease state can be accomplished simply by transforming a biosensing synthetic circuit into the cell of interest.

An example of such an approach was demonstrated by the use of a cell classifier genetic circuit that distinguished between HeLA and non-HeLa cells (Xie, Wroblewska, Prochazka, Weiss, & Benenson, 2011). The genetic circuit was able to interrogate the intracellular environment for the expression levels of six endogenous miRNA species and computed a "diagnosis". The diagnosis was computed by the use of a genetic logic circuit implementing RNAi gene silencing mechanism and the result was transduced to the observer by the expression of a fluorescent reporter. Alternatively, a therapeutic response was initiated which involved the expression of a proapoptosis gene. In another example, a biosensor system named CaspaseTracker was used to identify *Drosophila melanogaster* embryos that recovered from apoptotic phenotype (Tang, Tang, Fung, & Hardwick, 2015). To achieve this, the researchers used a synthetic genetic circuit that monitored intracellular caspase activity and directed the expression of fluorescent markers to report current caspase activity. Additionally, the circuit used a memory switch to direct the expression of another fluorescent marker to permanently label cells and all its progeny that recovered from caspase activity.

1.4 Research motivation and objectives

Introduction Overview

Biosensors are small, compact analytical devices that can specifically detect an analyte of interest and report the analyte concentration in a sample in a qualitatively, semi-quantitatively or quantitatively manner. The design of biosensor devices makes them suitable for medical diagnostic applications for use at the point of care, and blood glucose biosensor technology is a notable example. The compact size of biosensor devices allows for their use at a hospital or in doctor's office, at home or in the field, and their integrated design allows for use by non-specialised personnel.

Synthetic biology is a biotechnology discipline that aims for the design of synthetic biological systems and organisms with a strong adherence to engineering principles. The design of biological systems underlined by engineering principles allows for the transfer of knowledge between projects and can result in significant cost reduction for project development. There has been a large number of synthetic biology biosensor designs demonstrated in the scientific literature that utilised various types of sensing elements coupled to various types of transducer/reporter modules to specifically analyse samples for a target molecule of interest. Nevertheless, almost all synthetic biology biosensors leverage on genetically encoded design to enable lower costs of manufacturing, and the integration of engineering principles into the design to enable lower costs of prototype development.

Research Motivation

The potential of synthetic biology for biosensor design has been explored extensively but the attention on developing designs that detect protein analytes has been very limited. This is despite the fact that proteins are a molecular class which form a significant portion of the known pool of medical biomarkers, and hence such a technology could potentially be very useful for medical diagnostic applications.

The challenge for developing biosensor devices of a synthetic biology design for the detection of protein analytes is two-fold. One obstacle is the lack of suitable sensor elements and the second obstacle is the challenges posed for analyte transport to the sensor elements of the device. For other biosensor applications synthetic biologists have borrowed sensor elements from nature, but there is an absence/scarcity of naturally

available sensors that recognise protein biomarkers of interest, and even less any that exhibit the specificity required for medical diagnostic applications. In addition, strategies that involve the development of sensing elements from directed evolution strategies of surface receptor scaffolds or transcription factors will probably not be suitable for any target analyte desirable and up to date such strategies have mostly been utilised for small molecule targets. As regarding the second obstacle of the challenge, for medical diagnostic applications the analyte of interest is added exogenously to the biosensing system. The most commonly used biosensing platform in synthetic biology, whole cell biosensors, requires the transport of the analyte across the cell membrane to be recognised by intracellular sensing elements and this has been possible only for small molecules, hydrophobic chemicals or ions. Alternatively, the extracellular detection of the analyte of interest is a compelling approach, but coupling detection to a transduction module has been very challenging to engineer. Another interesting approach to the transport challenge is the use of cell-extract systems that lack the membrane barrier and these have been utilised for the detection of viral RNA for medical diagnostic application.

Research Objectives

The purpose of this research was to investigate the application of synthetic biology principles to the design of biosensor solutions for medical diagnostic applications, with particular focus on the detection of protein class biomarker molecules. In particular, the aim of the project was to develop one or more biosensor designs from a synthetic biology perspective for the specific detection of protein analytes, and demonstrate proof of concept experiments for prototypes of these technologies.

In order to enable the development of the proposed solution, a set of specifications were established that guided the design of the technology. The first specification set was that the technology developed should feature a sensor element that can be evolved *in vivo* or *in vitro* for highly specific recognition of any protein analyte of choice as required for medical diagnostic applications. The second specification set was that the biosensor design would need to incorporate a strong element of modularity, and especially in the coupling of the sensor element to the transducer, to allow for easy prototyping of biosensor devices. Importantly, the designed technology should be implemented in a suitable biosensing platform that would tackle the problem of analyte transport. Also, the biosensor design

would need to be genetically encoded and implemented in a bacterial system to allow for low cost of production.

A list of research objectives was set that aided this project towards the goal of developing a synthetic biology biosensor technology for medical diagnostic applications with a focus on protein biomarker detection. These were:

- To develop one or more synthetic biology biosensor designs that satisfied the aforementioned specifications in terms of sensor element choice and addressed the issue of analyte transport to the sensing element
- To construct a prototype biosensor device according to the design proposed, and test the functionality of the prototype device for protein biomarker detection
- To investigate the design space of the biosensor design with the aim of identifying strategies that would allow the tailoring of future devices for specific applications
- To investigate the modularity of biosensor design for sensor element incorporation

Chapter 2

2 Materials and Methods

This chapter describes the materials and methods for experiments presented in chapters 3 & 4. Section 2.1 of this chapter provides the general molecular biology procedures used throughout this research project, while section 2.2 provides the materials and methods specific for chapter 3 experiments and section 2.3 the ones specific for chapter 4 experiments. Additionally, extended material and methods can be found in Appendix A that details buffer recipes, chemicals, molecular biology kits, instruments, software tools, plasmid construct sequences etc.

2.1 General molecular biology materials and methods

2.1.1 Bacterial strains

Table 2.1 lists the *E. coli* strains and their genotype utilised in this research for cloning, protein expression and production of cell-extract procedures.

Table 2.1: Bacterial strains

Name	Genotype
E. coli DH10B	F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara leu) 7697 galU galK rpsL nupG λ–
E. coli BL21(DE3)	fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5

2.1.2 Antibiotic selection

Antibiotics were used for the selection of *E. coli* cells successfully transformed with plasmid constructs or for maintaining plasmids constructs in transformed *E. coli* cells. Antibiotics were prepared as stock solutions of 50mg/ml Ampicillin in 50% ethanol solution in ddH_2O, and 34mg/ml Chloramphenicol in 100% ethanol. Stock solutions were diluted 1000-fold to obtain working concentrations.

2.1.3 Test digest of DNA plasmid constructs

Test digestions were carried out to verify correct plasmid construct upon receipt from other sources or after DNA plasmid construction procedures. The test digest reactions were set up with 100 – 500 ng plasmid DNA, 0.5μl of the appropriate restriction enzyme(s), 2μl of buffer system appropriate for restriction enzyme(s) used as suggested by the restriction enzyme manufacturer and the reaction volume was made to 20μl with ddH_20. Test digest reaction was then incubated for 30 – 60 minutes at 37°C. Upon completion, test digestion reaction results were visualised by DNA agarose gel electrophoresis.

2.1.4 DNA agarose gel electrophoresis

DNA agarose gel electrophoresis was carried out to determine the sizes of DNA fragments by comparing them to a DNA ladder standard. For test digest reactions, 100 – 500 ng DNA was loaded on the agarose gel. Gels were made of 0.8% agarose in TAE buffer (w/v). The gel also included SybrSafe stain (for gel extraction procedures) or GelRed (for test digest experiments) at 10,000x dilution factor. To assist sample loading in the gel, samples to be analysed were mixed with DNA sample loading buffer (6x) at 5:1 volume ratio. Agarose gels were run at 100V until the bromophenol dye front reached approximately 2/3 of the length of the agarose gel. The resulting band pattern was visualised and documented using a Gel Doc XR system. The DNA ladder standard used was 1kb Plus DNA ladder.

2.1.5 Plasmid construction procedures

Plasmid constructs were assembled from individual DNA fragments that contained the desired biological parts of interest. The DNA assembly process produces circular plasmid DNA that is able to independently replicate in *E. coli* host. The DNA fragments for plasmid construction were obtained from one of the following sources: i) PCR reactions, Ii) restriction enzyme digestions and iii) annealing of chemically synthesized oligonucleotides. Then, the DNA fragments were joined together in a ligation reaction and transformed in competent *E. coli* cells to isolate correctly assembled plasmids.

DNA fragments obtained from Polymerase Chain Reaction (PCR)

PCR was used to amplify DNA sequences for use in the construction of new plasmid constructs. PCR reactions were carried out using Phusion® High-Fidelity PCR Master Mix

with HF Buffer using the protocol described by the manufacturer. Oligonucleotide primers were ordered from IDT. The reactions were set in 50 μl total reaction volume, at 0.5μM final oligonucleotide concentration and by using 1-10 ng plasmid DNA as template. Upon PCR reaction completion, template DNA was removed by DpnI digestion. DpnI digestion reactions were carried out by adding 0.5 μl DpnI restriction enzyme to the PCR reaction mix and incubation for 30 minutes at 37°C.

Next, the DNA fragment of interest was obtained at high purity levels by gel extraction. This procedure was carried out to remove DNA fragments resulting from non-specific binding of primers in PCR. The PCR derived/DpnI treated DNA sample was analysed by DNA electrophoresis (section 2.1.4) and the DNA fragment that matched the expected fragment size was cut out of the gel. DNA fragments of size > 4,000 bp were purified using the QIAquick Gel Extraction Kit and DNA fragments of size < 4,000 bp were purified using the MinElute Gel Extraction Kit. The extraction procedure from the gel slab was carried out as proposed by the extraction kit manufacturer. Finally, the purified DNA fragment was quantified using the Nanodrop 1000 spectrophotometer.

DNA fragments obtained from PCR lack a 5' phosphate group. PNK phosphorylation reaction was carried out on DNA fragments for the addition of a 5' phosphate group that would enable the joining of DNA fragments in blunt end ligation reaction. PNK phosphorylation reactions were carried out using 100 ng of purified DNA fragment, 2 μl T4 DNA Ligase Reaction Buffer (10x) and 1 μl T4 Polynucleotide Kinase made to 20 μl total volume using ddH_20. The reaction was incubated for 30 minutes at 37°C which was followed by enzyme inactivation for 20 minutes at 65°C. Alternatively, purified DNA fragments from PCR were digested with restriction enzymes (restriction digestion reactions are described in the following section) to produce compatible ends for stick end ligation reaction. Following restriction enzyme digestion reaction, the DNA fragments were purified using the QIAquick PCR Purification Kit.

DNA fragments obtained from restriction enzyme digestion

Restriction enzyme digestion was used to produce DNA fragments with compatible ends for DNA assembly of plasmid constructs. DNA plasmids or PCR product DNA fragments were used as substrates for restriction enzymes. Reactions were set up with 2 μg DNA, 1 μl

restriction enzyme(s) of choice, 5 μl of the appropriate restriction enzyme buffer system as suggested by the manufacturer and made to a total volume of 50 μl with ddH_2O. Reactions were incubated for 2 – 3 hours at 37°C. Following digestion reaction, the DNA fragments were purified by gel extraction as described previously. Finally, the purified DNA fragment was quantified using the Nanodrop 1000 spectrophotometer.

DNA fragments obtained from oligonucleotide annealing

Pairs of chemically synthesized oligonucleotides of complementary sequences were annealed to produce DNA fragments with compatible ends for DNA assembly. This method was used when the desired DNA fragments were of small size <100 bp, but long enough to prohibit construction by PCR methods. Reactions were set up using 10 μl of each oligonucleotide of 100 μM stock concentration, 4 μl T4 DNA Ligase Reaction Buffer (10x), 1 μl T4 Polynucleotide Kinase made to 40 μl total reaction volume with ddH_2O. The reaction was incubated for 1 hour at 37°C. Next, the reaction volume was made to 400 μl with ddH_2O and heated to 100°C for 5 minutes. Finally, the reaction was left to cool to room temperature for 2 hours.

Ligation Reactions

DNA fragments obtained from any of the methods aforementioned were used in ligation reactions for DNA assembly into circular plasmids. Ligation reactions were set up with 50 ng of vector DNA fragment and insert DNA fragment mass that was equal to 2x the number of moles of vector DNA, 2 μl T4 DNA Ligase Reaction Buffer (10x) and 1 μl T4 DNA ligase to a total reaction volume of 20 μl. Ligation reactions were incubated for 30 minutes at room temperature in the case of sticky end ligation reactions or for 2 hours at room temperature in the case of blunt end ligation reactions.

Transformation of competent E. coli cells

Following DNA assembly in ligation reactions, bacterial transformations were carried out to isolate and identify correctly assembled plasmids. 50μl electrocompetent cells of *E. coli* DH10B strain were thawed for 5 minutes on ice and then were added 5-15μl of the ligation reaction mix. Next, the cell suspension was added to pre-chilled electroporation cuvettes (1mm gap) and mixed gently by pipetting. Then, the cell suspension was electroporated, and immediately after electroporation the cell suspension was added 950μl

LB-media and transferred into a 1.5 ml Eppendorf tube. The suspension was incubated for 1 hour at 37°C with shaking at 250 rpm. Dilutions of the suspension were then plated on LB-agar plates supplemented with appropriate antibiotic for the selection marker of the transformed plasmid construct to obtain colonies from single cells. The plates were incubated overnight at 37°C or until visible colonies were obtained.

2.1.6 Plasmid DNA isolation from bacterial cell cultures

E. coli cells transformed with plasmid DNA construct were used for plasmid DNA amplification and isolation. LB-medium (5-10 ml) supplemented with appropriate antibiotic were inoculated from single colonies from transformation plates. These bacterial cell cultures were grown overnight at (14-16 hours) at 37°C with shaking 250rpm and then plasmid DNA was isolated using the QIAprep Spin Miniprep Kit as described by the kit manufacturer. Finally, the purified plasmid DNA was quantified using the Nanodrop 1000 spectrophotometer.

2.1.7 DNA Sequencing Reactions

All DNA plasmids were sequence verified by DNA sequencing. DNA sequencing reactions were carried out by Source BioScience.

2.1.8 Glycerol Stock preparation of *E. coli* cell lines

Cell cultures in LB-medium (5ml) inoculated from a single colony were grown overnight (14-16 hours) at 37°C with shaking 250rpm with the addition of the appropriate antibiotic. The following morning, 0.5ml cell culture was mixed with 0.5ml sterile glycerol solution (40% glycerol) in a sterile Eppendorf tube and placed immediately at -80°C freezer.

2.2 Materials and methods for Chapter 3

2.2.1 Plasmid constructs for Chapter 3

Table 2.2 lists all plasmid constructs utilised in experiments in chapter 3 of this document. The table provides the plasmid construct name, a short description of the genetic elements encoded in each plasmid construct as relevant to this research and the origin of the plasmid construct. The plasmid construct sequences are provided in the appendices of this document.

Table 2.2: Plasmid constructs for Chapter 3

Plasmid Construct	Description	Origin / constructed
His-L_GFP /pPROEX-Htb	Expression of GFP protein with an N-terminal His-tag for easy protein purification under the control of a IPTG-inducible-promoter. Amp^R	Provided by Christ Hirst
12L-GFP_2	Expression of GFP_2 protein (GFP dimer with 12 amino acids peptide linker) with an N-terminal His-tag for easy protein purification under the control of a IPTG-inducible-promoter. Amp^R	(Hung El, 2014)
pNK301	Expression of GFP_2 protein (GFP dimer with 7 amino acids peptide linker) with an N-terminal His-tag for easy protein purification under the control of a IPTG-inducible-promoter. Amp^R	This work
pNVgfp	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Plac promoter. Cam^R	(Salema et al., 2013)
pNK302	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Bba_J23105 promoter and designed RBS TIR strength 12K. Cam^R	This work
pNK303	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Bba_J23106 promoter and designed RBS TIR strength 12K. Cam^R	This work

pNK304	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Bba_J23101 promoter and designed RBS TIR strength 12K. Cam^R	This work
pNK305	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Bba_J23105 promoter and designed RBS TIR strength 25K. Cam^R	This work
pNK306	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Bba_J23106 promoter and designed RBS TIR strength 25K. Cam^R	This work
pNK307	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Bba_J23101 promoter and designed RBS TIR strength 25K. Cam^R	This work
pNK308	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Bba_J23101 promoter and designed RBS TIR strength 50K . Cam^R	This work
pNK309	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Bba_J23101 promoter and designed RBS TIR strength 50K. Cam^R	This work
pNK310	Expression and surface display of an anti-GFP nanobody by a surface display vector based on intimin protein. Expression under the control of Bba_J23101 promoter and designed RBS TIR strength 50K. Cam^R	This work
pNVfib1	Expression and surface display of an anti-fibrinogen nanobody by a surface display vector based on intimin protein. Expression under the control of Plac promoter. Cam^R	Provided by Prof. Luis Ángel Fernández
pNVfib2	Expression and surface display of an anti-fibrinogen nanobody by a surface display vector based on intimin protein. Expression under the control of Plac promoter. Cam^R	Provided by Prof. Luis Ángel Fernández

2.2.2 Transformation of electrocompetent *E. coli* cells

E. coli cells were transformed with plasmid constructs for protein overexpression (*E. coli* BL21(DE3) strain) or expression of synthetic gene systems (*E. coli* DH10B strain). 50µl electrocompetent cells were thawed for 5 minutes on ice and then were added 1-10 ngl of known plasmid DNA. Next, the cell suspension was added to pre-chilled electroporation cuvettes (1mm gap) and mixed gently by pipetting. Then, the cell suspension was electroporated, and immediately after electroporation the cell suspension was added 950µl LB-media and transferred into a 1.5 ml Eppendorf tube. The suspension was incubated for 1 hour at 37°C with shaking at 250 rpm. Dilutions of the suspension were then plated on LB-agar plates supplemented with appropriate antibiotic for the selection marker of the transformed plasmid construct to obtain colonies from single cells. The plates were incubated overnight at 37°C or until visible colonies were obtained.

2.2.3 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE technique was carried out to analyse the constituents of protein samples which separate on the gel according to their size to charge ratio. Protein sample was prepared by adding an equal volume of Laemmli sample loading buffer (2x) and heated to 100°C for 5 minutes. Then, the prepared protein samples were loaded on a 4–20% Mini-PROTEAN® TGX™ Precast Protein Gel in TGS running buffer system. Electrophoresis was carried out at 200V until the dye front had reached end of the gel. Gel was stained using SimplyBlue Safe stain following the procedure recommended by the manufacturer in the product's user manual. Protein standard used was PageRuler™ Plus Prestained Protein Ladder. Electrophoresis results were documented on a Gel Doc XR system.

2.2.4 GFP and GFP_2 overexpression and protein purification

GFP and GFP_2 proteins were recombinantly expressed in an *E. coli* strain optimised for high protein expression, purified and stored appropriately for use in experimental procedures.

Plasmid constructs encoding devices for the overexpression of: i) a his-tagged monomeric GFP (His-L_GFP/pPROEX-Htb construct), ii) a his-tagged dimeric GFP joined by a 7 amino acid long linker peptide [GFP_2] (pNK301 construct) and iii) a his-tagged dimeric GFP

joined by a 12 amino acid long linker peptide [GFP_2] (12L-GFP2 construct) were used to transform electrocompetent cells of *E. coli* BL21(DE3) strain (section 2.2.2). Bacterial cell cultures prepared from single colonies on transformation plates were used to make glycerol stocks for each of the aforementioned cell line. These glycerol stocks were used to inoculate LB-media (10 ml volume) supplemented with ampicillin antibiotic and cultured overnight at 37°C with shaking at 250 rpm. The next morning, 1L LB-media supplemented with ampicillin antibiotic were inoculated from the overnight cultures at volume ratio of 1/100 times. The prepared cultures were then grown to an optical density value ($O.D._{600}$) of 0.6 at 600 nanometers at 37°C with 250 rpm shaking and then induced with 1mM IPTG. Following induction the cultures were grown overnight at 20°C. The next morning cultures were spun down to pellet cells for 20 minutes at 4,000 rcf and supernatant was removed.

Following protein overexpression, the protein purification procedures of nickel affinity chromatography and size-exclusion chromatography were used to obtain pure samples of the GFP proteins. The pelleted *E. coli* cell cultures induced for protein expression were resuspended in lysis buffer at a volume 1/100 times the initial culture volume. The cell suspensions were then added lysozyme (1mg/ml) and benzonase (50 units/ml) and incubated for 1.5 hour at room temperature with shaking. This was followed by freezing of the cell suspensions for 20 minutes at -80°C and then thawing the frozen suspensions for 2 hours at room temperature. Cell lysis was completed by the sonication of the cell suspensions (kept on ice) for 7 minutes at 50% amplitude (10 seconds pulse on, 20 seconds pulse off) using a Vibra Cell VCX 130 model sonicator. Then, the cell lysates were cleared from membrane debris by spinning the cell lysates for 30 minutes at 10,000 rcf. The cleared lysates were then mixed with NTA-Ni slurry (1 ml slurry / 4 ml cleared lysate) and incubated for 1 hour at 4°C with gentle shaking. Finally, the slurry mixed samples were loaded and immobilised onto a plastic column where they were washed 2 times with washing buffer (4 ml) and eluted 6 times with elution buffer. The washed and eluted fractions were run on a SDS-PAGE gel for confirming isolation of protein of interest. The eluted fractions with high mass of protein of interest, as determined by SDS-PAGE gel electrophoresis (section 2.2.3), were pooled together and passed through a gel filtration column (Superdex 75, 3, 000 – 75,000 Da separation range) at a flow rate of 1ml/minute to achieve high protein purity.

Protein purity of the gel filtration fractions was confirmed by SDS-PAGE gel electrophoresis (section 2.2.3) and pure fractions for protein of interest were pooled together.

Pooled protein samples from gel filtration purification were buffer exchanged with storage buffer (50 mM NaH_2PO_4 (pH8.0), 150 mM NaCl,) using protein concentrator sizes of 3,000 Da MWCO for GFP and 30,000 Da MWCO for GFP_2. The procedure for GFP_2 (7 amino acid peptide linker) was aborted as it showed no fluorescence under UV-excitation. Extinction coefficients used for protein quantification on Nanodrop 1000 spectrophotometer were 27,850 $M^{-1}cm^{-1}$ for GFP and 49,740 $M^{-1}cm^{-1}$ for GFP_2 (12 amino acid peptide linker construct) as calculated using ProtParam software tool (Wilkins et al., 1999). The concentration of purified protein samples were 2.42 mg/ml GFP (80 μM) and 1.05 mg/ml GFP_2 (18 μM). Subsequently the protein samples were stored at -80°C in PCR tube aliquots which were thawed for further experiments with the protein samples.

2.2.5 Fluorescence of purified GFP and GFP_2 samples

The fluorescent signal from equimolar samples of purified GFP and GFP_2 samples was measured to determine folding state of the engineered GFP_2 model biomarker.

Stock concentrations of purified GFPs were made to 1.8 μM by dilution into storage buffer solution (50 mM NaH_2PO_4 (pH8.0), 150 mM NaCl,). A 384-welll microplate was loaded with 50 μl samples of storage buffer solution, 1.8 μM purified GFP and 1.8 μM purified GFP_2. A microplate reader was used to measure the fluorescence signal emitted from the samples upon excitation. Samples excitation took place at 485nm and emitted signal was recorded at 520nm at room temperature with voltage gain setting of 1000. The fluorescent signal recorded for the storage buffer sample (background signal) was subtracted from the reading recorded for the GFP and GFP_2 samples. The fold-increase fluorescent signal reported was calculated by diving the background corrected signal for the GFP_2 sample by the background corrected signal for the GFP sample.

2.2.6 Cell toxicity assay

Cell toxicity assay was carried out to investigate any effects affecting the cell growth rate of *E. coli* cells upon expression of the cell-surface display vector protein carrying the anti-GFP nanobody (NVgfp protein construct).

Electrocompetent cells of *E. coli* DH10 cell strain were transformed with the pNVgfp construct and transformed cell lines were glycerol stocked (methods described in sections 2.2.2 and 2.1.8). Overnight cultures of *E. coli* cell lines were prepared by inoculating cells from the glycerol stock in antibiotic supplemented LB-medium (5 ml). The cultures were grown overnight at 37°C with shaking 225rpm. The following morning, the cultures were diluted 100-fold by inoculating fresh LB-medium from the overnight culture cell suspension overgrown under the same conditions to an $O.D._{600}$ value of 0.5 (3 hours – 3.5hours). A volume of 95µl of cell culture was then transferred to the wells of a 96-well microplate and induced with 5µl of IPTG at appropriate concentrations in triplicates. Final IPTG concentration of the cultures volume were 0mM, 0.1mM, 0.25mM, 0.5mM, 0.75mM and 1mM. The microplate was then incubated for 180 minutes at 30°C with shaking at 500rpm in a POLARstar microplate reader. Cell growth was monitored by measuring the $O.D._{600}$ value of the cell cultures. Readings were recorded at regular intervals of 10 minutes. The mean background absorbance value of LB medium was subtracted from the optical density values of samples.

2.2.7 NVgfp protein expression, cell localisation and nanobody functionality

The cell-surface display vector protein localisation at the *E. coli* cell surface and functionality of the displayed anti-GFP nanobody were assayed by live cell fluorescent labelling and recording of cell fluorescence.

Bacterial cell cultures of *E. coli* DH10 strain transformed with the pNVgfp construct were induced for NVgfp protein expression. Overnight cultures were prepared by inoculating antibiotic supplemented LB-medium (5ml) with cells from the glycerol stock of the aforementioned cell line. The cultures were grown overnight at 37°C with shaking at 225rpm. The following morning, the cultures were diluted 100-fold by inoculating fresh antibiotic supplemented LB-medium with cell suspension from the overnight culture. The fresh cultures were then grown for 2 hours at 37°C with shaking at 225rpm. Then, the cell culture was induced with 0mM, 0.1mM, 0.25mM, 0.5mM and 1mM and culturing continued for another 3 hours at 30°C with shaking at 225 rpm. Control sample of *E. coli* DH10 strain cells were treated in the same manner but cell cultures were prepared without the use of antibiotic.

Following NVgfp protein expression, the cells were labelled with anti-myc Ab or purified GFP, and cell fluorescence was recorded by flow cytometry. A volume of cell culture equivalent to 100 µl divided by the $O.D._{600}$ value of that culture were washed 3 times with PBS and were resuspended in 50 µl PBS. The cell suspension was added 50 ng Myc-Tag (9B11) Mouse mAb (PE Conjugate) (50µg/ml stock) or 1µl of purified GFP (80µM stock) and was incubated for 30 minutes at room temperature. Then, the labelled cells were washed again 3 times with PBS to remove excess label molecules and resuspended in 100µl PBS solution.

The cell fluorescence of the labelled cells was measured using a BD FACScan flow cytometer. The cells labelled with anti-myc ab were excited using 561nm laser and the cell fluorescence was measured with a 650 nm bandpass filter (FL5 channel). The cells labelled with purified GFP were excited using a 488nm laser and the cell fluorescence was measured with a 530nm bandpass filter (FL1 channel). The FlowJo software Vx0.7 was used for flow cytometry data analysis. Recorded events were gated for *E. coli* cells at forward scatter 108-609 au and side scatter 8 – 34 au. Then, gated DH10B population was divided into two subgroups; a negative fluorescence subgroup and a positive fluorescence subgroup according to the 99^{th} percentile cell fluorescence value of the control sample. Then, the mean cell fluorescence value was determined for the subpopulation with highest cell count and its coefficient of variance.

2.2.8 Quantification of displayed anti-GFP nanobodies

This methodology was used to quantify the number of anti-GFP nanobodies functionally displayed per cell.

E. coli DH10 strain transformed with the pNVgfp construct were induced with 1mM IPTG for NVgfp protein expression as previously described (section 2.2.7). Following expression, the growth medium was removed from the cell culture by pelleting cells and resuspending them in in PBS solution to make a cell suspension of $O.D._{600}$ value equal to 1. This cell suspension was left overnight at 4°C. The following morning, 200 µl of cell suspension were placed in individual 1.5 ml Eppendorf tubes each with increasing concentrations of purified GFP (80µM stock) as stated in the experimental results. The cell suspensions were incubated with GFP for 30 minutes at room temperature for binding, and

unbound/excess GFP was removed by washing the cells once with PBS and reconstituted in PBS solution.

The labelling procedure was followed by flow cytometry analysis on a BD FACScan flow cytometer that measured cell fluorescence using a 488 nm laser excitation source and detection was carried out with a 530nm bandpass filter (FL1 channel). FlowJo software was used for the data analysis. Using the software, recorded events were gated for forward scatter values 10 – 100 and side scatter values 20 – 100 that represented the cell population. The software was used to calculate mean cell fluorescence based on the gated cell population. The mean fluorescence value and the standard deviation was calculated for 3 technical repeats of the flow cytometry procedure for each GFP concentration datapoint. Finally, the Prism software was used to analyse the data by non-linear regression using receptor binding saturation models. Initially, the cell fluorescence data from flow cytometry were transformed to values of sensor occupancy. First step was to subtract the mean autofluorescence value (cell fluorescence of 0 picomoles GFP samples) from all data. Then, 0 picomoles GFP sample was fixed to 0% occupancy, while average fluorescence for 20-30 pmoles GFP was fixed to 100% occupancy. Finally, receptor binding saturation model: one site – specific binding with Hill slope was chosen among the candidate saturation models for its highest R square value in goodness of fit. The software, using the model selected, identified the GFP concentration for 50% sensor occupancy.

2.2.9 Levels of anti-GFP nanobody display in biosensor device variants

Flow cytometry was used to measure cell fluorescence of GFP-labelled biosensor device variants that correlates to the levels of anti-GFP nanobody displayed for each device variant.

E. coli DH10B strain cells transformed with the pNVgfp construct (prototype device) or one of the pNK302-pNK310 constructs (device variants v02-v10) were cultured for NVgfp protein expression and GFP labelled as described in section 2.2.8. The NVgfp protein expression protocol was slightly modified for device variants v02-v10. Following the initial 2 hour outgrowth the device variants were further cultured at 37°C for 3 hours with no IPTG induction and as they use constitutive promoters. Also, the cell suspension for all biosensor devices were labelled 160 picomoles of purified GFP for saturation of binding sites and not

with a range of GFP concentrations. Lastly, control sample of *E. coli* DH10 strain cells were treated in the same manner but cell cultures were prepared without the use of antibiotic.

The measurement of cell fluorescence of the prepared samples was carried out on an Attune NxT acoustic focusing cytometer. Cells bound with GFP were excited using a 488nm blue excitation laser and detection was carried out with the 530/30nm default filter. Flow cytometry data analysis was carried out in FlowJo software Vx0.7 tool. Recorded events were gated using the ellipse gate function to select for single cells at forward scatter (FSC-H) values 629 – 23,000 and side scatter (SSC-H) values 6,300 – 58,000. These values were chosen for providing a large coverage of the cell population events recorded. Then, the software tool was used to calculate cell fluorescence using the mean fluorescence function for each sample. The standard deviation of the mean cell fluorescence for each sample was calculated using Prism software tool using measured mean fluorescence values from 3 technical repeats of the flow cytometry measurements.

2.2.10 Plate agglutination assay

Plate agglutination assays were carried out to test whole cell biosensor devices displaying antibody fragments for their ability to form agglutinated complexes in the presence of appropriate concentrations of their target antigen protein molecule.

Agglutination assays were carried out using cell suspensions that were washed and resuspended in PBS solution at $O.D._{600}$ value of 1, kept overnight at 4°C and that have been previously cultured or induced for the expression of NVgfp protein construct (as appropriate for each device). 200 μl of cell suspension were added in each of the wells of a 96-well, round bottom microplate supplemented with a 50 μl of appropriate concentration protein sample as dictated by the experimental design of the particular experiment. Then, the microplate was incubated o/n (~18 hours) at room temperature and shaking (90 rpms).

Experimental results were documented using a mobile phone camera in one or more of the following photography shot set ups: i) microplate placed on top of a dark surface, ii) microplate placed on top of a blue LED transilluminator without the use of blue light filter and iii) microplate was placed on top of a blue LED transilluminator and the screen for blue light absorption was placed on top of the microplate.

2.2.11 Simulation of the agglutination reaction

The simulation of the agglutination reaction carried out in this research was based on a mathematical model developed and implemented in MATLAB software in collaboration (Hung El, 2014). Appendix B provides relevant parts from (Hung El, 2014) that includes the derivation of the bacterial agglutination reaction mathematical model, the model simplification, the parameter values used and the Matlab script for simulating the agglutination reaction. The script includes input parameter values from this study where these are different from (Hung El, 2014).

2.2.12 Fibrinogen stock solutions

Human Fibrinogen Plasminogen at stock concentration of 13.77mg/ml –as delivered by the supplier- was aliquoted in 0.5 ml Eppendorf tubes and stored at -80°C until needed. For each experiment a new aliquot was used.

2.3 Materials and methods for Chapter 4

2.3.1 Plasmid constructs for Chapter 4

Table 2.3 lists all plasmid constructs utilised in experiments in chapter 4 of this document. The table provides the plasmid construct name, a short description of the genetic elements encoded in each plasmid construct as relevant to this research and the origin of the plasmid construct. The plasmid construct sequences are provided in the appendices of this document.

Table 2.3: Plasmid constructs for Chapter 4

Plasmid Construct	Description	Origin / Constructed
pNK400	Encodes a GFP expression device. J23101 promoter, B0034 RBS, E0040 GFP gene, B0015 terminator, pSB1C3 plasmid backbone. Cam^R	Provided by Simon Moore
pNK401	Encodes a GFP expression device. T7 promoter, B0030 RBS, E0040 GFP gene, B0015 terminator, pSB1A2 plasmid backbone. Used as the control device in experiments. Amp^R	This work
pNK402	Encodes the prototype thrombin RNA-switch device. Derived from pNK401 with the addition of an anti-thrombin RNA-aptamer sequence. Amp^R	This work
pNK403 - 409	Encode the thrombin RNA-switch device variants devices v02-v08. Derived from pNK402 by mutating the stem helix that supports the anti-thrombin RNA-aptamer. Amp^R	This work
pNK410 - 412	Encode the VEGF RNA-switch device variants devices v01-v03. Derived from pNK401 by PCR methods. Amp^R	This work
pNK413 - 416	Encode the toggle RNA-switch device variants devices v01-v03. Derived from pNK401 by PCR methods. Amp^R	This work

2.3.2 Chemically synthesized RNA-aptamers

Table 2.4 lists the RNA-aptamers used in this studied that were chemically synthesized as RNA oligonucleotides by IDT.

Table 2.4: Chemically synthesised RNA-aptamers

Name	Sequence	Modification
Anti-thrombin RNA-aptamer	AAUUUGGAAGAUAGCUGGAGAACUAACCAAAUU (33 RNA bases)	5' Biotin
Anti-VEGF RNA-aptamer	GGAAGCUUGAUGGGUGACACACGUCAUGCCGAGCU (35 RNA bases)	5' Biotin
Toggle25t RNA-aptamer	GGGAACAAAGCUGAAGUACUUACCC (25 RNA bases)	None

2.3.3 In vitro transcription

In vitro transcription reactions were carried out to produce pure mRNA templates of RNA-switch devices. The *in vitro* transcription reactions used purified linear DNA template that was prepared by PCR (method described as part of section 2.1.5) and gel extracted (method described as part of section 2.1.5). Primers used in PCR reactions for the production of linear DNA template are listed in Table 2.5. The MEGAscript® T7 Transcription Kit was used to produce mRNA from the linear DNA template by using the methodology described by the manufacturer of the kit. Mass of linear DNA template used in the *in vitro* transcription reactions was 500ng and total reaction volume was 20 µl. The reactions were incubated for 2.5 hours at 37°C, followed by 1 hour DNAse treatment. Upon completion, mRNA was purified and isolated from the reaction mixture using the MEGAclear™ Transcription Clean-Up Kit. The procedure was carried according to the methodology suggested by the manufacturer of the kit with an mRNA elution step carried out using RNase treated ddH_2O. The concentration of the purified mRNA was quantified using the Nanodrop 1000 spectrophotometer.

Table 2.5: Primers used for *in vitro* transcription procedures

Primer name	Sequence	Information

VF2	tgccacctgacgtctaagaa	Forward primer used in pair with all the reverse primers. Binds upstream of T7-promoter of pNK401 plasmid and derivatives.
GFP ~150nts REV	gacagaaaatttgtgcccattaac	Reverse primer that binds within the GFP sequence. Used in pair with VF2 primer to produce DNA template coding for ~150 nucleotides long partial sequence mRNA
GFP ~250nts REV	tagtgacaagtgttggccatgg	Reverse primer that binds within the GFP sequence. Used in pair with VF2 primer to produce DNA template coding for ~250 nucleotides long partial sequence mRNA
B0010 terminator REV	ctctagtagagagcgttcaccg	Reverse primer that binds end of B0010 terminator part. Used in pair with VF2 primer to produce full length biosensor device mRNA

2.3.4 RNA gel electrophoresis

RNA gel electrophoresis was carried out to determine the size and purity of *in vitro* transcribed mRNA product. The mRNA products analysis was carried out in continuous PAGE system and the polyacrylamide percentage used for each experiment was adjusted according to the expected size of the mRNA product analysed. Also, the mRNA products were analysed under denaturing conditions to remove RNA secondary structures that could affect mRNA mobility in the gel. Urea was used as the denaturing agent in the RNA gels. mRNA samples were prepared for RNA gel electrophoresis by mixing the samples with RNA sample loading buffer, heating for 5 minutes at 80°C and placed immediately on ice. Then, 25 – 200 ng RNA from the ice cold samples was loaded on the RNA gel and gel electrophoresis was carried out for 25 minutes at 180 V at 4°C. Upon completion, the RNA gel was stained for 10 minutes in 2x GelRed solution dissolved in water. Gel electrophoresis results were visualized and recorded on a Gel Doc XR system. The RNA gel recipe, the buffer system recipe and the sample loading buffer recipe are provided in the appendices.

2.3.5 Proteins stock solution preparation

Thrombin from human plasma was received lyophilized from saline sodium citrate buffer (pH 6.5) and was reconstituted in ddH_2O to stock solution concentration of 2.165 μM. VEGF165 human recombinant, expressed in *E. coli* was received lyophilized with no additives and was reconstituted in ddH_2O to stock solution concentration of 2.62 μM. Supplier Information for can be found in Appendix A.

2.3.6 Binding reactions

Binding reactions were set up to allow mixtures of interacting molecules of RNA and protein to equilibrate. The RNA samples were prepared for binding reactions by unfolding them for 5 minutes at 80°C and then allowed to fold at room temperature. Binding reactions were carried out as following: i) RNA samples that included the anti-thrombin RNA-aptamer sequence were incubated for 30 minutes at 4°C in 1x HEPES buffer; ii) RNA samples that included the anti-VEGF RNA-aptamer were incubated for 15 minutes at 37°C or 30 minutes at 4°C (as stated in the experiment) in 1x PBS and iii) RNA samples that included the toggle25 RNA-aptamer sequence were incubated for 10 minutes at 37°C or 30 minutes at 4°C (as stated in the experiment) in 1x HEPES buffer. The protein concentration added to the binding reactions was as stated for the particular experiment.

2.3.7 Electrophoretic Mobility Shift Assay (EMSA)

EMSA was carried out to investigate interactions between RNA and protein molecules. RNA species used in EMSA experiments were: i) chemically synthesized RNA oligonucleotides or ii) purified mRNA template from *in vitro* transcription. Protein samples utilised were purchased as purified lyophilised proteins that had been reconstituted in solution (section 2.3.5).

The RNA and protein species under investigation were incubated in binding reactions (as described in section 2.3.6) before EMSA analysis. Next, the samples were added 2μl DNA Sample Loading buffer (6x) and loaded on the EMSA gel for analysis. Gel electrophoresis for EMSA analysis was carried out at constant current of 0.1 Amp, at 4°C until dye front reached 2/3 of the length of the gel. Following electrophoresis, to visualise nucleic acids the gel was stained in 2x GelRed solution for 10 minutes at room temperature and experimental results

were documented on a Gel Doc XR system. Then, to visualise protein molecules the gel was washed briefly with ddH_2O and stained in SyproRuby solution overnight at room temperature. The next morning, the gel was destained using Destain solution for 1 hour at room temperature and results were documented on a Gel Doc XR system. The signal intensity images acquired using the transilluminator were then overlaid and pseudocolored using ImageJ software. The recipe for the EMSA system utilised is provided in the Appendices. Various buffer systems were used for the set up and running of each EMSA experiment, and the suitability of each buffer system for the particular experiment was determined empirically for its ability to maintain RNA/protein interaction in the gel system. Additionally, a range of acrylamide percentage PAGE gel systems were used to set up the EMSA gel systems, and this was according to the size of the RNA species analysed in the particular experiment. All the EMSA gel parameters are provided in the experimental results section for each experiment individually.

2.3.8 Cell-extract expression assay using a commercial S30 kit

Cell-extract expression assays were carried out to investigate the functionality of RNA-switch devices under various conditions. The assay reactions were prepared using the S30 T7 High-Yield Protein Expression System. The reaction mix was set up with 9 μl T7 S30 Extract, 10 μl S30 Premix plus and made to 25 – 27.5 μl total reaction volume by mixing with the experimental samples derived from binding reactions (section 2.3.6). This mix was prepared on ice and then loaded in the wells of pre-chilled (-20°C) 384-well microplate. Due to the cell-extract being very viscous, the loaded microplate was centrifuged for 2 minutes at 2,000 rcf. At room temperature The microplate was also sealed with a breathable membrane to avoid evaporation. Then, the loaded microplate was loaded in a POLARstar microplate reader where the cell-extract reactions were incubated for 90 minutes at 37°C. GFP expression in samples was monitored at regular time intervals by recording the emitted fluorescent signal using a 485-12nm excitation filter and a 510-10nm emission filter.

2.3.9 TX-TL *E. coli* cell extract production

A protocol was implemented for the production of *E. coli* cell-extract in-house for use in cell-extract expression assays. The protocol was implemented as described by (Sun et al., 2013) with a few modifications. The *E. coli* cell strain BL21(DE3) was used instead of the

Rosetta strain for producing the cell mass. Additionally, on day 3 the cell culture was induced with 1mM IPTG at $O.D._{600}$ value of 0.2 for T7-phage RNA-polymerase expression. Furthermore, on day 4 the cell lysate was produced by the sonication method rather than lysing the cells using bead beating. For sonication, the resuspended cell suspension was aliquoted in 1 ml fractions in 2 ml Eppendorf tubes and sonicated for 1.5 minutes (2 seconds pulse on, 2 seconds pulse off) on ice at 65% amplitude setting using the Vibra Cell VCX 130 model sonicator (130 Watt ultrasonic processor). The sonicated samples were then consolidated and cell debris was removed by centrifution at 16,000 rpm for 15 minutes at 4°C.

2.3.10 Cell-extract expression assay using homemade cell-extract

Cell-extract expression assays were carried out to investigate the functionality of RNA switch devices under various conditions. The assay reactions were prepared using the cell-extract made in-house using *E. coli* BL21(DE3) cell strain (section 2.3.9). The reaction mix was set up with 3.3 µl cell-extract, 4.2 µl reaction buffer (2.5mM Magnesium Glutamate, 100mM Potassium Glutamate) and made to 10 µl (or more where necessary) total reaction volume by mixing with the experimental samples derived from binding reactions (section 2.3.6). This mix was prepared on ice and then loaded in the wells of pre-chilled (-20°C) 384-well microplate. Due to the cell-extract mix being very viscous, the loaded microplate was centrifuged for 2 minutes at 2,000 rcf at room temperature. The microplate was also sealed with a breathable membrane to avoid evaporation. Then, the loaded microplate was loaded in a POLARstar microplate reader where the cell-extract reactions were incubated for 90 minutes at 37°C. GFP expression in samples was monitored at regular time intervals by recording the emitted fluorescent signal using a 485-12nm excitation filter and a 510-10nm emission filter.

2.3.11 Data analysis for cell-extract assays

Raw data from cell-extract assays were analysed as following unless otherwise stated. Error bars represent standard deviation when technical repeat data were available. For calculating fold change in GFP expression, the following equation was used: fold change = (fluorescence value for condition 1 – background fluorescence value) / (fluorescence value for condition 2 – background fluorescence value), where background fluorescent was the

signal of the cell-extract mix for that experiment. Data analysis in section 4.2.6 for GFP expression cell-extract assays incorporated an additional calculation, where the background corrected fluorescence values for GFP expression were divided by the concentration of the purified mRNA template for each device variant (measured on the spectrophotometer in µg/ml). This extra transformation normalised GFP expression for input mRNA template as 2.5µl of purified mRNA template was added to cell-extract reaction without regard to template concentration to streamline the process and reduce time of experimental setup.

Chapter 3

3 Biosensor design: Nanobody & whole cell chassis

3.1 Background Information

3.1.1 Antibodies and antibody fragments

Antibodies are large glyocoprotein molecules produced and used by the immune system to recognise foreign agents entering the body and eliminate them. An antibody is able to recognise an area on the target molecule through non covalent interactions such as ionic and hydrogen bonds, van der Waal's forces and hydrophobic interactions (Padlan et al., 1989). Using this binding event, the antibody tags the foreign agent entering the body for attack by other parts of the immune system or neutralises it directly by blocking some of its essential function (George, 2000). The target molecule recognised by an antibody molecule is commonly referred as the antigen while the specific region recognised by the antibody on the antigen molecule is referred as the epitope site. Also, the specific region of the antibody that binds the epitope site is referred as the paratope region. The paratope region of antibodies can be made to recognise a wide variety of antigens with high specificity by a combinatorial process called V(D)J recombination (Z. Li, Woo, Iglesias-Ussel, Ronai, & Scharff, 2004). The high degree of specificity towards antigen recognition along with the development of technologies that allow antibody production with specificity for any target of choice established antibodies and immunoassays as a leading technology in the development of medical diagnostic testing assays (Siddiqui, 2010).

Conventional antibodies –such as those of human origin- are made of four polypeptide chains, two heavy chains (H) and two light chains (L), which are held together by both noncovalent interactions and disulphide bonds (Figure 3.1) (George, 2000). Structurally, antibody molecules chains are divided into constant domains (C), which have the same sequence among antibody molecules of the same subclass, and variable domains (V) which are found on the top extremities of the Y-shape structure and show sequence variability. Most of the variability in the V domain is confined to three regions termed complementarity-determining regions (CDRs) (Kabat, Wu, & Bilofsky, 1978). The three CDRs

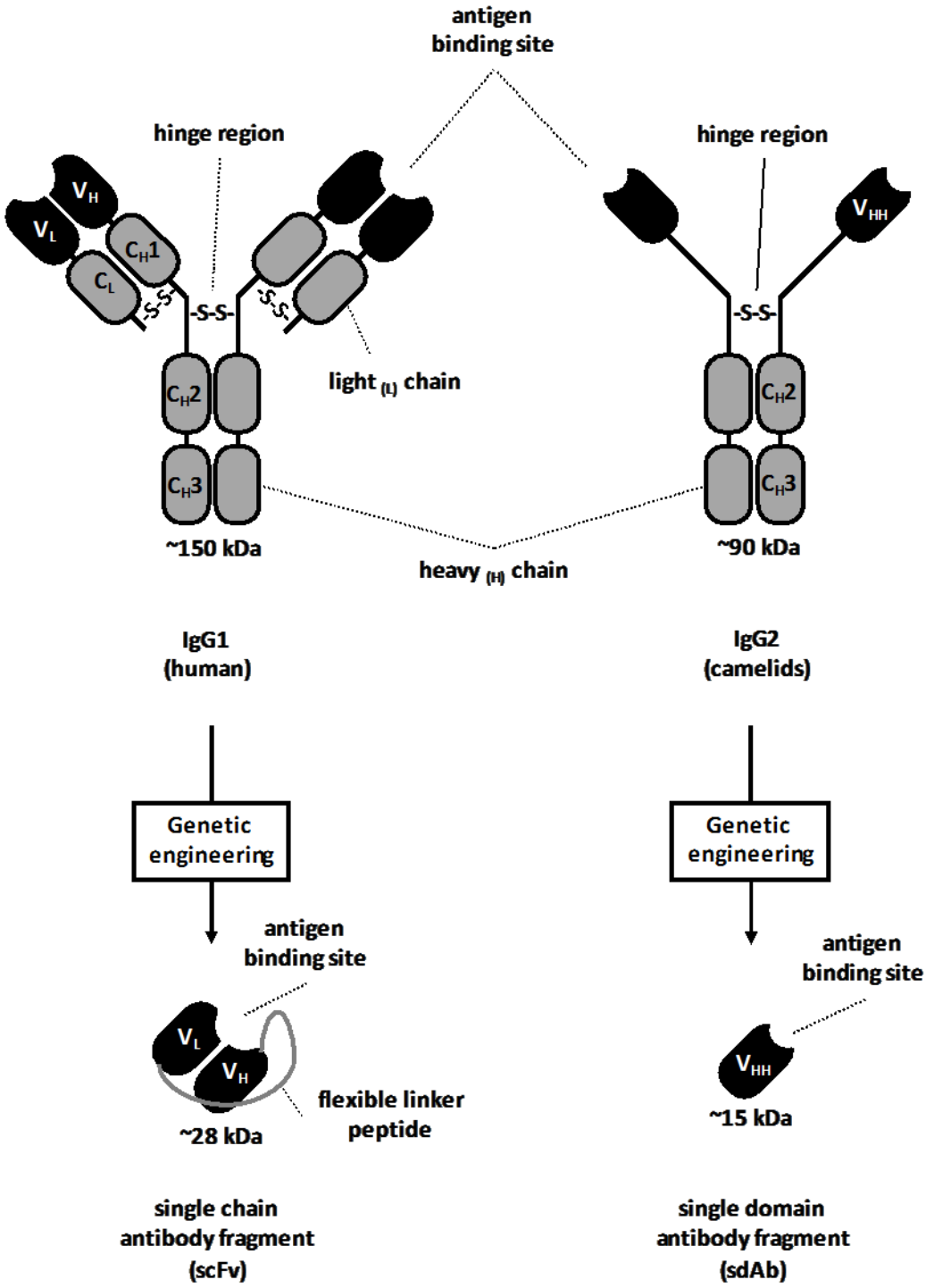


Figure 3.1: Structure of antibody molecules. Conventional antibody versus atypical antibody of *Camelidae*, and their genetically engineered antibody fragments.

from the V_H domain and the three CDRs from the V_L domain come together in the three-dimensional space to form the antigen (or paratope) binding site. In contrast to this convectional antibody structure, types of antibodies found in the sera of Camelidae (Hamers-Casterman et al., 1993), are completely devoid of light chain polypeptide and lack the C_H1 domain of the heavy chain (Figure 3.1). In these heavy chain antibodies (HCAbs), the variable N-terminal domain, named V_{HH} is solely responsible for the formation of the antigen binding site. The The V_{HH} domain includes three CDRs for the formation of the antigen binding site, as do the conventional V_H domains, but to compensate for the lack of the three CDRs provided by the V_L domain in conventional antibodies, two of the CDR loops of the V_{HH} domain are longer to provide sufficiently large antigen interacting surface (Nguyen, Hamers, Wyns, & Muyldermans, 2000). The extra flexibility of the longer loops, which is counterproductive for binding is usually dealt with a disulfide bond between the two CDR loops (Govaert et al., 2012) although in llama V_{HH} this occurs less frequently (Harmsen et al., 2000).

A number of technologies have been developed over the years for the on demand production of a high affinity antibody recognising a target antigen of choice with high specificity. Traditionally, target specific antibodies were isolated by the injection of the antigen in an animal, such rabbit, goat, sheep among others, which elicited an immune system reaction that resulted in the activation and clonal expansion of antigen-reactive B lymphocytes (Ritter, 2000). Polyclonal antibodies were then purified from the serum of the animal. Alternatively, hybridoma technology was developed for the continuous production of monoclonal antibodies in cell culture (Ritter, 2000). Activated B-lymphocytes were collected and fused to immortal myeloma cells, and a single parent hybridoma cell line was isolated that produced antibodies of single clonality that recognised a specific epitope on the antigen (Kohler & Milstein, 1975). Recombinant antibodies have also been produced by cloning antibody heavy and light chain genes (Orlandi, Gussow, Jones, & Winter, 1989) from antibody producing cell types to produce chimeric antibodies (Morrison, Johnson, Herzenberg, & Oi, 1984), humanised antibodies (Jones, Dear, Foote, Neuberger, & Winter, 1986) as well as genetically engineered antibody fragments (Glockshuber, Malia, Pfitzinger, & Pluckthun, 1990). For the selection of high affinity antibodies of single clonality by by-passing of hybridoma technology and animal immunisation, technologies were developed to

replicate the immune system selection process *in vitro* by phage display (Clackson, Hoogenboom, Griffiths, & Winter, 1991). In the aforementioned study, combinatorial libraries of antibody genes were genetically engineered into single chain antibody fragments and displayed on a minor phage coat protein to select high affinity antibodies.

For some uses such as for diagnostic applications and for some therapeutic applications the recruitment effector capabilities of the antibody which are mediated through the Y-shape stem Fc domain are not needed, and instead antibody fragments that include just the antigen binding arms of the Y-shape molecule are utilised (Verma & Boleti, 2000). Genetic engineering techniques have been used for the fusion of the V_H and the V_L domains of the conventional antibody structure into a single polypeptide chain by a flexible linker domain (Glockshuber et al., 1990). Such engineered constructs have been termed single chain antibody fragments (scFv) (Figure 3.1) and can be expressed in low cost protein expression systems such as *E. coli* (Glockshuber et al., 1990). For the expression of scFvs in *E. coli* two strategies have been usually employed; the first strategy being the expression of the antibody fragment in the periplasm space of *E. coli* cell (Glockshuber et al., 1990) and the second strategy being the cytosolic expression of the antibody fragment in inclusion bodies and their refolding *in vitro* (Lee, Park, Park, & Kwak, 2002). These strategies are necessary because scFVs are often poorly soluble and are prone to aggregation (Holliger & Hudson, 2005). In contrast, recombinant antibody fragments from Camelidae HCAbs, that have naturally evolved to be single domain, have been shown to express well and in a soluble form in *E. coli* expression system (Arbabi Ghahroudi, Desmyter, Wyns, Hamers, & Muyldermans, 1997). These recombinant antibody fragments derived from the variable heavy chain domain (V_{HH}) of HCAbs (Figure 3.1) have been termed single domain antibodies (sdAbs) or nanobodies.

In this chapter, a sdAb/nanobody that recognises an epitope site on the GFP protein molecule was utilised as the sensor element in developing a synthetic biology design biosensor device for protein detection. This anti-GFP nanobody was developed previously by others (Rothbauer et al., 2006) by the immunisation of *Lama pacos* with purified GFP. The nanobody was isolated from the animal by: i) harvesting the lymphocyte cells, ii) cloning the antibody V_{HH} sequences to form a library and iii) panning the phage displayed library of antibody fragments (Rothbauer et al., 2006). Following nanobody isolation,

surface plasmon resonance measurements were used to calculate its dissociation constant for GFP at 0.23 nM (Rothbauer et al., 2006). Additionally, this chapter also describes the use of two anti-fibrinogen nanobodies that were used as sensing elements for the development of fibrinogen synthetic biology biosensor devices. These anti-fibrinogen nanobodies were developed by the immunisation of *Camelus dromedarius* with purified fibrinogen (Campuzano et al., 2014), and were isolated by a methodology similar to the one previously outlined for the anti-GFP nanobody.

3.1.2 Engineering surface display in *E. coli*

Surface display is the presentation of a molecule of interest on the outside of the cell by its attachment to the outermost cell layer from where the molecule can interact with the extracellular environment. *E. coli* cells are enveloped by a double membranous structure separated by the periplasmic space (Ruiz, Kahne, & Silhavy, 2006). The inner membrane is directly in contact with the cytoplasm and the periplasm, whereas the outer membrane separates the periplasmic space from the extracellular environment. Thus, surface display systems in *E. coli* have been engineered to attach on the outer membrane of the cell (van Bloois, Winter, Kolmar, & Fraaije, 2011). This gives the displayed molecule free access to reagents, substrates or molecules without the need for them to cross the double membrane barrier. Biotechnological applications of surface display technologies have been applied for biocatalysis (Baek, Han, Lee, & Lee, 2010), protein library screening (Binder, Matschiner, Theobald, & Skerra, 2010), vaccine development (Leclerc, Charbit, Martineau, Deriaud, & Hofnung, 1991), antibody library screening and bioremediation (Zhang et al., 2004) among others.

The most frequently utilised systems for the surface display of polypeptides on the *E. coli* surface include members of the autotransporter group of proteins, for their ability to display large and complex proteins (van Bloois et al., 2011). The organisational structure of the autotransporter proteins with modular domains allows for the straightforward exchange of the displayed polypeptides for different applications (Celik et al., 2012; Rutherford & Mourez, 2006). The autotransporter polypeptide is essentially organised in three domains: the N-terminal signal sequence, the passenger domain and the translocation domain at the C-terminus (Henderson, Navarro-Garcia, Desvaux, Fernandez, & Ala'Aldeen, 2004). The N-terminal signal peptide targets the polypeptide to the inner membrane for its further export to the periplasm. The passenger domain is the one that confers specific functionality to the autotransporter protein and this is the domain that is usually exchanged for a protein of interest in biotechnology applications. The C-terminal translocation domain is the one that facilitates the translocation of the passenger domain through the outer membrane. It adopts a beta-barrel tertiary structure that embeds in the outer membrane, while it is connected to the passenger domain though a short linker sequence of helical secondary structure that spans the pore of the barrel (van den Berg, 2010). Autotransporter proteins

make use of Type V protein secretion pathway of Gram-negative bacteria and more specifically the type Va autotransporter system (Henderson et al., 2004). The classical (type Va) autotransporter protein biogenesis pathway is diagrammatically presented in Figure 3.2. The signal peptide of the autotransporter polypeptide targets the protein to the Sec translocon machinery located in the inner cytoplasmic membrane and which transfers the unfolded polypeptide into the periplasmic space (Natale, Bruser, & Driessen, 2008). In the periplasm the polypeptide is maintained in the unfolded state and undergoes processes of quality control by its interaction with chaperones SurA and Skp and chaperone/protease DegP (Leo, Grin, & Linke, 2012). Then, the autotransporter polypeptide, most often through its interaction with the Bam complex in the outer membrane, adopts the beta-barrel tertiary structure which embeds in the outer membrane and transfers the passenger domain through it to achieve surface display (Costa et al., 2015).

Attempts have been made to merge bacterial surface display technology and antibody engineering technology. In most cases, studies used other surface display systems than the autotransporter systems described above. Early reports for the bacterial cell surface display of engineered antibodies used lipoprotein vectors, such as the PAL lipoprotein for the display of an anti-lysozyme scFv (Fuchs, Breitling, Dubel, Seehaus, & Little, 1991) in *E. coli* and the Lpp-OmpA display system for the display of an anti-digoxin scFv. (Francisco, Campbell, Iverson, & Georgiou, 1993). The Lpp-OmpA system was later utilised for the selection of a high affinity anti-digoxin scFv antibody from a small light chain library using flow cytometry screening (Daugherty, Chen, Olsen, Iverson, & Georgiou, 1998). A different display vector based on the INP lipoprotein was used for the display of an anti-c-myc scFv (Bassi, Ding, Gloor, & Margaritis, 2000). A radically different approach was taken with the Anchored Periplasmic expression (APEx) technology (Harvey et al., 2004). The scFV was anchored to the periplasmic side of the inner bacterial membrane and then the outer membrane was stripped away to allow ligand access to antibodies. Finally, only more recently autotransporter systems were used for the surface display of antibody fragments, as in the case of the IgA protease autotransporter that was used for the display of scFv and

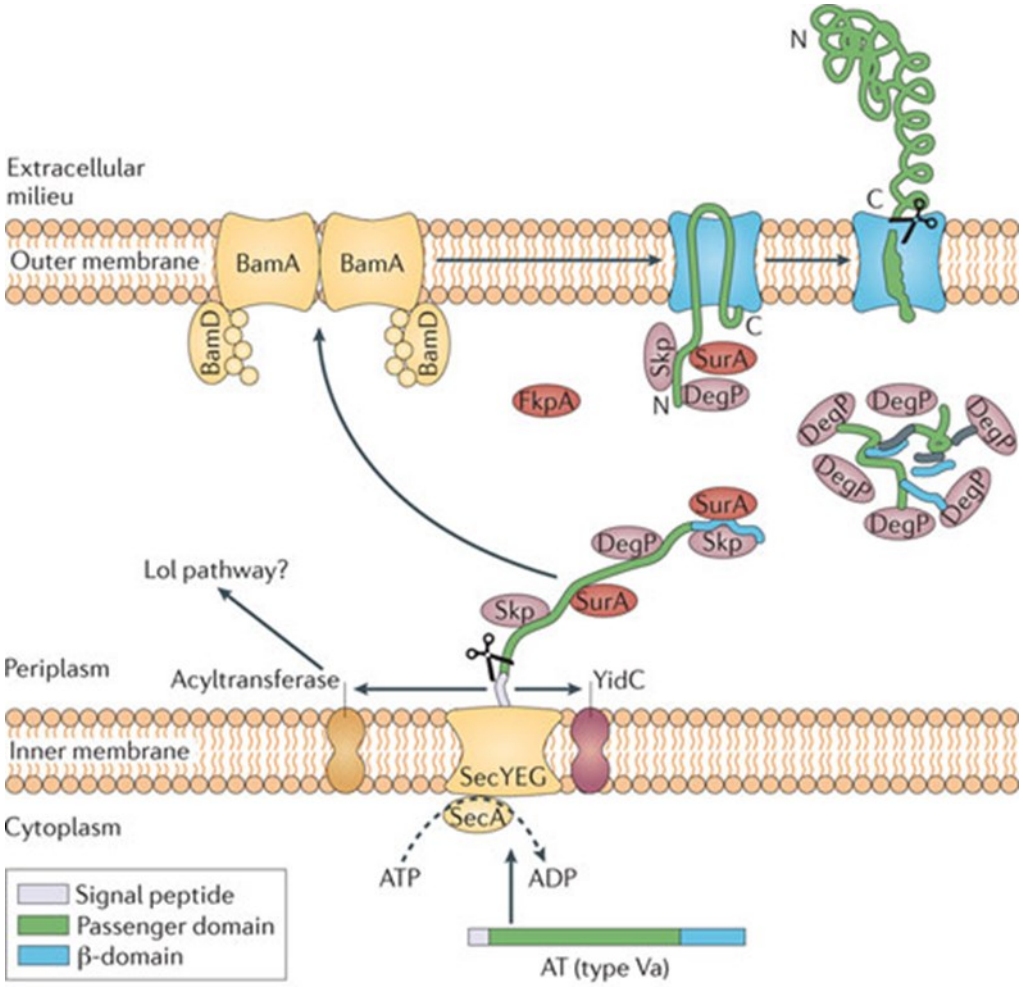


Figure 3.2: Classical (type Va) autotrasporter protein biogenesis pathway. The type Va autotransporter polypeptide is essentially organised in three domains: the N-terminal signal sequence, the passenger domain and the translocation domain at the C-terminus. It uses the Type V secretion pathway of *E. coli* to display (or secrete) its passenger domain to the cell surface. The signal peptide of the autotransporter polypeptide targets the protein to the Sec translocon machinery located in the inner cytoplasmid membrane and which transfers the unfolded polypeptide into the periplasmic space. In the periplasm the polypeptide is maintained in the unfolded state and undergoes processes of quality control by its interaction with chaperones SurA and Skp and chaperone/protease. Then, the autotransporter polypeptide, most often through its interaction with the Bam complex in the outer membrane, adopts the beta-barrel tertiary structure which embeds in the outer membrane and transfers the passenger domain through it to achieve surface display. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Microbiology (Leyton, Rossiter, & Henderson, 2012), copyright (2012).

sdAb antibody fragments (Veiga, de Lorenzo, & Fernandez, 2004).

In this chapter, an autotransporter display vector based on the intimin protein was used for the surface display of sdAbs in *E. coli* cell host. Intimin, encoded by the eaeA gene, is a virulence factor located in the outer membrane protein of enteropathogenic (EREC) and enterohemorrhagic *E. coli* that attaches the bacterium to the intestinal epithelial cells (Luo et al., 2000). The protein includes a N-terminal signal sequence with a LysM domain, a β-barrel domain that is followed by three immunoglobulin-like domain and a c-terminal c-type lectin domain (Fairman et al., 2012). The modular domain organisation of the intimin protein with a signal peptide, a translocation and a passenger domain matches that of autotransporter proteins, but the translocation and passenger domain order is inverted. Investigations into the secretion mechanism of Intimin showed that the protein follows the same rules that apply for autotransporter secretion through the Type V secretion pathway, and thus Intimin is classified as a Type Va (autotransporter) secretion system protein or a novel Type Ve secretion system (Oberhettinger et al., 2012). Intimin has been recombinantly engineered as a surface display vector and has been used for the display of an anti-GFP nanobody and to enable the screening an immune library (Salema et al., 2013). This display system was utilised in this project for the display of sdAbs.

3.1.3 The agglutination reaction as an immunologic diagnostic technique

The agglutination reaction is a chemical phenomenon in which interactions between particles in a mixture cause the particles to stick together and form a bigger complex (Price, 2001). In the particular case that the agglutination reaction is driven by specific interactions, as in the instance of antibody/antigen interaction, the phenomenon can be exploited for the development of diagnostic tests. Diagnostic immunological tests are well established in clinical settings and are used for the identification of microorganisms (Ellis & Sobanski, 2000), in blood typing (Mujahid & Dickert, 2015) and virus detection (Tortora, Funke, & Case, 2012) among others. Figure 3.3 panels A & B show the visual output of a special type of agglutination reaction, hemagglutination, where the agglutination of red blood cells by virus particles enables the identification of virus strains(Eisfeld, Neumann, & Kawaoka, 2014).

A requirement for an agglutination reaction is that the species of particles involved in the reaction have multiple sites of interaction for each other (Price, 2001). This principle can be demonstrated with a system of two interacting species, species A and species B. In this hypothetical example, species A is a bacterial cell covered with a large number of surface antigens and species B takes the form of a bivalent antibody with specificity for the antigen presented by species A. Thus, species B is able to crosslink two species A particles forming a larger complex. Under favourable conditions, more species B particles can attach to the A-B-A sandwich complex by binding to the multivalent A particle, followed by the attachment of more A particles and so on. This will eventually lead to a large complex formation.

The concentration of species A antigens relative to the concentration of species B are important factors in determining the progression of an agglutination reaction. The relationship between the concentration of two species that is generally applicable to any immunoagglutination reaction is described by the Heidelberger-Kendall curve (Figure 3.3) (Price, 2001). The curve can be divided into 3 zones; the middle zone that is called the equivalence zone represents the species concentration where cross-linking and agglutination is at the maximum, while the zones on either side of the equivalence zone represent species concentrations where cross-linking is reduced due to excess concentration of one or the other species. For example, in the hypothetical example previously described:i) when there is excess concentration of bacterial antigens, then a large

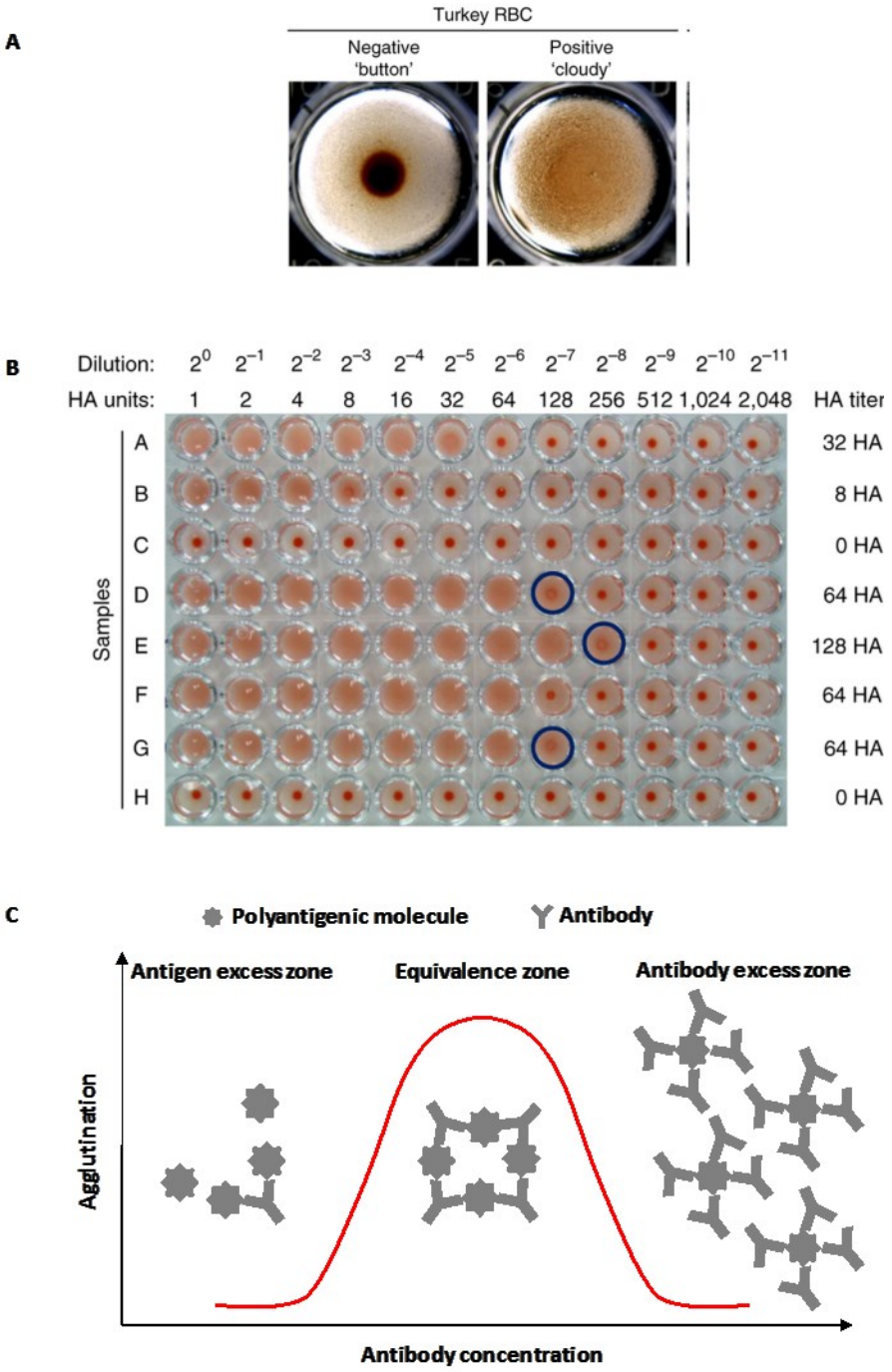


Figure 3.3: Agglutination assay. A Close range image of negative and positive results of hemagglutination assay in U-bottom microplate wells. **B** Plate hemagglutination assay for the identification and quantification of Influenza A virus strains. **C** The Heidelberger-Kendall curve generally applicable to agglutination reactions. [For parts A & B: Reprinted by permission from Macmillan Publishers Ltd: Nature Protocols(Eisfeld et al., 2014), copyright 2014.]

number of the bacterial cells will not be agglutinated because of shortage of cross-linking antibody molecules and ii) when there is excess of antibody molecules, antigens of bacterial cell will be saturated by antibodies and no cross-linking will occur as there would be no available antigen for binding.

Agglutination can be detected by visual observation or by light scattering techniques such as turbidimetry and nephelometry (Newman, Henneberry, & Price, 1992). For the detection of agglutination by visual observation, it has been determined that the species of particles used in the agglutination reaction needs to be of large size in the range of hundreds of nanometers to allow sensitivity of assay tests in the order of picograms (Molina-Bolívar & Galisteo-González, 2005). Antibodies or antigens are frequently attached to polystyrene polymer particles (latex particles) for enhancing the detection of antibody/antigen complexes in immunological tests known as latex agglutination tests (Ortega-Vinuesa & Bastos-Gonzalez, 2001). Various methodologies have been employed for the attachment of antibodies to the latex particles including physical adsorption or covalent linkage (Molina-Bolívar & Galisteo-González, 2005) but this often leads to colloidal instability of the latex particles and non-specific agglutination (Ortega-Vinuesa & Bastos-Gonzalez, 2001).

This chapter investigated the feasibility of developing an immunological test based on the principle of agglutination using engineered *E. coli* cells. The *E. coli* cells surface displayed recombinant antibody fragment molecules and thus constituted one of the multivalent species of the agglutination reaction. Also, the *E. coli* cell with size dimensions of 0.84μm in diameter and 2.83μm in length (P. Y. Liu et al., 2014) enabled the visual detection of agglutination test results.

3.2 Experimental Results

3.2.1 Biosensor design

A biosensor design was developed for the detection of protein analytes with the use of nanobody technology and *E. coli* cell chassis. Proteins are large biomolecules with charged and/or hydrophilic surfaces that are unable to cross the plasma membrane barrier. To enable the specific and sensitive detection of protein analytes at the extracellular milieu, the developed design incorporates *E. coli* whole cells with surface displayed nanobody molecules that serve as the sensor elements of the biosensor device.

The biosensor design relies on the presence of multiple sites of interaction between the engineered *E. coli* cells and the protein analyte molecule. The display of multiple copies of a nanobody sensor element on the surface of *E. coli* cell would constitute the cell multivalent towards its target analyte. This biosensor design also necessitates the analyte molecule to be multivalent towards the nanobody sensor. Analytes that would be suitable targets compatible with this biosensor design are molecules that include two or more epitope sites recognised by the nanobody sensor element; for example homodimeric proteins or other proteins with repeated domains. Alternatively, any other analyte that does not meet this specification could be targeted by this biosensor design with the use of *E. coli* cells that display two or more nanobodies with different epitope specificities for that target.

In the presence of the target analyte of interest in the extracellular milieu, biosensor devices of *E. coli* cells with displayed nanobodies specific for that target would be able to bind the analyte molecules. Due to the multivalency of both the biosensor *E. coli* cell and the analyte molecule, a second *E. coli* cell would be able to bind the same analyte molecule. This would result in a sandwich complex of cell/analyte/cell. Successive sandwich complexes formation in the reaction, as well as complex formation between formed sandwich complexes should result in large aggregates of bacteria and analyte molecules. The formation of bacterial aggregates would indicate the positive result of the test for analyte detection. The aggregates would form the observable output detected by the user by visual examination or other method. In the absence of analyte bacterial cells will remain as individual units. The proposed mechanism and proposed test assay workflow for this biosensor design are illustrated in Figure 3.4.

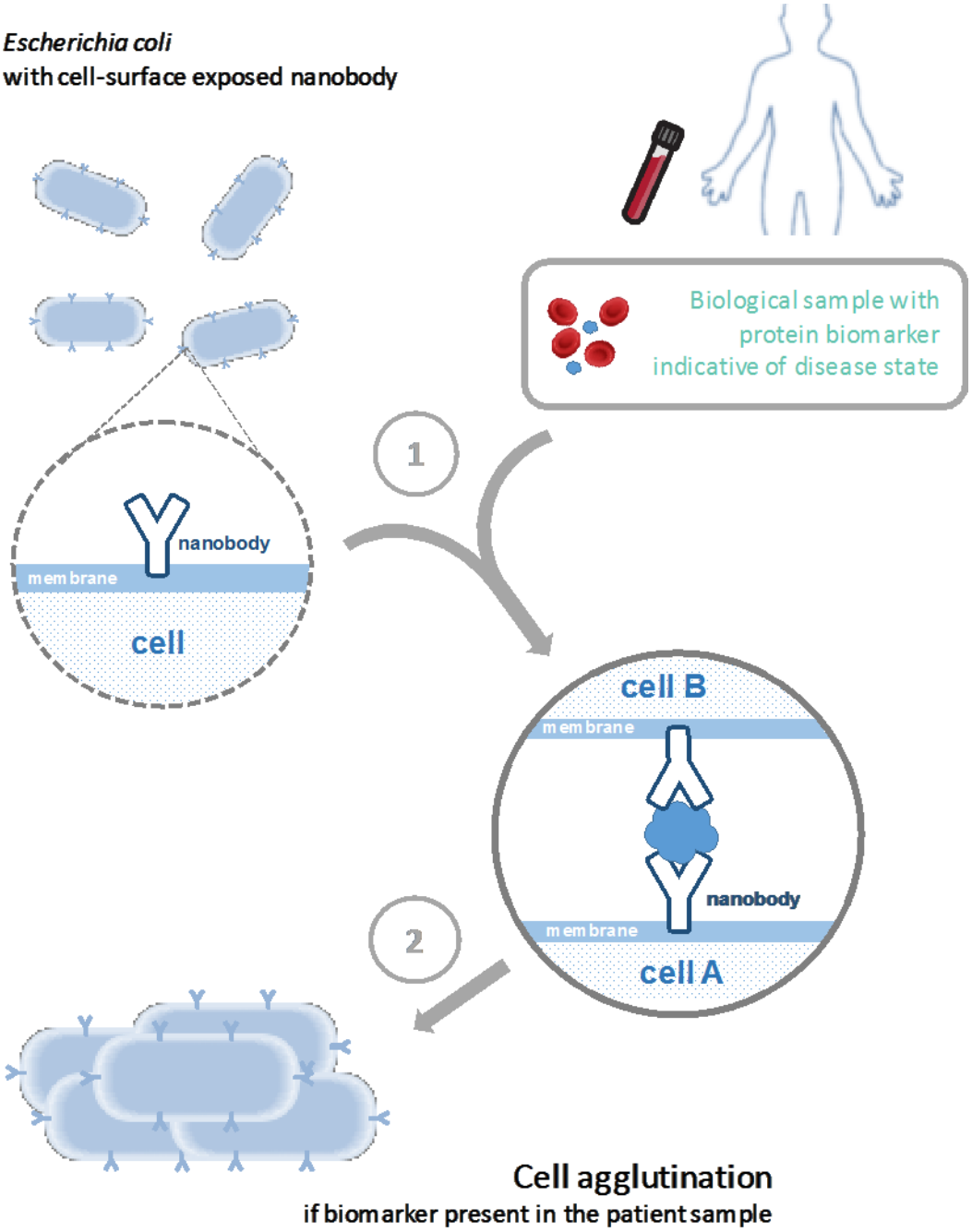


Figure 3.4 Mechanism and test assay work flow of nanobody/whole-cell biosensor design. This schematic illustrates the principles of the biosensor design. *E. coli* cells with cell surface displayed nanobodies expressed from designer DNA plasmids are mixed with biological sample (eg. blood) that can include protein biomarker indicative of a disease state. In the presence of the biomarker, this causes the crosslinking of *E. coli* cells in solution in a sandwich format by nanobody recognition of two epitopes on protein surface. Extensive cross-linking between cells and protein molecules leads to cell agglutination (or aggregation) that changes the visual appearance or optical properties of the test sample, and thus reporting test result to the user.

3.2.2 Prototype design and device construction

For testing the proposed biosensor design, a prototype device was designed and constructed. The prototype device would surface display an anti-GFP nanobody that would serve as the sensor element of the device. The anti-GFP nanobody would be displayed by the NVgfp display vector (Salema et al., 2013). This vector display system was chosen as it is one of the few systems reported in the literature that successfully surface displayed a recombinant antibody fragment in *E. coli* cell chassis. Also, the fluorescence of the target GFP molecule made experimental testing convenient in standard molecular biology lab facilities. The protein selected to be the model analyte for testing the biosensor design was dimeric GFP. This is a genetically engineered fusion protein of two monomeric GFP molecules that provides two identical epitope sites per protein molecule and thus exhibits multivalency. The host chassis in which the biosensor device design would be implemented was the *E. coli* cell chassis. The *E. coli* cell would act as the production platform of the anti-GFP nanobody. Additionally, the aggregation of *E. coli* cells would enable the reporting of the analyte recognition event to the user of the biosensor, effectively acting as the output or reporter module of the device.

The NVgfp surface display system employed by the prototype design is presented in Figure 3.5 (Salema et al., 2013). The intimin β-barrel domain that serves as the carrier domain is located in the N-terminus and facilitates the incorporation of the construct into the outer membrane of *E. coli*. The passenger domain, that incorporates the anti-GFP nanobody (Vgfp), is located in the C-terminus of the fusion protein. The display system also encodes for a N-terminal signal peptide that targets the protein construct to the sec-secretion pathway for translocation to the periplasmic space. Additionally, the construct includes a LysM domain that is predicted to bind the peptidoglycan layer, an Ig-like domain that may help the nanobody folding and display, and an E-tag and C-terminal myc-tag to facilitate detection if needed.

The plasmid map of the pNVgfp construct is presented in Figure 3.5 and shows that the NVgfp expression is regulated by an IPTG inducible Plac promoter. This pNVgfp construct was supplied by Prof. Luis Ángel Fernández of the National Center for Biotechnology (CNB), Madrid,

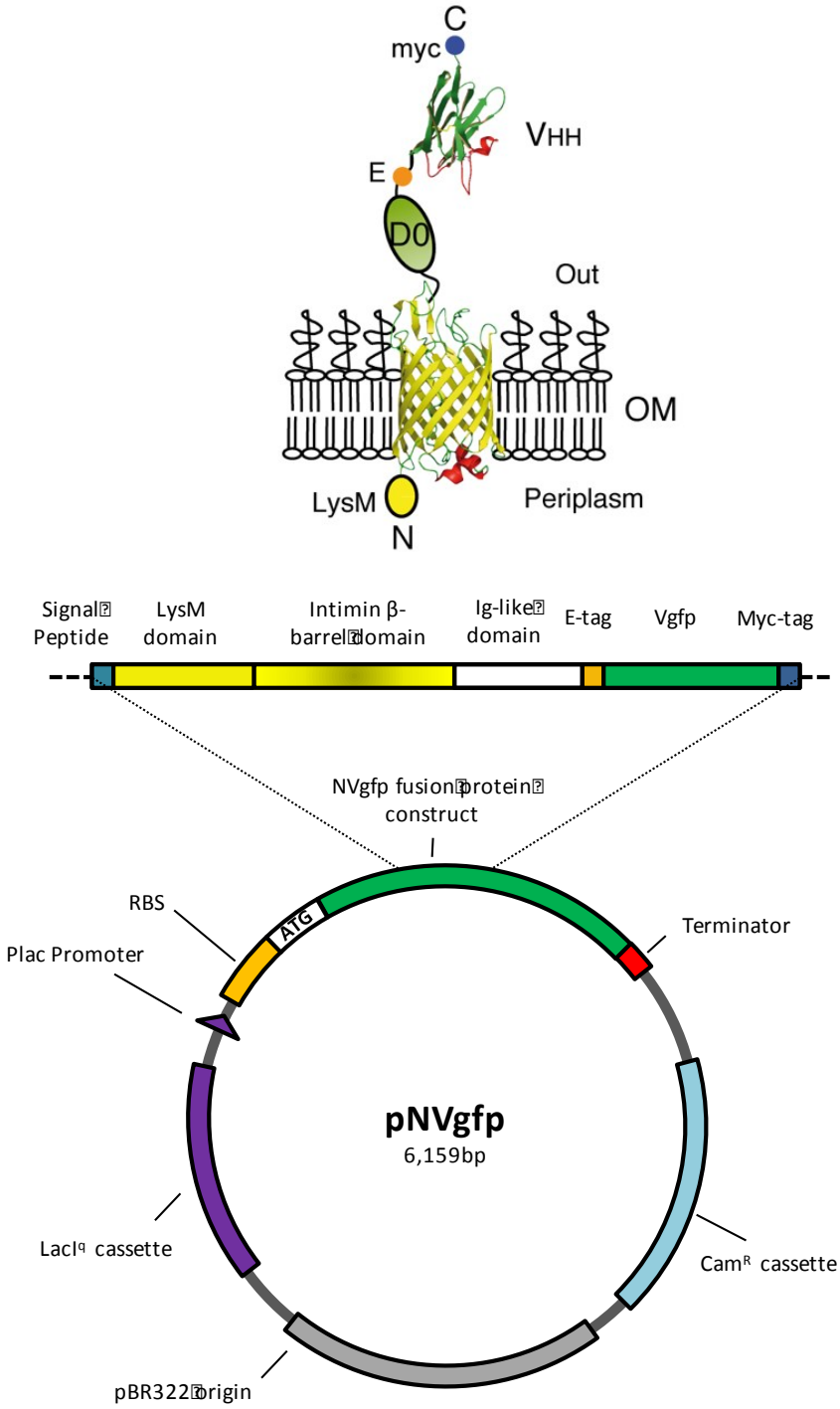


Figure 3.5: NVgfp display vector and the pNVgfp construct plasmid map. (Bottom) Plasmid map of the pNVgfp DNA construct with an emphasis on the NVgfp display vector constituent parts. (Top) Schematic of the NVgfp display construct and its predicted localization in the outer membrane of *E. coli* cell chassis. Top schematic adapted from (Salema et al., 2013).

Spain. Upon receipt, the plasmid construct size was verified by test digest (Figure 3.6) and the coding sequence of the display system was verified by DNA sequencing. The plasmid construct was transformed in *E. coli* DH10B cell strain to produce the prototype biosensor device.

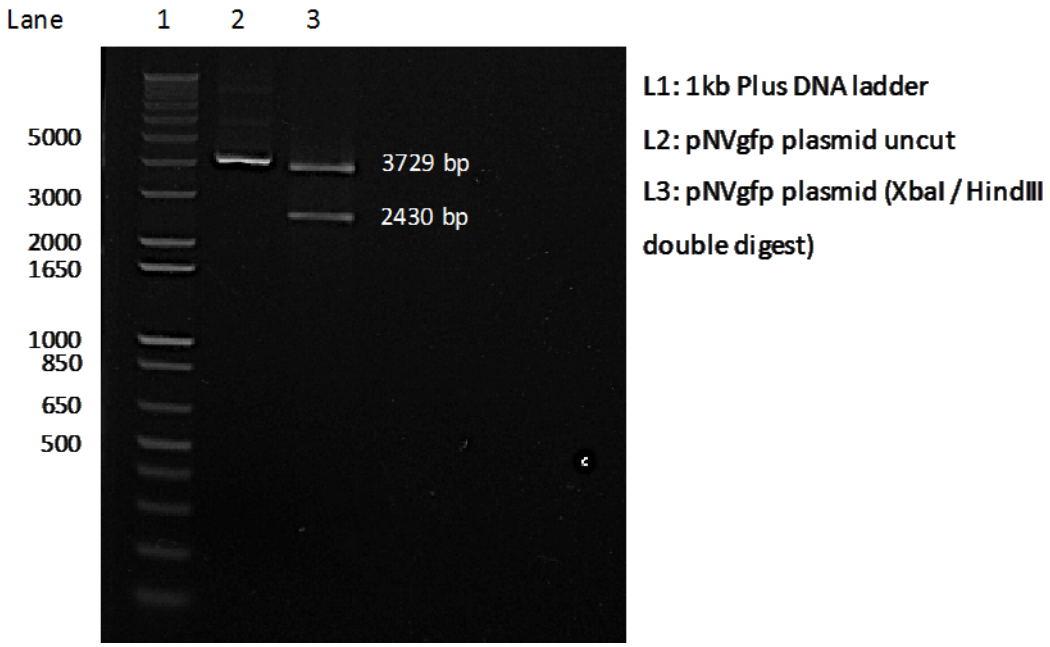


Figure 3.6: Test digest of pNVgfp DNA construct. Test digest of the pNVgfp DNA construct (6159 bp) that encodes the NVgfp display construct fusion protein (Figure 3.5). DNA plasmid was digested with restriction enzymes XbaI and HindIII, followed by agarose gel DNA electrophoresis.

3.2.3 Characterisation of the pNVgfp construct in *E. coli* (DH10B) cell chassis

Previously, the pNVgfp construct that encodes the anti-GFP nanobody sensor element display vector was transformed into *E. coli* DH10B cell strain for nanobody production and cell surface display. However, various studies in the literature have reported that the expression of proteins targeted to the outer membrane often cause membrane destabilisation, leakage or cell toxicity (Georgiou et al., 1996; Stathopoulos, Georgiou, & Earhart, 1996). Additionally, many complications could arise from the heterologous expression of engineered protein fusion constructs, such as protein misfolding, incomplete formation of disulphide bridges and non-functional protein display among others. To identify suitable conditions for anti-GFP sensor element display by the prototype device, a series of experiments were designed to characterise the pNVgfp construct and NVgfp display vector expression in DH10B *E. coli* cell chassis.

Testing for cell toxicity

Initially, an experiment was designed that investigated any cytotoxic effects of the anti-GFP nanobody expression to the host chassis. Such effects can be monitored by the measurement of the growth rate of the host organism. This is because the expression of a recombinant protein that interferes with normal proliferation and homeostasis of the host would result in slower growth rate (Doherty, Connolly, & Worrall, 1993; Dong, Nilsson, & Kurland, 1995; Rosano & Ceccarelli, 2014).

E. coli DH10B cells transformed with the pNVgfp plasmid construct were induced for NVgfp protein expression and the cell culture density was monitored over time. Exponentially growing cells were induced with 0, 0.1, 0.25, 0.5, 0.75 and 1 mM IPTG inducer to activate the Plac promoter for the expression of various protein levels of NVgfp surface display vector. NVgfp expression was carried out at 30°C.

The cytotoxicity assay for NVgfp expression in *E. coli* DH10B cell strain host showed negligible effects on the growth rate of the host cell (Figure 3.7). The mean growth rate of bacterial cultures for different IPTG treatments was calculated for the time interval of 180-210 minutes (Figure 3.7B). One way ANOVA test was used to test the null hypothesis that different IPTG treatments do not result in different growth rate values. The P-value was calculated to be 0.6036. This suggested that the difference in the mean growth value (as a proxy marker to cytotoxicity) between different inducer concentration treatments was not significant and thus 1mM IPTG concentration was used in subsequent experiments.

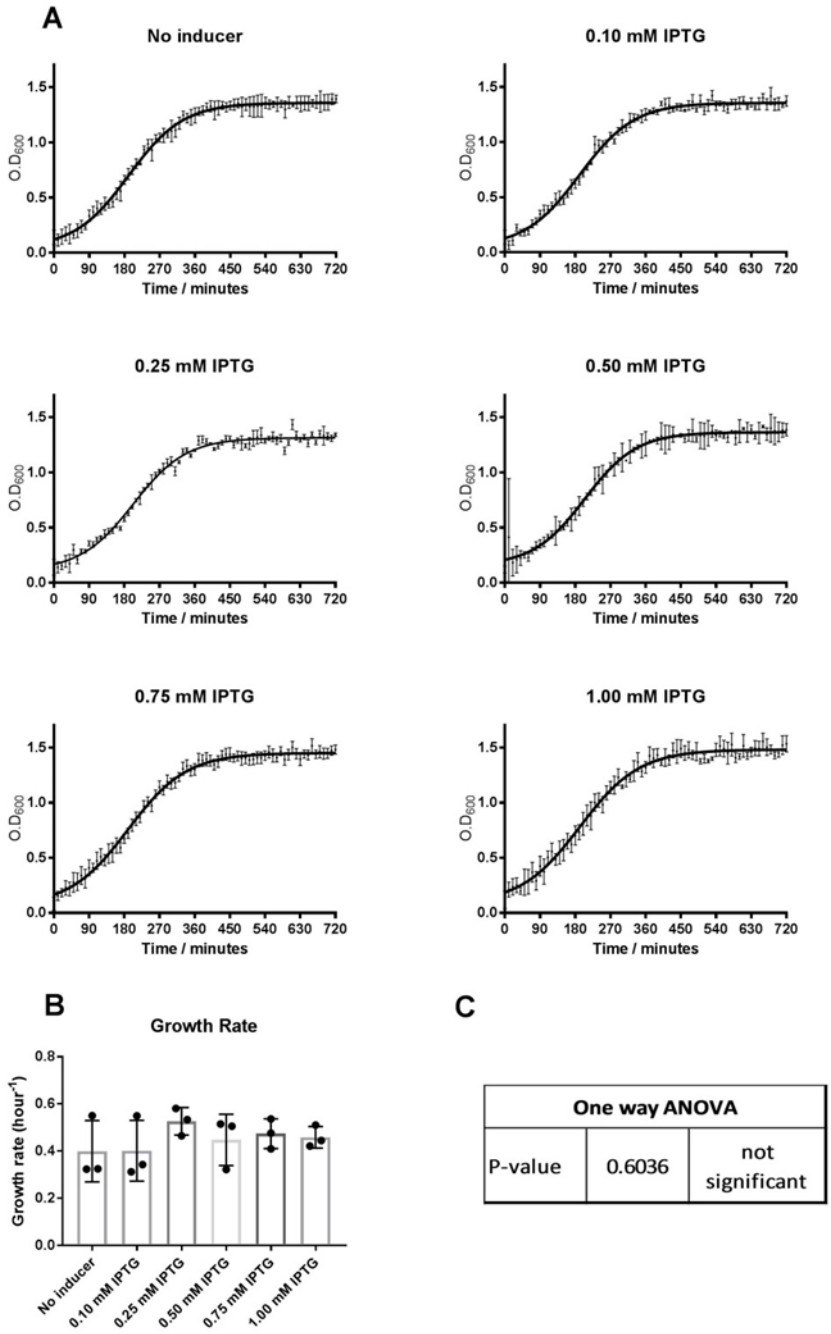


Figure 3.7: Cell toxicity assay for NVgfp protein expression in *E. coli* DH10B cell strain.A Growth curves of *E. coli* cells induced for the expression of NVgfp surface display construct with various IPTG concentrations. *E. coli* DH10B cells transformed with the pNVgfp construct. NVgfp protein expression is under the control of the Plac inducible promoter on pNVgfp plasmid. Protein expression was carried out at 30°C and 500rpm and cell culture density was monitored over time by absorbance measurements at 600nm. Error bars represent standard deviation of 3 technical repeats.**B** Bar graph of mean growth rate of cell cultures for IPTG concentrations. **C** One way ANOVA test for determining the significance of difference in the mean growth rate.

Testing localisation and functionality of the sensor element

Next, an experiment was designed to investigate the localisation of the expressed anti-GFP nanobody and its functionality as a sensor element when displayed at the cell surface of the host. Correct localisation and folding of the NVgfp protein would have resulted in the projection of the anti-GFP nanobody and its C-terminal myc tag to the extracellular milieu (Figure 3.5), where it can be detected by antibody labelling against the myc-tag. Additionally, the display of a functional anti-GFP nanobody at the cell surface should allow for GFP protein labelling of the *E. coli* cells.

Cells of the *E. coli* DH10B strain cell line harbouring the pNVgfp construct were induced for NVgfp protein expression and subsequently labelled with anti-myc antibody conjugated with a fluorescent molecule and purified GFP. Specifically, cultures of this cell line were induced with 0, 0.1mM, 0.25mM, 0.5mM and 1mM of IPTG inducer, and protein expression was carried out at 30°C. Subsequently, cell suspensions in PBS derived from the cell cultures were labelled with antibody and purified GFP separately and whole cell fluorescence was determined for each label by flow cytometry. To account for cell autofluorescence, a control sample was used of *E. coli* DH10B cells that did not carry any plasmid construct.

Flow cytometry analysis suggested that the NVgfp display protein localised at the *E. coli* cell surface and displayed functional anti-GFP nanobodies (Figure 3.8). In Figure 3.8 top, samples induced with IPTG concentrations 0.1-1mM showed an increase in the measured cell fluorescence as compared to the uninduced sample at 0 mM IPTG inducer. The increased cell fluorescence resulted from the binding of fluorescent anti-myc antibodies to myc-tags cell surface displayed by the NVgfp display protein. Since antibody molecules cannot penetrate the cell membrane, the increased cell fluorescence suggests expression and localisation of the NVgfp protein to the *E. coli* cell surface. In Figure 3.8 bottom, samples induced with IPTG concentrations 0.1-1mM showed an increase in the measured cell fluorescence as compared to the uninduced sample at 0 mM IPTG inducer. The increased cell fluorescence resulted from the binding of fluorescent GFP molecules recognised by the anti-GFP nanobodies. These results demonstrated the functional display of the nanobody molecules by the surface localising NVgfp protein. In both experiments, the control cell line of *E. coli* DH10B cells showed similar cell fluorescence to the uninduced pNVgfp harbouring

cell line that suggested that there was no or minimal leaky expression of NVgfp protein from the Plac expression system. Additionally, it was observed that cell fluorescence was similar in all samples despite the use of different IPTG inducer concentrations. This suggested that all samples displayed same amounts of anti-GFP nanobody and NVgfp protein.

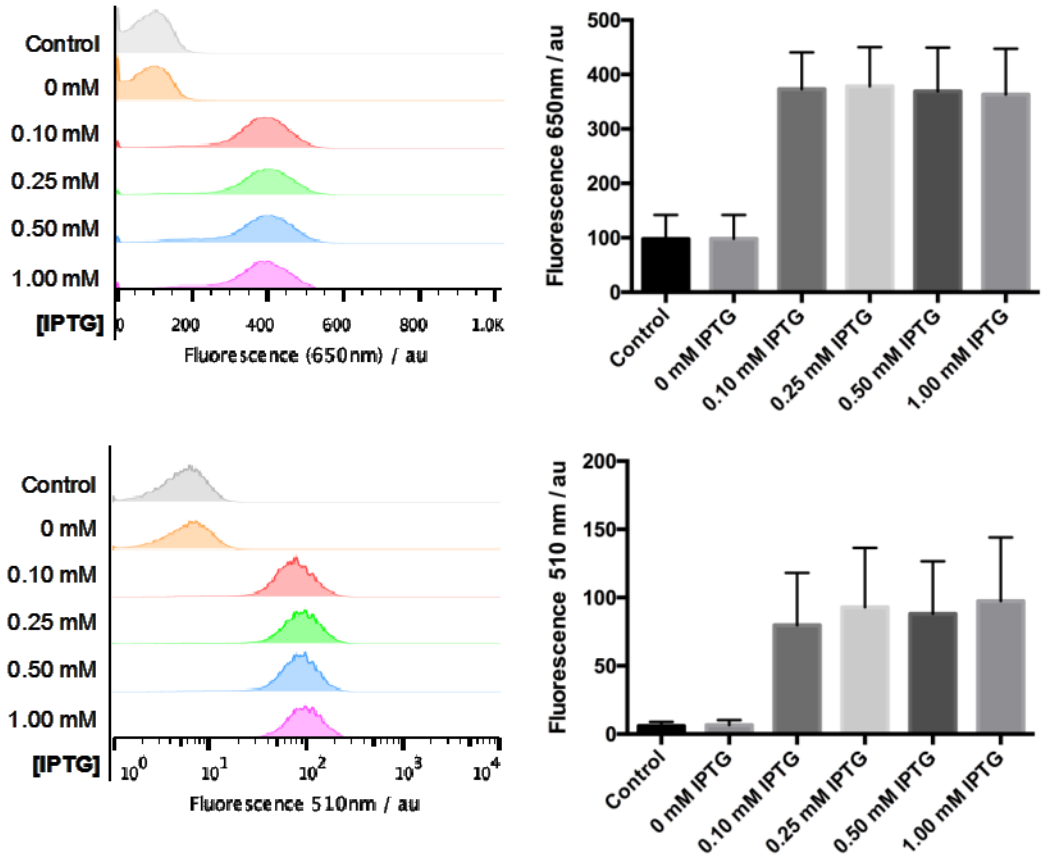


Figure 3.8: Flow cytometry analysis of *E. coli* DH10B cells harbouring pNVgfp. (Top) Cells labelled with anti-myc Ab/fluorescent conjugate(PE) to detect cell surface localisation of the NVgfp protein. Cell fluorescence histogram of labelled cells following induced expression of NVgfp at various inducer concentrations. Bar chart of the calculated mean cell fluorescence for each sample. (Bottom) Cells labelled with purified GFP to assess NVgfp mediated displayed anti-GFP nanobody functionality. Cell fluorescence histogram of labelled cells following induced expression of NVgfp at various inducer concentrations. Bar chart of the calculated mean cell fluorescence for each sample. In both experiments cells were labelled for 30 minutes at room temperature. Cell fluorescence was measured by flow cytometry at 650nm and 510nm respectively. Cells carried the pNVgfp plasmid and NVgfp expression was carried out at 30°C under the control of Plac promoter. Reported values derived from one experimental repeat and error bars represent the coefficient of variance for each sample.

Quantification of anti-GFP sensor elements displayed

Following the successful demonstration of functional anti-GFP nanobodies displayed on *E. coli* cell surface, an experiment was designed to attempt the quantification of displayed sensor elements to further characterise the prototype device. It was reasoned that the number of anti-GFP nanobodies displayed could be calculated by titrating known amounts of purified GFP concentrations against a constant number of *E. coli* cells. Then, the number of displayed antibodies could be derived from the point where increasing GFP concentrations do not result in increased cell fluorescence due to saturation of nanobody binding sites. A similar approach for the quantification of scFv displayed on *E. coli* cells have been reported previously (G. Chen, Cloud, Georgiou, & Iverson, 1996) as well as for the quantification of displayed enzymes (Wendel, Fischer, Martínez, Seppälä, & Nørholm, 2016).

The experiment was carried out using NVgfp expressing cells labelled with purified GFP and their cell fluorescence was measured using flow cytometry. The *E. coli* DH10B strain cells transformed with the pNVgfp construct were induced for NVgfp protein expression. Following protein expression, the cells were separated out from the growth medium of the cell culture and were re-suspended in PBS to an $O.D._{600}$ value of 1. 200µl of cell suspension were incubated with 0, 1, 2, 5, 10, 15, 20, 25 and 30 pmoles of purified GFP. Excess unbound GFP was washed away and cell fluorescence was measured by flow cytometry.

Flow cytometry analysis enabled the quantification of displayed anti-GFP nanobodies in *E. coli* DH10B cells transformed with the pNVgfp plasmid (Figure 3.9). NVgfp cells showed cell fluorescence signal saturation at concentrations equal or more than 5 pmoles of GFP for 200µl cell suspension at $O.D._{600}$ = 1 (top Figure 3.9). To enable specific estimation of the total number of displayed nanobodies, the data were analysed by non-linear regression and fitted with a receptor binding saturation model (Hulme & Trevethick, 2010) in Prism software (bottom Figure 3.9). Using this model, the 50% occupancy was estimated to be achieved at 1.13 pmoles purified GFP. This value corresponds to 5.65nM GFP concentration in 200µl cell suspension volume. Given that the dissociation constant of the anti-GFP nanobody for its GFP target is 0.23nM (Rothbauer et al., 2006), the assumption was made that at 5.65nM GFP concentration (that is 25-fold higher than the Kd value of the anti-GFP nanobody) almost all (<95%) of the sensor elements were ligand occupied. This assumption, allowed for calculating that each cell displays 6,750 functional anti-GFP sensor elements

given that the number of cells in 1ml of $O.D._{600}$ = 1 is $1x10^9$ cells (Moran, Phillips, & Milo, 2010).

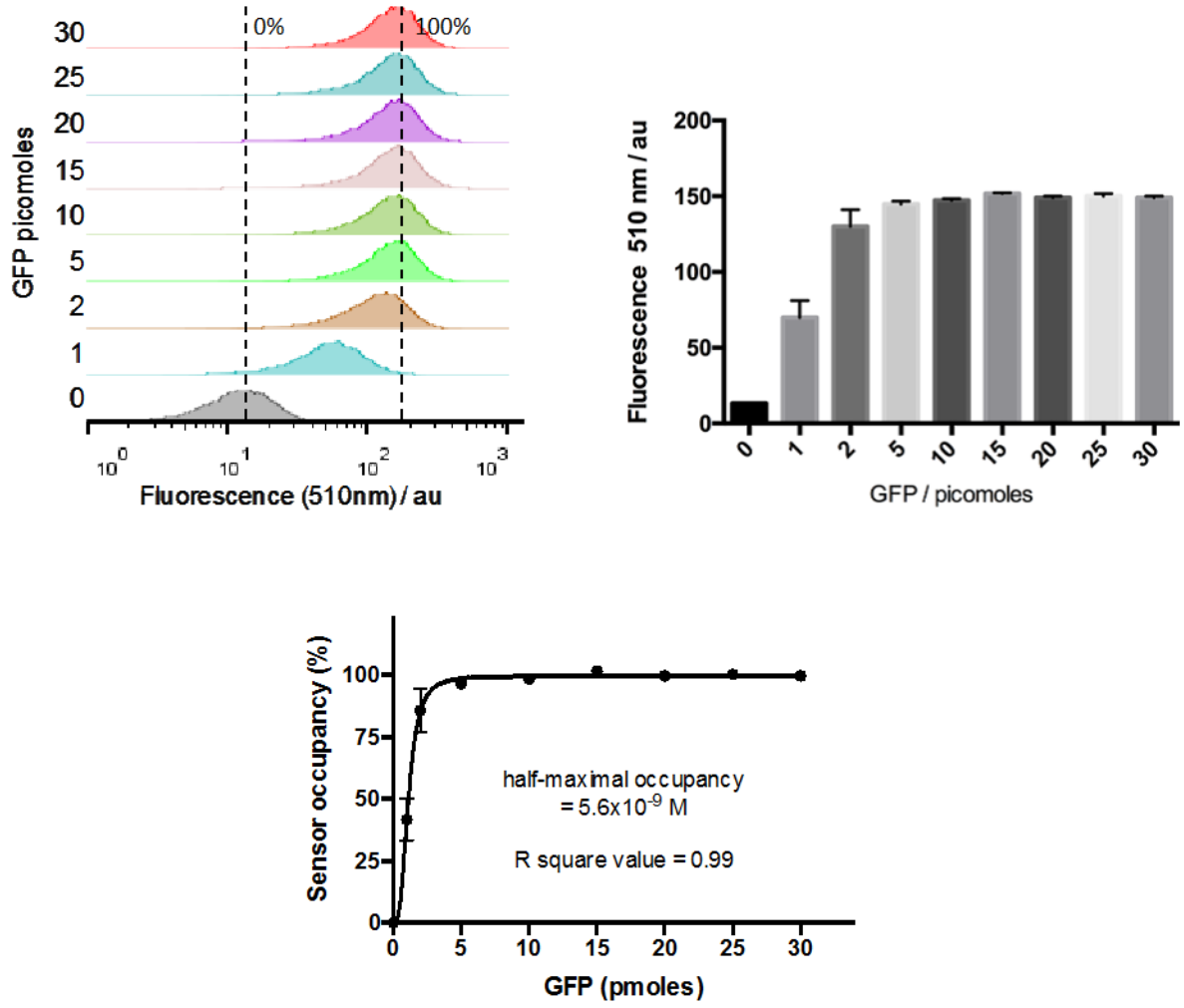


Figure 3.9: Quantification of anti-GFP nanobodies displayed on prototype device. *E. coli* DH10B strain cells surface displaying the anti-GFP nanobody labelled with increasing concentrations of purified GFP. (Top left) Overlaid plots of flow cytometry data showing cell fluorescence when labelled with purified GFP (0 - 30 picomoles). (Top right) Bar chart of mean cell fluorescence of cell populations as determined by flow cytometry and analysis in FlowJo software. (Bottom) Scatterplot of mean cell fluorescence from flow cytometer data plotted against the GFP concentrations. Cell fluorescence labelled with 0 picomoles GFP was fixed to 0% occupancy, while average fluorescence for 20-30 pmoles GFP was fixed to 100% occupancy. Data were analysed using non-linear regression and fitted with a receptor binding saturation model in Prism software to calculate 50% sensor occupancy. Cells carry the pNVgfp plasmid and NVgfp expression was carried out at 30°C under the control of Plac promoter. Error bars represent standard deviation of 3 technical repeats.

Summary

In this series of experiment, the prototype biosensor device of *E. coli* DH10B strain cells harbouring the pNVgfp plasmid was thoroughly characterised. Induction of protein expression in the prototype device was shown to result in the localisation of the NVgfp display construct to the bacterium cell surface. The surface localised NVgfp construct was shown to functionally display anti-GFP sensor elements that were able to bind purified GFP protein in the extracellular milieu. The number of displayed anti-GFP nanobodies was measured to be 6,750 per *E. coli* cell. In addition, the expression of the NVgfp protein construct in the prototype device was determined to not cause serious cytotoxic effect to the host chassis.

3.2.4 Development of the model biomarker protein: GFP_2

Previous experiments demonstrated that *E. coli* DH10B cells transformed with the pNVgfp construct were able to cell-surface display multiple anti-GFP nanobody sensor elements and bind GFP molecules in the extracellular milieu (Figure 3.8, Figure 3.9). According to the principles of agglutination reaction, both the sensor element of the device and the analyte molecule should provide multivalency towards each other (Price, 2001). Analyte molecule multivalency could result from the presence of 2 epitope sites on the analyte molecule recognised by two antibody sensor elements with different epitope specificities. Alternatively, the analyte molecule multivalency could result from epitope presence on repeated domains in the analyte structure, for example in homodimeric proteins. With this latter strategy in mind, protein engineering was used to develop a fusion protein of two GFP monomers that would provide 2 epitope sites for the anti-GFP nanobody.

A dimeric GFP (GFP_2) was constructed by cloning in tandem two GFP DNA sequences joined by a peptide linker domain. The linker peptide sequence used was twelve amino acid long (G H G T G S T G S G S S), and it was chosen as it has been used previously by others to protein engineer an EGFP dimer fusion protein (Y. Chen, Wei, & Muller, 2003). Standard molecular biology procedures were used to construct the plasmid DNA for the GFP_2 design, and the fusion protein was recombinantly expressed in *E. coli* and purified. The purified sample of the expressed protein showed bright fluorescence when observed by naked eye.

An experiment was carried out to determine the folding state of the engineered GFP_2 protein. This was necessary as previous attempts to engineer GFP_2 with shorter linker sequences resulted in non-fluorescent versions that were determined to be misfolded according to circular dichroism analysis (data not shown). It was reasoned that a properly folded GFP_2 protein that incorporates two properly folded and functional fluorescent chromophores would exhibit twice the fluorescence signal when excited as compared to a monomeric GFP protein at an equimolar concentration. For the experimental setup monomeric GFP and GFP_2 were recombinantly expressed *E. coli* and purified to high purity as determined by SDS-PAGE electrophoresis (Figure 3.10 panel A) and protein concentration was quantified by a spectrophotometer. Then, purified proteins were diluted to 1.8 μM concentration and a fluorescence plate reader was used to measure the fluorescence signal

of the samples at 510nm. Background fluorescence was accounted for by the use of an equal volume solution of the buffer system.

The fluorescent intensity of the signal from the GFP_2 samples was almost double the intensity of the signal from the GFP samples (Figure 3.10 panel B). Fluorescence signal for GFP_2 protein was measured to 121, 022 au, while fluorescent signal for GFP protein was measured to 66,733 au. This represents a change of 1.81-fold increased fluorescent signal for GFP_2 as compared to GFP monomer for equimolar protein concentrations. This result suggested that a large percentage of the GFP_2 fusion protein folded to its native structure.

The successful engineering of the GFP2 model biomarker that is divalent for its sensor element anti-GFP nanobody allowed for further experiments that investigated the proposed biosensor design (Section 3.2.1).

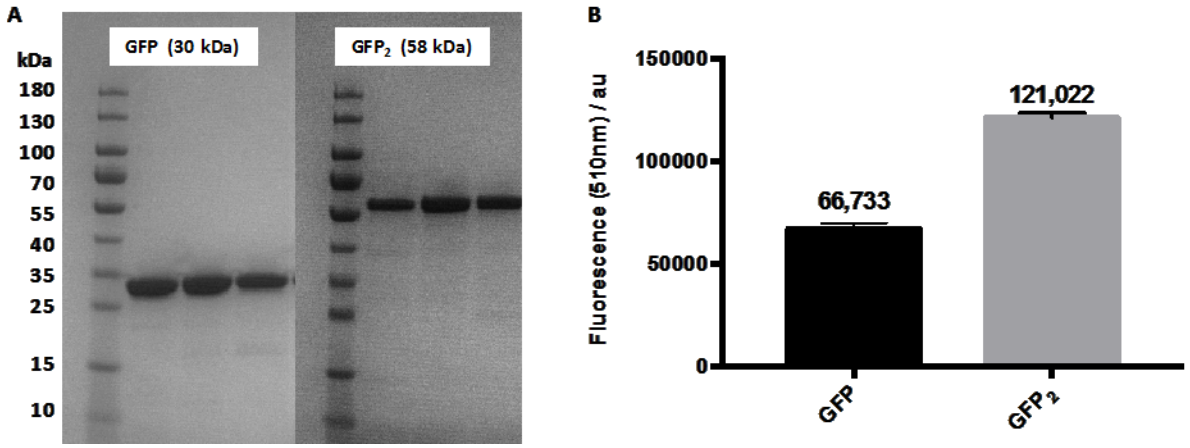


Figure 3.10: GFP fluorescence versus GFP_2 fluorescence. A SDS-PAGE analysis of eluted fractions from size exclusion chromatography follwing nickel-affinity chromatography purification. **B** Bar chart of fluorescence signal emitted by monomeric GFP and GFP_2 fusion protein, showing properly folded protein exhibiting twice the fluorescence signal from monomeric GFP. Proteins were purified, protein concentrations were made equimolar at 1.8 μM. Measurement were carried out on a microplate reader and error bars represent standard deviation of 3 technical repeats.

3.2.5 Simulation of the bacterial agglutination reaction

An attempt was carried out to describe the dynamics of the agglutination reaction in a mathematical model that would enable simulating the agglutination reaction *in silico*. The development of the model was important in aiding the identification of the relevant analyte concentrations at which the wet-lab experiments were to take place for the validation of the prototype biosensor design.

In a previous study, a kinetic model of the agglutination reaction was developed that was experimentally verified by the binding of divalent antibodies to polyvalent *B. subtilis* cells (Dolgosheina, Karulin, & Bobylev, 1992). Due to the similarity of the antibody/cell system to the biosensor design proposed here, the developed model was used as a framework to build a mathematical model that would describe the dynamics of the biosensor. The mathematical model for the bacterial agglutination reaction and its Matlab implementation used here were developed in collaboration (Hung El, 2014).

In short, the mathematical model developed to describe the dynamics of the agglutination reaction considered the reaction as a system of finite number of cells that could cluster together under certain conditions to form larger aggregates. Individual cells, each one of them considered as a cluster, were modelled to respond to the presence of GFP_2 ligand concentrations by the formation of clusters of larger sizes (aggregates) but nevertheless reducing the total number of clusters in the system. This reduction of clusters in the system is referred as the normalised aggregation value. The aggregation reaction was mathematically described by a system of three equations. The first equation, describes the time evolution of the number of clusters in the system by considering the density of bound GFP_2 ligands, the density of available/unbound nanobody receptors and the total number of ways of constructing/deconstructing clusters. The second equation describes the evolution of free GFP_2 ligands in the system via the binding/unbinding of the ligand to the nanobody receptors and the third equation models the evolution of bound GFP_2 ligands density on the surface of a cell. The model was then simplified into a system of two ordinary differential equations and was implemented in Matlab for solving and analysis. The full description and derivation of the mathematical model, the model simplification, the Matlab scripts and the parameter values used are available in Appendix B.

An experiment was carried out *in silico* to investigate the GFP_2 ligand concentrations that could result in bacterial agglutination. The previously determined number of nanobodies surface displayed per bacterial cell of 6,750 (Section 3.2.3, Figure 3.9) was used as an input parameter for the reaction simulation. The model also requires an input parameter for the number of *E. coli* cells present in the agglutination reaction. A wet-lab experiment was designed to investigate the cell concentration that would produce a visibly detectable cell pellet in the absence of bacterial agglutination. *E. coli* DH10B cells transformed with the pNVgfp construct were made into cell suspensions of $O.D._{600}$ value of 2 and two-fold dilutions thereof and added in a round bottom well plate. No GFP_2 was added to the cell suspensions. It was determined that 200µl of cells at $O.D._{600}$ = 1 produced a clearly visible cell pellet when they were allowed to settle overnight (Figure 3.11). Cell pellets were visible for cell suspension down to an $O.D._{600}$ value of 0.25. Thus, the value of $0.2x10^9$ (200µl cell suspension at $O.D._{600}$ value = 1) was used as an input parameter for the number of bacterial cells in the reaction simulation. The ligand concentration range explored for agglutination was 100 femtomoles to 1 nanomole (500 picoMolar to 5 microMolar for 200µl reaction volume) GFP_2 concentration. The lower boundary of the concentration range was chosen to be more than 2 times the dissociation constant value of nanobody binding to GFP (Rothbauer et al., 2006). The higher boundary was chosen to allow for testing a large space of concentrations values. Finally, the reaction was modelled for a time span of 18 hours to resemble an overnight incubation of the test assay.

The results of the simulation suggested a relatively wide range of GFP_2 concentrations at which bacterial agglutination is predicted to occur (Figure 3.12). The range of ligand amount at which the bacteria were predicted to agglutinate was 100 femtomoles to 2 picomoles GFP_2, while the highest value of normalised agglutination was predicted to occur at approximately 1 picomole GFP_2. At this ligand concentration, the value of normalised aggregation was found to be 0.7, that suggests that the existence of a number of bacterial clusters remaining in the system. Shortly after its maximum value, the normalised aggregation value decreases steeply towards zero aggregation. It is possible that this abrupt drop in aggregation is the result of a complex relationship between factors such as the saturation of nanobodies by the ligand binding and the fast kinetics of ligand binding versus the much slower cell cross-linking dynamics.

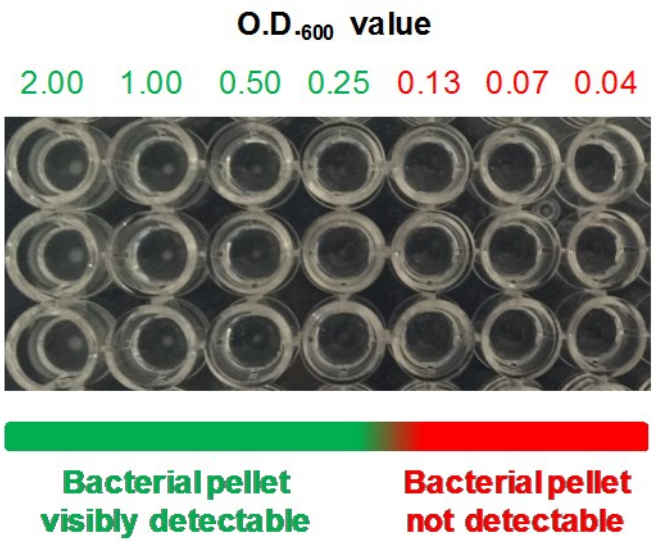


Figure 3.11: Bacterial cell pellet as a visual output for bacterial agglutination reactions. Image of a round bottom well microplate in which 200µl *E. coli* cell suspensions at various $O.D._{600}$ values were left to settle to cell pellets overnight. Cell suspensions were of *E. coli* DH10B strain cells transformed with the pNVgfp construct.

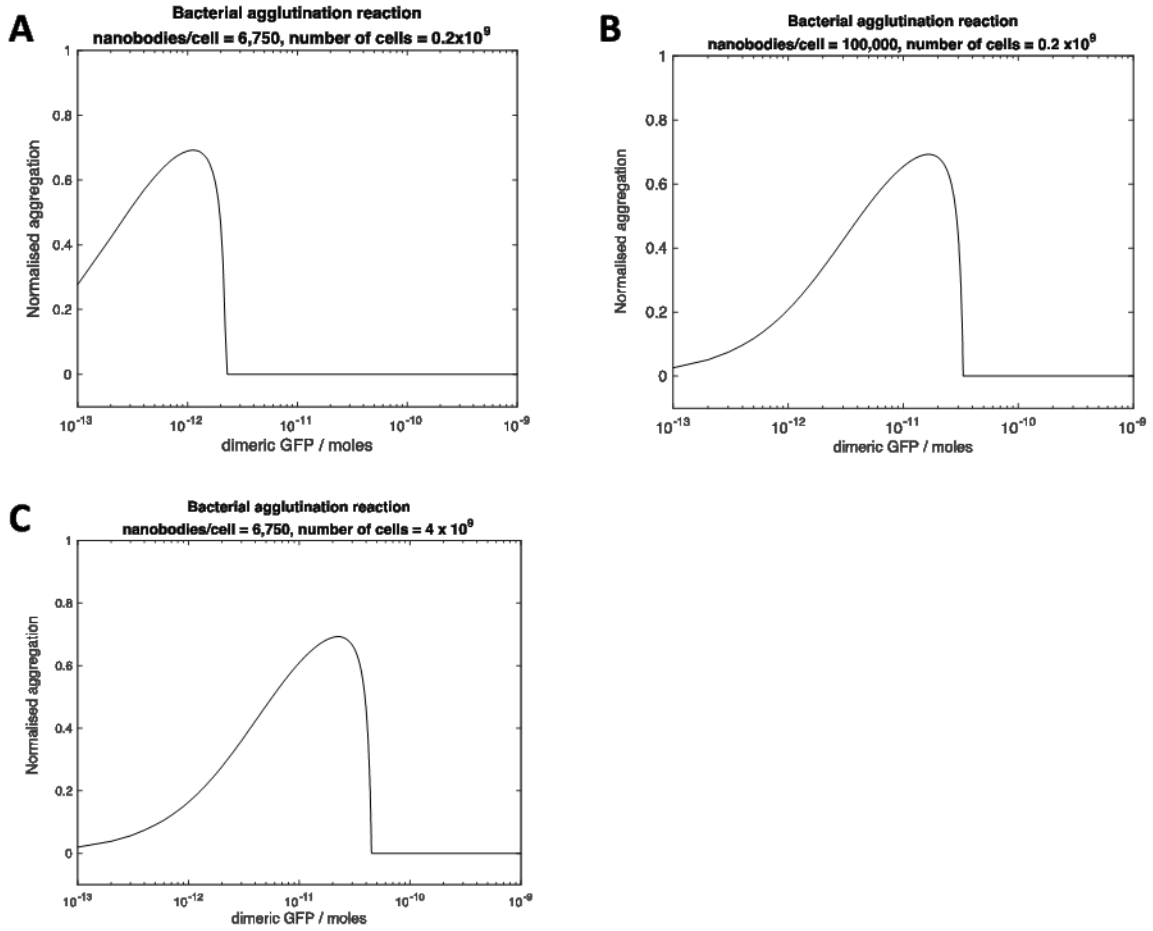


Figure 3.12: *In silico* simulation of bacterial agglutination reaction. Graphs show the normalised aggregation or agglutination of bacterial cells over a range of GFP_2 ligand molecule concentrations over 18 hours. Normalised aggregation "translates" to the level of reduction of individual bacterial cells or the reduction of number of clusters in the system. **A** Simulation of the agglutination reaction with input parameters for 200µl cell suspension at $O.D_{600}$ = 1 of *E. coli* DH10B cells transformed with pNVgfp construct. **B** Simulation of the agglutination reaction for a biosensor device with higher number of displayed anti-GFP nanobodies than the device in panel A. **C** Simulation of the agglutination reaction for *E. coli* DH10B cells transformed with pNVgfp construct in a denser cell suspension (more cells in reaction) than in panel A. Simulations and plots were generated with Matlab software.

Next, the bacterial agglutination reaction model was used to study the system further. First, the input parameter for the number of nanobodies displayed by each bacterial cell was set to 100,00 rather than the experimentally derived number of 6,750 for *E. coli* DH10B cells transformed with pNVgfp. The simulation results suggested that an increase in the number of nanobodies displayed per cell shifts the region of the graph that exhibits high values for normalised aggregation towards a higher range of GFP_2 ligand concentrations (Figure 3.12 panel B). Then, the input parameter for the number of nanobodies per cell was set back to 6,750 and the number of cells parameter was increased to 4×10^9. As previously, the simulation results suggested that an increase in the number of cells participating in the agglutination reaction shifts the agglutination positive region towards a higher range of GFP_2 ligand concentrations (Figure 3.12 panel C). Both input parameters, the number of nanobodies displayed per cell and the number of bacterial cells participating in the reaction, determine the total number of nanobody sensing elements in the system. Thus, the model suggested that increasing the total number of nanobodies displayed would enable analyte detection at higher concentrations and *vice versa*.

3.2.6 Testing prototype biosensor device for agglutination

Previous experiments demonstrated that *E. coli* DH10B cells transformed with the pNVgfp construct were able to cell-surface display multiple anti-GFP nanobody sensor elements and bind GFP molecules in the extracellular milieu (Figure 3.8, Figure 3.9). Also, a fusion protein GFP_2 was engineered (Figure 3.10) with two epitope sites recognised by the anti-GFP nanobody and was expected to provide divalency for that sensor element. Thus, the GFP_2 could serve as the model analyte for testing the proposed biosensor design (section 3.2.1) in a bacterial agglutination reaction. It was expected that in the case of successful agglutination of bacterial cells by the GFP_2 molecules, this will result in an extended network of cells. Given enough time, this extended network of cells should settle at the bottom of a well to form a membrane-like structure when the experiment is carried out in round bottom microplate (Figure 3.3 panel A). In contrast, in the absence of bacterial cell agglutination the cells would float freely in solution until they eventually settle at the bottom of the well as a pellet (Figure 3.3 panel A). Thus, the presence of a bacterial pellet in a well would indicate a negative result for bacterial cell agglutination and serve as the visual output of the test.

Next, an experiment was carried out for testing bacterial cell agglutination in the presence of GFP_2 model protein analyte. The assay was carried out with *E. coli* DH10B cells transformed with the pNVgfp construct and induced for anti-GFP nanobody cell surface display. The analyte concentration range tested was 1.15µM GFP_2 and two-fold dilution thereof. Additionally, the assay included a no analyte condition. The assay was set up in round bottom well microplates with biosensor cell suspensions at an $O.D._{600}$ value of 1. To account for non-specific protein interactions that might result in bacterial agglutination, a negative control experiment was set up with GFP added to the cell suspensions. Experimental results were documented with photographs of microplates back-illuminated on a blue box with and without a filter for green fluorescence visualisation.

The agglutination assay showed GFP_2 mediated bacterial cell agglutination (Figure 3.13). Images of the microplate following overnight incubation with GFP_2 analyte show absence of cell pellet formation for cell suspensions added with 4.5nM up to 72nM of GFP_2 analyte. Inspection of these wells revealed a membranous-like structure formation

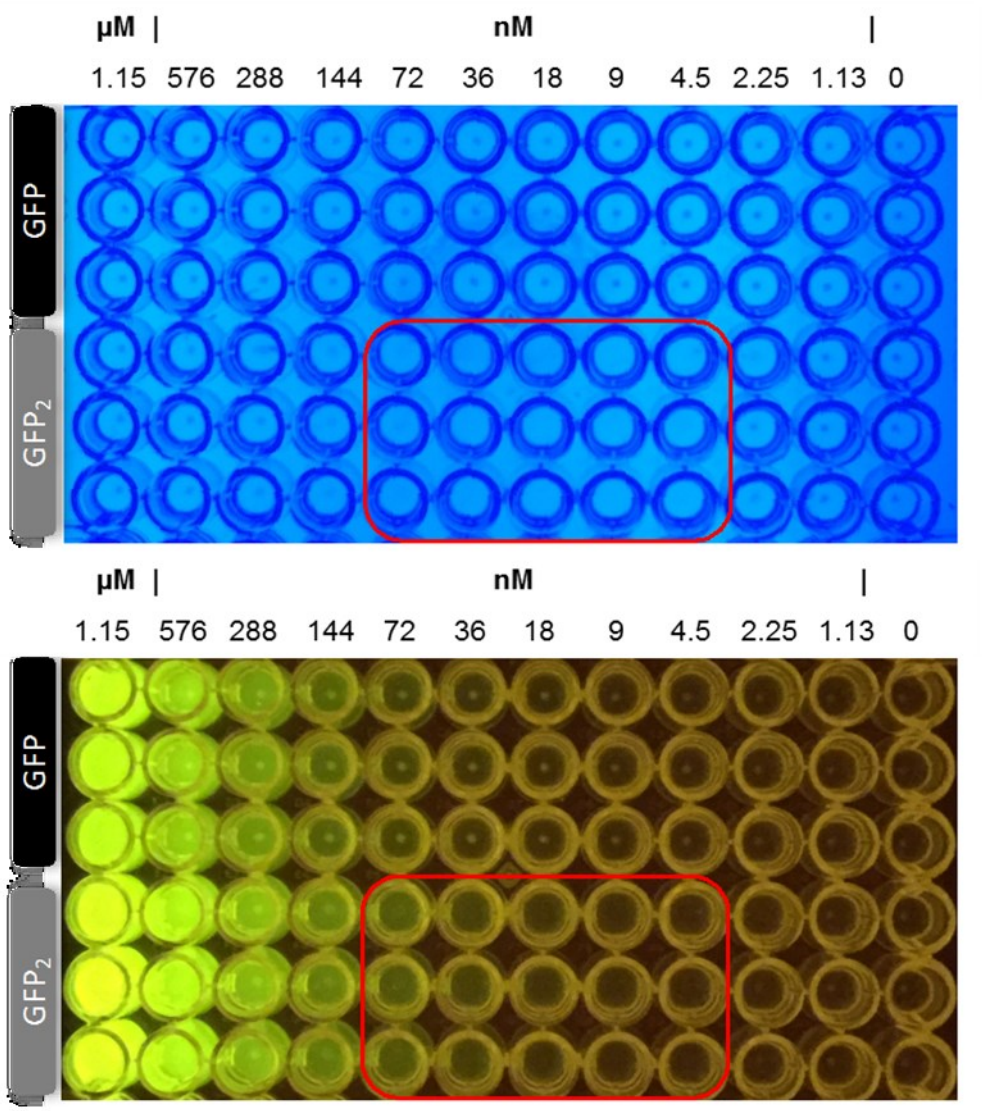


Figure 3.13: Plate agglutination assay using the designed prototype biosensor device. Images of a round bottom well microplate with cell suspensions of *E. coli* DH10B cell strain surface displaying the anti-GFP nanobody that were left to settle overnight with added monomeric GFP or GFP_2. Microplate wells which are absent of cell pellet formation indicate a positive result for bacterial agglutination. The red box highlights the agglutination positive region. In the top image, the microplate was back-illuminated with visible blue light. In the bottom image, the microplate was illuminated similarly but with the application of a GFP filter. 200µl cell suspensions at $O.D._{600}$ = 1 were added with 50µl analyte solution to a total assay volume of 250 µl. *E. coli* DH10B cell strain were transformed with the pNVgfp construct and and induced for NVgfp expression.

lining the bottom of the wells. This structure was hard to document photographically or with light microscopy due to its transparency but it was clearly visible to the naked eye. The region of agglutination consistent phenotype was centred around 18nM of GFP_2 analyte concentration, while microplate wells at the border of the region (2.25nM / 4.5nM concentrations and 72nM / 144nM concentrations) exhibited a mixed phenotype of cell pellet formation and membranous-like structure formation. Cell pellet formation was observed for all other GFP_2 concentartions. In contrast, negative control agglutination reactions added with monomeric GFP showed cell pellet formation. This suggested that bacterial cell agglutination was a result of specific interactions between bacterial displayed anti-GFP nanobodies and multivalent GFP_2 molecules and not a result of agglutination due to non-specific protein interactions.

Experimental wet lab results from the agglutination assay were compared to the *in silico* predictions of the model. In the wet lab assay, the agglutination positive region centred on the 18nM analyte concentration sample (4.5 picomoles of GFP_2 ligand) (Figure 3.13) while the *in silico* model predicted that maximum agglutination of bacterial cells would take place at 1 picomole GFP_2 ligand (Figure 3.12 panel A). The two results were not in complete agreement. However, the model probably incorporates many of the important characteristics of the system since the error is small compared to the parameter space investigated that spanned analyte amounts over 4 orders of magnitude. One possible explanation for the gap between wetlab and *in silico* results might be the degradation of the GFP_2 analyte by contaminating proteases into monomeric GFP molecules during the overnight incubation at room temperature.

Summary

In summary, the experiments described successfully demonstrated proof of concept for the biosensor design proposed (3.2.1). The prototype device of *E. coli* DH10B cells displaying anti-GFP nanobody was used in an agglutination reaction assay for the detection of GFP_2 anlayte. The anti-GFP nanobodies displayed by the bacterial cells were able to bind GFP_2 molecules extracellularly. The divalency of GFP_2 molecules allowed the formation of an extended network of bacterial cells. This extended network of cells could be visibly distinguished from non-agglutinated cells that formed a cell pellet at the bottom of the well. More specifically, the prototype biosensor device exhibited specific detection of GFP_2 at 4.5-

72nM analyte range of concentrations visualised by aggregate formation in a total assay volume of 250µl.

3.2.7 Engineering the analytical characteristics of the agglutination assay

Previously, an *in silico* investigation using the bacterial agglutination reaction model suggested that an increased number of displayed anti-GFP nanobodies per cell would result in bacterial agglutination at higher amounts of GFP_2 analyte (Figure 3.12). In an attempt to modulate the analytical characteristics of the assay, the prototype device was modified to produce device variants. The device variants would employ a combination of promoter and ribosome binding site (RBS) genetic elements to control NVgfp protein expression to facilitate the engineering of devices that display different levels of anti-GFP sensor elements on the *E. coli* cell surface.

The Anderson promoter collection was used for selecting promoter elements that would control the expression levels of NVgfp (Registry_of_Standard_Biological_Parts, 2016). The Anderson collection is a family of constitutively expressing *E. coli* promoters isolated from a small combinatorial library. The collection includes 20 promoter parts named J23100 to J23119. These promoters have been characterised for their relative expression strength against a reference standard. On the other hand, the software tool RBScalculator was used for the selection of RBS parts that would control protein production from the transcribed mRNA (Salis et al., 2009).

For the construction of the GFP_2 binding biosensor device variants, three promoter elements and three RBS parts were utilised to engineer nine variant devices. The promoter parts used for device variants construction were the Anderson collection elements J23105, J23106 and J23101 with relative expression strength of 0.24, 0.47 and 0.70. The RBS parts utilised for device construction were designed using the RBScalculator v2.0 algorith with targeted translation initiation rates (TIRs) of 12,500, 25,000 and 50,000 relative units. Each promoter element was combined with each RBS parts and cloned upstream of the NVgfp protein CDS of the pNVgfp DNA construct. The resulting DNA constructs were named pNK302-pNK310 and were transformed in *E. coli* DH10B cells to produce biosensor devices v02 – v10 as summarised in Table 3.1. The original biosensor device from hereafter is referred as prototype device or device v01.

Table 3.1: Biosensor device variants for GFP_2 analyte detection. Device variants incorporate different combinations of promoter and RBS elements to achieve varied levels of anti-GFP nanobodies display. The control elements influence expression of NVgfp display vector protein that displays the anti-GFP nanobody on the cell surface of the cell.

Biosensor device variant	Promoter element[a]	Predicted relative promoter strength	Designed RBS element[b]	Predicted relative RBS strength
prototype	P_{lac}	-	expression vector RBS	-
device v02	J23105	weak	12K	weak
device v03	J23106	medium	12K	weak
device v04	J23101	strong	12K	weak
device v05	J23105	weak	25K	medium
device v06	J23106	medium	25K	medium
device v07	J23101	strong	25K	medium
device v08	J23105	weak	50K	strong
device v09	J23106	medium	50K	strong
device v10	J23101	strong	50K	strong

[a] Anderson promoter library elements (Registry_of_Standard_Biological_Parts, 2016)

[b] Designed RBS elements using RBS calculator (Salis et al., 2009)

Next, an experiment was carried out in which the device variants were tested for the levels of displayed anti-GFP nanobody and for their analytical characteristics in agglutination assays. The device variants v02 – v10 were cultured for NVgfp expression and labelled with saturating amounts of purified GFP. After removal of excess GFP, the cell fluorescence of each device variant was measured by flow cytometry to determine levels of anti-GFP sensor element display. Then, the device variants were tested for bacterial agglutination in a microplate agglutination assay. The GFP_2 biomarker concentrations tested in agglutination reactions were 144 picomoles and two-fold dilutions thereof to 141 femtomoles. Additionally, a test condition of no added GFP_2 was included as well. Agglutination test results were documented by photographs of the microplate back illuminated with blue light for easier visualisation of test results. *E. coli* DH10B strain cells were used as the negative control for the experiments. The negative control served the purposes of: i) benchmarking background cell fluorescence in flow cytometry measurements and ii) controlled for non-specific effects of GFP_2 analyte in the agglutination test. The prototype device of *E. coli* DH10B cells transformed with the original pNVgfp construct was used as the positive control for the experimental setup.

Mean cell fluorescence of labelled devices, as determined by flow cytometry analysis, showed varied levels of nanobody elements display between variants (Figure 3.14). Device v07, that incorporated a combination of the J23101 high strength promoter and the 25K TIR medium strength RBS part, exhibited the highest mean cell fluorescence value of 7,910 au. Device v01, that incorporated the combination of the J23105 low strength promoter and the 12K TIR low strength RBS part, exhibited the lowest mean cell fluorescence value of 658 au. After accounting for background fluorescence as determined by measuring the autofluorescence signal from the control sample, the change in mean cell fluorescence between device v07 and device v01 was ~19-fold. All other device variants showed mean cell fluorescence values within these two extremes.

The mean cell fluorescence data were used to identify suitable strategies for easier forward engineering of biosensor devices. To this end, the engineered device variants were separated into three groups, and each group included device variants that incorporate an identical RBS element. This allowed the comparison of experimentally measured activity of a

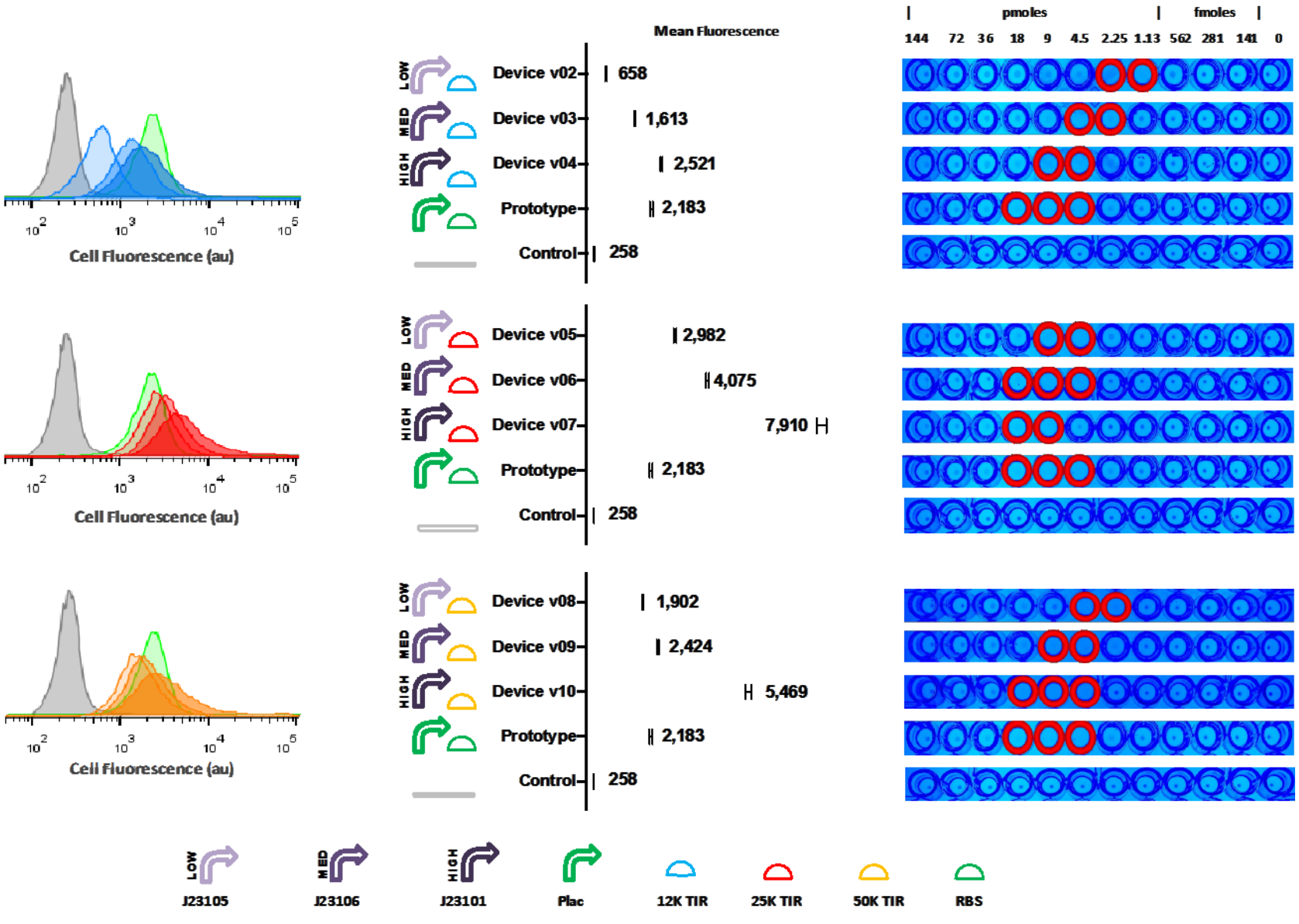

Mean Fluorescence
pmoles
fmoles
144 72 36 18 9 4.5 2.25 1.13 562 281 141 0
Cell Fluorescence (au)
Device v02 658
Device v03 1,613
Device v04 2,521
Prototype 2,183
Control 258
Device v05 2,982
Device v06 4,075
Device v07 7,910
Prototype 2,183
Control 258
Device v08 1,902
Device v09 2,424
Device v10 5,469
Prototype 2,183
Control 258
LOW
MED
HIGH
J23105
J23106
J23101
Plac
12K TIR
25K TIR
50K TIR
RBS

Figure 3.14: Biosensor device variants mean cell fluorescence and agglutination assays. (Left) Flow cytometry data for anti-GFP nanobody displaying device variants v01 – v10 labelled with saturating amounts of purified GFP (Data analysed by FlowJo software). Device variants incorporate promoter and RBS elements as indicated by the schematic to the left of the bar chart graphs. (Middle) Bar charts of mean cell fluorescence for device variants v01 – v10 labelled with saturating amounts of purified GFP. Error bars represent standard deviation for 3 technical repeats (data analysed by FlowJo software and graphed by Prism software). (Right) Images of 96-wells microplates showing bacterial agglutination assay results for device variants v01 – v10 added with purified GFP_2 analyte. Bacterial agglutination positive phenotype is highlighted by the use of a red circle. *E. coli* DH10B cells were used as the experimental negative control both for flow cytometry measurements and for the bacterial agglutination assay. The protyotype biosensor device is also referred as device v01. Visible blue light was used to back-illuminate microplates for better contrast and results were documented using a mobile phone camera.

specific promoter element versus its reported relative strength value (Table 3.2). As previously mentioned, the Anderson collection promoter elements J23105, J23106 and J23101 are reported to exhibit relative promoter strength activities of 0.24, 0.47 and 0.70 units relative to the reference standard promoter. If one sets the J23105 promoter element as the reference standard, then relative promoter strength of these elements translates to 1, 1.96 and 2.91 for J23105, J23106 and J23101 respectively. The levels of cell-surface displayed nanobodies, as inferred from recorded cell fluorescence values, showed that the promoters elements showed weak, medium and strong transcriptional activity as expected. For all three groups of device variants, the weak strength promoter J23105 exhibited the lowest levels of nanobody expression, while device variants incorporating the strong J23101 promoter showed the highest levels of nanobody expression. However, the recorded fold change in nanobody expression between promoter elements did not accurately follow the reported relative strength values. For example, it was expected that the transcriptional output from device v03 and device v04 promoters and would be 2-fold and 2.9-fold as high in comparison to the transcriptional output of device v04 promoter. Instead, experimentally derived data showed that devices v03 and v04 expressed nanobody levels of 3.4-fold and 5.7-fold higher in comparison to device v02.

A similar analysis was carried out with respect to the RBS elements used for the construction of device variants. The engineered devices were again separated into three groups, and each group included device variants that incorporated an identical promoter element. This allowed the comparison of the experimentally measured translation levels versus its designed TIR values (Table 3.3). As previously mentioned, the RBS elements were designed using the RBS calculator software for target TIR values of 12,500, 25,000 and 50,000 for RBS elements 12K, 25K and 50K respectively. The levels of cell-surface displayed nanobodies, as inferred from recorded cell fluorescence values, showed that the RBS parts did not behave as expected. In all three groups of device variants the 25K RBS elements exhibited higher translation activity than the 50K RBS element counter to the intended design specification.

Table 3.2: Comparison of reported relative promoter activity versus measured activity. Anderson collection promoters J23101, J2305 and J23106 of reported relative promoter activity of 0.70, 0.24 and 0.47 respectively were used for NVgfp protein expression from plasmid constructs in *E. coli* DH10B cells. Measured fold change data were derived from biosensor device variants labelled with saturating amounts of purified GFP and measured by flow cytometry for mean cell fluorescence.

Group	Biosensor device variant	Promoter element	Reported relative strength[a]	Measured fold change [b]
12K RBS group	device v02	J23105	1.00	1.0
	device v03	J23106	1.96	3.4
	device v04	J23101	2.91	5.7
25K RBS group	device v05	J23105	1.00	1.0
	device v06	J23106	1.96	1.4
	device v07	J23101	2.91	2.8
50K RBS group	device v08	J23105	1.00	1.0
	device v09	J23106	1.96	1.3
	device v10	J23101	2.91	3.2

[a] Normalised against J23105 promoter element measure relative expression strength = 0.24

[b] Normalised against mean cell fluorescence (adjusted for cell autofluorescence) of J23105 promoter incorporating device variant of same RBS element

Table 3.3: Comparison of designed RBS TIR activity versus measured activity. RBS calculator designed RBSs of translation initiation rates (TIRs) of 12,500 (12K), 25,000 (25K) and 50,000 (50K) were used for NVgfp protein expression from plasmid constructs in *E. coli* DH10B cells. Measured fold change data derived from biosensor device variants labelled with saturating amounts of purified GFP and measured by flow cytometry for mean cell fluorescence.

Group	Biosensor device variant	RBS element	Designed TIR strength[a]	Measured fold change[b]
J23105 promoter group	device V02	12K	1.00	1.00
	device V05	25K	2.00	6.8
	device V08	50K	4.00	4.1
J23106 promoter group	device V03	12K	1.00	1.0
	device V06	25K	2.00	2.8
	device V09	50K	4.00	1.6
J23101 promoter group	device V04	12K	1.00	1.0
	device V07	25K	2.00	4.7
	device V10	50K	4.00	3.2

[a] Normalised against 12K RBS element designed TIR value = 12K

[b] Normalised against mean cell fluorescence (adjusted for cell autofluorescence) of 12K RBS incorporating device variant of same promoter element

Nevertheless, in all three groups the 12K RBS element exhibited lower activity than both the 25K and 50K RBS element as designed. In conclusion, the analysis results suggested that for future designs the forward engineering of biosensor devices by the use of the Anderson promoter library is a more reliable strategy the use of RBS parts designed by the RBScalculator.

Visual inspection of the agglutination assays microplates revealed that the varied levels of nanobody displayed by device variants resulted in biosensing assays that exhibited various agglutination patterns for the range of GFP_2 concentrations tested (Figure 3.14). In addition, the agglutination assays results correlated to the flow cytometry analysis results. For example, device v07 exhibited the highest mean cell fluorescence when measured previously with flow cytometry, and also showed bacterial agglutination at the highest GFP_2 analyte concentration among the device variants. More specifically, device v07 agglutinated at the 36nM -72nM concentration range of GFP_2 analyte and even exhibited significant levels of aggregation at 144nM GFP_2 concentration sample well. Similarly, device v02 that exhibited the lowest mean cell fluorescence signal when measured with flow cytometry, it agglutinated at the GFP_2 concentration range of 4.5nM – 9nM that are the lowest analyte concentration values at which bacteria exhibited agglutination phenotype among device variants. A comparison between the GFP_2 analyte concentrations aggregating device v07 and those analyte concentrations aggregating device v02 bacterial cell suspensions reveals a change in the analytical range of similar magnitude to the 19-fold increase of recorded mean cell fluorescence for those devices. The rest of the device variants exhibited positive agglutination regions at various GFP_2 concentrations between the 4.5nM and 75nM analyte concentration range. In contrast, there was a clear cell pellet formation for all conditions tested for the control sample of DH10B *E. coli* cell suspension. This suggested that bacterial agglutination was a direct consequence of the NVgfp protein expression and anti-GFP nanobodies display rather than a non-specific cell aggregation effect.

Summary

In these experiments, the prototype biosensor device for the detection of GFP_2 was modified to produce another nine biosensor device variants for the same target analyte. The device variants incorporated different promoter and RBS elements pairs that resulted in the

display of various levels of nanobody molecules at their cell surface. The varied levels of nanobody displayed modulated the analytical characteristics of the biosensing assay and achieved detection of GFP_2 at a range of analyte concentrations. Among the devices tested, the lowest concetration of GFP_2 analyte detected by bacterial agglutination was 4.5nM (01.13 picomoles). Additionally, the use of a promoter library tool for fine tuning the levels of nanobody display by bacterial cells was determined to be a more reliable strategy for the forward engineering of biosensor devices rather than RBS design.

3.2.8 Testing the biosensor design for sensor element modularity

Previous experiments successfully demonstrated proof of concept of a novel biosensor design for bacterial mediated agglutination tests (Section 3.2.1) by prototyping a device for the detection of GFP_2 analyte (Figure 3.13).To investigate the biosensor design modularity for sensor element incorporation, an experiment was designed to test a new prototype device that incorporates an alternative nanobody into the surface display vector protein.

Nanobodies that specifically recognise fibrinogen protein have been reported previously in the literature (Campuzano et al., 2014; Salema & Fernandez, 2013). Fibrinogen is a hexameric protein containing two sets of three different chains (α, β, and γ), linked to each other by disulphide bonds (Everse, Spraggon, Veerapandian, Riley, & Doolittle, 1998). Thus, a fibrinogen molecule exhibits divalency, with two identical epitope sites available for recognition by the anti-fibrinogen nanobody (Campuzano et al., 2014; Salema & Fernandez, 2013). This feature makes fibrinogen a suitable target for further testing the bacterial agglutination test developed here. Moreover, analytical diagnostic tests that measure fibrinogen levels can be of significant value. Fibrinogen is a plasma protein and low plasma levels are associated with increased risk of excessive bleeding (G. D. Lowe, Rumley, & Mackie, 2004). Additionally, abnormal fibrinogen levels in the blood have been associated with increased risk of cardiovascular diseases (Kannel, Wolf, Castelli, & D'Agostino, 1987) and fibrinogen has been reported to be an important determinant of the metastatic potential of circulating tumour cells (Palumbo et al., 2000).

Two biosensor devices were constructed for the detection of fibrinogen protein concentrations using the bacterial agglutination test developed here, by the transformation of *E. coli* DH10B strain cells with the pNVfib1 or the pNVgfp2 construct. The pNVfib1 and pNVgfp2 constructs were supplied by Prof. Luis Ángel Fernández of the National Center for Biotechnology (CNB), Madrid, Spain. These plasmids encode the NVfib1 and NVfib2 fusion proteins that are identical in structure to the NVgfp protein previously described (Figure 3.5) except than the anti-GFP nanobody was replaced by two anti-fibrinogen nanobodies (Campuzano et al., 2014). The expression of NVfib1 and NVfib2 protein constructs are under the control of the P_{lac} promoter and anti-fibrinogen nanobody expression in *E. coli* DH10B cells was carried out as previously described for the pNVgfp construct. Bacterial microplate

agglutination assays were prepared as previously described with commercially purchased fibrinogen protein purified from human plasma. The range of fibrinogen concentration tested in agglutination reactions were 250 picomoles and two-fold dilutions thereof to 244 femtomoles. Additionally, a test condition of no added fibrinogen was included in the assay. The *E. coli* DH10B cell line transformed with the pNVgfp construct that expresses the anti-GFP nanobody construct was used as the experimental negative control. This was used to control against any non-specific effects of fibrinogen to the cells. Agglutination test results were documented by photographs of the microplate back illuminated with blue light for easier visualisation.

The agglutination assay showed fibrinogen mediated bacterial cell agglutination (Figure 3.15). Images of the assay microplates showed absence of cell pellet formation for all fibrinogen concentrations tested -except at no fibrinogen condition- for both biosensor devices of *E. coli* DH10B cells transformed with pNVfib1 and pNVfib2 constructs. In regard to the biosensor device transformed with the pNVfib1 construct, a careful visual inspection of the assay microplate revealed some degree of cell pellet formation at 960 pM fibrinogen analyte concentration. In contrast to devices of *E. coli* DH10B cells transformed with pNVfib1 and pNVfib2 constructs, negative control reactions with *E. coli* DH10B cells transformed with pNVgfp construct showed bacterial pellet formation at all fibrinogen concentrations. This suggested that the addition of fibrinogen did not cause bacterial agglutination non-specifically. Rather, the widespread agglutination observed for the anti-fibrinogen biosensor devices was due to specific interactions between the anti-fibrinogen nanobodies displayed by the bacterial cells and fibrinogen protein molecules. Nevertheless, bacterial agglutination at such a wide range of fibrinogen protein concentrations was surprising. The particular mechanisms that resulted to this phenotype could not be established but possible explanations could take into account the particular nature of the fibrinogen protein and its propensity to form higher order structures when is processed by proteases to produce fibrin protein.

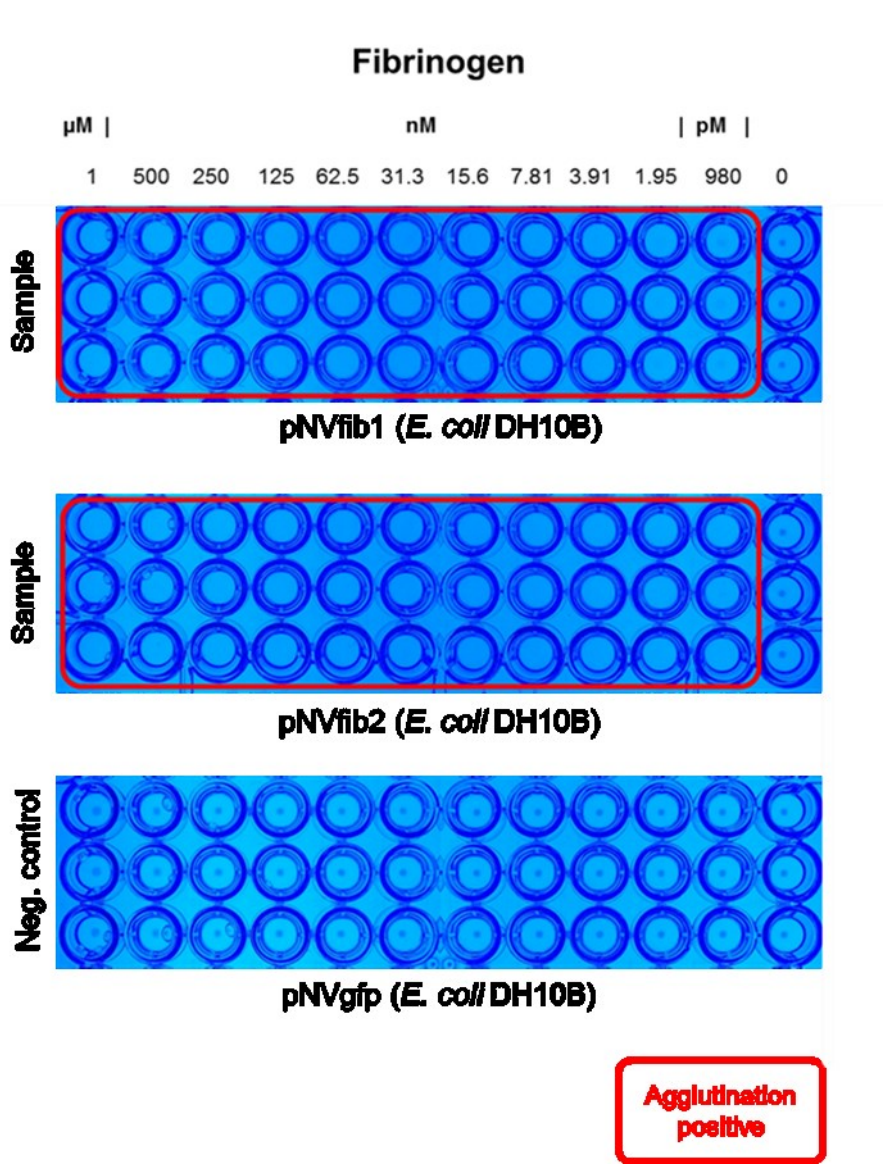


Figure 3.15: Bacterial cell agglutination of anti-fibrinogen biosensor devices. Images of round bottom well microplates in which cell suspensions of whole cell biosensor devices were incubated with various concentrations of purified fibrinogen protein concentrations. The biosensor devices were *E. coli* DH10B cells transformed with the pNVfib1 or pNVfib2 constructs and induced for cell surface display of anti-fibrinogen nanobodies (Vfib1 and Vfib2); or transformed with the pNVgfp plasmid and displayed the anti-GFP nanobody (Vgfp) and used as the negative control. Cell suspensions ($O.D._{600}$ = 1) were added fibrinogen protein (total reaction volume = 250µl) and cells were allowed to settle overnight. Negative results for bacterial agglutination resulted in pellet formation, while bacterial agglutination resulted in lack of cell pellet. Images show microplates back-illuminated with blue light for easier visualisation of cell pellet. NVfib1, NVfib2 and NVgfp expression was carried out at 30°C by IPTG induction.

Summary

In this experiment the biosensor design modularity was tested for sensor element incorporation. Two new biosensor devices were constructed that surface displayed two different anti-fibrinogen nanobodies. Cell suspensions of these two biosensor devices were used in a microplate agglutination assay for the detection of purified fibrinogen blood protein biomarker. Both devices successfully detected the target analyte by the formation of aggregated clumps of bacterial cells down to 960pM of fibrinogen analyte (244 femtomoles). This output is very similar to what was observed for the prototype biosensor device that surfaced displayed anti-GFP nanoby and was used to detect the GFP_2 model target analyte. This suggests that the biosensor design is modular for sensor element incorporation. However, the fibrinogen biosensor devices exhibited bacterial agglutination at a much wider range of target analyte concentrations in comparison to the GFP_2 prototype device and all its variants that exhibited bacterial agglutination at a more restricted range.

3.3 Technical Discussion

This chapter describes the development of a novel biosensor design for the detection of protein biomarkers with intended use in medical diagnostic applications. The design uses nanobodies as the sensing elements of the biosensor device to leverage the high specificity and affinity that this class of sensing elements can exhibit for their target molecule. The nanobodies are genetically encoded in *E. coli* cells, and the cells make use of a display vector to express and display the nanobody molecules on their cell surface. The cell-surface display of the nanobody molecules allows the cells to detect analytes in the extracellular environment. In the presence of the multivalent target analyte in the extracellular environment, the simultaneous binding of the analyte molecule to more than one bacterial cell crosslinks the cells. Overtime, extensive cross-linking between cells and analyte molecules results in large aggregates of bacterial cells. The formation of aggregates or the lack of them is the output of this bacterial agglutination test. The information of the presence/absence of the target analyte molecule in the sample is thus reported to the user of the test that is able to visibly observe the test results.

The aforementioned design was tested by building a prototype biosensor device for the detection of GFP_2. For the construction of the prototype device, *E. coli* DH10B cells were transformed with the pNVgfp construct for the expression and surface display of the anti-GFP nanobody. Experiments with a different cell host chassis, *E. coli* BL21(DE3) strain, failed to surface display the anti-GFP nanobody. This could be attributed to the different genotype of the strain which lacks a number of outer membrane proteases which may be of importance to the processing of surface displayed proteins.

The prototype device cells were tested for induced cell toxicity upon NVgfp protein expression by monitoring the growth curves of induced cell cultures. Results showed negligible cell toxicity to the host chassis by the expression of the NVgfp construct (Figure 3.7). This method for measuring cell toxicity was preferred over other methods for its easy implementation and use of standard molecular biology lab equipment. However, this method provided information only for the overall phenotypic fitness of the host cell. Alternatively, other methods could have been employed for studying the effects of heterologous protein expression

to the host chassis. The burden imposed to the cell host could have been determined by the use of an internal reference standard, such as a co-expressed fluorescent protein reporter (Ceroni, Algar, Stan, & Ellis, 2015). This method would have enabled the assessment of whether cell toxicity arises from the excessive use of the cell's shared pool of resources (ex. ribosomes, RNA polymerases) by the synthetic construct. Even more, cell toxicity caused specifically by the destabilisation of the outer membrane of the host could be studied by other methods such as leakage assays (Georgiou et al., 1996).

Next, the surface localisation of the NVgfp display vector and the functionality of the anti-GFP nanobody were assessed by flow cytometry. Whole cells were labelled with fluorescently conjugated antibodies that targeted the NVgfp construct and the increase in cell fluorescence demonstrated the successful localisation of the display vector protein to the outer membrane of the *E. coli* cell (Figure 3.8). In addition, whole cells were labelled with purified GFP and the increased cell fluorescence as a result of GFP binding by the anti-GFP nanobody demonstrated the functional display of the nanobody molecules. Also, this experiment provided evidence for successful binding of target molecules in the extracellular milieu. Cell labelling with flow cytometry is a straightforward and easy technique for the detection of surface localising proteins in bacteria when labelling agents are readily available for the target protein or if the protein under investigation can be genetically engineered to display a peptide tag (ex. myc tag, his tag). Other methods exist that could have been used to demonstrate the presence of a protein in the bacterial outer membrane, such as cell fractionation with SDS-PAGE/western blotting (Georgiou et al., 1996) or protease accessibility assays (Salema et al., 2013), but these are significantly more laborious.

The biosensor device was further characterised by the quantification of the number of nanobodies displayed on the cell surface per cell by the use of flow cytometry. The number of displayed nanobodies was found to be 6,750 per cell (Figure 3.9). A previous study that used the same display vector system incorporating a different nanobody found 7,800 copies of the pNVhh display vector per bacterial cell (Salema et al., 2013) by employing a western blotting based methodology. The flow cytometry based method offered the advantage of accounting for any possible non-functionally displayed nanobodies as opposed to the western blot based

methodology. However, it necessitated the use of a fluorescent analyte molecule, but in future biosensor designs this limitation could be overcome by chemical/biological labelling of the analyte molecule or the adoption of a competitive binding assay format (G. Chen et al., 1996).

To enable the testing of the proposed biosensor design in an agglutination assay format, a GFP_2 model analyte was genetically engineered that was expected to exhibit divalency for the anti-GFP nanobody. The engineered GFP_2 analyte was purified and was assessed for correct folding of both its constituent GFP chromophores by comparing its fluorescence signal to equimolar amounts of monomeric GFP protein. The results of this functionality based assay suggested that a large percentage of the GFP_2 fusion protein folded to a functional structure. Nevertheless, circular dichroism spectroscopy can be used to provide further evidence that would demonstrate the correct folding of GFP_2 by comparing its secondary structure against the native secondary structure of the GFP monomer (Pelton & McLean, 2000).

A kinetic mathematical model of the agglutination reaction and its algorithmic implementation were used to aid in the identification of a range of GFP_2 analyte concentrations to be tested in wet lab experimentation for the agglutination assay. The agglutination reaction simulation suggested a range of GFP_2 analyte concentrations where bacterial agglutination would likely manifest (Figure 3.12). Input parameters for the model were derived from experimental data, the literature and by estimation (Hung El, 2014). The number of nanobodies displayed by each cell were derived from experimental work carried out specifically for this system (Figure 3.9). Nevertheless, this number is the average number of nanobodies per bacterial cell and variance between the population members may need to be considered for increased accuracy of the model. The kinetics of the GFP and nanobody association and dissociation were acquired from the literature (Rothbauer et al., 2006) as well as the number of cells in cell suspension of the chosen $O.D._{600}$ value. Alternatively, the number of cells at a particular $O.D._{600}$ value for the particular *E. coli* DH10B strain can be calculated by counting colony forming units. Other parameters with values not available in the literature -such as the dynamics of aggregation interaction among others- were given arbitrary values and in the future the model in combination with experimental methods (Herrmann et al., 2015) could be used to define them. Alternatively to the mathematical model approach, the analyte

concentrations at which bacterial cells agglutinate could be defined by experimental trial and error procedures.

Following, the prototype biosensor device for the detection of GFP_2 analyte was tested in a microplate agglutination assay format. The prototype device showed specific bacterial agglutination at specific GFP_2 analyte concentrations (Figure 3.13) and thus successfully demonstrated proof of concept for the proposed biosensor design (Figure 3.4). The positive output of the agglutination test appeared in the form of a membranous structure deposited in the bottom of the microplate well and was easily observed by naked eye. The visible output can be a desirable feature for diagnostic applications at the point-of-care or at resource limited settings as it requires no instrumentation to relay the test results to the user. This claim is supported by the widespread use of latex agglutination technology (Molina-Bolívar & Galisteo-González, 2005). Importantly the positive test results were easily distinguished from agglutination reactions with negative results for the target analyte (Figure 3.13). In its current format, the agglutination test results were qualitative or semi quantitative at best but microfluidic methods exist that could provide quantitative test results (Herrmann et al., 2015).

Next, the prototype device was modified by genetic engineering to produce another nine functional device variants with varied analytical ranges for GFP_2 detection (Figure 3.14). The varied ranges were a result of different expression levels of surface displayed nanobody molecules per cell for each device variant. This variation was achieved by the use of a library of promoter elements and designed RBS elements. The promoter elements and RBS elements were used in various combinations to control expression levels of the NVgfp construct that incorporates the anti-GFP nanobody. The variation in the number of displayed nanobodies per bacterial cell between device variants was demonstrated directly by flow cytometry and indirectly by agglutination assays (Figure 3.14). The difference in the levels of displayed nanobodies achieved between the lowest and highest expressing biosensor device variant was 19-fold as determined by flow cytometry. This variation in the levels of nanobody displayed allowed the detection of varied levels of GFP_2 analyte in agglutination assays within a similar analytical range. Alternatively, the analytical range of the agglutination test for GFP_2 detection could have been varied by controlling the number of cells participating in the agglutination

reaction. This could have been achieved by modulating the density of the cell suspension in the microplate wells or the volume of the cell suspension added. It is possible that a combination of the genetic engineering approach and adjusting the number of cells participating in the agglutination reaction could produce test assays capable of detecting analyte concentrations across a very wide analytical range.

Additionally, the construction and testing of device variants facilitated the identification of strategies for more reliable forward engineering of biosensor devices of this type. The device variants were constructed by the utilisation of genetic elements that enabled the fine tuning of protein expression (Table 3.1). These were promoter elements from the Anderson collection (Registry_of_Standard_Biological_Parts, 2016) and designed RBS elements by the RBScalculator tool (Salis et al., 2009). Data analysis of experimental data from flow cytometry measurements regarding the levels of displayed nanobodies suggested that the biosensor devices can be forward engineered more reliably with the use of promoter library tools rather than RBS design (Table 3.2; Table 3.3). The reason for the more consistent functionality of the promoter elements could be their modularity in respect to the downstream coding gene sequence. Current technologies for RBS design probably have not yet achieved such a degree of functional modularity.

The modularity of the biosensor design for the incorporation of alternative sensor elements was investigated by the construction of two prototype biosensor devices for the detection of fibrinogen analyte. The demonstration of bacterial aggregation in agglutination assays for the detection of fibrinogen analyte successfully demonstrated the modularity of the biosensor design (Figure 3.15). The test results suggested that the bacterial agglutination was due to specific interactions between anti-fibrinogen nanobodies and fibrinogen protein rather than non-specific cell aggregation. Further evidence to support the modularity of the biosensor design towards sensor element incorporation could be provided by demonstrating cell-surface display of the anti-fibrinogen nanobody and specific recognition of fibrinogen protein. These experiments could be carried out with methods discussed previously (cell labelling) or other methods such as bacterial adhesion to coated abiotic surfaces (Pinero-Lambea et al., 2015).

Chapter 4

4 Biosensor design: RNA-aptamer & *E. coli* cell-extract

4.1 Background Information

4.1.1 RNA-aptamers

RNA-aptamers are oligonucleotides sequences that can form defined scaffold structures to recognise target molecules. The word aptamer is derived from the the latin word “aptus” which means to fit and the Greek word ‘’meros’’ which means part. Typically, aptamers are 15 – 45 nucleotides long (5 – 15 kDa) and can recognise the target molecule with affinities down to nanomolar and sub-nanomolar range, rivalling affinities achieved by monoclonal antibodies. Also, they are capable of very specific interactions and can discriminate among closely related targets. In the case of the theophylline aptamer, this RNA-aptamer displays an affinity for Theophylline ligand that is 10,000 times stronger than its affinity for caffeine, despite the fact that caffeine differs from theophylline just by one methyl group (Jenison et al., 1994). RNA-aptamers can be developed to recognise any chemical target such small ligands as in the case of ATP (Sassanfar & Szostak, 1993), amino acids such as arginine (Yang, Kochoyan, Burgstaller, Westhof, & Famulok, 1996), oligosaccharides such as the tobramycin antibiotic (Y. Wang & Rando, 1995) and proteins such as the thrombin protease. (Long, Long, White, & Sullenger, 2008).

RNA-aptamers are able to recognise their target through various types of binding interactions including hydrogen bonding, stacking interactions, electrostatic interactions, shape complementarity and hydrophobic contacts. These interactions have been observed in structural studies that helped elucidate the basis for the high selectivity of RNA-aptamers. For example, the molecular basis for the high selectivity of the aforementioned theophylline binding RNA-aptamer against the chemically similar caffeine has been demonstrated in an NMR structural study (Zimmermann, Jenison, Wick, Simorre, & Pardi, 1997). In this study, the solution structure of the complex showed a binding pocket formed by the RNA-aptamer that locked theophylline in a complex network of hydrogen bonding and stacking interactions. At

the same time a cytosine base in the binding pocket conferred the selectivity against caffeine, as caffeine accommodation in the binding pocket would require the breaking of favourable hydrogen bonding and steric clash between the methyl group of the caffeine and the base.

RNA-aptamers can be developed *in vitro* and selected for high affinity and selectivity towards any target molecule by Systematic Evolution of Ligands by Exponential Enrichment (SELEX), and such protocols have been initially described in the early 1990s (Ellington & Szostak, 1990; Tuerk & Gold, 1990). A SELEX protocol for the development of RNA-aptamer starts with the generation of a library of random RNA sequences (~10^{14}) transcribed *in vitro* from a DNA library. This pool of sequences is then selected (or evolved) in a 4 phases cycle until an aptamer meeting the specifications required is isolated (Figure 4.1). During complex formation phase 1, the target molecule is incubated with the pool of sequences for RNA/target complex formation. In partition phase 2, the RNA/target complex is partitioned from unbound members of the pool of library sequences. The non-binder RNA sequences are removed from the process. In the elution phase 3, the RNA binders are eluted from the target molecule and in amplification phase 4, the enriched population of binders is reverse transcribed and amplified to undergo further rounds of selection. Since the first publications describing SELEX, the process was developed and refined further. Initially, the SELEX procedure was optimised for eliminating binders adsorbed to the matrix used to immobilise the target (Ellington & Szostak, 1992) and then for counter-selection of aptamers that show affinity for similar targets (Jenison et al., 1994). Later, SELEX protocols were developed for particular applications (Aquino-Jarquin & Toscano-Garibay, 2011) and additionally the selection procedure was automated (Eulberg, Buchner, Maasch, & Klussmann, 2005).

In this chapter, a design was developed that incorporates RNA-aptamers as the sensing element of a biosensing technology for the detection of protein biomarkers. A SELEX derived anti-thrombin RNA-aptamer (X. M. Liu et al., 2004; X. M. Liu et al., 2003) was used for the initial proof-of concept experiments. Then, toggle-25t RNA-aptamer (White et al., 2001) and anti-VEGF RNA-aptamer (Jellinek, Green, Bell, & Janjic, 1994) were used to investigate the modularity of the design for sensing element incorporation. These RNA-aptamers were selected from the literature as they were used successfully by more than one research group and their target proteins were available commercially in purified form. Following, background information is provided for each of the aforementioned RNA-aptamers.

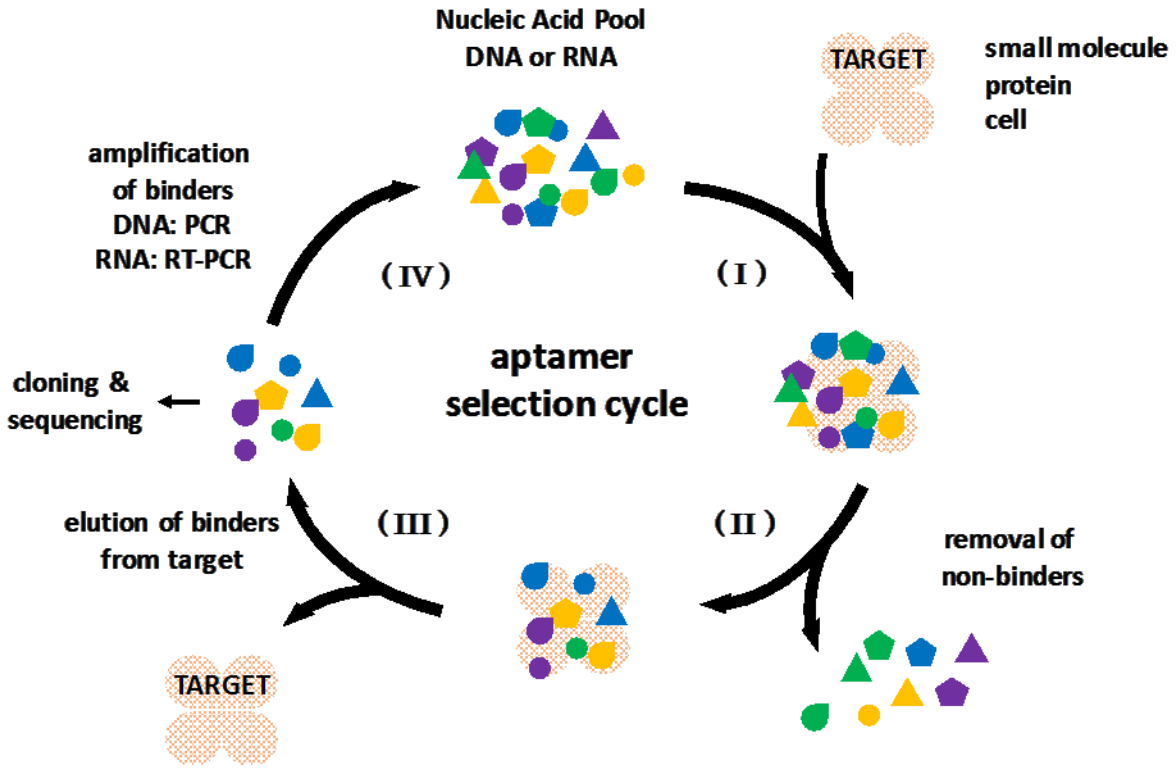


Figure 4.1: Systematic Evolution of Ligands by Exponential Enrichment (SELEX). Schematic of the *in vitro* evolution procedure. For the evolution of RNA-aptamers, the procedure starts with the generation of a library of random RNA sequences (~10^{14}) transcribed in vitro from a DNA library. Then, this pool of sequences is selected in a 4 phases cycle of (i) complex formation, (ii) partition, (iii) elution and (iv) amplification, until an aptamer meeting the specifications required is isolated.

Anti-thrombin RNA-aptamer

This RNA-aptamer was isolated with SELEX procedure for specific binding to bovine thrombin protein (X. M. Liu et al., 2003) (Figure 4.2). The specificity of the *in-vitro* evolved oligonucleotide binder was demonstrated by electrophoretic mobility shift assay, and the affinity of the anti-thrombin RNA-aptamer for its target protein was found to be 164nM by nitrocellulose filter binding assay (X. M. Liu et al., 2003). Further work on the isolated RNA-aptamer managed to identify a minimal RNA-aptamer region that is sufficient for target recognition (X. M. Liu et al., 2004). Although initially developed for discrimination between bovine and human isoforms of thrombin protein, this RNA-aptamer has been reported to bind human thrombin protein as well (Y. Li, Lee, & Corn, 2007).

Toggle-25t RNA-aptamer

Toggle-25t RNA-aptamer is also specific for thrombin protein binding and it was developed by a specialised SELEX procedure, namely Toggle SELEX (White et al., 2001). In this procedure, the pool of oligonucleotides in SELEX phase (i) was toggled between human thrombin and porcine thrombin in successive rounds to isolate a RNA-aptamer that recognized a conserved region of the thrombin protein isoforms. The *in vitro* evolved RNA-aptamer (Figure 4.2), was found to bind human thrombin with a dissociation constant of 0.54nM in the aforementioned study and porcine thrombin with a dissociation constant of 1.4nM as determined by nitrocellulose filter binding assay. Additionally, a study crystallised the complex and showed that the secondary structure of the toggle-25t RNA-aptamer bound to the human thrombin was a stem loop structure with an internal bulge (Long et al., 2008).

Anti-VEGF RNA-aptamer

A number of anti-VEGF RNA-aptamers were selected in an effort to develop potent and specific VEGF antagonists and were identified to belong to six different families of consensus sequences (Jellinek et al., 1994). Of these anti-VEGF ligands, a RNA-aptamer family 2 member which was truncated to its minimal sequence and named ligand 44t (Figure 4.2) showed a biphasic binding to its target protein with dissociation constants of 0.48nM and 82nM as determined by nitrocellulose filter binding assay.

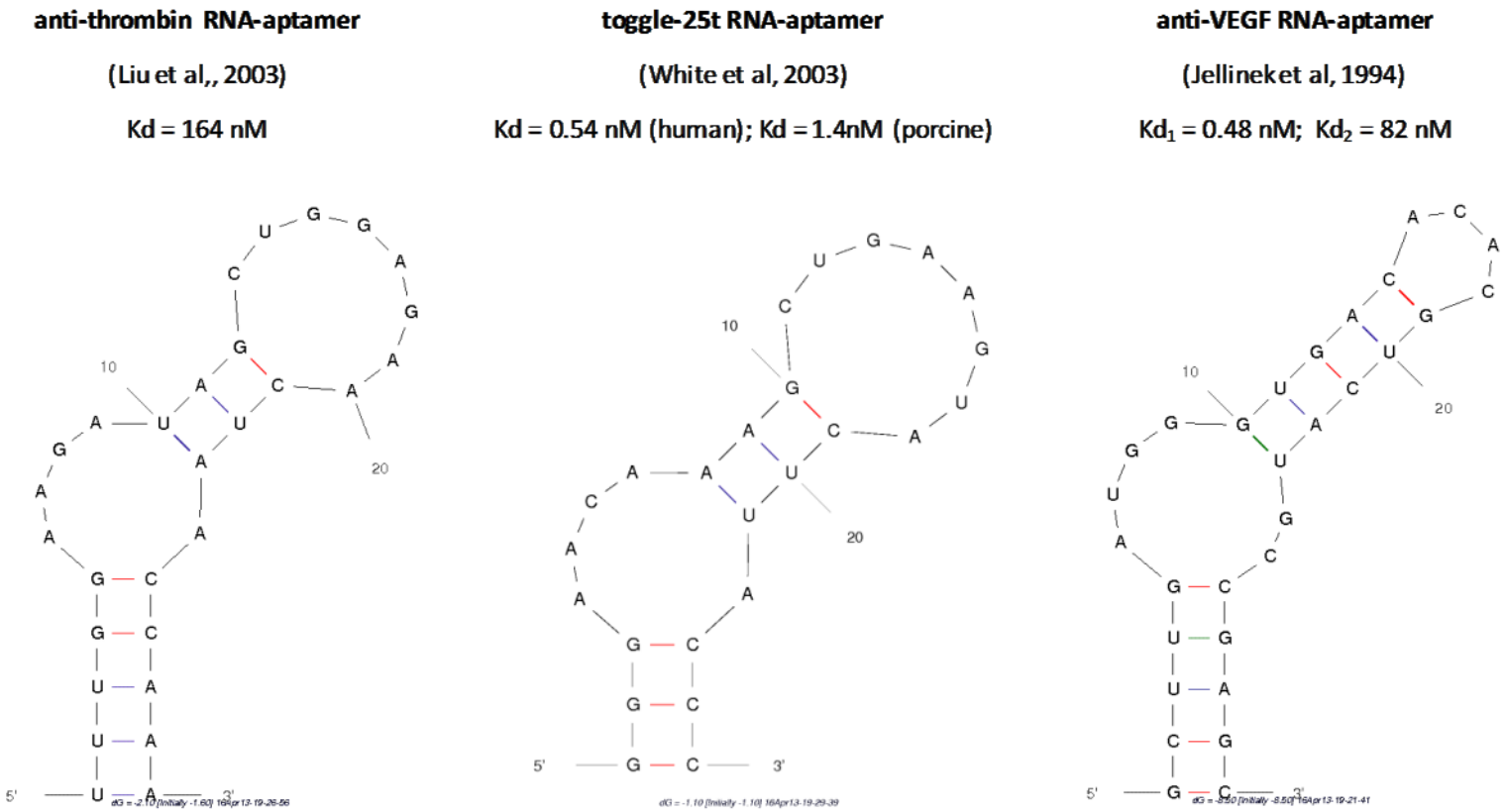


Figure 4.2: Secondary structure of RNA-aptamers. Schematics of the predicted secondary structure of RNA-aptamers used in this chapter and dissociation constants for their target. The schematics were generated using mFold software (Zuker, 2003).

4.1.2 Bacterial translation initiation and translation initiation rate prediction

The information stored in the DNA –cell's long term information storage medium- is transcribed into mRNA that contains information in the form of amino acid coding sequences (CDS) before translated into proteins that constitute the molecular machinery of the cell. The ribosome, a large ribonucleoprotein complex, is the main player of the translation process which process is conceptually divided into four phases: i) translation initiation, ii) elongation, iii) termination and iv) ribosome recycling (Ramakrishnan, 2002). During translation initiation phase, the ribosome assembles on the mRNA molecule and proceeds to the elongation phase where the polypeptide chain is synthesized. Then, at the termination phase newly synthesized protein is released from the ribosome upon encounter of a stop codon, and then follows the ribosome recycling phase where the ribosomal subunits dissociate and release from the mRNA.

The process of translation initiation is highly regulated with a large number of biomolecular factors involved (Laursen, Sorensen, Mortensen, & Sperling-Petersen, 2005). Other than the bacterial ribosome -that is composed of a large 50S subunit and a small 30S subunit (Cate, Yusupov, Yusupova, Earnest, & Noller, 1999)-, the process of translation initiation involves the aminoacylated and formylated initiator $tRNA^{fMeth}$, initiation factor (IF1), initiation factor (IF2), initiation factor (IF3) and the mRNA molecule. The process initiates with the binding of the IF1 and IF3 to the assembled ribosome 30S subunit, and the action of the IF3 promotes the dissociation of the assembled ribosome into its two subunits. Following dissociation, initiator tRNA, IF2 and the mRNA molecule associate with the 30S ribosomal subunit. In the newly formed complex, $tRNA^{fMeth}$ base pairs with the start codon of the CDS on the mRNA and the Shine Delgarno (SD) sequence of the mRNA base pairs with the anti-SD sequence on the 16S rRNA of the 30S ribosomal subunit (Shine & Dalgarno, 1974; Yusupova, Yusupov, Cate, & Noller, 2001). Then, a multistep process results in the formation of a stable initiation complex in which IF1 and IF3 are ejected and IF2 stimulates the association of the 50S ribosomal subunit to the complex. During the process a GTP molecule is hydrolysed to GDP with IF2 released from the complex and this results in the formation of the 70S initiation complex ready to enter the elongation phase of the translation process. The translation initiation process is diagrammatically illustrated in Figure 4.3.

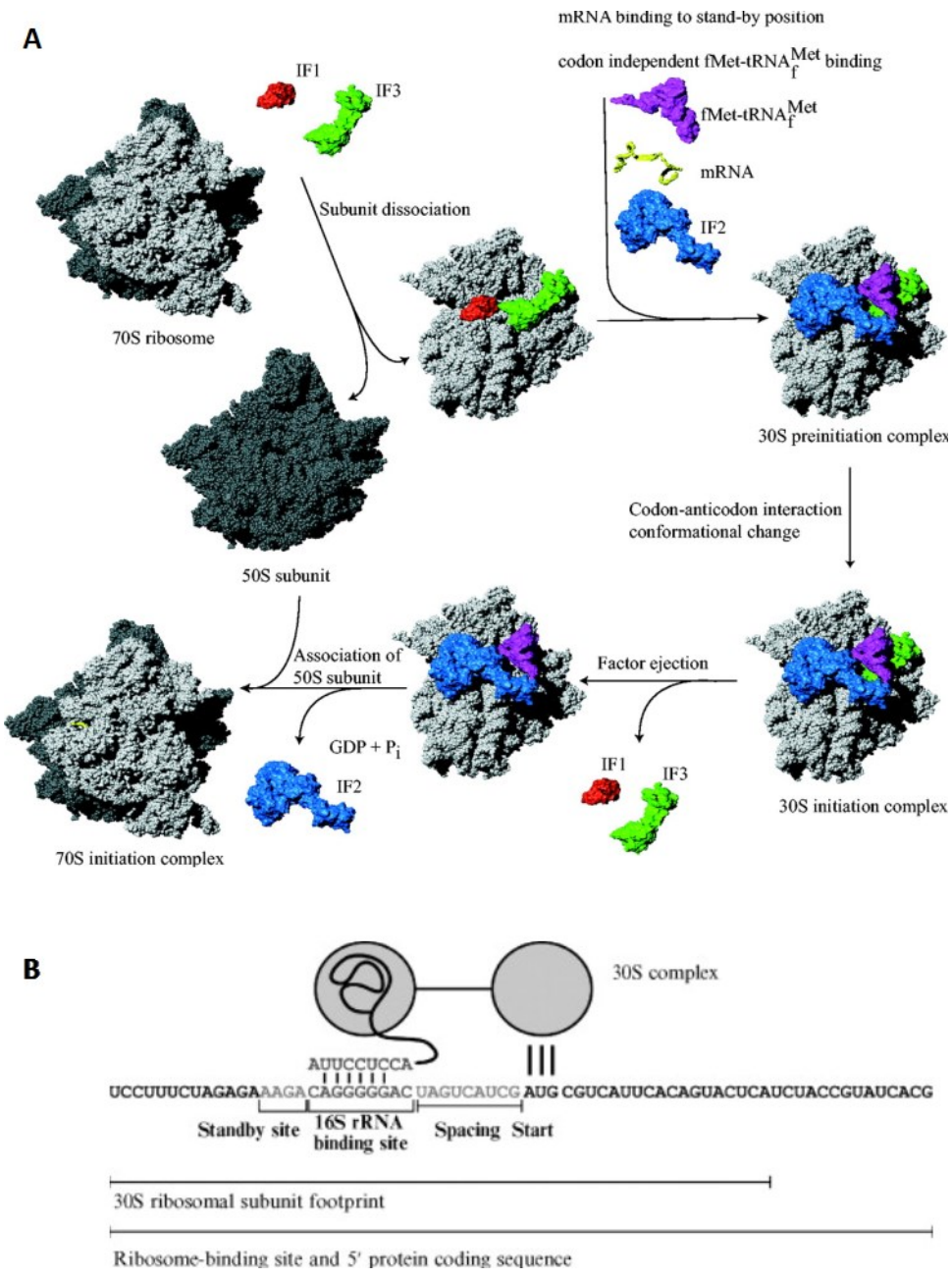


Figure 4.3: Translation initiation in bacteria. A Schematic of the translation initiation pathway in bacteria incorporating structural information of biomolecular factors involved in the process from experimental data. The pathway results in the formation of the 70S initiation complex that is able to proceed to the elongation phase of the translation process. Reproduced with permission from (Laursen et al., 2005). **B** The ribosome binding site as defined by the RBScalculator v01. Reprinted by permission from Elsevier: Methods in Enzymology (Salis, 2011), copyright 2011.

The ribosome footprint region is the length of mRNA that is accommodated and protected by the ribosome when the two biomolecules form a complex. Assuming the formation of the initiation complex, this covers the area from the -35 mRNA position in reference to the AUG-start codon and extending to the +19 position downstream the start codon (Salis, 2011). This area covers a mRNA length of 54 nucleotides and the ribosome footprint region was determined initially by nuclease protection assays (Steitz, 1969), as well as with probing with hydroxyl radicals (Hüttenhofer & Noller, 1994) and confirmed by crystallographic studies (Yusupova et al., 2001). Traditionally, biologists referred to the Shine-Delgarno sequence located generally 8 base pairs upstream the start codon as the ribosome binding site (RBS). However, more recently the definition of the RBS as a genetic element for the control of translation initiation has been expanded to include more elements of the ribosome footprint region as well as elements outside of it (Figure 4.3) (Salis, 2011).

The development of technologies that allow reliable and precise control of the translation initiation process is of great interest for their useful application in research and biotechnology. Examples of such software tools include the RBScalculator (Espah Borujeni, Channarasappa, & Salis, 2014; Salis et al., 2009), RBS designer (Na & Lee, 2010) and UTR designer (Seo et al., 2014). The RBScalculator -the most popular RBS design tool among synthetic biologists- is able to reverse engineer and forward engineer RBS elements by quantifying the translation initiation rate (TIR). The calculator uses a model based on thermodynamics, where the TIR is proportional to the difference in the free energy (ΔG) between the secondary structure folded mRNA and the assembled mRNA/ribosome initiation complex (Salis et al., 2009). Components of the model to calculate ΔG values include the minimum free energy secondary structure of the RBS part as determined by secondary structure prediction tools (discussed in section 4.1.3), free energy of hybridisation for SD sequence and anti-SD of 16S rRNA, free energy of hybridisation for the initiation tRNA to the AUG codon and energetic penalties for imperfect spacing between SD and AUG start codon. In addition, the updated version of the tool, RBScalculator 2.0, incorporates thermodynamic contributions from standby sites far upstream of the RBS genetic part (Espah Borujeni et al., 2014).

In this chapter, a biosensor design was developed in which an RNA-aptamer situated within the ribosome footprint region aimed to control the translation initiation rate of mRNA encoding a reporter protein. In addition, the RBScalculator software tool was used in an attempt to forward engineer the signal output of constructed biosensing devices.

4.1.3 Ligand inducible control of bacterial translation initiation

In biological systems, a change in the state of the system such as the depletion of a nutrient or the availability of a new energy source results in the change of the gene expression profile of that organism in order to maximise chances of survival. The information carriers of changing conditions most often take the form of ligand molecules that are able to interact with the gene expression machinery of cells. In bacteria, control of protein synthesis at the translation level is most frequently encountered at the phase of translation initiation, as the ribosome has the ability to disrupt the secondary structure of mRNA by intrinsic helicase activity once it enters the elongation phase (Takyar, Hickerson, & Noller, 2005). The chemical nature of ligands that act upon the process of translation initiation is diverse and includes RNA ligands (Franch, Gultyaev, & Gerdes, 1997), small molecule metabolite ligands (Winkler, Nahvi, & Breaker, 2002) and proteins (Peabody, 1990).

Ligand-inducible RNA-switches are conceptually divided into two modules; the mRNA platform for the expression of the gene of interest and a RNA-aptamer domain for the specific recognition of the ligand of interest (Mandal & Breaker, 2004; Montange & Batey, 2008; Serganov & Nudler, 2013). The mRNA platform consists of the protein coding sequence with its start codon (most commonly AUG) and the 5'UTR (untranslated region) that includes expression control elements such as the Shine-Delgano sequence and standby site. These two modules work together so that during ligand binding the information of target recognition by the RNA-aptamer is relayed to the mRNA platform by a mechanism that alters the availability of translation initiation region for ribosome entry. Among the systems studied for the control of translation initiation by ligand inducible translational level RNA-switches, two are the most common mechanisms encountered for controlling ribosome access to the ribosome binding site and formation of the initiation complex: i) ligand dependent availability of the translation initiation region elements (SD sequence and AUG start codon) by secondary structures (base pairing) or ii) direct competition of the ligand between the ribosome for translation initiation region binding that is sterically mutually exclusive. These two control mechanism are schematically presented in Figure 4.4.

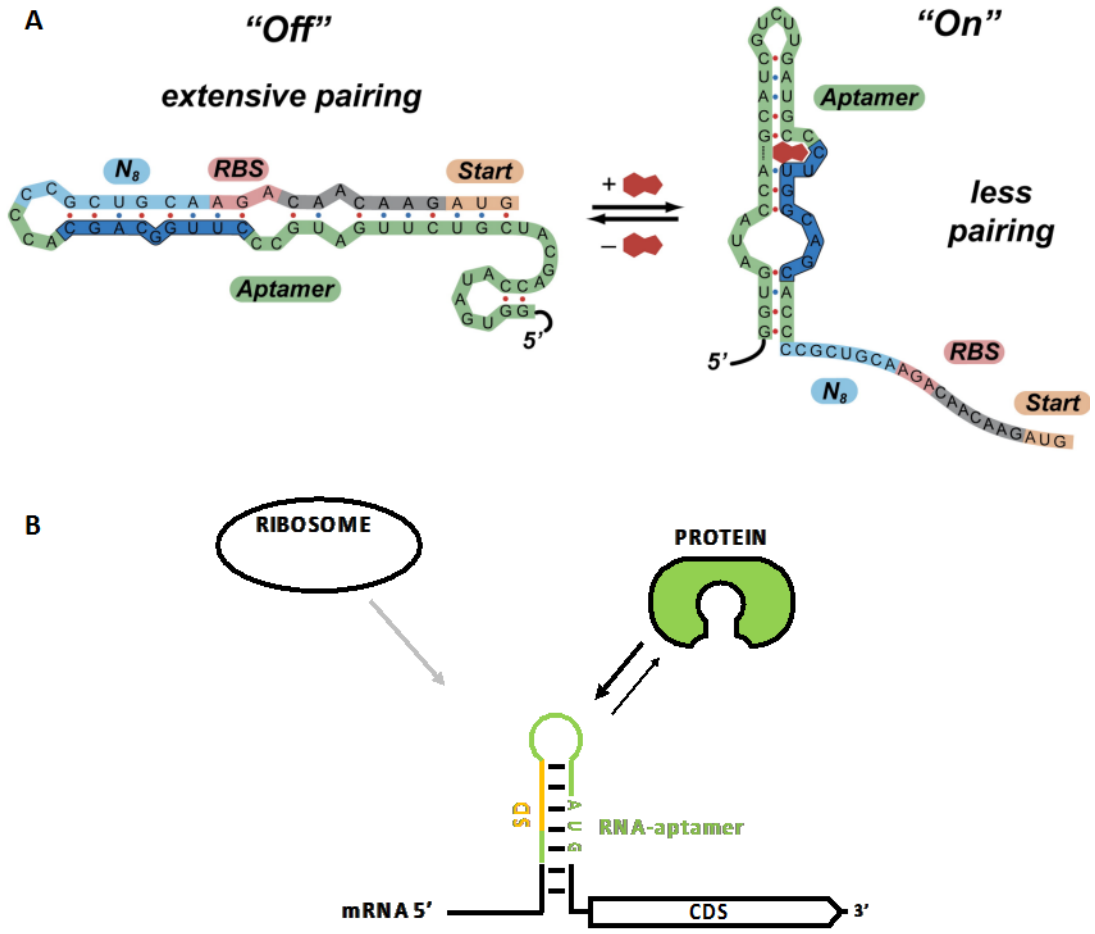


Figure 4.4: Mechanisms of ligand inducible control of translation initiation. A Control of translation initiation by ligand binding through induced secondary structure rearrangements. In absence of ligand (red shape), extensive base pairing between RNA-aptamer bases and translation initiation region (including RBS (or SD sequence) and AUG start codon) bases prevent ribosome entry. In presence of ligand, secondary structure rearrangements result in single stranded translation initiation region that is a more energetically favoured structure for ribosome entry and translation initiation. Reprinted from Chemistry & Biology, 14/2, Lynch, Desai, Sajja, & Gallivan, A High-Throughput Screen for Synthetic Riboswitches Reveals Mechanistic Insights into Their Function, 12., Copyright (2007), with permission from Elsevier. **B** Control of translation initiation by ligand through direct competition against the ribosome for translation initiation region binding. In absence of ligand, the translation initiation region is free for ribosome entry and translation initiation. In the presence of the ligand (green protein), the ligand sterically inhibits entry of the ribosome to the translation initiation region. The schematic presented resembles the translational operator site of MS2 virus replicase gene control by MS2 protein (Peabody, 1990).

The first of the aforementioned mechanisms is commonly encountered for small molecule metabolite ligands. Binding of the metabolite to the RNA-aptamer domain causes secondary structure rearrangements to the mRNA molecule that change the availability of the translation initiation region by the formation/breaking of base pairing (Figure 4.4 panel A) (Mandal & Breaker, 2004). Such RNA-switches for small molecule metabolites are referred as riboswitches and are ubiquitous across the bacterial kingdom (Serganov & Nudler, 2013). Researchers have also explored methodologies for engineering synthetic riboswitches by library and rational engineering approaches. For example, the Gallivan lab used high throughput screening to isolate a theophylline riboswitch from a library of mutants generated by randomising part of the translation initiation region (Lynch et al., 2007). This riboswitch in the absence of theophylline inhibited gene expression by extensive base pairing of the nucleotides of the theophylline binding RNA-aptamer with the elements of the translation initiation region. In contrast, in the presence of theophylline the riboswitch developed a more energetically favourable secondary structure that allowed translation initiation and protein expression by the formation of single stranded translation initiation region. On the other hand, the Salis lab demonstrated a rational engineering approach for riboswitch design (Espah Borujeni et al., 2014; Salis et al., 2009). In aforementioned research, the functionality of the RBScalculator algorithm was extended to enable the automated forward engineering of ligand activated riboswitches for theophylline, tetramethylrosamine, fluoride, dopamine, thyroxine and 2,4-dinitrotoluene based on thermodynamic models.

The second mechanism for ligand induced control of translation initiation involves the direct competition of the ligand and ribosome for binding to the translation initiation region in a mutually exclusive manner (Figure 4.4 panel B). This mechanism has been most commonly observed in bacteriophage derived mRNA sequences and in the regulation of ribosomal protein production in *E. coli* (Babitzke, Baker, & Romeo, 2009). For example, the coat protein of the MS2 phage virus has been shown to be able to recognise an RNA-aptamer domain (usually called translational operator in this context) in the translation initiation region of the viral replicase gene and repress its expression. In particular, the MS2-coat protein was shown to recognise a hairpin structure that its sequence includes both the Shine-Delgano sequence and

the AUG start codon of the replicase gene (Peabody, 1990). Binding of the coat protein effectively blocks entry of the ribosome, with the two molecules being in direct competition for this mutually exclusive binding site. Other than the naturally occurring examples, there has been a demonstration of a synthetic RNA-switch that utilised this mechanism of control (Saito et al., 2010). In this synthetic RNA-switch, a natural RNA motif able to recognise the L7Ae protein (analogues function to an RNA-aptamer) was incorporated into a GFP expression platform mRNA directly downstream of the AUG start codon. Thus, the L7Ae protein directly competed against the ribosome for the translation initiation region and the L7Ae protein presence successfully repressed GFP expression.

In this chapter, a biosensor design was developed based on the mechanism of direct competition between ligand and ribosome. Based on this biosensor design, RNA-switch devices were constructed that incorporated SELEX derived RNA-aptamers in the translation initiation region of an mRNA expression platform. Incorporation of the RNA-aptamer in the mRNA-platform encoding a reporter protein aimed to transform the platform into an analyte controlled switch of a gene expression.

4.1.4 RNA secondary structure prediction

RNA has diverse roles in cellular life including the regulation of gene expression (Kortmann & Narberhaus, 2012), immunity (Horvath & Barrangou, 2010) and catalysis (Scott & Klug, 1996) amongst others for which the structure of the RNA molecule plays a significant role. The primary sequence of RNA initially folds into a secondary structure and then to a tertiary structure. The secondary structure of the RNA is a very important determinant to the final structure of the molecule and thus its function, and methodologies have been developed to predict the secondary structure computationally.

The principle of Free Energy Minimisation is one of the approaches used for the modelling and prediction of the secondary structure of RNA molecules (Mathews & Turner, 2006). According to this principle, if the secondary structure of an RNA molecule was solely dependent on the thermodynamics, then the primary sequence of RNA could be folded to its secondary structure by finding the base pairing combination for this molecule that results in the largest energy release at a particular temperature. Studies have suggested that thermodynamics is a major determinant of RNA secondary structure, as a 770 nucleotides long RNA sequence can be folded with tools that apply free energy minimisation to predict up to 70% of the experimentally determined base pairs (Mathews et al., 2004).

Starting from an RNA sequence in an unfolded state, the sequence can be folded to its minimum energy secondary structure state with the release of Gibbs free energy (ΔG) equal to the energy difference between the two states. This energy difference can be calculated from the summation of free energies of individual RNA structure motifs and modelled by the nearest neighbour method for which the free energy parameters have been determined experimentally in calorimetric experiments (Mathews et al., 2004; Xia et al., 1998). Software tools for RNA secondary structure prediction use dynamic programming algorithms to efficiently identify the minimum free energy (Mathews, 2006).

Nevertheless, factors such as the incompleteness of the thermodynamic models used for the secondary structure predictions, experimental uncertainties and limitations of nearest-neighbour models can lead to secondary structure predictions that are not biologically relevant

(Mathews, 2006). To tackle this, methods were developed that predict a number of suboptimal secondary structure folds for a given RNA sequence that have energies at the close vicinity of the minimum energy fold structure (Zuker, 1989). Thus these methods can provide alternative structures that can potentially be of biological significance.

In this chapter, the nucleic acid folding software tools mFold (Zuker, 2003) and RNAstructure (Mathews, 2004) were used for the prediction of secondary structures and free energy values for RNA-aptamers and for much longer mRNA sequences that incorporated the RNA-aptamers. The correct folding of RNA-aptamers into their particular secondary structure was investigated for constructed RNA devices in relation to the ability to bind their target ligand.

4.1.5 Transcription/translation cell-free technologies

Cell-free systems are technologies developed for the *in vitro* reconstitution of cellular processes of interest. The *in vitro* reconstitution can offer a number of advantages to the user by reducing the complexity of living cell systems. For example, a cell-free system can be used for process optimisation towards a particular application without concerns for cell viability and growth. Additionally, the absence of a cell membrane barrier allows the user direct access to the molecular machinery and enzymatic reactions, as well as the opportunity for direct sampling and monitoring of the chemical reactions of the system.

There have been two different approaches for the *in vitro* reconstitution of the transcription/translation processes of living cells. One approach attempts the reconstitution of the protein synthesis from purified components of the cellular machinery. A notable success of this approach was the reconstitution of the prokaryotic *E. coli* translational machinery and its coupling to a bacteriophage RNA polymerase to produce a transcription/translation cell-free system that could produce mg/ml mass of expressed protein (Shimizu et al., 2001). In this study, the initiation, elongation and termination factors, the 20 aminoacyl-tRNA synthases and the methionyl-tRNA transformylase, T7 RNA-polymerase, and ribosomes were his-tagged, overexpressed in *E. coli* expression host and purified, and finally reconstituted to a fully-functional state *in vitro*. The other approach for the *in vitro* reconstitution of the transcription/translation processes is with crude cell extract. Cell-extract systems have been derived from *E. coli* (Kigawa et al., 2004), wheat germ (Takai & Endo, 2010), rabbit reticulocytes (Pelham & Jackson, 1976), insect cells (Ezure et al., 2006), yeast (Iizuka, Najita, Franzusoff, & Sarnow, 1994), cancer cells (Mikami, Kobayashi, Masutani, Yokoyama, & Imataka, 2008) and hybridoma (Mikami, Kobayashi, & Imataka, 2010) among others.

E. coli cell-extract technology has been widely adopted due to its simplicity, efficiency and inexpensiveness, as *E. coli* can be fermented in large quantities using low cost media and gives the highest yield of expressed protein production (Caschera & Noireaux, 2014). To achieve high levels of protein expression, the cell-extract technologies initially utilised the T7 RNA polymerase but more recently protocols have been described for cell-extract technologies that

make use of the endogenous *E. coli* RNA-polymerase to produced equivalent amounts of expressed protein (Shin & Noireaux, 2010).

Even more, a streamlined protocol has been described for the in-house production of *E. coli* cell-extract at a 98% cost reduction as compared to similar systems commercially available (Sun et al., 2013). This technology is branded as TX-TL cell-extract, and the protocol for the production involves three phases. Initially, *E. coli* cells are cultured for the production of the molecular machinery biomass from inexpensive carbon sources. Then, *E. coli* cells are lysed to release the molecular machinery from the bounds of the cell membranes, and finally the cell-extract is clarified and conditioned for cell-extract reactions. Following cell-extract production, cell-extract reactions are prepared by the addition of amino acids and tRNA molecules to support high yields of protein synthesis, as well as with an energy source for the chemical reactions. The primary energy source supplied is ATP but this is quickly depleted. For this, secondary energy sources are supplied to the reaction, in the form of 3-phosphoglycerate or other metabolic intermediates (Caschera & Noireaux, 2014; Kim & Swartz, 2001) along with co-factors such as NADH and acetyl-CoA, that can act through the glycolytic pathway for the regeneration of ATP that results to longer reaction times and higher protein yields (Kim & Swartz, 2001). In addition, the TX-TL cell-extract technology is supported by mathematical models that can aid predictable design of biological circuits with parameters determined for its constituent biomolecules (Karzbrun, Shin, Bar-Ziv, & Noireaux, 2011; Siegal-Gaskins, Noireaux, & Murray, 2013).

In this chapter, a commercially available *E. coli* S30 cell-extract system and the TX-TL cell-extract system were used for the implementation of a novel biosensing platform design for the detection of protein biomarkers. The use of the cell-extract systems allowed the direct access of the analyte protein molecules to the sensing element of the biosensor device without obstruction by the cell membrane barrier. Additionally, the cell-extract systems acted as the hardware component of the biosensing platform facilitating the information relay from the sensing element to the transduction module and the synthesis of the reporter protein.

4.2 Experimental Results

4.2.1 Biosensor design

A RNA-switch biosensor design was developed for the detection of protein analytes in *E. coli* cell-extract system. The biosensor design included a RNA-aptamer domain to serve as the sensing element of the device for the specific recognition of the analyte molecule of interest. The output module of the device was formed by a mRNA expression platform which coded the expression of a reporter protein gene. The device was predicted to have the ability to act like a molecular switch through the action of a relay mechanism that transduced the information of RNA-aptamer/analyte binding to the output module.

The biosensor design employed the principle of steric hindrance (or direct competition) as the mechanism through which the switch function of the biosensor would be achieved (Figure 4.5). The RNA-aptamer was placed at a strategic position on the mRNA expression platform. In particular, the RNA-aptamer was placed immediately downstream of the AUG start codon of the reporter protein CDS and in close proximity to the SD sequence, within a region frequently referred as the ribosome footprint region (Salis, 2011). These sequences at this region are the initial contact sites between the ribosome and the mRNA during the process of translation initiation. Thus, in the absence of the protein analyte of interest, ribosomes in the cell-extract were predicted to be able to associate with the ribosome footprint region of the mRNA and initiate translation of the reporter protein. In the presence of the protein analyte of interest, the RNA-aptamer was predicted to bind and recruit the protein molecule at the ribosome footprint region of the mRNA. Since binding of the protein molecule and the ribosome is mutually exclusive through steric hindrance, the analyte presence was predicted to repress translation initiation and reporter protein expression.

In summary, this design couples together two processes to produce a RNA-switch biosensor device for the detection of a protein biomarker. A RNA-aptamer domain is used for the specific recognition of the analyte of interest and this binding event is coupled to the process of reporter protein expression through the action of a ribosome steric hindrance relay mechanism.

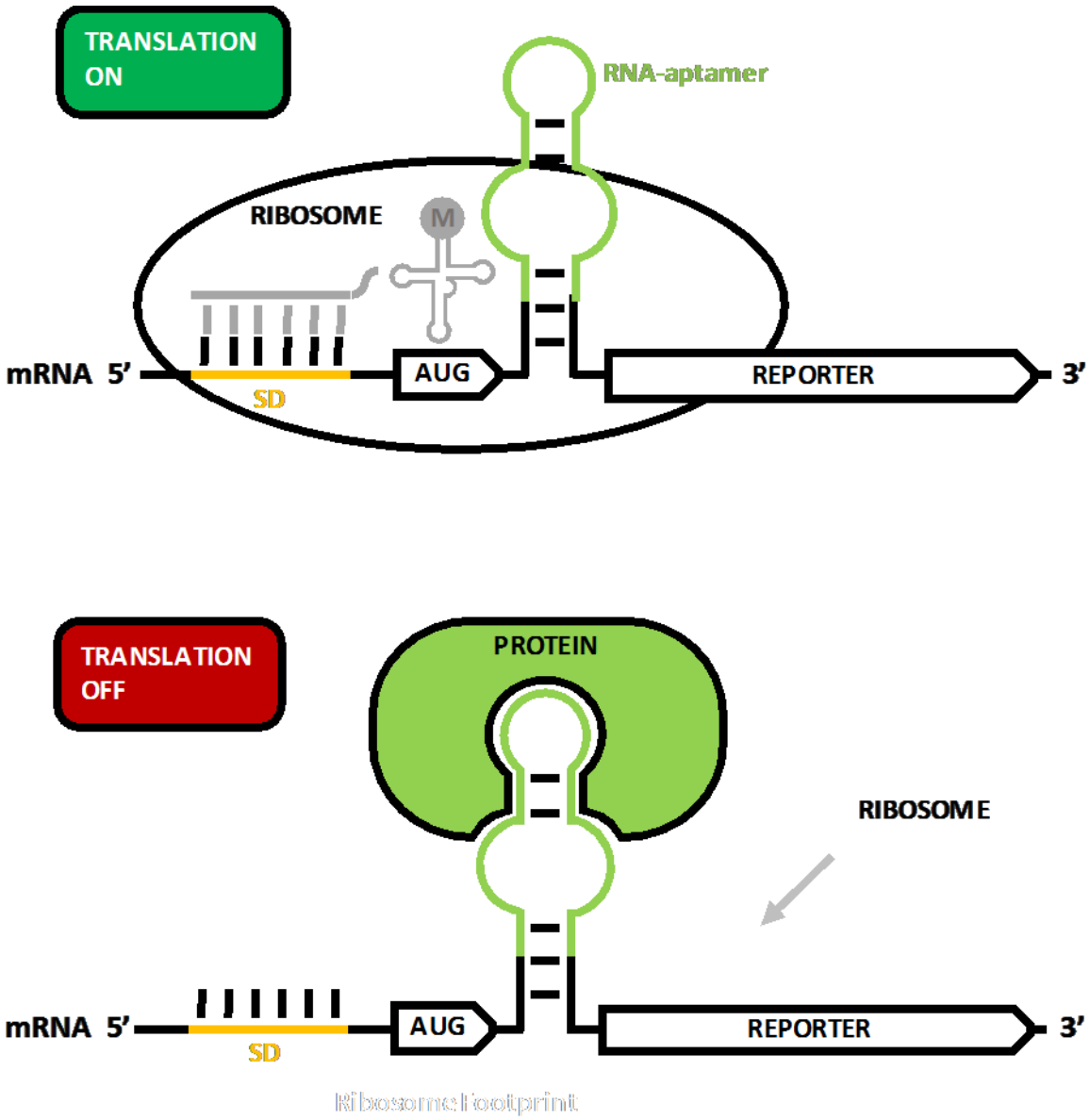


Figure 4.5: RNA-switch biosensor design. The schematic illustrates the RNA-switch device design developed. The physical form of the device is that of an mRNA molecule which incorporates a RNA-aptamer molecule downstream of the AUG start codon in addition to the Shine-Dalgarno (SD) sequence and reporter protein coding sequence. The top schematic illustrates the "Translation on" state of the device in which the ribosome associates with the SD sequence and start codon to initiate translation initiation and reporter protein expression. The bottom schematic illustrates the "Translation off" state of the device where protein binding to the RNA-aptamer sensor element sterically hinders ribosome access to the ribosome footprint region of the mRNA. This prevents translation initiation and reporter protein expression.

4.2.2 Prototype design and construction

For the testing of the proposed biosensor design, a prototype device was constructed for the detection of human thrombin protein. Thrombin was chosen as the analyte of interest as it is the most frequently used target molecule in proof-of-concept experiments of novel aptamer-based affinity sensor designs (Deng et al., 2014). Additionally, thrombin protein samples purified from human plasma are commercially available and at low cost. Lastly, thrombin being a bloodstream protein serves as an ideal model system for developing a biosensor design for medically relevant biomarkers. A RNA-aptamer sequence specific for the binding of the thrombin protein (Figure 4.2) was chosen as the sensor element of the biosensor device and the GFP gene was to serve as the reporter protein of the output module.

The prototype biosensor device was encoded on a plasmid DNA construct. Initially, a plasmid DNA construct was assembled that encoded a GFP expression device (Figure 4.6 top panel). This construct, named pNK401, did not incorporate an anti-thrombin RNA-aptamer and it would serve as the negative control device in subsequent experiments. The construct was assembled by restriction enzyme cloning using the following standard biological parts: the T7 promoter, the Bba_B0030 RBS element, the Bba_E0040 GFPmut3b gene sequence and Bba_B0015 double terminator part (incorporates Bba_B0010 and Bba_B0012 terminators). The GFP expression device was placed in a pSB1A2 vector backbone. Then, the pNK401 construct was utilised for the construction of the prototype thrombin RNA-switch device. The DNA sequence of anti-thrombin RNA-aptamer (X. M. Liu et al., 2003) was incorporated directly downstream of the ATG start codon by PCR and this plasmid construct was designated the code pNK402 (Figure 4.6 bottom panel). Successful construction of plasmid constructs pNK401 and pNK402 was verified by restriction enzyme test digestion and agarose gel electrophoresis analysis (Figure 4.7), and by DNA sequencing reactions.

The physical form of the RNA-switch biosensor device is one of purified mRNA transcript that was produced by *in vitro* transcription and purification. The pNK401 and pNK402 constructs were used as template DNA in PCR reactions utilising a primer pair to produce linear

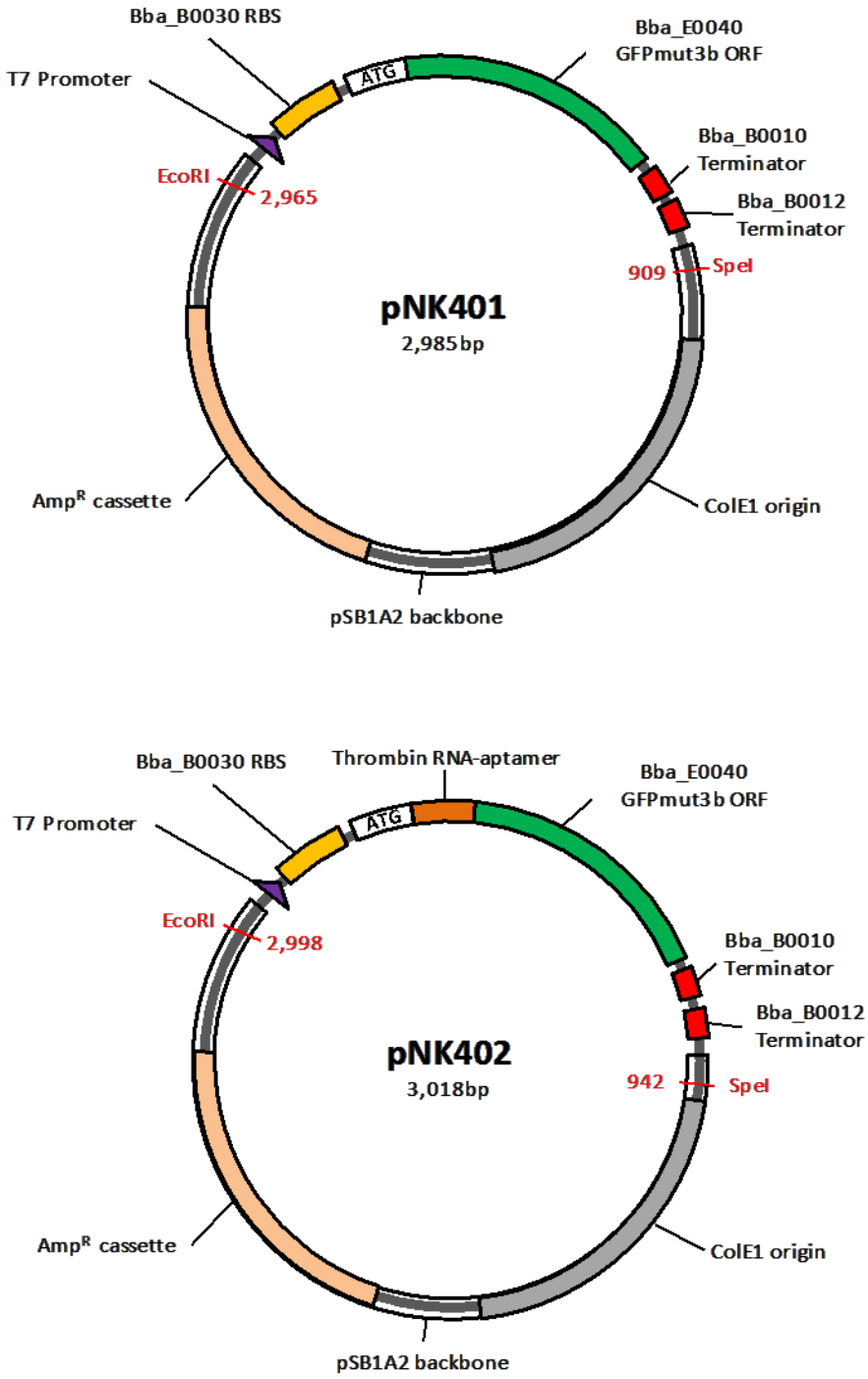


Figure 4.6: Plasmid maps of DNA constructs pNK401 and pNK402. The illustrated plasmid constructs both incorporate the T7 promoter, the Bba_B0030 RBS genetic part, the Bba_E0040 GFPmut3b part gene sequence and Bba_B0015 double terminator part (incorporates Bba_B0010 and Bba_B0012 terminators) in a pSB1A2 plasmid backbone. pNK402 also incorporates a thrombin RNA-aptamer (reference) directly downstream of ATG-start codon of GFP ORF.

dsDNA fragments. These dsDNA fragments included the encoded control device and the prototype thrombin switch device respectively and were used in *in vitro* transcription reactions to produce the desired mRNA transcripts. Correct sizes of mRNA transcript size and product purity were analysed by RNA gel electrophoresis on a denaturing 5% PAGE gel system (Figure 4.8) and results were found to be satisfactory for further experimentation.

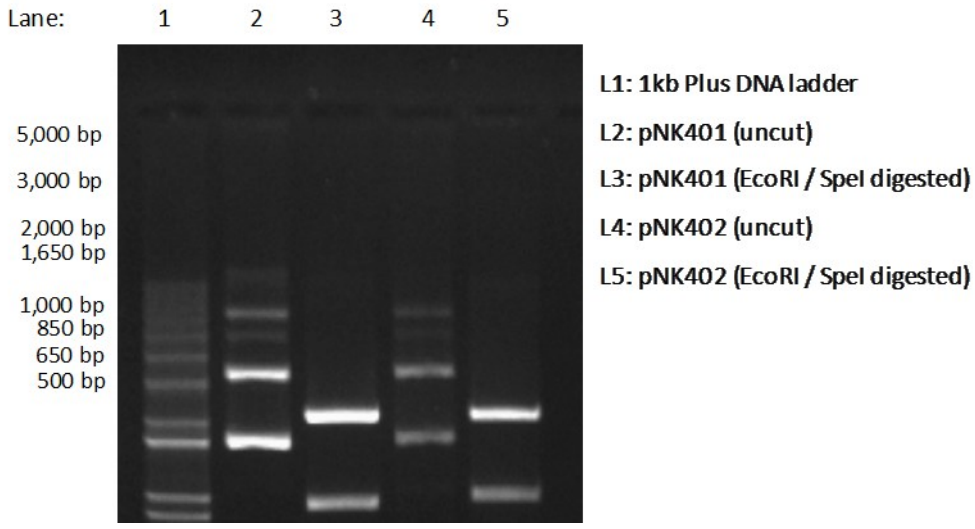


Figure 4.7: Test digest of plasmid constructs pNK401 and pNK402. Test digestion of plasmid constructs pNK401 & pNK402 with EcoRI and SpeI restriction enzymes resulted in two DNA fragments for each plasmid that migrate similarly to the 2,000 bp and 1,000 bp ladder bands. Uncut plasmid constructs produced multiple bands.

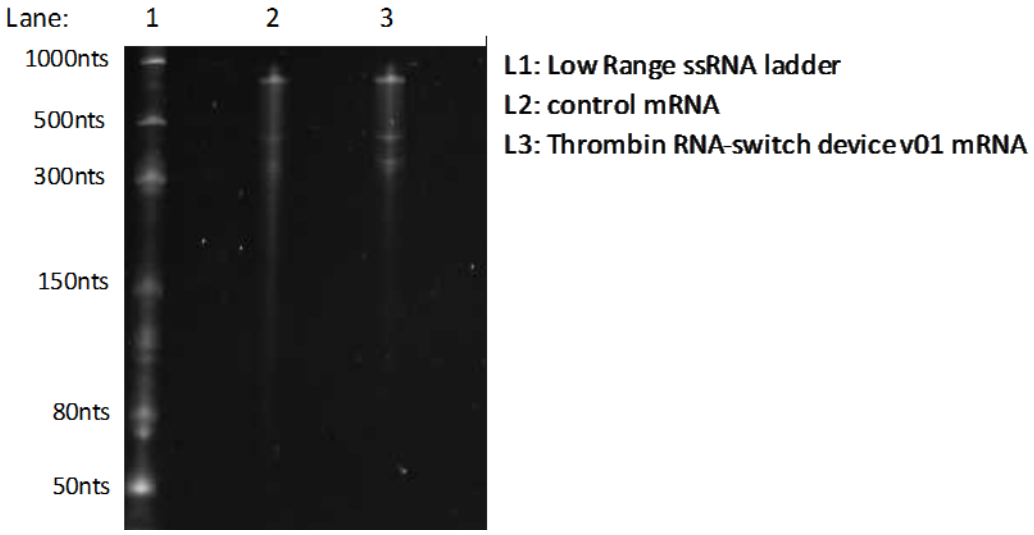


Figure 4.8: RNA electrophoresis of control device and prototype thrombin switch device. *In vitro* transcribed, purified control device mRNA and thrombin switch device mRNA were analysed by RNA-electrophoresis. Dominant mRNA band for the *in vitro* transcribed and purified mRNA samples migrate slightly further than the 1000nts band of the RNA ladder. Expected sizes for the *in vitro* transcribed mRNA products were 848nts for the control device and 881nts for the prototype thrombin switch device. RNA gel analysis carried out on a 5% PAGE / TBE-urea denaturing gel.

4.2.3 Investigating RNA-switch functionality of the prototype device

The RNA-switch functionality of the prototype device, as proposed by the biosensor design, was tested in cell-extract expression assays. Purified mRNA of the prototype thrombin switch device was added to cell-extract reactions and GFP expression from the mRNA template was monitored under different conditions. The commercial S30 *E. coli* cell extract kit was used to set up the cell-extract reactions in this series of experiments. GFP expression was monitored by recording the fluorescence signal emitted from the cell-extract reactions over time using a microplate reader.

Testing for thrombin mediated repression of GFP expression

An experiment was designed to test the RNA-switch functionality of the prototype thrombin switch device in the absence or presence of thrombin protein. According to the proposed design, cell-extract reactions of the prototype thrombin switch device would result in GFP reporter protein expression in the absence of thrombin protein. In contrast, cell-extract reactions of the prototype device would not produce GFP in the presence of thrombin protein. As a result, in the presence of thrombin the fluorescent signal emitted from the reaction would be lower that than the signal emitted from the reaction in the absence of thrombin.

An experiment was setup in which two binding reactions were prepared in 8 μl total reaction volume; in the first condition 9.2 picomoles of the prototype thrombin switch device mRNA was incubated with 6 picomoles of thrombin protein, and in the second condition 9.2 picomoles of the prototype thrombin switch device mRNA was incubated in the absence of thrombin. The binding reactions were incubated for 30 minutes at 4°C and then samples were added to the cell-extract reaction mix. Then, the prepared cell-extract reactions were transferred to a microplate which was loaded onto a microplate reader to monitor GFP expression. In addition to the two conditions tested, a negative control experiment was set up as well. In the negative control reactions, the prototype device mRNA was replaced by the control device mRNA that did not incorporate the anti-thrombin RNA-aptamer. This negative control experiment was set up to control for non-specific effects caused by the addition of thrombin protein to the cell-extract reactions and/or for non-specific binding of the thrombin

to the mRNA template. Lastly, in order to account for the background fluorescent signal emitted by the cell-extract mix, a cell-extract reaction was setup in which no mRNA was added but otherwise treated similarly.

Experimental results of the cell-extract repression assay for the prototype device showed repression of GFP expression in the presence of thrombin (Figure 4.9). The reaction profile of the cell-extract reaction for the prototype device showed a gradual build-up of fluorescence signal levels over time due to GFP expression for both reaction conditions, but in the absence of thrombin analyte the rate of GFP expression was much higher (top right panel). The reaction profile for the control device showed that both in the absence and presence of thrombin the rate of GFP expression was very similar (top left panel). From these control reactions, it was determined that the addition of thrombin protein in the cell-extract reaction did not adversely affect protein expression capabilities of the cell extract (for the concentration tested) and that non-specific binding of thrombin to mRNA template did not affect GFP expression from the mRNA expression platform. A comparison of the endpoint fluorescent signal (90 minutes time point) from the prototype device cell extract reactions showed that the endpoint signal was significantly lower in the presence of thrombin than in the absence of thrombin (bottom left panel). More specifically, the endpoint fluorescent signal from the prototype device cell-extract reaction in the presence of thrombin was at 43% as compared to the signal from the cell reaction in the absence of thrombin protein (bottom right panel). This decrease in endpoint fluorescence was not observed for the control device cell extract reactions.

Testing specificity of prototype device towards intended target

To investigate the specificity of the prototype device for the analyte of interest, an experiment was designed that compared GFP expression in cell-extract repression assay in the presence of thrombin or VEGF analytes. It was expected that GFP expression would decrease only in the presence of thrombin since the thrombin switch device incorporated a specific anti-thrombin RNA-aptamer, while the GFP expression should remain at basal levels in the presence of VEGF protein.

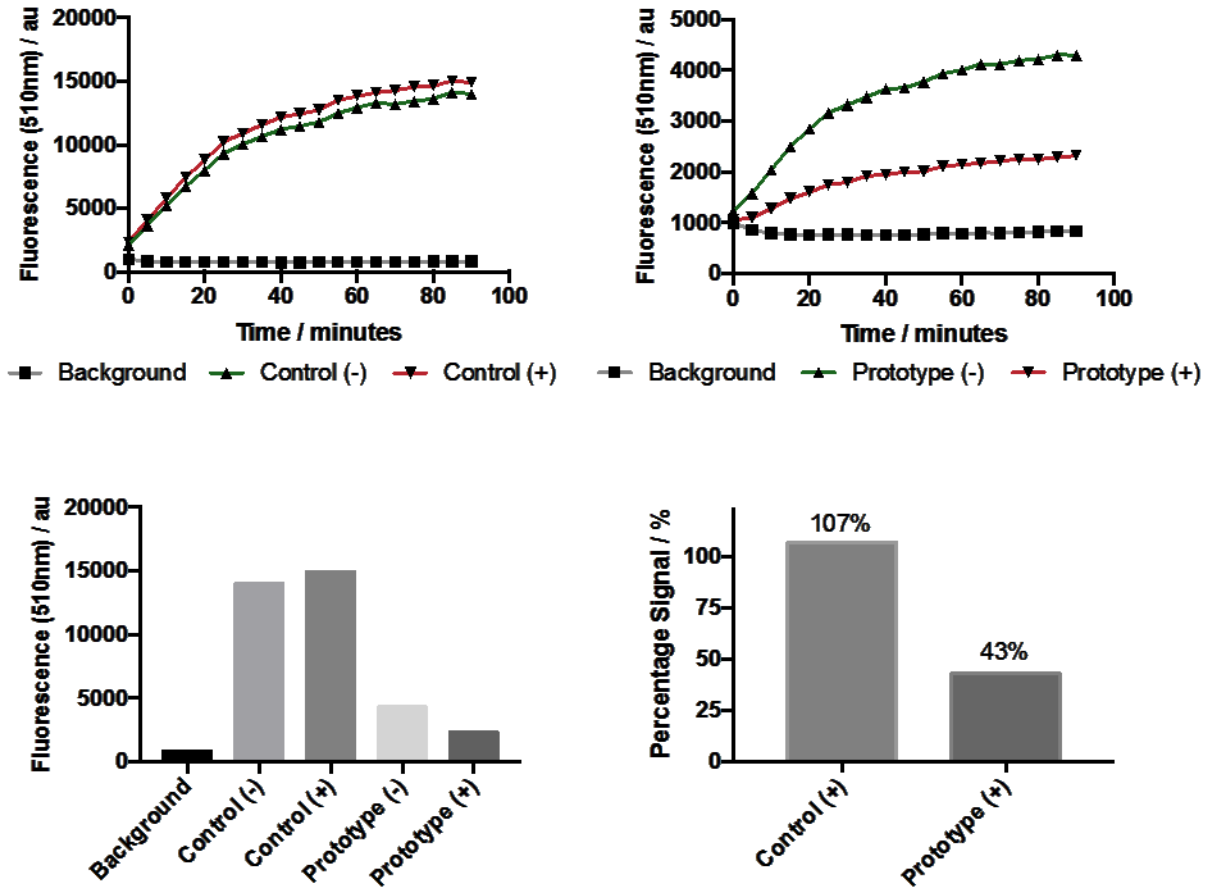


Figure 4.9: Cell-extract repression assay for prototype device. Purified mRNA of thrombin switch device and control mRNA (mRNA template identical to thrombin switch device that encodes translation of GFP gene but lacks the sequence of the anti-thrombin RNA-aptamer) were incubated in binding reactions in the absence (-) or presence of 6 picomoles thrombin (+) and added in cell-extract reactions and GFP expression was monitored. Top panels show the reaction profile and GFP expression for the control cell-extract reactions (top left) and the prototype device cell-extract reactions (top right). Bottom left panel shows a bar chart of the endpoint fluorescence signal (90 minutes) of the cell-extract reactions. Bottom right panel shows a bar chart of the percentage endpoint fluorescence signal from cell-extract reactions in the presence of thrombin protein as compared to their respective reactions in the absence of thrombin protein, following background signal subtraction. Reported values represent one experimental repeat. Background signal is the fluorescence signal of cell-extract.

An experiment was set up in which binding reactions were prepared in 8 μl total reaction volume with the prototype device mRNA under two conditions; 6.3 picomoles of device mRNA was incubated in the presence of 6.7 picomoles thrombin or in the presence of 6.7 picomoles VEGF. The binding reactions were incubated for 30 minutes at 4°C and then samples were added to the cell-extract mix. Then, the prepared cell-extract reactions were transferred to a microplate which was loaded in the microplate reader to monitor GFP expression. To benchmark the basal GFP expression of the device in the absence of any potential repressor protein, a control cell-extract reaction was set up in which the prototype thrombin switch device mRNA was pre-incubated with no protein analyte added. Also, to account for the background fluorescence of the cell-extract itself, a cell-extract reaction was set up in which no mRNA was added.

Experimental results of the cell-extract specificity assay for the prototype device showed specific repression of GFP expression by thrombin analyte (Figure 4.10). The top panel of Figure 4.10 shows the reaction profile of the specificity assay which clearly demonstrated that VEGF did not repress GFP expression from the prototype device. The bottom right panel of Figure 4.10 shows the endpoint fluorescence signal (90 minutes time point) for each reaction condition. The recorded fluorescent signal suggested that similar levels of GFP protein were expressed when no protein biomarker was added and when VEGF protein was added to the reactions. In contract, the endpoint fluorescence signal in the presence of thrombin protein suggested strong repression GFP expression as fluorescence signal recorded was close to the background fluorescence signal of the cell-extract mix. In particular, following background signal subtraction, it was calculated that 6.7 picomoles of thrombin repressed GFP expression to 11% of basal GFP expression signal from the prototype device as opposed to no repression (112% of basal signal) for 6.7 picomoles of VEGF (bottom right panel).

Investigation of the input / output characteristics of the prototype device

An experiment was designed to investigate the input/output characteristics of the prototype thrombin switch device, by carrying out a cell-extract assay that monitored GFP expression of cell-extract reactions in the presence of increasing thrombin analyte concentrations.

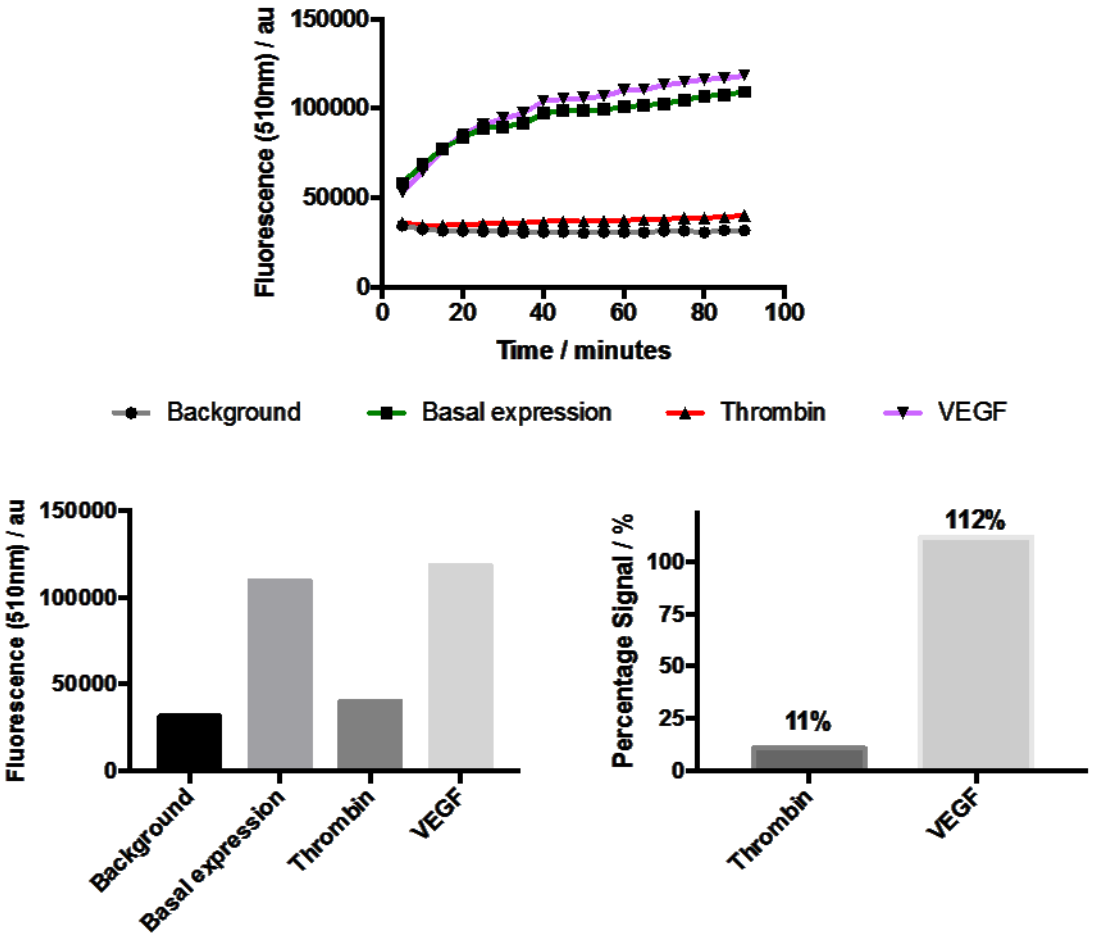


Figure 4.10: Specificity assay for prototype device. Purified mRNA of thrombin switch device was incubated in binding reactions in the absence of any analyte (basal expression), in the presence of 6.7 picomoles of thrombin or in the presence of 6.7 picomoles VEGF and then samples were added in cell-extract reactions. Top panel shows the reaction profile and GFP expression of the cell-extract reactions for each condition. Bottom left panel shows a bar chart of endpoint fluorescent signal (90 minutes' time point) of the cell-extract reactions for each condition that showed that GFP expression repression was specific the target analyte thrombin. Bottom right panel shows a bar chart of the percentage endpoint fluorescence signal in the presence of an analyte protein as compared to the basal expression, following background signal subtraction. Reported values represent one experimental repeat. Background signal is the autofluorescence signal of the cell-extract mix.

In this experiment, binding reactions were set up (8 μl total volume) with 6.3 picomoles of purified thrombin switch device mRNA and 0, 2.2, 4.5 or 6.7 picomoles of thrombin protein. The binding reactions were incubated for 30 minutes at 4°C and then samples were added to the cell-extract mix. Then, the prepared cell-extract reactions were transferred to a microplate which was loaded in the microplate reader to monitor GFP expression. To account for the background fluorescence of the cell-extract, a cell-extract reaction was setup in which no mRNA was added.

Experimental results showed a linear and inversely proportional relationship between thrombin protein concentrations and GFP expression from the prototype thrombin switch device (Figure 4.11). The reaction profile for GFP expression was similar for all reactions but GFP accumulation occurred at a slower rate for cell-extract reactions with higher thrombin protein concentrations (left panel). The endpoint fluorescence signal (90 minutes time point) and the thrombin protein concentration in binding reactions [2.2 pm → 275 nM, 4.5 pm → 563 nM and 6.7 pm →838 nM] were plotted in a scatterplot to determine the input / output characteristics relationship (right panel). The data showed a good fitting (R square value= 0.99) for a linear trendline in an inversely proportional relationship between increasing thrombin protein concentration and emitted fluorescent signal for the protein concentration range tested.

Summary RNA-switch functionality of prototype thrombin switch device

In summary, the prototype biosensor device constructed, thrombin switch device, has been demonstrated to function as a ligand-inducible RNA-switch in cell-extract reactions, in which thrombin protein analyte represses expression of GFP reporter protein. The ligand-inducible repression exhibited by the device was shown to be specific for the intended target and GFP expression was shown to be linear and inversely proportional for thrombin protein at the concentration range 275 nM - 838 nM.

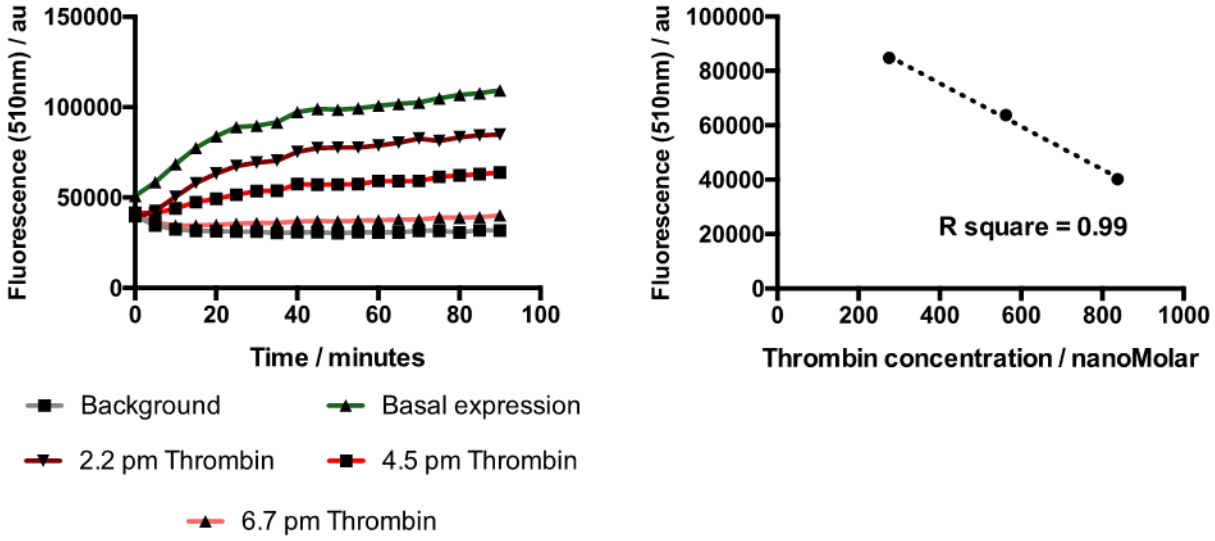


Figure 4.11: Input / output characteristics of prototype device. Purified mRNA of prototype thrombin switch device was incubated in binding reactions in the presence of 0 (basal expression), 2.2, 4.5 or 6.7 picomoles of thrombin protein and added in cell-extract reactions. Left panel shows the reaction profile and GFP expression of cell extract reactions in the presence of various thrombin protein concentration. Right panel shows the transfer plot for the endpoint fluorescent signal (90 minutes) against the tested thrombin protein concentrations (275nM, 563 nM and 838 nM in binding reaction mix) that exhibited a linear, inversely proportional relationship between the input (thrombin) / output (fluorescence) characteristics for the prototype device for the stated range. Reported values represent one experimental repeat. Background signal is the fluorescence signal of the cell-extract.

4.2.4 Investigating thrombin binding to prototype device by EMSA

A series of experiments were designed to investigate thrombin protein binding to the thrombin switch device prototype by Electrophoretic Mobility Shift Assay (EMSA). Previous experiments investigated thrombin protein binding to the prototype biosensor device using cell-extract expression assays (Section 4.2.3.). Experimental results from these assays showed repression of GFP expression in the presence of thrombin protein, indicating protein binding to the RNA biosensor molecule. However, the observed repression in GFP expression was a result of two biological processes: 1) protein binding/recognition by the RNA-aptamer domain and 2) decreased rate of translation initiation. To allow for studying the binding/recognition process of the system independently of the translation initiation process, the EMSA technique was utilised.

EMSA is a technique for detecting protein/nucleic acid interactions using gel electrophoresis based on the observation that the electrophoretic mobility of a nucleic acid/protein complex is typically less than that of the free nucleic acid (Hellman & Fried, 2007). In this series of experiments, after the completion of the gel electrophoresis step of the EMSA assay, the EMSA gels were first stained for RNA species visualisation and results were documented, and subsequently stained for protein species visualisation and documented. Thus, two gel images were produced for each EMSA gel which they were later overlaid using ImageJ software. Using the ImageJ software, the signal resulting from staining for RNA was pseudo-coloured green and the signal resulting from staining for protein was pseudo-coloured red.

EMSA of anti-thrombin RNA-aptamer (X. M. Liu et al., 2003)

Initially, an EMSA experiment was carried out using chemically synthesized anti-thrombin RNA-aptamer and purified thrombin protein. This experiment aimed to validate the EMSA as a suitable assay system for studying the interaction between these biomolecules, given that the specific interaction between these two species has been demonstrated previously by others (Y. Li et al., 2007).

For this experiment, binding reactions were set up using thrombin protein and anti-thrombin RNA-aptamer at protein:RNA molar ratios of: 0:1, 0.48:1, 0.95:1, 1.43:1, 1.91:1 and

2.38:1. The binding reactions were carried out for 30 minutes at 4°C. Following completion of the binding reactions, the samples were loaded on an 8% PAGE / HEPES-Imidazole buffered (pH = 7.55) gel system for EMSA analysis. To account for any non-specific binding of thrombin protein to the negatively charged RNA molecules, binding reaction were set up with thrombin protein and a chemically synthesized anti-VEGF RNA-aptamer (Jellinek et al., 1994). These samples served as the experimental negative control. Additionally, a thrombin protein sample was loaded on the EMSA gel in order to control for possible co-migration of free protein with the RNA species.

Experimental results showed specific interaction between the anti-thrombin RNA-aptamer and thrombin protein (Figure 4.12). The image of the EMSA gel recorded following staining for RNA visualisation (right top panel Figure 4.12) shows a band shift of the anti-thrombin RNA-aptamer in gel lanes 2-7. In lane 2, in which the binding reaction sample loaded contained only the anti-thrombin RNA-aptamer, a single RNA band is observed (from hereafter referred as the basal band). In lanes 3-7, which were loaded with samples including the anti-thrombin RNA-aptamer with increasing concentrations of thrombin protein, a second, shifted band is observed in addition to the basal band. The signal intensity of the shifted band increases towards gel lanes of samples with higher protein:RNA molar ratios. In contrast, the signal intensity of the basal band decreases towards gel lanes of samples with higher protein:RNA molar ratios. Gel lanes 8-12 were loaded with the binding reaction samples of the control experiment that included the anti-VEGF RNA-aptamer. In these lanes there was only a single basal band for all thrombin protein concentrations.

The right bottom panel of Figure 4.12 shows an image of the EMSA gel following staining for protein visualisation. Fluorescent bands appeared in lanes 3-7. These are the gel lanes that were loaded with binding reaction samples including anti-thrombin RNA-aptamer with thrombin protein. The signal intensity of the bands is higher for gel lanes that were loaded with samples including higher concentrations of thrombin protein. In gel lane 7, the lane well also exhibited some fluorescent signal similarly to gel lane 1 that was loaded just with protein sample. From this, it was speculated that the signal in the gel well of lane 7 results from excess, unbound protein in the binding reaction samples. Similarly, for gel lanes 8-12, that were loaded

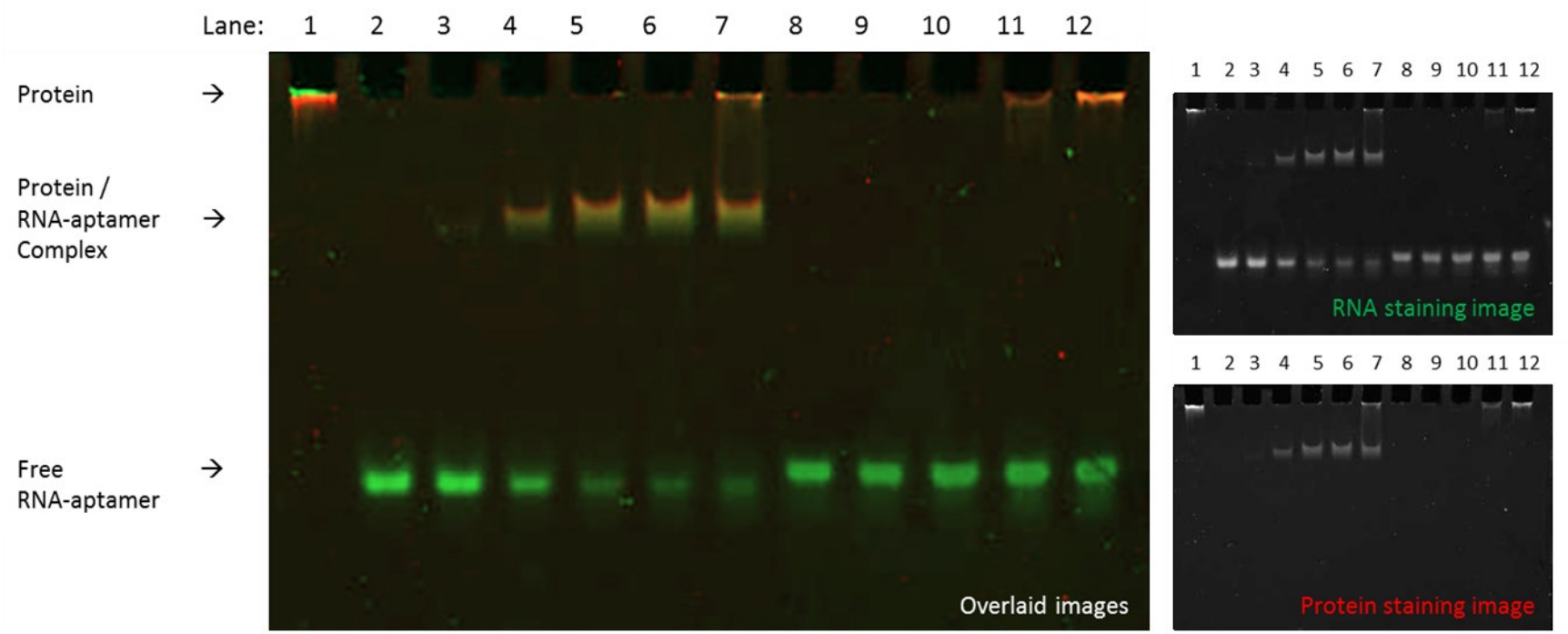


Figure 4.12: EMSA of anti-thrombin RNA-aptamer. Chemically synthesized anti-thrombin RNA-aptamer (lanes 2 – 7) and chemically synthesized anti-VEGF RNA-aptamer (lanes 8 – 12) -used as control- were incubated in binding reactions with thrombin protein purified from human plasma at protein / RNA molar ratios of 0.00, 0.48, 0.95, 1.43, 1.91 and 2.38 and loaded on gel for EMSA analysis. Lane 1 was loaded with a sample that just included thrombin protein. Gel images show shifted bands of protein / RNA complex formation for anti-thrombin RNA-aptamer / thrombin protein pair. Binding reactions were carried out for 30 minutes at 4°C. The EMSA gel system used was 8% PAGE / HEPES-Imidazole buffered (pH = 7.55) and electrophoresis was carried out at 4°C. Gel was first stained with GelRed dye and imaged, and then stained with SybroRuby dye and imaged. Images were then overlaid using ImageJ software.

with samples including anti-VEGF RNA-aptamer with thrombin protein, the only signal observed was in the wells of gel lanes 11-12.

When the gel images for RNA staining and protein staining were overlaid using the ImageJ software, the bands of the protein stained gel overlapped with the shifted bands on the RNA stained gel for lanes 3 – 7. In conclusion, the experimental results showed thrombin protein binding to the anti-thrombin RNA-aptamer and specific recognition as no interaction was observed between anti-VEGF RNA-aptamer and the thrombin protein. This experiment validated EMSA as a suitable assay system for studying the binding between thrombin protein and constructs that incorporate the anti-thrombin RNA-aptamer.

EMSA of prototype device

Subsequently, an EMSA experiment was carried out to investigate the binding of thrombin protein to the prototype thrombin switch device. As stated previously, EMSA experiments were designed to study the binding/recognition process of the biosensor device independent of the translation initiation process and the previous EMSA experiment with chemically synthesized anti-thrombin RNA-aptamer validated the EMSA technique and its setup for such an investigation.

In this experiment, *in vitro* transcribed and purified prototype thrombin device mRNA was incubated with increasing concentration of thrombin protein. The mRNA was incubated in binding reactions at protein/RNA molar ratios of: 0:1, 0.46:1, 0.92:1, 1.38:1, 1.84:1 and 2.30:1, and the binding reactions were carried out for 30 minutes at 4°C. Following completion of the binding reactions, the samples were loaded on a 3.5% PAGE / HEPES-Imidazole buffered (pH = 7.55) gel system for EMSA analysis. To account for non-specific binding of the thrombin protein to the RNA, binding reactions were set up with thrombin protein and control mRNA that did not incorporate the anti-thrombin RNA-aptamer and loaded on another EMSA gel.

Results from this EMSA experiment did not show band shift for the thrombin switch device (Figure 4.13). Staining for RNA, showed a single band present in gel lanes 1-6 of the thrombin switch device gel. The same pattern of a single base band was observed for all lanes (1-6) of the control device gel as well. Staining for protein showed a single band in lanes 2-6

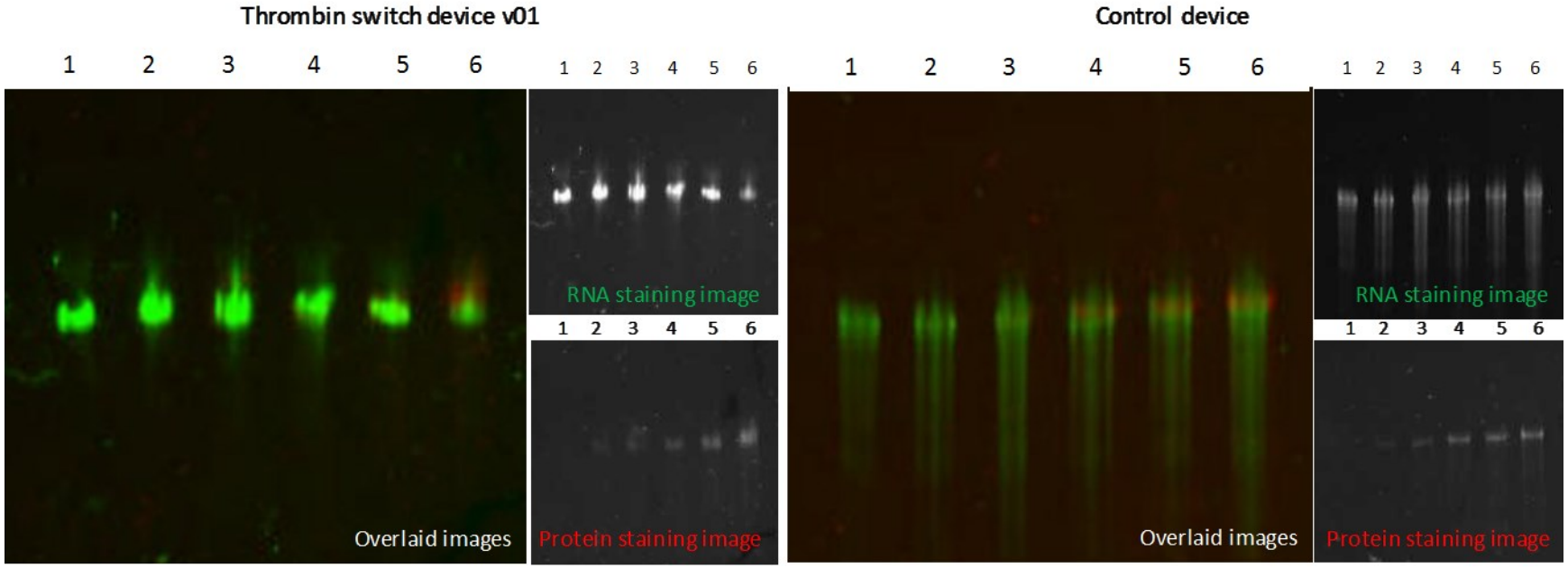


Figure 4.13: EMSA of prototype device. Purified mRNA of thrombin switch device v01 and Control device were incubated in binding reactions with thrombin protein purified from human plasma at protein/RNA molar ratios of: 0:1, 0.46:1, 0.92:1, 1.38:1, 1.84:1 and 2.30:1 and loaded on gels for EMSA analysis. Gel images did not show any shifted bands for either device. Binding reaction were carried out for 30 minutes at 4°C. EMSA gel system used was 3.5% PAGE / HEPES-Imidazole buffered (pH = 7.55) and gels were run at 4°C. Gels were first stained with GelRed dye and imaged, and then stained with SybroRuby dye and imaged. Images were then overlaid using ImageJ software.

which increases in intensity towards higher thrombin concentration for the thrombin switch device gel. Same band pattern appeared on the control device gel. When gel images were overlaid using ImageJ software, the protein bands on the gel images overlaid with the RNA bands for both mRNA constructs.

These results contradicted observations from cell-extract expression assays of previous experiments (Section 4.2.3), in which thrombin protein inhibited GFP expression from the prototype device mRNA in cell-extract assays. One possible explanation for the absence of RNA band shift in Figure 4.13 for samples including the prototype device mRNA with thrombin protein, is that the size of the thrombin protein (~37 kDA) is very small compared to the size of the biosensor device mRNA molecule (~280 kDa). Thus, binding of the protein to the mRNA did not significantly alter the electrophoretic mobility of the mRNA and in combination with the low resolving capability of a 3.5% PAGE gel resulted in the absence of band shifting. The co-localisation on the gel of the prototype device mRNA bands and thrombin protein bands suggested an interaction between the two species. Nevertheless, the same observation was true for the control mRNA gel, so that it was not possible to determine whether the interaction was specific to the incorporated anti-thrombin RNA-aptamer or non-specific towards the RNA molecule.

EMSA of partial sequences of the prototype device

Previous EMSA experiments highlighted the strengths and limitations of this technique for assessing thrombin protein binding to its RNA interaction partner. The EMSA assay that utilised the chemically synthesised anti-thrombin RNA-aptamer established the validity of this assay for studying the thrombin protein binding event directly, independently from the process of translation initiation and protein expression. However, the EMSA assay with the full-length prototype device mRNA appeared to be limited for assessing the binding event, probably due to the low resolution of the gel system employed and due to non-specific binding of thrombin protein to the RNA molecule. In order to tackle the limitations encountered and allow for studying the binding event using the EMSA assay, two model systems of the prototype device were investigated. These model systems were shorter mRNA transcripts of the full length prototype device that included all the functional elements of the prototype that allow the

studying of the binding event as well as most of the mRNA context effects. At the same time, these model systems were shorter in length to allow their study in an EMSA gel with higher resolving capacity, and also aimed to minimise non-specific binding of thrombin protein to the RNA molecules.

Two model systems were investigated: i) a ~150 nucleotide system for which an mRNA was produced that included the 5'UTR region, the ATG start codon, the incorporated anti-thrombin RNA-aptamer and the first 84 ribonucleotides of the GFP coding sequence of the prototype device and ii) a ~250 nucleotides model system for which an mRNA was produced that included the 5'UTR region, the ATG start codon, the incorporated anti-thrombin RNA-aptamer and the first 184 ribonucleotides of the GFP coding sequence of prototype device; These model systems of partial sequences included the thrombin binding RNA-aptamer incorporated into the mRNA sequence, and thus were believed to capture most of the context effects of the local environment and its effects on the RNA-aptamer binding properties. These mRNA molecules were produced by *in vitro* transcription, purified and the correct product was verified by RNA gel electrophoresis.

For the EMSA assay experiment, binding reactions were set up including purified mRNA of the prototype device partial sequences incubated with increasing concentrations of thrombin protein. The mRNA was kept at a constant concentration, while thrombin protein was added in the binding reaction at protein/RNA molar ratios of: 0:1, 0.46:1, 0.92:1, 1.38:1, and 1.84:1. Following incubation for 30 minutes at 4°C, the binding reaction samples were loaded on the gel for EMSA analysis. The gel system used in this experiment was a 5% PAGE / HEPES-Imidazole buffered (pH = 7.55) gel. To account for non-specific binding of thrombin protein to the mRNA molecules, negative control reactions were set up with purified mRNA of the equivalent partial sequence region of the control device construct. As previously, these mRNA constructs were identical to the prototype device partial sequences constructs apart from the absence of the anti-thrombin RNA-aptamer.

EMSA analysis of the 150 nucleotides long model system showed thrombin protein binding to the prototype device partial sequence mRNA (Figure 4.14). In this EMSA gel, lanes 1-

5 were loaded with binding reactions samples that included the prototype device partial sequence mRNA, while lanes 6-10 were loaded with the negative control binding reaction samples that included the control device partial sequence mRNA. In Figure 4.14, the top right panel shows an image of the EMSA gel following staining for RNA visualisation. In this image, the gel lanes 1-5 showed RNA band shift in the presence of thrombin protein. Gel lane 1, that included just the mRNA molecules, shows a single basal band which is shifted in the presence of thrombin protein as exhibited by lanes 2-5. This shift is mostly obvious in lanes 3-5, while gel lane 4 shows a strong basal band signal. This could be attributed to pipetting error during the loading of thrombin protein to the binding reaction mixture. In contrast, negative control gel lanes 6 -10 did not show a clear RNA band shift. Nevertheless, lanes 7-10 showed a very diffused band of increasing signal intensity trailing the base bands towards higher thrombin protein concentrations. This was attributed to non-specific binding of thrombin protein to the RNA molecules that is also supported by the observation that the shifting of the RNA bands increases in magnitude at higher thrombin protein concentrations in the sample rather than a consistent shift at a particular position on the gel which is characteristic of specific binding. In Figure 4.14, the bottom right panel shows the image of the EMSA gel following staining for protein visualisation. This image shows formation of distinctive bands in gel lanes 3-5, while gel lanes 8-10 exhibited the diffuse bands of increasing intensity towards lanes of higher thrombin protein concentrations that were discussed previously in this paragraph.

When images for RNA staining and protein staining were overlaid (left panel Figure 4.14) the sharp bands of the protein gel in lanes 3 – 5 overlapped with shifted bands of the RNA gel. Also the diffuse protein bands in lanes 7 – 10 overlapped with the diffused RNA bands. In conclusion, the evidence from the EMSA analysis support the specific recognition and binding of thrombin protein to the mRNA incorporated anti-thrombin RNA-aptamer for the ~150 nucleotide long partial sequence of the prototype device mRNA.

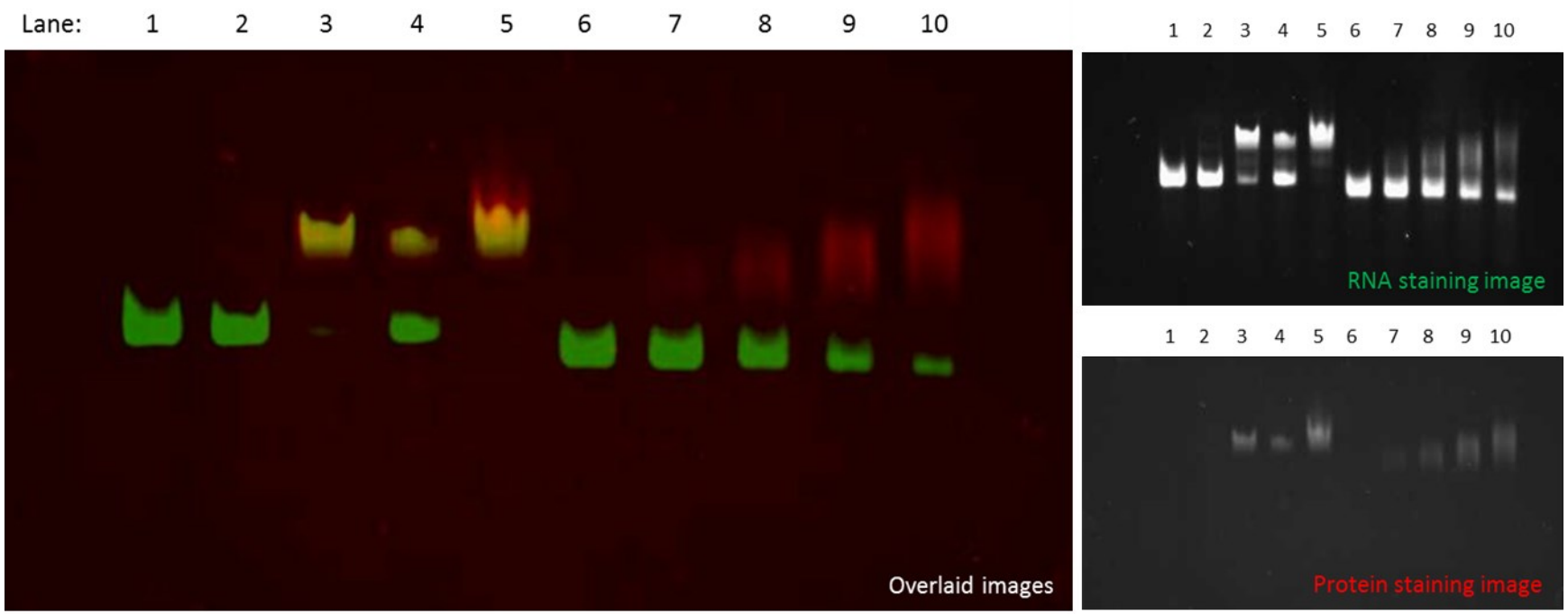


Figure 4.14: EMSA of the prototype device partial sequence (~150 nucleotides). Purified mRNA of ~150 nucleotides long partial sequences of prototype device (lanes 1 – 5) and purified mRNA of the corresponding region of control device -used as control- were incubated in binding reactions with thrombin protein purified from human plasma. Binding reactions were set up at protein/RNA molar ratios of: 0:1, 0.46:1, 0.92:1, 1.38:1, and 1.84:1 and loaded on gel for EMSA analysis. Gel images show shifted bands of RNA / protein complex formation for prototype device including samples. Binding reactions were carried out for 30 minutes at 4°C. EMSA gel system used was 5% PAGE / HEPES-Imidazole buffered (pH = 7.55) and gel was run at 4°C. Gel was first stained with GelRed dye and imaged, and then stained with SybroRuby dye and imaged. Images were then overlaid using ImageJ software.

EMSA analysis of the 250 nucleotides long model system showed thrombin binding to the prototype device partial sequence mRNA as well (Figure 4.15). In this EMSA gel, lanes 1-5 were loaded with binding reactions samples that included the prototype device partial sequence mRNA while lanes 6-10 were loaded with the negative control binding reaction samples that included the control device partial sequence mRNA. In Figure 4.15, the top right panel shows an image of the EMSA gel following staining for RNA visualisation. The image shows two RNA bands for lane 1 in which sample applied included just the mRNA and no protein. One possible explanation for this could be that the mRNA molecule adopted two different conformations that exhibit different electrophoretic motilities. In lanes 2 – 5 a third shifted band appeared, that was of stronger signal intensity toward gel lanes loaded with samples of higher thrombin protein concentrations. At the same time the bottom basal band's signal intensity decreases towards gel lanes loaded with samples of higher thrombin protein concentration. The basal band completely disappeared in lanes 4 and 5. The top basal band that persisted in all gel lanes regardless of the thrombin protein concentration. A possible explanation for this observation is that the second RNA conformation adopted by the mRNA molecules did not favour protein binding. Lanes 6 – 10, that were loaded with samples of control device derived mRNA, two RNA bands were apparent that decreased in intensity with higher thrombin protein concentrations. In contrast to lanes 2 – 5 where a sharp shifted band appeared, in lanes 7-10 a very diffused band of increasing signal intensity starts to form trailing the basal bands towards higher thrombin protein concentrations probably due to non-specific binding to thrombin protein.In Figure 4.15, the bottom right panel shows the image of the EMSA gel following staining for protein visualisation. This image shows formation of distinctive bands in gel lanes 3-5, while gel lanes 8-10 exhibited the diffuse bands of increasing intensity towards lanes of higher thrombin protein concentrations that were discussed previously. When images for RNA staining and protein staining were overlaid (left panel Figure 4.14) the sharp bands of the protein gel in lanes 3 – 5 overlapped with shifted bands of the RNA gel. Also the diffuse protein bands in lanes 7 – 10 overlapped with the diffused RNA bands. In conclusion, the evidence from the EMSA analysis support the specific recognition and binding of thrombin

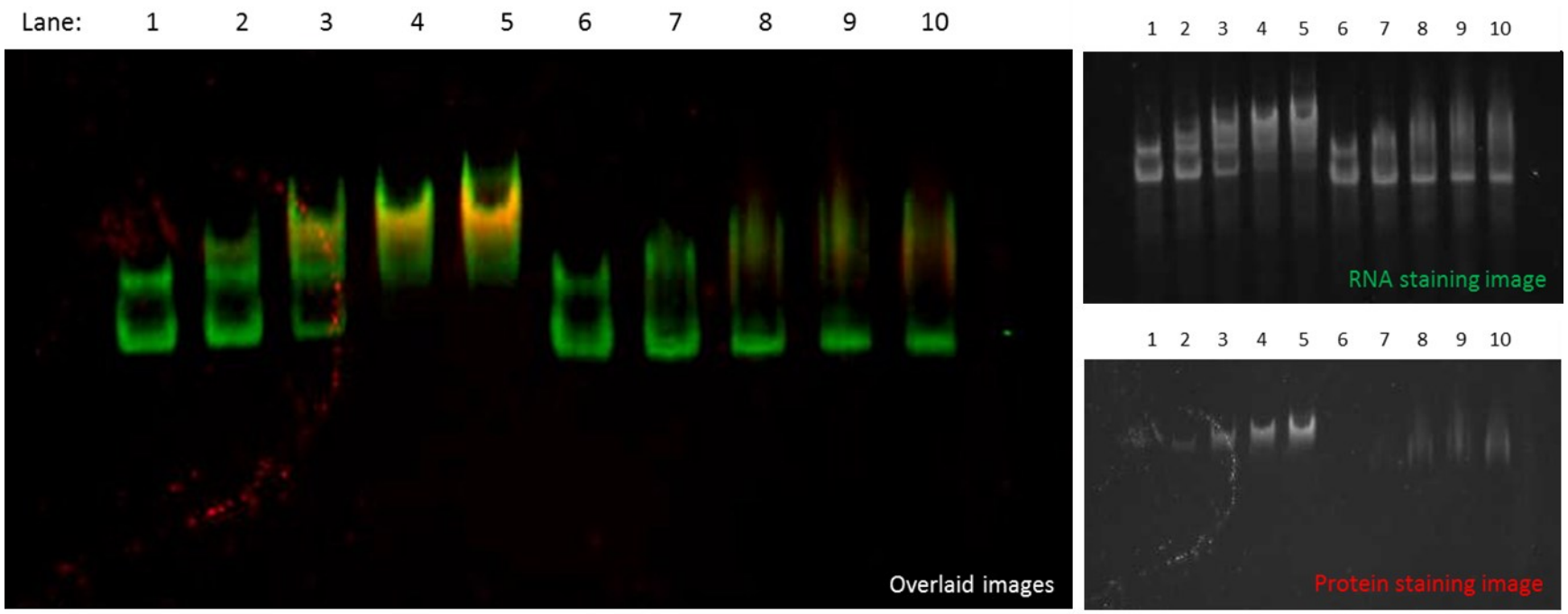


Figure 4.15: EMSA of the prototype device partial sequence (~250 nucleotides). Purified mRNA of ~150 nucleotides long partial sequences of prototype device (lanes 1 – 5) and purified mRNA of the corresponding region of control device -used as control- were incubated in binding reactions with thrombin protein purified from human plasma. Binding reactions were set up at protein/RNA molar ratios of: 0:1, 0.46:1, 0.92:1, 1.38:1, and 1.84:1 and loaded on gel for EMSA analysis. Gel images show shifted bands of RNA / protein complex formation for prototype device including samples. Binding reactions were carried out for 30 minutes at 4°C. EMSA gel system used was 5% PAGE / HEPES-Imidazole buffered (pH = 7.55) and gel was run at 4°C. Gel was first stained with GelRed dye and imaged, and then stained with SybroRuby dye and imaged. Images were then overlaid using ImageJ software.

protein to the mRNA incorporated anti-thrombin RNA-aptamer for the ~250 nucleotide long partial sequence of the prototype device mRNA.

Summary of EMSA for thrombin switch device

In summary, this set of experiments investigated the binding/recognition process of thrombin protein by the prototype device mRNA directly by the use of the EMSA technique. EMSA techniques allowed the study of binding event of the system independent of the translation initiation and protein expression processes that were employed in previous experiments. Initially, the EMSA technique was validated for the study of this particular interaction system by the use of the anti-thrombin RNA-aptamer. Subsequently, EMSA was carried out using the full-length prototype thrombin device but the binding event could not be observed. This was due to assay limitations such as poor gel system resolution for mRNA molecules of such a large size and due to non-specific binding of the thrombin protein to the RNA molecules. Instead, two model systems of the prototype device were utilised of sizes ~150 and ~250 nucleotides that allowed for higher resolving capabilities gel systems and exhibited less non-specific binding to thrombin protein. Using these partial sequences in EMSA experiments, the binding event between the thrombin protein and the mRNA molecule was clearly demonstrated and observed.

4.2.5 Testing a low-cost cell-extract system for the biosensor device

Experimental assays of cell-extract reactions carried out thus far made use of a commercially available *E. coli* S30 Extract System from Promega Corporation. These experiments enabled proof-of-concept experiments that validated the functionality of the prototype device in a high quality cell-extract system (Section 4.2.3). However, the cost of the Promega Corporation product per cell-extract reaction was found to be prohibitive for further experimental scale-up. Recently, a protocol has been described that allows the preparation and execution of an efficient *E. coli* based transcription-translation (TX-TL) cell-extract expression system, which compared to similar commercial systems showed a 98% cost reduction (Sun et al., 2013). To reduce experimental costs for further testing of the RNA-switch technology and to serve as the platform technology and powerhouse of reporter protein expression for biosensor applications, a homemade cell-extract was produced based on the aforementioned protocol.

Development of a homemade TXTL cell-extract system

For the in-house development of a TXTL cell-extract system, the *E. coli* BL21(DE3) cell strain was utilised instead of the *E. coli* Rosetta2 cell strain described in the aforementioned protocol (Sun et al., 2013). The *E. coli* BL21(DE3) cell strain incorporates a genomic integration of a prophage carrying the bacteriophage T7 RNA-polymerase. This polymerase protein is frequently used for achieving high yields of protein expression when used in combination with the T7 promoter element (Rosa, 1979; Tabor & Richardson, 1985). Thus, the cell-extract produced here should be able for protein expression from recombinant DNA plasmid systems that makes use of: i) promoter elements recognised by the endogenous *E. coli* polymerase or ii) promoter elements that utilise a bacteriophage T7 promoter element.

The newly produced *E. coli* cell extract was tested for protein expression in cell-extract expression reactions by the use of two plasmid constructs. For the purpose of testing gene expression by the endogenous *E. coli* RNA-polymerase, the plasmid construct pNK400 was utilised that incorporated the constitutive synthetic promoter element Bba_J23101. For the purpose of testing gene expression by the T7 bacteriophage RNA-polymerase, the plasmid construct pNK401 (Figure 4.6) was utilised that incorporated a T7 promoter element. Both

plasmid constructs employed coded for the expression of GFP reporter protein. To account for the autofluorescent signal from the cell-extract, the background fluorescence was measured by the use of negative control reaction samples with no plasmid construct added to them.

Experimental results showed that the prepared cell-extract could express GFP in cell-extract reactions by the use of both the endogenous *E. coli* RNA-polymerase and the phage derived T7 RNA-polymerase (Figure 4.16). The reaction profile was typical for cell-extract reactions, where following the start of the experiments there is a gradual build-up of expressed GFP protein and slowly over time the reactions reached a plateau as the cell-extract system becomes inactive and protein expression ceases. The cell-extract reactions that utilised the plasmid construct incorporating the T7 bacteriophage promoter exhibited higher levels of GFP expression as compared to the cell-extract reactions that utilised the plasmid construct incorporating the synthetic bacterial promoter. More specifically, the endpoint fluorescent signal (480 minutes) was 26-fold above background signal fluorescence for the former plasmid construct, while the signal was 2.2-fold above background signal for the latter construct. The different levels of GFP protein expression were attributed to the different strength of the promoter elements, to the different capacity of the cell-extract system for transcription initiation by the endogenous or bacteriophage RNA-polymerase or a combination of the two aforementioned factors. In conclusion, the prepared cell-extract system batch was found to be functional and able to drive protein expression from recombinant DNA plasmid systems.

For the purpose of optimising the conditions under which cell-extract reactions were carried out, an experiment was set up that investigated various magnesium glutamate (Mg-glutamate) and potassium glutamate (K-glutamate) salt concentration buffer systems for the cell-extract reactions. The salt concentrations tested were: 0 – 6 mM magnesium glutamate and 20 – 140mM potassium glutamate. The effect of the buffer salt composition on levels of protein expression was monitored by recording endpoint GFP fluorescence (480 minutes) using a microplate reader. The experiment investigated buffer composition for both endogenous RNA-polymerase and phage-derived T7 RNA-polymerase mediated protein expression through the use of the pNK400 and pNK401 plasmid systems as previously described in this section.

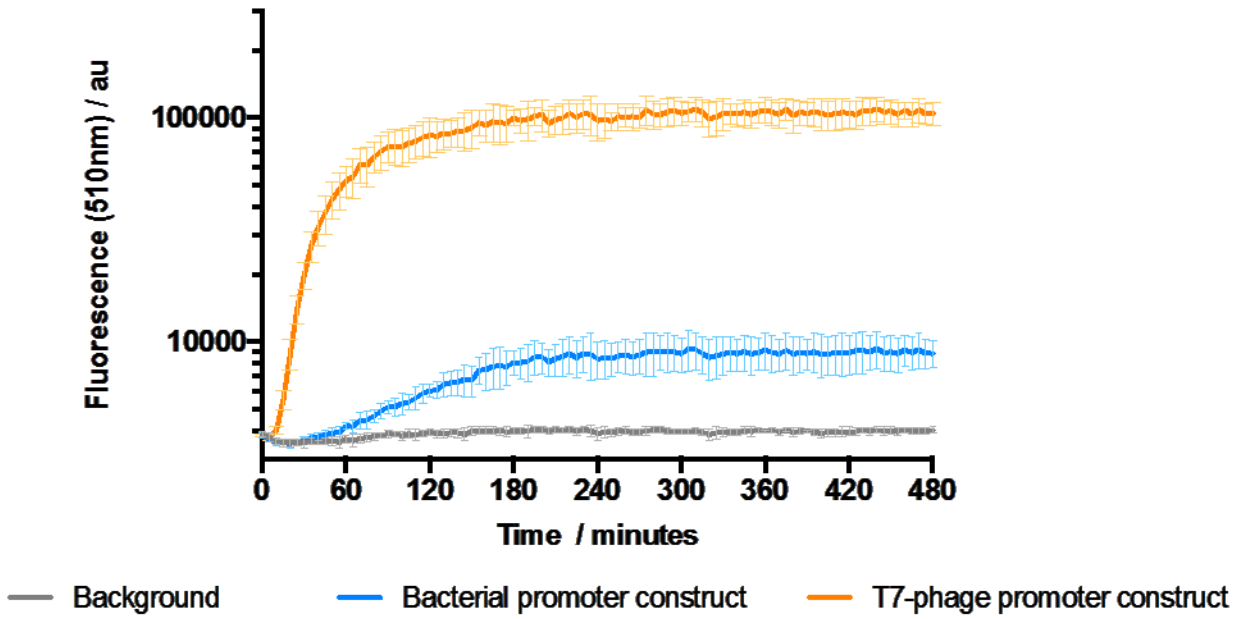


Figure 4.16: GFP expression in *E. coli* BL21(DE3) strain cell-extract system. GFP expression of plasmid construct pNK400 (incorporated a bacterial promoter Bba_J23101) and pNK401 (incorporated the consensus T7-bacteriophage promoter) when used in cell-extract reactions utilising an in-house prepared cell-extract batch derived from *E. coli* BL21(DE3) cell strain. Cell-extract reactions that did not include a plasmid construct were used to benchmark the autfluorescence signal of the cell-extract mix. Error bars represent standard deviation of 3 technical repeats. Cell-extract reactions were carried out at 10nM plasmid concentration (10µl total reaction volume) at 37°C and fluorescence recorded at 510nm.

Experimental results showed that buffer salt concentrations had a significant effect on the protein expression capabilities of the *E. coli* cell-extract (Figure 4.17). For the salt concentration conditions tested, comparing of the best performing condition versus the worst performing condition showed a: i) 7.6-fold difference for expression from the endogenous *E. coli* polymerase and ii) 4-fold difference for expression from the T7 RNA-polymerase. The optimal salt concentrations for protein expression mediated by the endogenous RNA-polymerase were 5mM Mg-glutamate and 80mM K-glutamate (Figure 4.17 top panel), while optimal salt concentration for the T7-phage RNA-polymerase were 2mM Mg-glutamate and 120mM K-glutamate (bottom panel Figure 4.17). At these salt concentrations, the fluorescent signal over background was 23-fold for protein expression mediated by the endogenous RNA-polymerase and 27-fold for the T7-phage RNA-polymerase. These values compared favourably to protein expression values of 2.2-fold and 26-fold over background signal obtained for the respective RNA-polymerase in pre-optimisation condition for cell-extract reactions, especially for expression mediated by the endogenous RNA-polymerase.

Testing the prototype device in TX-TL cell-extract

Section 4.2.3 demonstrated the functionality of the prototype biosensor device for the detection of thrombin in the commercially available *E. coli* S30 Extract System from Promega Corporation. In this section, a low cost cell-extract system was prepared and demonstrated to be capable of protein expression from DNA plasmid constructs. A fusion of these two technologies could facilitate the production of a biosensor technology of a very low cost per test reaction, a desirable characteristic for medical diagnostic testing at the point-of-care.

An experiment was carried out to test the functionality of the prototype thrombin switch device in cell-extract reactions using the prepared cell-extract system. Purified prototype device mRNA (2.2 picomoles) was incubated in binding reactions of 0 or 4.5 picomoles of thrombin protein for 30 minutes at 4°C and then samples were added to cell-extract reactions. Fluorescent signal from the cell-extract reactions was recorded over a 180 minutes period to monitor expression of GFP protein. To account for the background fluorescence signal of the cell-extract mix, a control reaction was set up in which no mRNA was added.

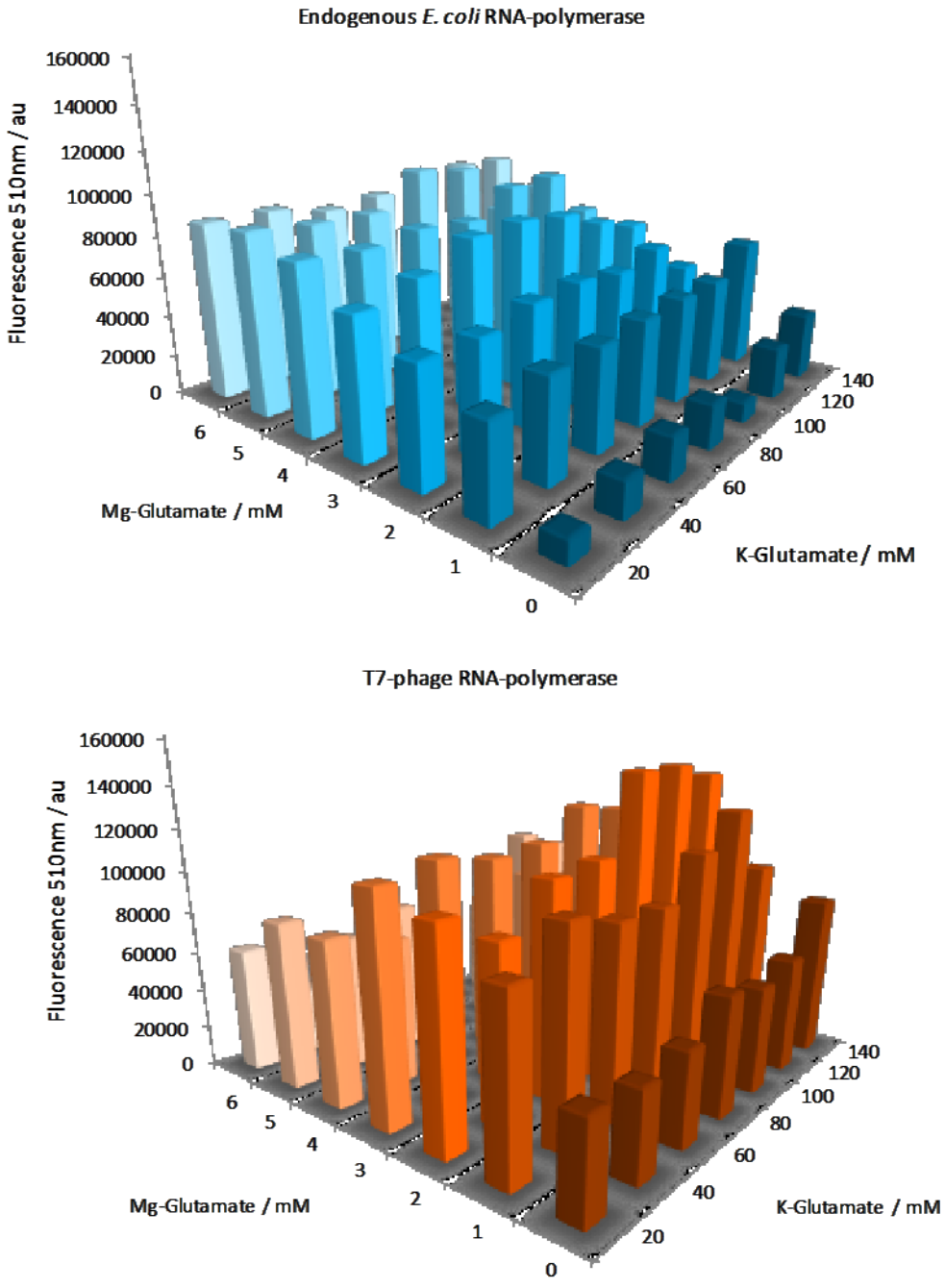


Figure 4.17: Buffer salts concentration optimization for cell-extract reactions. Cell-extract reaction conditions were optimised for magnesium glutamate and potassium glutamate to achieve high protein expression. Experiment investigated protein expression from: i) bacterial promoter system (top graph) using the pNK400 plasmid and ii) T7 phage promoter system (bottom graph) using the pNK401 plasmid construct. Protein production was determined for GFP expression at endpoint fluorescence (480 minutes).of reactions carried out at 37°C. Reported values represent one experimental repeat.

Experimental results confirmed that the thrombin switch device retained its functionality as a biosensor device for the detection of thrombin protein in the prepared cell-extract system (Figure 4.18). The top panel of Figure 4.18 shows the cell-extract reactions profile over time where samples without thrombin protein exhibited higher rates of GFP expression in comparison to samples in which thrombin protein was added. The left bottom panel of Figure 4.18 shows the cell-extract reaction samples endpoint fluorescence signal (180minutes), while the right bottom panel of Figure 4.18 shows the calculated percentage fluorescence signal following background signal correction. The endpoint fluorescence signal data were used to calculate the percentage fluorescence signal, and showed a mean 67% drop in fluorescence signal for samples added with thrombin protein in comparison to the mean fluorescence signal of samples not added any thrombin protein.

Summary of the low cost cell-extract platform technology section

In summary, a low cost cell-extract system was prepared and optimised for high levels of protein expression to be used for cell-extract reactions as an alternative platform to costlier commercial systems. The prototype biosensor device developed for the detection of thrombin was tested in the prepared cell-extract system and was shown to retain its biosensing functionality as previously demonstrated. Thus, experiments that follow utilised this alternative cell-extract system to drive reporter protein expression in cell-extract assays.

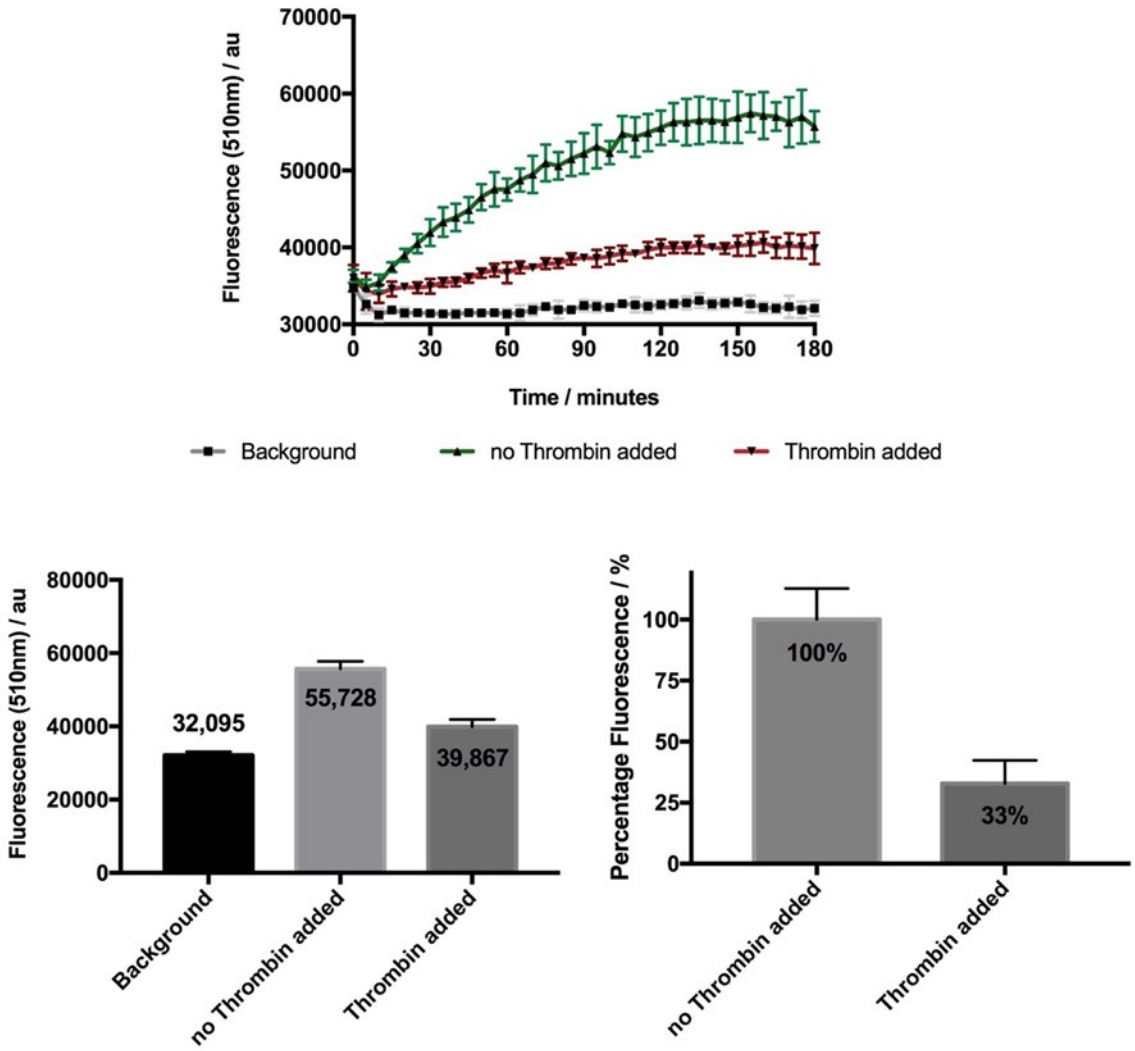


Figure 4.18: TXTL cell-extract repression assay for prototype thrombin device. Purified prototype thrombin switch device mRNA was incubated in binding reactions with either no added thrombin or with 4.5 picomoles thrombin protein and added to cell-extract reactions. Top panel shows the cell-extract reactions GFP expression profile over time. Bottom panel shows endpoint fluorescence signal (180 minutes) of cell-extract reactions (left) and the calculated percentage fluorescent signal following background signal correction (right). Cell-extract reactions were carried out at 37°C and fluorescent signal from GFP expression was monitored over time at regular intervals. Background samples was not added any mRNA. Error bars represent standard deviation of 3 technical repeats.

4.2.6 Strategies for modulating the biosensing assay characteristics

In previous experiments, cell-extract reactions carried out for testing the prototype biosensor device exhibited low levels of fluorescence signal generation for basal GFP expression as compared to background fluorescence (Figure 4.9, Figure 4.18.). It was speculated that the low signal/noise ratio was a result of low levels of GFP protein expression as compared to the autofluorescent signal of the cell-extract system mix. In the cell-extract reactions, the level of GFP production is assumed to be a function of: 1) the capacity of the cell-extract system for protein expression, 2) the amount of mRNA template added to the reaction mix and 3) the efficiency of the translation process from the mRNA transcripts. As previously discussed in section 4.1.2, the secondary structure stability of the mRNA at the translation initiation region is an important factor that determines translation initiation efficiency and levels of protein expression. With the purpose of exploring the design space for modulating the biosensing assay characteristics, a series of experiments were designed that investigated the effect of secondary structure stability around the RNA-aptamer region on the prototype device mRNA.

Construction of device variants of the prototype device

A PCR-based method was used to create a library of thrombin switch devices variants of the prototype device, hereafter referred also as thrombin switch device v01. The pNK402 construct (Figure 4.6), encoding the thrombin switch device v01, was used as the DNA template in a PCR reaction for the introduction of random nucleotides substitutions at 4 specified positions on the DNA code. These DNA positions corresponded to nucleotides in the helical sem that supports the formation of the anti-thrombin RNA-aptamer secondary structure in the transcripted biosensor device (Figure 4.5). The resulting DNA library from the PCR-based method was transformed in *E. coli* cells to facilitate isolation of individual library members. For the isolated library members, their plasmid DNA was extracted and sequenced. This produced a small library of thrombin switch devices that differed in the stability of the helical stem structure that supports the formation of the RNA-aptamer secondary structure.

The library of thrombin switch devices DNA sequences was examined for the stability of RNA secondary structure using a nucleic acid folding software application. Out of the library

members folded *in silico* using the mFold web platform (Zuker, 2003), 7 device variants were selected that exhibited lower secondary structure stability than the prototype thrombin switch device v01 (Table 4.1). The input mRNA sequence used for folding analysis included the anti-thrombin RNA-aptamer, the helical stem region, the AUG start codon and the first sequence amino acid codon of the GFP gene sequence. The predicted stabilities of the device variants spanned the deltaG range of -2.9 to -7.3 kcal/mole, while the prototype thrombin switch device v01 deltaG value was predicted to be -8.8 kcal/mole. The device variants were named thrombin switch devices v02 – v08 as specified in Table 4.1.

Table 4.1: Thrombin switch device variants predicted deltaG values for incorporated RNA-aptamer. Secondary structure folding analysis for thrombin switch device variants incorporating the anti-thrombin RNA-aptamer. Prediction was carried out using the mFold webserver (Zuker, 2003) using presented sequences as the input variable .

Name	Sequence [1]	deltaG (kcal/mole)
Prototype device v01	... ATG AAT TTG GAA GAT AGC TGG AGA ACT AAC CAA ATT CGT ...	-8.8
device v02	... ATG aaT TTG GAA GAT AGC TGG AGA ACT AAC CTA Ata CGT ...	-4.4
device v03	... ATG agT TTG GAA GAT AGC TGG AGA ACT AAC CAA Aca CGT ...	-6.2
device v04	... ATG agT TTG GAA GAT AGC TGG AGA ACT AAC CAA Ata CGT ...	-4.7
device v05	... ATG ggT TTG GAA GAT AGC TGG AGA ACT AAC CAA Aac CGT ...	-6.6
device v06	... ATG tgT TTG GAA GAT AGC TGG AGA ACT AAC CAA Agg CGT ...	-7.3
device v07	... ATG atT TTG GAA GAT AGC TGG AGA ACT AAC CAA Agc CGT ...	-4.0
device v08	... ATG gaT TTG GAA GAT AGC TGG AGA ACT AAC CAA Acg CGT ...	-2.9

[1] N: nucleotides identical to prototype device sequence; n: nucleotides mutated

GFP expression of thrombin switch device variants

In silico folding analysis of the newly constructed device variants showed differences in the stability of the mRNA secondary stucture around RNA-aptamer region (Table 4.1). Additionally, the subsequences (sequence that is a constituent part of a larger sequence) used for the folding analysis aforementioned overlap extensively with the translation initiation region of biosensing devices. Thus, an experiment was designed to investigate GFP expression of thrombin switch device variants since the stability of the secondary structure of mRNA at the translation initiation region is an important determinant of the translation initiation rate (Salis, 2011).

In this experiment, the thrombin switch devices v01 – v08 were investigated for their strength of GFP expression in cell-extract reactions. Template mRNA used in this experiment was produced by *in vitro* transcription and purified, while correct product size verification and purity assessment was examined and validated by RNA gel electrophoresis. The mRNA template of thrombin switch device variants was added to cell-extract reactions and GFP expression was monitored overtime using a microplate reader. The experiment also included a negative control reaction with no added mRNA to account for background fluorescent signal from the cell-extract mix.

Experimental results from the protein expression assay showed that all thrombin-switch device variants exhibited higher levels of GFP expression as compared to the prototype device (Figure 4.19). The top panel of Figure 4.19 shows a bar chart of the endpoint fluorescence signal (90 minutes) from the cell extract reactions. Thrombin switch device v07 showed a 13.9-fold increase in GFP expression as compared to the original prototype device which was the largest change in expression strength among the device variants. Thrombin switch device v05 showed a 6.6-fold increase in GFP expression and it was the lowest expressing device amongst the variants tested. The experimental data of GFP expression of the thrombin switch devices v01 – v08 were used to investigate any possible correlation between deltaG of incorporated RNA-aptamer secondary structure (Table 4.1) and protein expression. These two variables were plotted on an xy-scatter plot and were fitted with an exponential trendline with a correlation coefficient R^2-value = 0.42 which did not suggest a strong correlation.

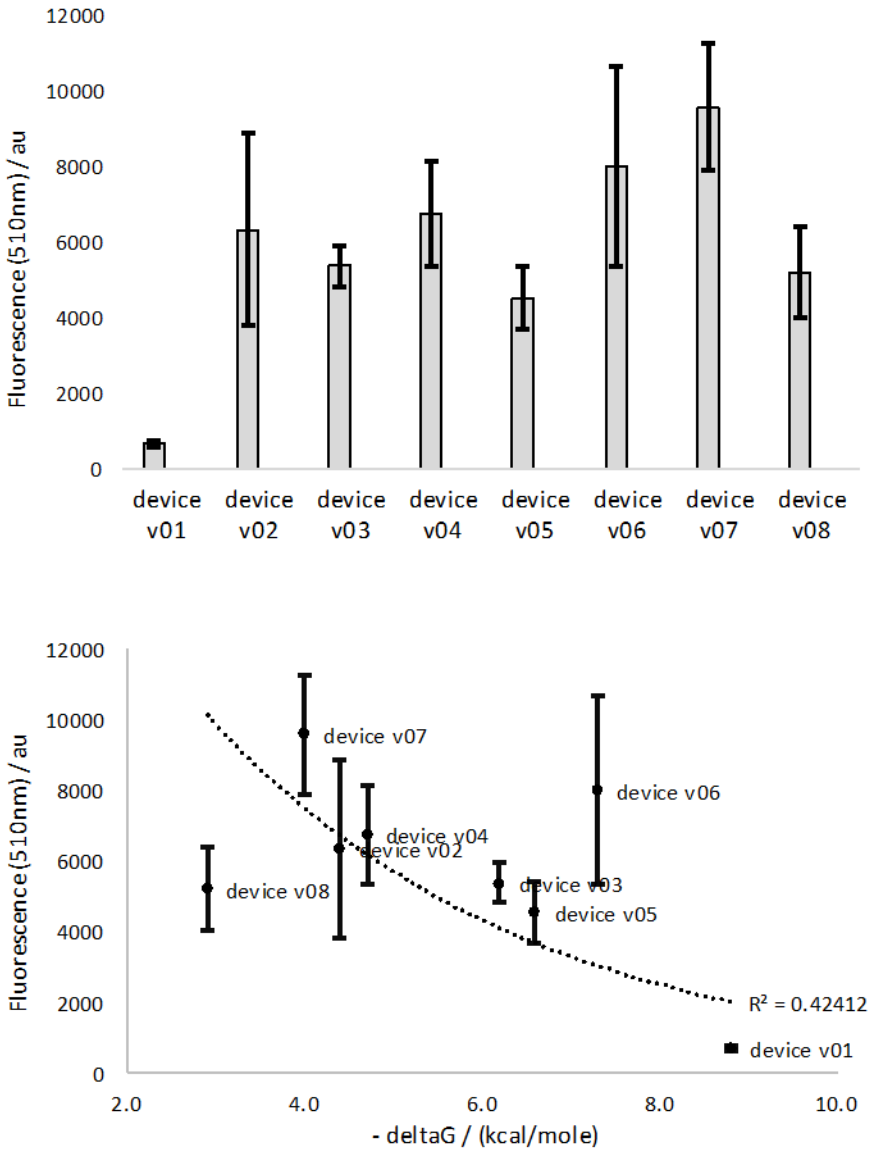


Figure 4.19: GFP expression of thrombin switch device variants in cell-extract reactions. Thrombin switch devices v01 – v08 mRNA in cell-extract reactions showed varied levels of GFP expression. (Top panel) Bar chart of endpoint fluorescence (90 minutes) from cell-extract reactions of thrombin switch devices measured at 510nm. (Bottom panel) XY-scatter plot with a trendline fitted on data of mRNA incorporated RNA-aptamer stability (-deltaG value as calculated by mFold) for devices v01 – v08 against experimentally determined levels of protein expression in cell-extract reaction. The trendline was fitted based on an exponential model of translation initiation for correlation analysis. Error bars represent standard deviation from 3 experimental repeats.

TIR prediction of thrombin-switch devices using the RBS calculator software tool

In the previous section, analysis of the experimental results did not show a strong correlation between the experimental results for GFP expression of thrombin switch devices and the predicted free energy of folding for subsequences of the RNA-aptamer local region (Figure 4.19 bottom panel). A strong correlation would have enabled the *in silico* design of biosensing devices by modulating the GFP expression strength of devices towards desired levels of expression. To this end, an experiment was designed to investigate whether the RBScalculator software tool (Salis, 2011; Salis et al., 2009) could be used to predict GFP expression levels of thrombin switch devices.

This experiment compared the experimental data for GFP expresion of thrombin switch devices v01 – v08 from cell-extract reactions to the predicted translation initiation rate (TIR) using the RBS calculator software tool. The TIR prediction was carried out using the RBScalculator v2.0 algorithm of the software tool. The input variables used for the TIR prediction were: i) the subsequence of each thrombin switch device mRNA sequence that included the first 85 nucleotides; starting from the GGG nucleotides at the end of the T7-promoter and ii) the *Escherichia coli* BL21(DE3) as organsim of choice and _ACCTCCTTA_ 16S anti-SD sequence. Predicted TIRs were compared against endpoint fluorescence signal data previously collected in protein expression cell-extract reactions (top panel of Figure 4.19).

RBScalculator predicted TIRs for thrombin switch devices v01 – to v08 are shown in Figure 4.20. The software tool predicted an identical TIR value for devices v01, v02 and v05 of 41,193 TIR au. The highest score was attained for device v08 with a TIR value of 61,763 au and lowest score for device v07 with a TIR value of 17, 517 au. Device v03 scored 20,049 au, device v04 35,991au and device v06 18,995 au. The predicted TIR values were plotted in a xy scatter plot against experimental data of GFP expression (Figure 4.20). A linear regression analysis found an R^2-value = 0.27, that suggested poor predictive ability by the RBScalculator model for these devices.

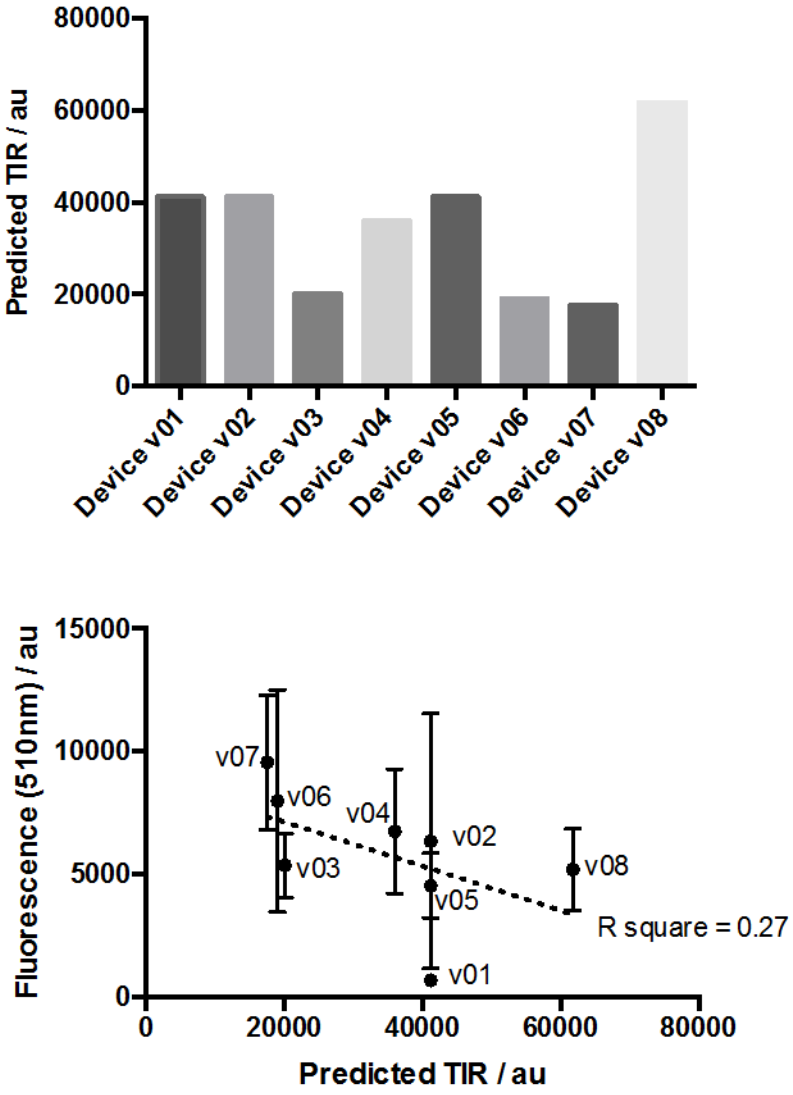


Figure 4.20: TIR prediction for thrombin switch device variants by RBScalculator. (Top) Bar chart of the predicted TIRs by RBScalculator v2.0. (Bottom) Scatter plot with a trendline fitted to data of predicted TIRs against experimentally derived GFP expression data in cell-extract reactions. The trendline was fitted based on an linear model of translation initiation for correlation analysis. Cell-extract reactions were carried out at 37°C and error bars represent standard deviation for 3 technical repeats.

Testing device variants for thrombin protein binding by EMSA

Section 4.2.4 described the development and validation of an EMSA setup for assessing thrombin protein binding to the prototype device (thrombin switch device v01). The EMSA assay allowed for the direct observation and validation of the interaction between the thrombin protein and the mRNA molecules of the thrombin switch device v01. The newly constructed thrombin switch devices (Table 4.1), include nucleotide substitutions at the helical stem that supports the formation of the anti-thrombin RNA-aptamer structure and these nucleotide substitutions could potentially affect the ability of devices to bind the thrombin protein molecule.

To investigate whether the thrombin switch device variants retained the ability for binding thrombin, an EMSA experiment was carried out as previously described for ~150 nucleotide long partial sequences of devices (section 4.2.4). Purified mRNA of thrombin switch devices v01 - v08 were produced by *in vitro* transcription and purification. Correct products and purity of samples were verified by RNA electrophoresis, in which the purified RNA migrated on the gel similarly to the 150 nucleotide RNA ladder band (Figure 4.21). Then, thrombin switch device variants were incubated in binding reactions under two conditions: either in the absence or in the presence of thrombin protein (protein:RNA molar ratio = 1:1) and loaded on gel for EMSA analysis. To account for protein:RNA non-specific interactions, control device mRNA (partial sequence) -which did not include an anti-thrombin binding RNA-aptamer sequence- was used as the experimental negative control (Figure 4.21). The thrombin switch device v01 mRNA (partial sequence) was used as the positive control for the experimental setup.

EMSA analysis showed thrombin protein induced band shift for all thrombin switch device variants (Figure 4.22). Bottom left panel of Figure 4.22 shows an image of the EMSA gel following staining for RNA-visualisation. This image shows that gel lanes loaded with binding reaction samples of device variants in which no thrombin protein was added produced single band (basal bands). The gel lanes loaded with samples incubated with thrombin produced a shifted band (or a band pattern of a shifted band and basal band RNA). This band pattern was in accordance with the band pattern of the positive control sample of prototype device mRNA, while the control samples did not produce any shifted bands. The bottom right panel of Figure

4.22 shows an image of the EMSA gel following staining for protein visualisation. The staining produced a single protein band in gel lanes loaded with binding reaction samples of all thrombin switch devices incubated in the presence of thrombin. In contrast, the protein band was absent from the corresponding gel lane of control device mRNA. When the two gel images were overlaid using ImageJ software, the protein bands overlapped with the shifted RNA bands. These EMSA results suggested that all thrombin switch device variants retained their ability to bind the thrombin protein molecule. The differences in the signal intensity of the shifted and basal bands for samples that were added thrombin protein could be due to differences in the stability of the RNA-aptamer structure or due to pipetting errors.

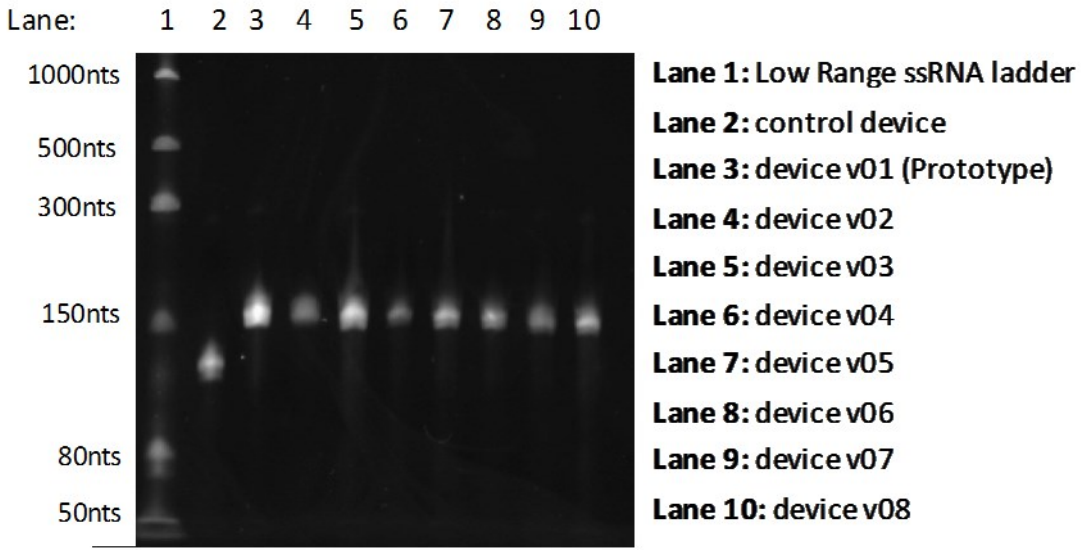


Figure 4.21: RNA electrophoresis of thrombin switch device variants partial sequences. RNA size determination for the purified mRNA thrombin switch devices v01 – v08 (152nts) and control device (119nts) partial sequences. Bands run at the expected sizes. RNA gel analysis carried out on a 5% PAGE / TBE-urea denaturing gel.

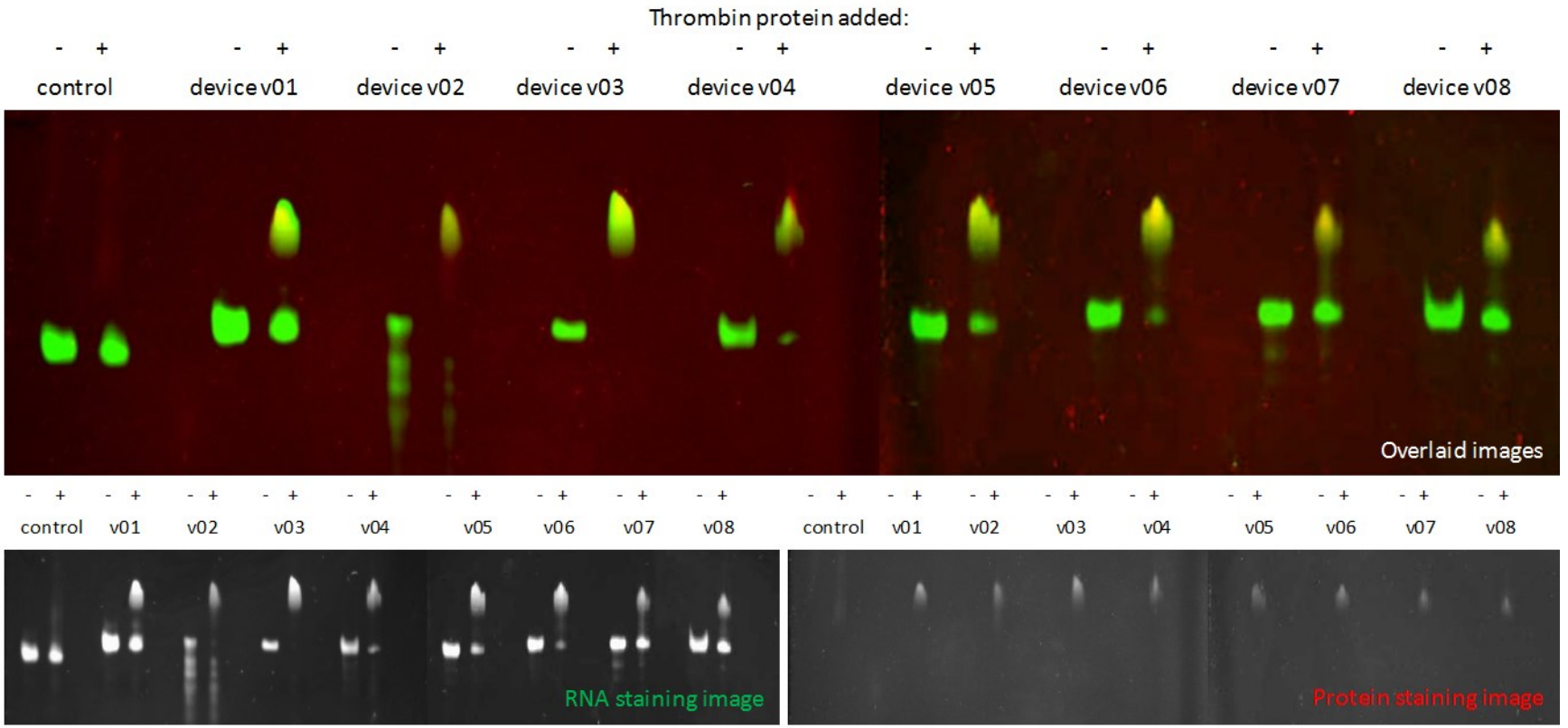


Figure 4.22: EMSA of the thrombin switch device variants. Purified mRNA samples of thrombin switch devices v01 – v08 partial (~150 nucleotides partial sequences) and of control device (corresponding partial sequence) were incubated in binding reactions in the absence (-) or presence (+) of thrombin protein (protein/RNA molar ratio of 1:1) and loaded on gels for EMSA analysis. Gel images show shifted bands of RNA/protein complex formation for all thrombin switch devices but not for control device. Binding reactions were carried out for 30 minutes at 4°C. EMSA gel system used was 5% PAGE / HEPES-Imidazole buffered (pH = 7.55) and the gel electrophoresis was carried out at 4°C. The gel was first stained with GelRed nucleic acid dye and imaged, and then stained with SybroRuby protein dye and imaged. The images were overlaid using ImageJ software.

Testing thrombin switch device variant in repression assay

Previous experiments showed that the thrombin switch device variants exhibited higher GFP expression capabilities than the prototype thrombin switch device v01 (Figure 4.19). Additionally, EMSA experiments determined that all the constructed device variants retained their ability to bind thrombin protein (Figure 4.22). These results suggested that the device variants, if functional as thrombin biosensing devices in cell-extract reactionions, should allow for the development of thrombin assays with different characteristics for analyte detection, such as dynamic range and/or limit of detection. The functionality of the device variants as thrombin biosensor devices in cell-ectract reactions was investigated using the thrombin switch device v04 variant in repression assays as previously (Figure 4.11). The reason for favouring the use of thrombin switch device v04 as opposed to other variants is that it exhibited more consistent behaviour in GFP expression assays in preliminary GFP expression tests.

In this experiment, thrombin switch device v04 was incubated in binding reacions at a range of thrombin protein concentrations and then added to cell-extract reactions. Binding reactions of 3.3 picomoles of device mRNA and 0, 1.5, 2.25 and 3 picomoles thrombin were incubated for 30 minutes at 4°C, and added to the cell extract reactions. The cell-extract reactions were incubated at 37°C and fluorescence signal was recorded over time. To account for any indirect effects of thrombin protein addition to the cell-extract system, an experimental negative control was included in which control device mRNA was treated identically to the sample reaction.

Experimental results showed thrombin protein mediated repression of GFP expression for thrombin switch device v04 (Figure 4.23). GFP expression of the thrombin switch device v04 decreases as thrombin concentration in the reaction mix increases and as compared to the basal expression values in the absence of thrombin. A trendline drawn through the points of the scatter plot revealed an inversely proportional linear relationship between GFP expression and thrombin protein concentration in the cell-extract reactions for the protein concentrations tested, with an R^2-value = 0.98. In contrast, for the control device the fluorescent signal remained constant for all thrombin protein concentrations tested.

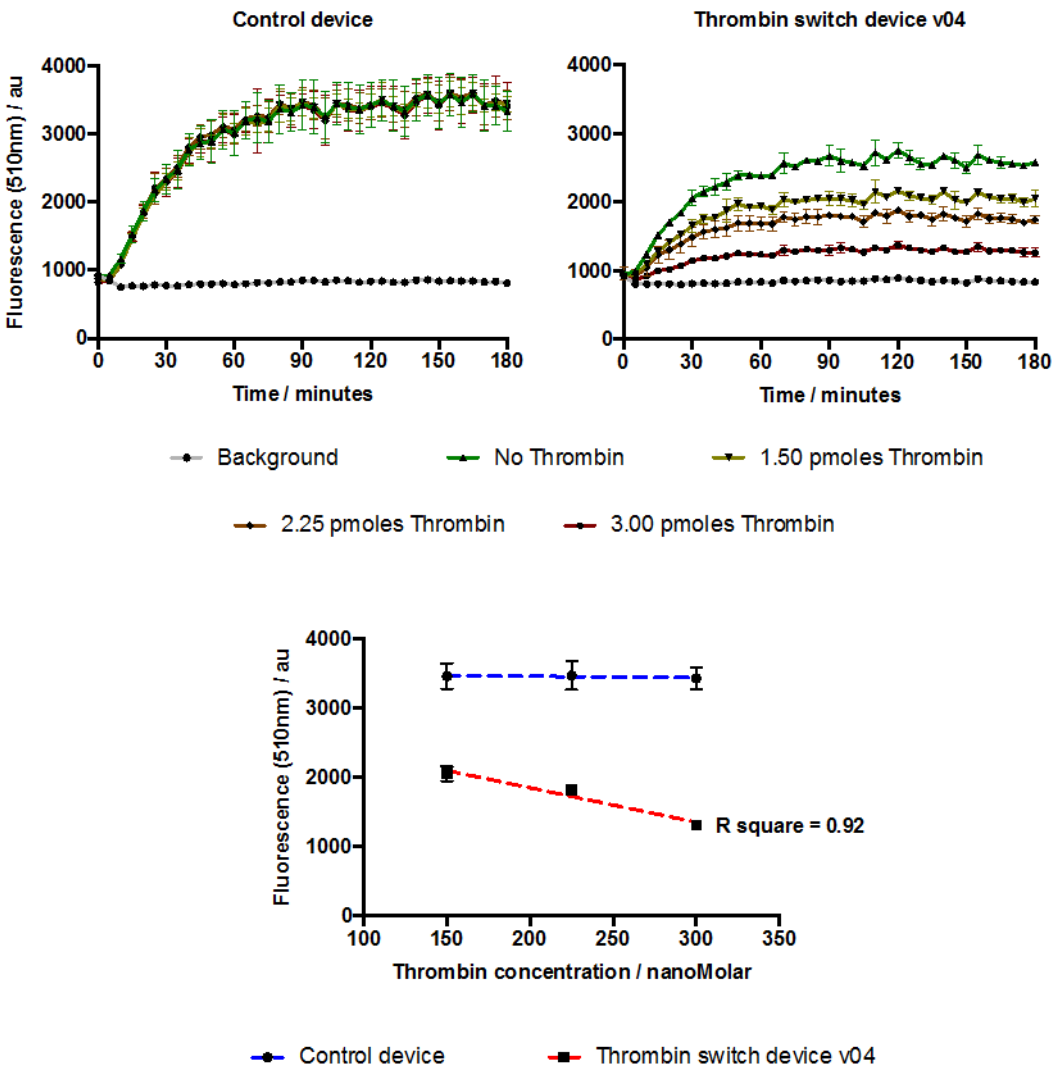


Figure 4.23: Input / output characteristics of thrombin switch device v04. (Top) Reaction profile of GFP expression cell-extract reactions for control device and thrombin switch device at the presence of various thrombin protein concentrations. Binding reactions were set up with purified mRNA of the respective device and 0, 1.5, 2.25 or 3 picomoles of thrombin protein and incubated for 30 minutes at 4°C . Samples were then added in cell-extract reactions and fluorescent signal was monitored over time at 37°C. (Bottom) Scatter plot of endpoint fluorescence (90 minutes) of from cell-extract reactions against thrombin protein concentrations (150, 225 and 300 nM). Error bars represent standard deviation of 3 technical repeats. Dotted lines represent linear regression analysis with 95% confidence limits and R square value of thrombin switch device v04 is noted on the graph.

Summary ofinvestigating the design space for engineering assay characterictics

In summary, during this set of experiments seven new variants of prototype thrombin switch device v01 were constructed for the purpose of investigating the design space of this RNA-switch technology for engineering the biosensing assay characteristics. The variants included nucleotide substitutions that were thought to provide lower secondary structure stability in a region that influences efficiency of translation initiation.

GFP expression assays in cell-extract reactions showed that all the device variant expressed higher levels of GFP as compared to the prototype device. Towards forward designing biosensing devices with high expression level capabilities, two approaches were tested for their ability to pedict protein expression levels *in silico.* The first approach attempted to build a model that related the secondary structure stability of a specified subsequence that included the local area of the RNA-aptamer to the GFP expression. The second approached utilised the RBScalculator software tool. Both approaches were found to be limited in their predictive capabilities.

The newly constructed thrombin switch devices were tested for their ability to bind thrombin by EMSA. EMSA showed that all device variants retained the ability to interact with thrombin protein despite exhibiting higher levels of GFP expression that were though to arise from a less stable stem helix supporting the structure of the incorporated RNA-aptamer.

Finally, one thrombin device variant was tested in a cell-extract thrombin biosensing assay. The device was able to detect the presence of thrombin protein by the repression of GFP expression. The relationship between the device GFP protein expression and thrombin protein concentration was linear over the analyte concentration range tested.

4.2.7 Investigating biosensor design modularity

An investigation was carried out to test the modularity of the RNA-switch biosensor design for the incorporation of other RNA-aptamer sensor elements. A device that is modular in design should be able to incorporate different sensor elements without detrimental effects to the functionality of the device. More specifically, the specific recognition of the target protein analyte and the differential reporter expression as a result analyte binding should hold true after the incorporation of an alternative RNA-aptamer sensor. If the designed device with the alternative sensor element exhibits behaviour that is fundamentally different from the behaviour of the prototype design, then the design is limited in its modularity. Experiments in this section, investigated the modularity of the biosensor design using an anti-VEGF RNA-aptamer and toggle-25t RNA-aptamer (Figure 4.2).

Anti-VEGF RNA-aptamer / VEGF protein

A SELEX derived RNA-aptamer that can specifically bind Vascular Endothelial Growth Factor (VEGF) protein biomarker (Jellinek et al., 1994) was selected for testing the modularity of the RNA-switch biosensor design. The primary sequence and the predicted secondary structure of the minimal anti-VEGF RNA-aptamer are presented in Figure 4.2. The target molecule for this RNA-aptamer is recombinant VEGF protein (165 amino acids form).

EMSA of the anti-VEGF RNA-aptamer

The interaction between the anti-VEGF RNA-aptamer and VEGF has been demostrated previously with the use of a nitrocellulose filter binding assay (Jellinek et al., 1994). An EMSA experiment was carried out using chemically synthesized anti-VEGF RNA-aptamer and purchased VEGF protein with the aim of validating the EMSA setup for this particular pair of interacting biomolecules.

In this experiment, the anti-VEGF RNA-aptamer was incubated in binding reactions with VEGF at protein:RNA molar ratios of 0:1, 0.6:1, 1.2:1 and 1.8:1. The binding reactions were carried out at 37°C for 15 minutes and then the samples were loaded on a gel for analysis. To account for non-specific binding of VEGF to RNA molecules, chemically synthesized anti-thrombin RNA-aptamer (previously utilised for the development of the thrombin prototype

RNA-switch device) was used as a negative control. Additionally, a sample of VEGF was added on the gel to track free VEGF mobility in the gel. The EMSA gel system used in this experiment was an 8% PAGE / TBE buffered (pH = 8.3) gel and electrophoresis was carried out at 4°C.

EMSA results showed VEGF induced shift of anti-VEGF RNA-aptamer bands (Figure 4.24). The image of the gel following staining for RNA visualisation (top right Figure 4.24) shows band shift for anti-VEGF RNA-aptamer (gel lanes 6-9). A single basal RNA band appeared in gel lane 6 that was loaded with the binding reaction sample with no VEGF added. Gel lanes 7–9 were loaded with binding reaction samples with added VEGF protein and showed a shifted RNA band in addition to the basal RNA band. For these gel lanes, higher concentration of VEGF in the samples resulted in shifted bands of greater signal intensity and basal bands of lower signal intesity towards gel lanes with higher protein concentration. In contrast, there was no induced shift of RNA bands for negative control gel lanes 2-5 that were loaded with binding reaction samples of anti-thrombin RNA-aptamer and VEGF protein. The image of the gel following staining for protein visulaisation (Figure 4.24 bottom right) shows a single protein band in each of lanes 7,8 and 9 at increasing signal intensity. All other gel lanes did not stain. However, the gel wells of lanes 1, 4, 5, 8 and 9 gave a strong signal. This signal most likely arises from unbound VEGF protein stuck in the well.

When the gel images for RNA staining and protein staining were overlaid using the ImageJ software, the protein signal bands overlapped with the shifted RNA bands in gel lanes 7-9. This shows the co-localisation of interacting anti-VEGF RNA-aptamer and VEGF protein in the gel. In conclusion, the experimental results showed VEGF protein binding to the anti-VEGF RNA-aptamer and the specific recognition by the RNA-aptamer, as no interaction was observed between anti-thrombin RNA-aptamer and VEGF protein. This experiment validated the EMSA setup as a suitable assay system for studying this system of interacting partners.

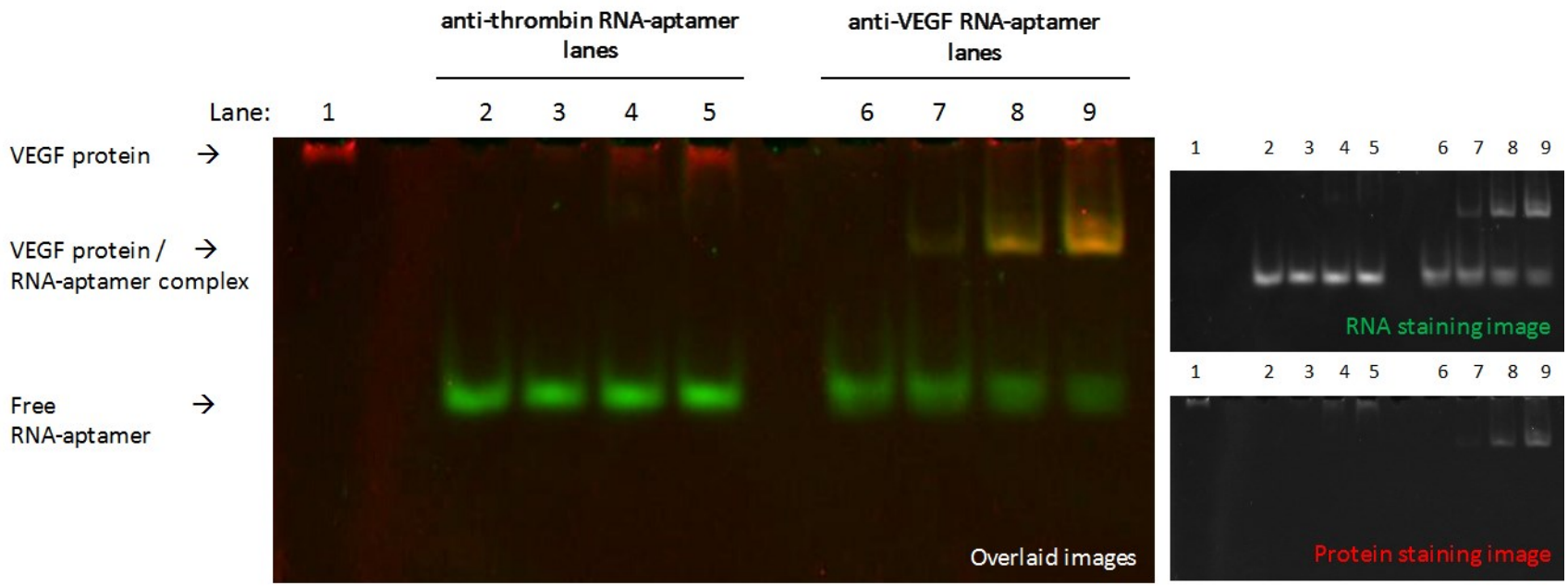


Figure 4.24: EMSA of the anti-VEGF RNA-aptamer. Image of EMSA gel showing band shift of anti-VEGF RNA-aptamer bands in the presence of VEGF protein. Chemically synthesised anti-VEGF RNA-aptamer was incubated in binding reactions with purified VEGF at protein:RNA molar ratios of 0:1, 0.6:1, 1.2:1 and 1.8:1. Following incubation, samples were loaded in gel (lanes 6 - 9). Negative control reactions were set up using chemically synthesised anti-thrombin RNA-aptamer and VEGF and loaded in the gel (lanes 2 - 5). Lane 1 was loaded with purified VEGF protein. Binding reactions were incubated at 37°C for 15 minutes. EMSA was carried out in an 8% PAGE/ TBE-buffer (pH8.3) gel system and electrophoresis was carried out at 4°C. The gel was first stained with GelRed dye for RNA visualisation and imaged, and then stained with SybroRuby protein dye and imaged. Images were overlaid using ImageJ software.

EMSA of the VEGF switch devices

Previous experiments validated the EMSA setup for studying VEGF binding to the anti-VEGF RNA-aptamer. Next, an EMSA experiment was designed to test whether the anti-VEGF RNA-aptamer retained its ability to bind VEGF protein when incorporated into the mRNA expression platform of the biosensor device design.

For this experiment, the anti-VEGF RNA-aptamer sequence was incorporated into the pNK401 construct to engineer VEGF RNA-switch biosensor devices plasmid constructs. Three different VEGF RNA-switch device variants were constructed according to the design specifications previously developed. These specifications are: i) incorporation of the VEGF RNA-aptamer directly downstream of the ATG start codon and ii) in-frame to the GFP reporter protein sequence. The device variants differed in the stability of the stem helix that forms the anti-VEGF RNA-aptamer (Table 4.2). The plasmid constructs of the VEGF switch devices v01-03 were verified by DNA sequencing.

An EMSA experiment investigated the ability of the designed VEGF switch device variants to bind VEGF, utilising partial sequences of devices as previously described (section 4.2.4). For this experiment, ~150 nucleotide long mRNA of devices was produced by *in vitro* transcription and purified. RNA gel electrophoresis was used to verify the size and purity of the mRNA products. The imaged RNA-electrophoresis gel showed that the RNA bands run at the expected size relative to the ladder standard, and that the samples were of sufficient purity (Figure 4.25).

Then, VEGF switch devices mRNA (partial sequences) was incubated in binding reactions in the absence or presence of VEGF. The binding reactions were set up at a protein:RNA molar ratio of 1.15:1, and reactions were incubated for 30 minutes at 4°C. Following incubation, the samples were loaded in a 5% PAGE / TBE buffered (pH = 8.3) gel system for EMSA analysis. To account for non-specific interactions between VEGF and RNA molecules, negative control reactions were set up using control device (no RNA-aptamer) mRNA (partial sequence).

Table 4.2: Designed VEGF RNA-switch devices. VEGF switch devices incorporating the anti-VEGF RNA-aptamer (Jellinek, 1994) into the mRNA reporter protein expression platform. Devices are presened with their associated deltaG as predicted for the subsequences presented here. Free energy prediction was carried out using the mFold tool.

Name	Sequence[1,2]	deltaG (kcal/mole)
device v01	...ATG gga A*Gc TTG ATG GGT GAC ACA CGT CAT GCC GAG C*tt CGT...	-15.6
device v02	...ATG ccc A*Gc TTG ATG GGT GAC ACA CGT CAT GCC GAG C*tg ggt CGT...	-21.6
device v03	...ATG cac A*Gt TTG ATG GGT GAC ACA CGT CAT GCC GAG C*tg ggt CGT...	-12.2

[1] N: nucleotide common to all devices; n: nucleotides in which devices differ
[2] *N or n*: minimal anti-VEGF RNA-aptamer sequence

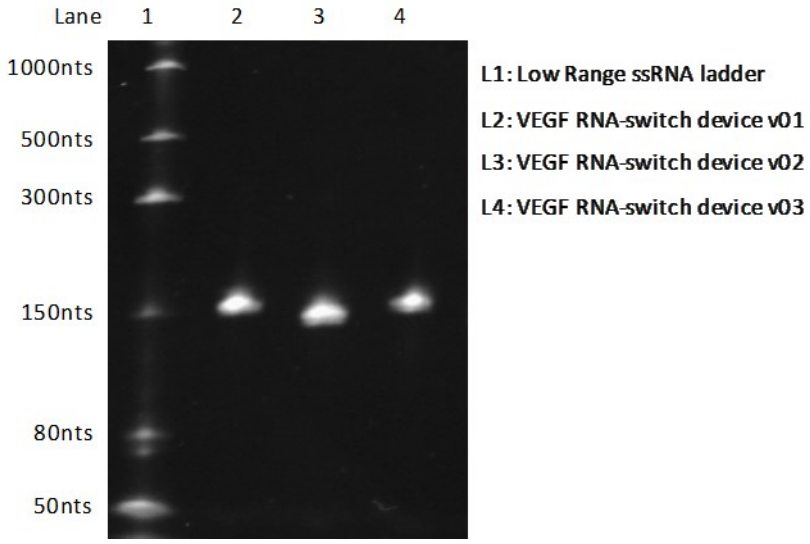


Figure 4.25: RNA-electrophoresis of designed VEGF switch devices (partial sequences). Image of RNA-electrophoresis gel for the VEGF switch devices v01 – v03 (~150nts) partial sequences. Analysis of *in vitro* transcribed mRNA was carried out on a 5% PAGE / TBE-urea denaturing gel system.

EMSA showed VEGF induced RNA band shift for all designed VEGF switch devices (Figure 4.26). Image at the top right corner of Figure 4.26 shows the EMSA gel following staining for RNA visualisation. In this image, gel lanes loaded with binding reaction samples of VEGF devices in the absence of VEGF protein show a single basal RNA band. On the contrary, gel lanes loaded with binding reaction samples of VEGF devices with added VEGF show a shifted RNA band in addition to the basal RNA band. Gel lanes loaded with the control device samples show a single basal band for both reaction conditions of absence and presence of VEGF. Additionally, the shifted RNA bands of VEGF switch devices did not produce sharp bands on the gel but rather broad signal bands. Shifted band of device v01 shows higher signal intensity in the lower part of the band. Shifted band of device v02 shows an even distribution of signal intensity. Finally, shifted band of device v03 shows higher signal intensity at the upper part of the band. The factors that resulted to these patterns could not be determined. The bottom right image of Figure 4.26 shows an image of the EMSA gel following staining for protein visualisation. In this image, protein bands appeared in gel lanes loaded with binding reactions samples in which the VEGF devices were incubated with added VEGF. No protein band appeared in the corresponding gel lane for the control device. Additionally, the signal distribution within the protein band matches the signal distribution of the shifted RNA bands.

When gel images were overlaid, the protein band signal overlapped with the signal of the shifted RNA bands of VEGF switch devices (composite image of Figure 4.26). In conclusion, the EMSA results suggested that the anti-VEGF RNA aptamer incorporated into the reporter expression platform retains its ability to bind its target VEGF protein. This was found to be true for all the designed VEGF switch devices tested.

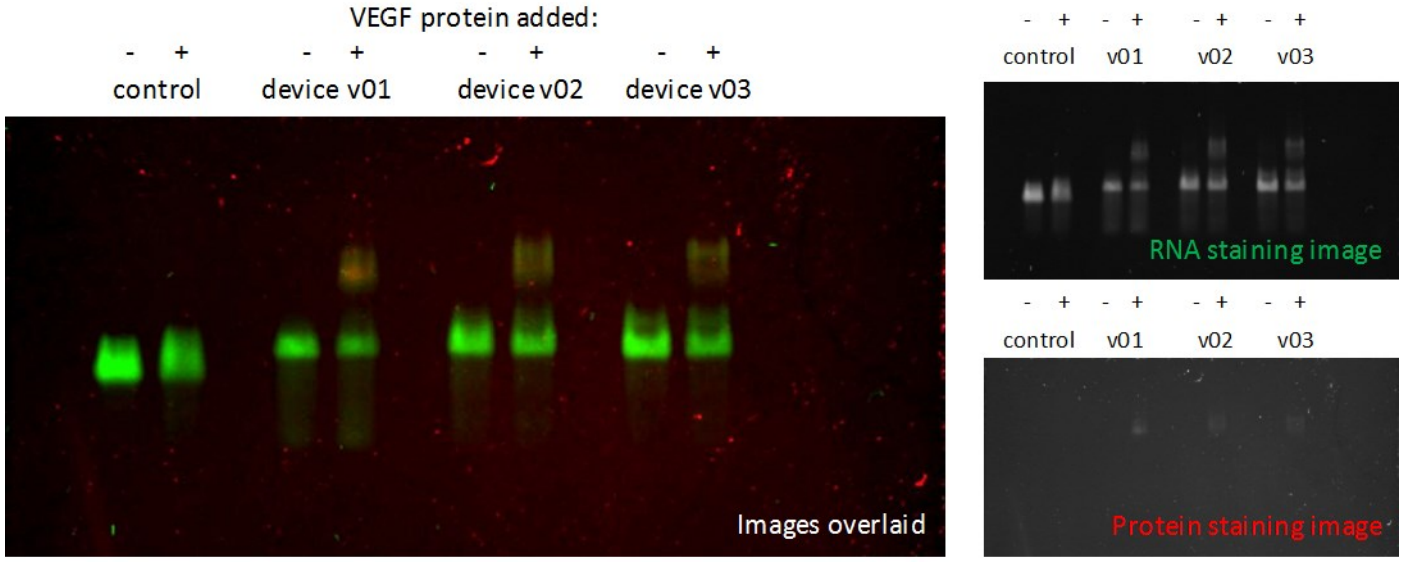


Figure 4.26: EMSA of designed VEGF switch devices. EMSA of VEGF switch devices v01-03 with purified VEGF protein. Purified mRNA of VEGF devices partial sequences (~150 nucleotides) were incubated in binding reactions in the absence (-) or presence (+) of VEGF protein (protein/RNA molar ratio = 1.15 : 1) and loaded on the gel for EMSA analysis. Binding reactions were carried out for 30 minutes at 4°C. EMSA gel system used was 5% PAGE / TBE-buffered (pH = 8.3) and the assay was run at 4°C. The gel was first stained with GelRed dye for RNA visualisation and imaged, and then stained with SybroRuby protein dye and imaged. Images were overlaid using ImageJ software.

GFP expression of VEGF switch devices

Previously, EMSA demonstrated that partial sequences of designed VEGF switch devices that incorporated the anti-VEGF aptamer were able to bind the VEGF (Figure 4.26). Next, an experiment was designed to test the GFP expression capacity of the VEGF switch devices in cell-extract reactions.

For this experiment, VEGF switch device v01-v03 full length mRNA was produced by *in vitro* transcription and purified. Then, more than 15 picomoles of devices mRNA (2.5µl -the maximum allowed volume- of mRNA stock for each device) was added to cell-extract reaction mix and incubated at 37°C. Fluorescent signal from the reaction samples was monitored over time. To account for background fluorescence of the cell-extract mix, cell-extract reactions were set up in which no mRNA added. Additionally, positive control reactions were set up with control device full length mRNA.

The fluorescent signal from VEGF switch devices cell-extract reactions did not rise above the background fluorescent signal (bottom Figure 4.27). This suggested that GFP protein was not expressed. Across the timeframe that the cell-extract reactions were monitored, the reaction profile showed that the fluorescent signal of the samples was similar to the background fluorescence signal as determined by the negative control reactions which did not include any mRNA template (top Figure 4.27). In contrast, the positive control reactions added with control device mRNA showed high levels of GFP expression. These results from the positive control samples eliminated the possibility of an incorrect experimental setup. The lack of GFP expression from the VEGF switch devices could be explained by inefficient translation initiation due to the imposition of very stable secondary structure by the anti-VEGF RNA-aptamer at the RBS region of the devices. Future experiments could investigate GFP expression from VEGF devices with less stable secondary structure around the anti-VEGF RNA-aptamer region.

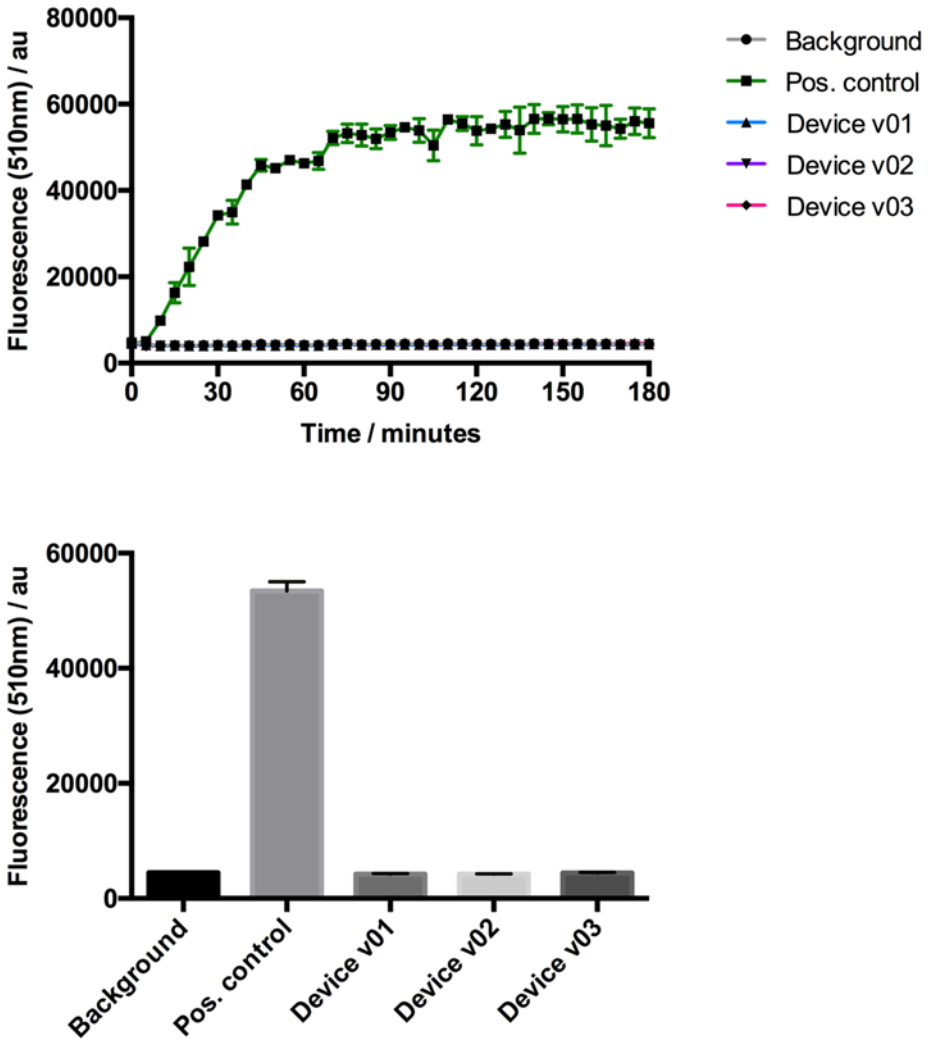


Figure 4.27: GFP expresion of designed VEGF switch devices. (Top) Reaction profile of cell-extract expression reactions of designed VEGF switch devices v01-v03. (Bottom) Bar chart of recorded fluorescent signal at 90 minutes for cell-extract reactions. Background samples included cell-extract reaction mix in which no template mRNA was added to determine aut-fluorescent signal from cell-extract mix. Positive control samples included control device mRNA (no RNA-aptamer) which was used as the experimental setup control. Cell-extract reactions were carried out at 37°C with purified mRNA. Error bars represent standard deviation of 3 technical repeats.

Summary for anti-VEGF sensor element / VEGF target protein pair

In summary, the anti-VEGF RNA-aptamer was utilised to investigate the modularity of the biosensor design for incorporating alternative sensor elements. Initially, an EMSA was carried out that demonstrated the utility of the assay to study the interaction between the anti-VEGF RNA-aptamer and VEGF protein. Following, three VEGF switch devices were constructed that incorporated the anti-VEGF aptamer into the GFP expression mRNA platform. Partial sequences of the constructed VEGF switch devices were tested with EMSA and were found to bind VEGF. However, the full length VEGF switch devices failed to express GFP in cell-extract reactions. It was speculated that the very stable RNA secondary structure imposed to the translation initiation region of the mRNA expression platform due to RNA-aptamer incorporation inhibited translation initiation. This sets limitations to the modularity of the RNA-switch biosensor design.

Toggle-25t RNA-aptamer / Human thrombin protein

A SELEX derived RNA-aptamer specific for the recognition of thrombin protein was selected to test further the modularity of the biosensor design. The nucleic acid primary sequence and the experimentally validated secondary structure (Long et al., 2008) are presented in Figure 4.2. The target molecule of the toggle-25t RNA-aptamer (White et al., 2001) is thrombin protein as is the case for the anti-thrombin RNA-aptamer (X. M. Liu et al., 2003) previously investigated. However, the two thrombin binding RNA-aptamers have different primary nucleotide sequence and binding affinities for their target (Figure 4.2). Also, the SELEX procedure for the selection of the toggle-25t RNA-aptamer used a pool of RNA oligonucleotides incorporating 2'-fluoropyrimidine nucleotides (White et al., 2001). Nevertheless, in this research the toggle-25t RNA-aptamer was synthesized to incorporate only "natural" 2'-hydroxyl group ribonucleotides.

EMSA of toggle-25t RNA-aptamer

To investigate the ability of the toggle-25t RNA-aptamer to specifically recognise thrombin, an EMSA experiment was carried out using chemically synthesized toggle-25t RNA-aptamer and purchased human thrombin protein.

For this experiment, the toggle-25t RNA-aptamer was incubated in binding reactions with thrombin at protein/RNA molar ratios of 0:1, 0.4:1, 0.7:1 and 1:1. Following 10 minutes incubation at 37°C, samples were loaded on EMSA 8% PAGE / HEPES-Imidazole buffered (pH = 7.55) gel system for analysis. To account for non-specific binding of thrombin to RNA molecules, chemically synthesized anti-VEGF RNA-aptamer (Jellinek, 1994) (previously utilised for the development of the VEGF specific RNA-switch devices) was used as a negative control.

EMSA analysis showed thrombin mediated band shift for the toggle25t RNA-aptamer bands (Figure 4.28). The image to the top right corner of Figure 4.28 shows the EMSA gel stained for RNA visualisation. Gel lanes 1-4 were loaded with the binding reaction samples of toggle-25t RNA-aptamer with increasing concentrations of thrombin. Gel lane 1 that was loaded with just the toggle25t RNA-aptamer shows a single band pattern. Gel lanes 2-4 that were loaded with toggle25t RNA-aptamer samples of increasing thrombin concentrations show a second shifted RNA band, while the signal intensity of the basal bands decreases towards higher concentrations of thrombin in these lanes. In contrast, negative control samples (gel lanes 5-8) that were loaded with samples of anti-VEGF RNA-aptamer and thrombin concentrations did not show any band shift. The image at the bottom right corner of Figure 4.28 shows an image of the EMSA gel stained for protein visualisation. Gel lanes 2-4 were the only ones that stained positive and formed signal bands for protein presence. When the two gel images were overlaid using the ImageJ software, the protein bands overlapped with the shifted RNA bands on the composite gel image. This showed the co-localisation of toggle-25t RNA-aptamer and thrombin protein in the gel.

In conclusion, the EMSA analysis suggested binding of the thrombin protein to the toggle-25t RNA-aptamer composed of 2-hydroxyl group nucleotides. The results also validated this EMSA gel system setup, which includes buffers and salt concentrations used, for studying this system of interaction partners.

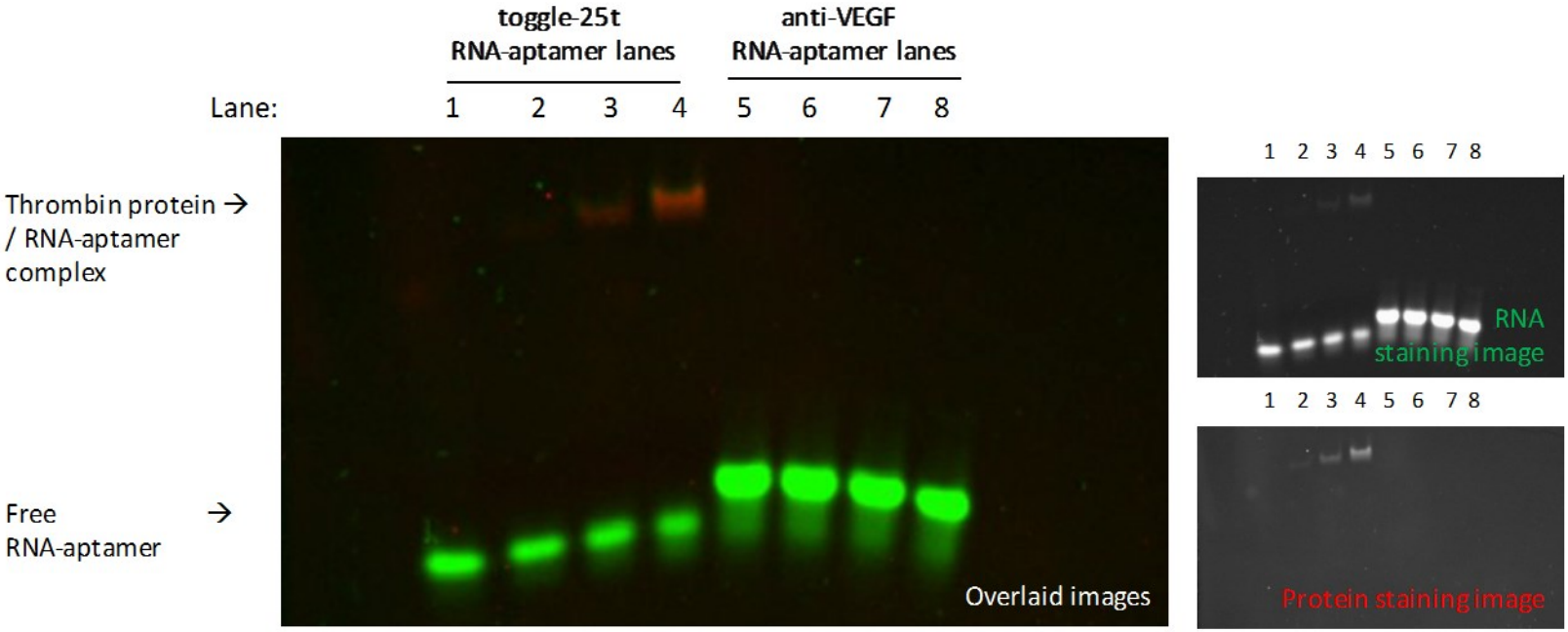


Figure 4.28: EMSA of toggle-25t RNA-aptamer. EMSA of chemically synthesized toggle-25t RNA-aptamer with purified thrombin protein. Binding reactions were set up at protein/RNA molar ratios of 0:1, 0.4:1, 0.7:1 and 1:1, and loaded on gel at lanes 1-4. Negative control reactions included chemically synthesized anti-VEGF RNA-aptamer (gel lanes 5 - 8). Binding reactions were incubated for 10 minutes at 37°C . EMSA gel system used was 8% PAGE/HEPES-Imidazole buffered (pH7.55). Following electrophoresis at 4°C , the gel was first stained with GelRed dye for RNA visualisation and imaged, and then stained with SybroRuby protein dye and imaged. Images were overlaid using ImageJ software.

EMSA of toggle switch devices

Previous experiments validated EMSA for studying the binding of thrombin protein to the toggle-25t RNA-aptamer. Next, an EMSA experiment was designed to test whether the toggle-25t RNA-aptamer retained its ability to bind thrombin when incorporated into the mRNA expression platform as proposed by the biosensor design.

For this experiment, the toggle-25t RNA-aptamer sequence was incorporated into the pNK401 construct to engineer toggle RNA-switch biosensor devices plasmid constructs. Four different toggle RNA-switch device variants were constructed according to the design specifications. The device variants differed in the stability of the stem helix that forms the toggle-25t RNA-aptamer (Table 4.3). Additionally, toggle switch devices v03-v04 included a nucleotide substitution in the toggle-25t RNA-aptamer sequence that has been shown to have negligible effects to the ability of the RNA-aptamer to recognise its target molecule (Long et al., 2008). Folding analysis showed that this nucleotide substitution favoured formation of the desired RNA-aptamer secondary structure as determined by mFold software and that was the reason for its incorporation in the RNA-aptamer sequence. Plasmid constructs of the toggle switch devices v01-04 were verified by DNA sequencing.

An EMSA experiment was carried out to investigate the ability of the designed toggle switch devices to bind thrombin protein, utilising partial sequences of devices as previously described (section 4.2.4). For this experiment, ~150 nucleotide long mRNA of devices was produced by *in vitro* transcription and purified. RNA gel electrophoresis was used to verify the mRNA products size and purity (Figure 4.29).

Then, the partial sequences of toggle switch devices mRNA were incubated in binding reactions in the absence or presence of thrombin protein. The binding reactions were set up at protein:RNA molar ratio = 0.95:1, and reactions were incubated for 30 minutes at 4°C. Following incubation, the samples were loaded in a 5% PAGE / HEPES-Imidazole buffered (pH = 7.55) gel system. To account for non-specific interactions, negative control reactions were set up using control device mRNA (partial sequence).

Table 4.3: Designed toggle RNA-switch devices. Toggle RNA-switch devices incorporating the toggle-25t RNA-aptamer (White et al., 2001) –into the reporter protein mRNA expression platform. Devices are presened with their associated deltaG as predicted for the subsequences presented here. Free energy prediction was carried out using mFold webserver (Zuker, 2003).

Name	Sequence[1,2]	deltaG (kcal/mole)
device v01	...ATG *GGG AAC AAA GCT GAA GTA CTT ACC* Ctt CGT...	-3.0
device v02	...ATG ccc tc*g GGA ACA AAG CTG AAG TAC TTA CCC* GGG CGT...	-10.6
device v03	...ATG ccc tc*G GGA ACA AAG CTG AAG aAC TTA CCC* ggg CGT..	-10.6
device v04	...ATG tcc tc*G GGA ACA AAG CTG AAG aAC TTA CCC* ggg CGT...	-8.2

[1] N: nucleotide common to all devices; n: nucleotides in which devices differ
[2] *N or n*: minimal Toggle-25t RNA-aptamer sequence

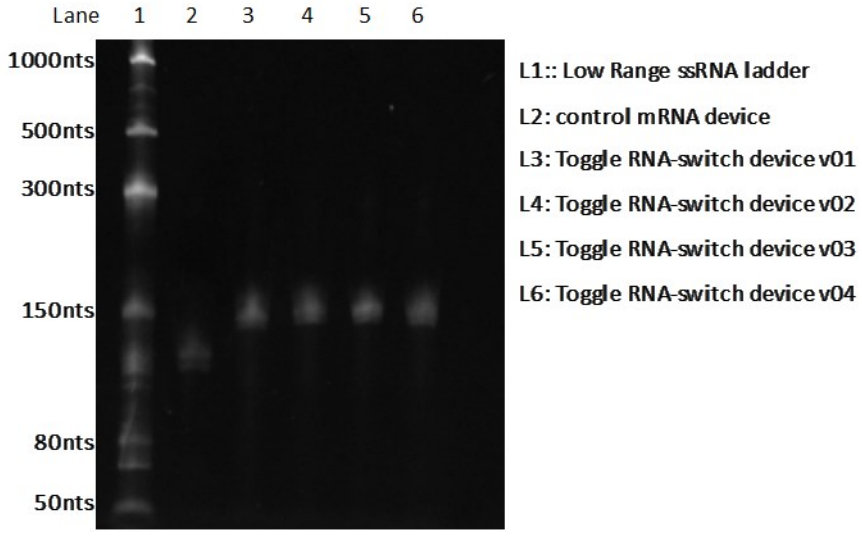


Figure 4.29: RNA electrophoresis of designed toggle switch devices (partial sequences). Image of RNA-electrophoresis gel for the toggle switch devices v01 – v04 (~150nts) partial sequences. Analysis of *in vitro* transcribed mRNA was carried out on a 5% PAGE / TBE-urea denaturing gel system.

EMSA analysis did not show thrombin mediated band shift for the toggle switch devices (Figure 4.30). The image at the top right corner of Figure 4.30 shows the EMSA gel stained for RNA visualisation. All designed toggle switch devices samples produced a single RNA band, both for gel lanes loaded with device mRNA or for gel lanes loaded with device mRNA in presence of thrombin. This was also the case for the negative control device samples. The lack of observed RNA band shift suggested that the designed devices did not bind thrombin. A second, weak RNA band appeared to be trailing the basal band in gel lanes loaded samples of toggle switch devices v02 and v03 with added thrombin protein. The factors that resulted in this pattern could not be determined. Bottom right of Figure 4.30 shows an image of the EMSA gel stained for protein visualisation but there was no protein staining detected or protein bands on the gel. This observation was consistent with the lack of shifted RNA bands when the gel was stained for RNA visualisation.

In conclusion, the results from the EMSA analysis suggested that the toggle switch devices did not bind thrombin protein. A possible explantion for this observation is that the incorporation of the toggle-25t RNA-aptamer into the GFP expression platform hinders the formation of the RNA-aptamer structure that is necessasy for thrombin protein binding.

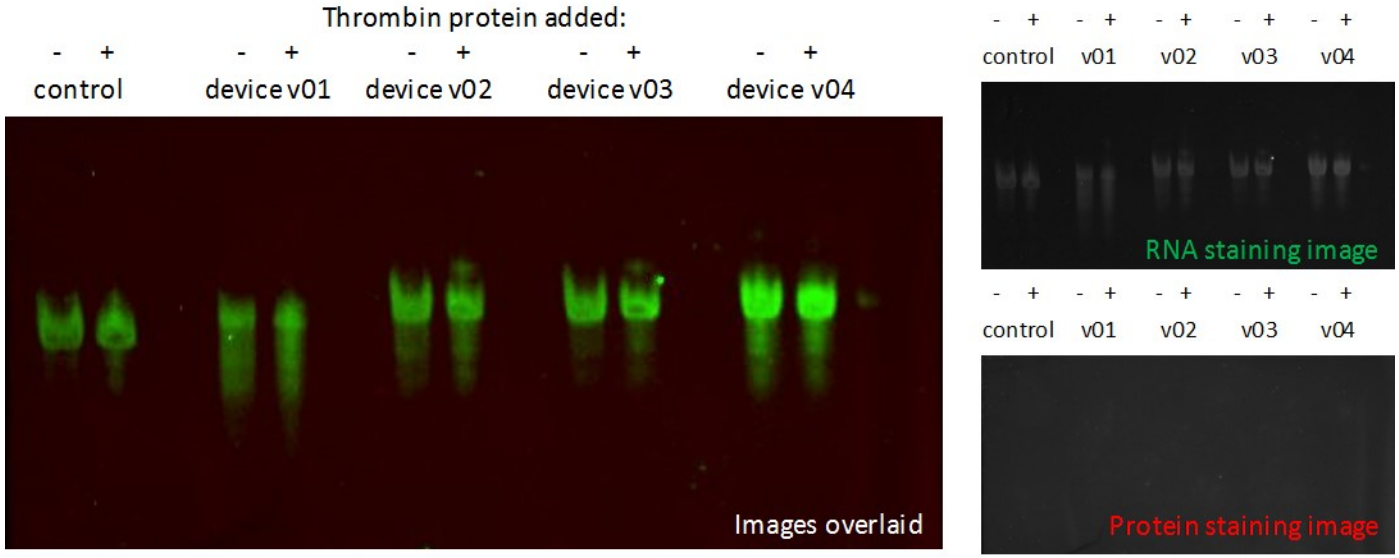


Figure 4.30: EMSA of designed toggle switch devices. EMSA of toggle switch devices v01-v04 with purified thrombin protein. Purified mRNA of toggle switch devices partial sequences (~150 nucleotides) were incubated in binding reactions in the absence (-) or presence (+) of thrombin protein (protein molar ratio/RNA = 0.95:1) and loaded on the gel for EMSA analysis. Binding reactions were carried out for 30 minutes at 4°C. EMSA gel system used was 5% PAGE / HEPES-Imidazole-buffered (pH = 7.55) and the assay was run at 4°C. The gel was first stained with GelRed dye for RNA visualisation and imaged, and then stained with SybroRuby protein dye and imaged. Images were overlaid using ImageJ software.

GFP reporter protein expression of toggle switch devices

In previous experiments, EMSA showed that the VEGF switch devices were able to bind VEGF protein (Figure 4.26), but subsequently a GFP expression assay demonstrated that these devices failed to express GFP in cell-extract reactions (Figure 4.27). The aforementioned results demonstrated some limitations regarding the modularity of the biosensor design. In regards to the toggle switch devices, EMSA showed that the designed devices did not bind their protein target (Figure 4.30). Next, an experiment was designed to investigate GFP protein expression of the designed toggle switch devices. It was speculated that this experiment could provide insides for developing design rules, although the designed toggle switch devices would not be useful for developing a thrombin biosensing assay since they lacked the ability to bind thrombin protein.

For this experiment, toggle switch device v01-v04 full length mRNA was produced by *in vitro* transcription and purified. Then, 7 picomoles of devices mRNA was added to cell-extract reaction mix and incubated at 37°C. Fluorescent signal from the reaction samples was monitored over time. To account for background fluorescence of the cell-extract intrinsic fluorescence, a cell-extract reaction was set up in which no mRNA was added.

Experimental results showed various levels of GFP expression for the designed toggle switch devices. (Figure 4.31). Higher levels of expression were exhibited by toggle switch device v01 that incorporated the toggle-25t RNA-aptamer supported by a stem helix that was the shorter among the stem helices of devices tested. Devices v02 and v03, which only differ by one base that lay within the primary sequence of the RNA-aptamer, exhibited almost identical GFP expression output which also was the lowest expression output between variants. Device v04 GFP expression was higher than that of v02 and v03, but significantly lower than that of v01 device.

In conclusion, the cell-extract GFP expression results suggested that the RNA-aptamer stem helix length and/or secondary structure stability could possibly be important factors to consider for the design of biosensor switch devices with high levels of GFP expression.

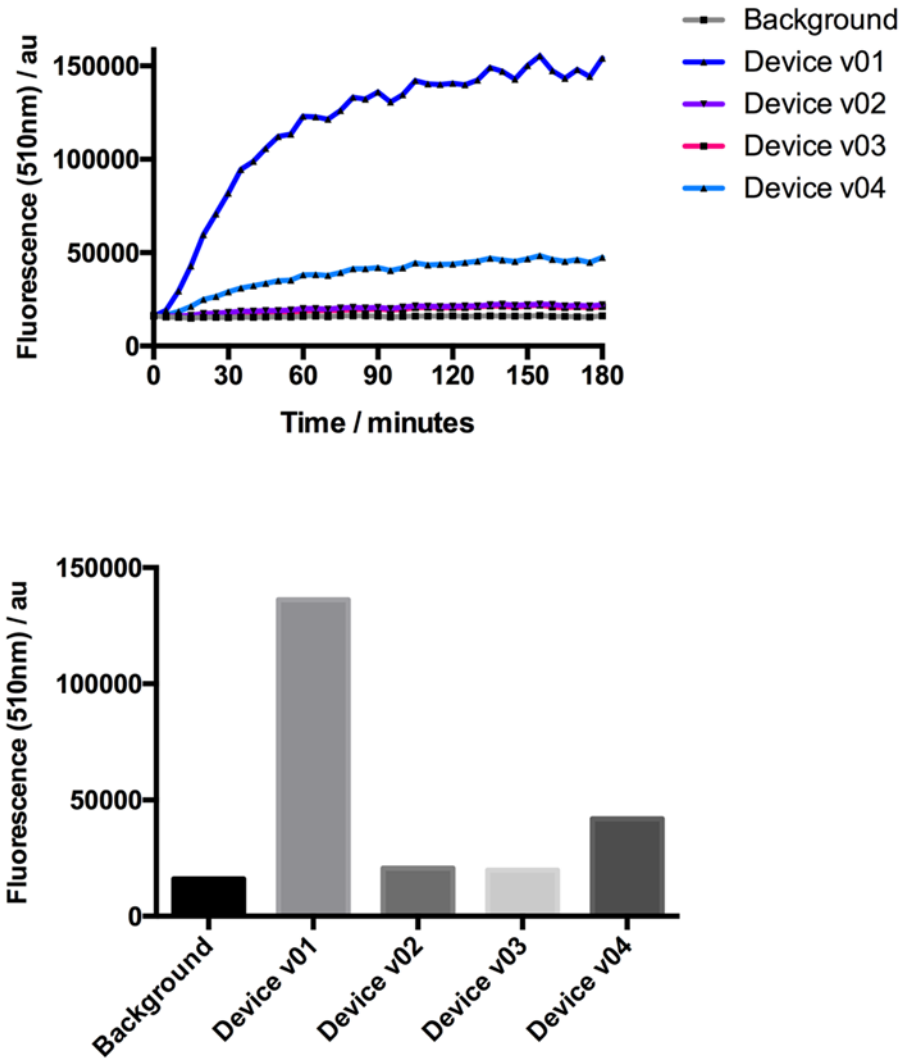


Figure 4.31: GFP expresion of designed toggle RNA-switch devices. (Top) Reaction profile of cell-extract expression reactions of designed toggle switch devices v01-v04. (Bottom) Bar chart of recorded fluorescent signal at 90 minutes for cell-extract reactions. Background samples included cell-extract reaction mix in which template mRNA was not added to determine aut-fluorescent signal from cell-extract mix. Positive control samples included control device mRNA (no RNA-aptamer) which was used as the experimental setup control. Cell-extract reactions were carried out at 37°C with purified mRNA. Reported values represent one experimental repeat. Background signal is the fluorescence signal of cell-extract.

Summary of experimental results for toggle-25t sensor element / thrombin target protein pair

In summary, the toggle-25t RNA-aptamer was utilised to investigate the modularity of the biosensor design for incorporating alternative sensor elements. Initially, an EMSA was carried out that demonstrated the interaction between the toggle-25t RNA-aptamer and thrombin protein. Then, four toggle switch devices were constructed that incorporated the toggle-25t RNA-aptamer into the GFP expression mRNA platform. Partial sequences of the constructed toggle switch devices were tested with EMSA and were found not be able to bind thrombin. It was speculated that this was due to context effects that adversely affect the formation of the RNA-aptamer structure when this was incorporated into the mRNA expression platform. This sets further limitations to the modularity of the RNA-switch biosensor design. On the other hand, GFP expression assays in cell-extract reactions showed that the designed toggle switch devices expressed GFP protein at various expression levels.

Summary of results for biosensor design modularity section

This set of experiments investigated the modularity of the RNA-switch biosensor design by the incorporation of two different RNA-aptamer sensor elements. Anti-VEGF RNA-aptamer was incorporated into the GFP expression mRNA platform to produce three different VEGF switch devices. Partial sequence models of these devices were shown to be able to bind VEGF protein through the incorporated RNA-aptamer. However, the designed devices were not able to express GFP in cell-extract reactions. This suggested that the incorporation of the anti-VEGF adversely affected the expression capabilities of the mRNA pltform. In contrast, toggle switch devices constructed by the incorporation of toggle-25t RNA-apttamer in the mRNA platform were able to express GFP in cell-extract reaction. However, these partial sequence models of these devices were not able to bind the target thrombin protein molecule. This suggested that the mRNA platform affected the ability of the incorporated RNA-aptamer to recognise its target. Thus, the biosensor design was founded to be of limited modularity since the constituent parts of the biosensing devices –the sensor and the transducer element- affected each other's functionality when assembled together. This highlighted the context dependency of assembled RNA parts and that design strategies are needed to minimise its effects.

4.2.8 *In silico* folding analysis of RNA-switch devices

Previously, EMSA analysis of the thrombin switch devices and VEGF switch devices (partial sequences) showed that the designed sequences were able to bind thrombin protein and VEGF protein respectively (Figure 4.22 ; Figure 4.26). In contrast, EMSA analysis of the toggle switch devices (partial sequences) showed that these designs failed to bind their taget molecule thrombin (Figure 4.30), although the un-incorporated toggle-25t RNA-aptamer was demostrated by EMSA to be able to bind thrombin protein (Figure 4.28). These results suggested that the incorporation of certain RNA-aptamers into the mRNA expression platform can result in loss of function for the RNA-aptamer.

The structure of the RNA-aptamer is an important determinant of its ability to bind the target molecule as it places the chemical groups its RNA bases at favourable positions in the 3-dimensional space to interact with the protein target molecule as discussed in section 4.1.1. The incorporation of RNA-aptamers into the mRNA expression platform brings the constituent nucleotides of the RNA-aptamer in close proximity to the constituent nucleotides of the mRNA platform. This can result in base pair interactions between the RNA-aptamer bases and the GFP expression platform bases. If the interactions between the RNA-aptamer bases with the bases of the platform are more favourable than the interactions between the RNA-aptamer bases themselves, this can potentially disrupt the formation of the desired RNA-aptamer secondary structure and lead to loss of its function. Thus, an experiment was designed to investigate the secondary structure of the designed switch devices by *in silico* folding analysis to determine whether the incorporated RNA-aptamers loss of function maybe result from a disruption of its secondary structure.

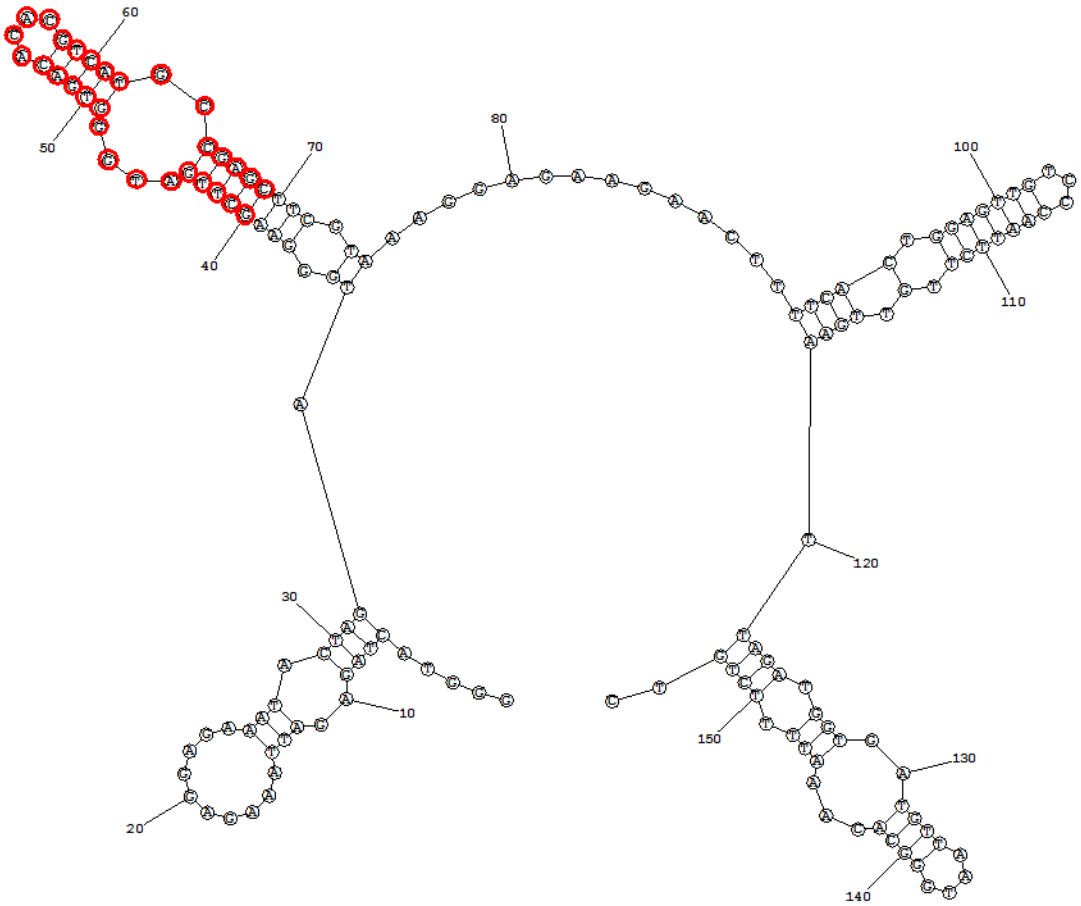
For this experiment, the designed RNA-switch devices were tested *in silico* for successful formation of the RNA-aptamer secondary structure of the RNA-aptamer they incorporate. Folding analysis was carried out using the ~150 nucleotide long partial sequences of the RNA-switch devices as this allowed to compare between *in silico* secondary structure predictions and EMSA experimental results. The RNAstructure software (Mathews, 2004) was the tool of choice for this experiment as its software package alowed computing prediction faster than webserver based tools. The Fold algorithm of the RNAstructure software was used to fold the RNA-switch

devices sequences to their lowest free energy conformation. The sequences analysed can be found in Appendix C. Then, the minimum energy structure was assessed for successful formation of the RNA-aptamer secondary structure as these RNA-aptamer structures are presented in Figure 4.2.

Table 4.4 presents a summary of the folding analysis results. Images of the minimum energy structure for each designed RNA-switch devices are available in Appendix C. Regarding the designed VEGF switch devices, *in silico* folding analysis predicted successful formation of the anti-VEGF RNA-aptamer secondary structure for all three devices. Figure 4.32 presents the predicted secondary structure of the VEGF switch device v02 mRNA as a representative example. These results were consistent with the experimental results from EMSA analysis which demostrated that VEGF protein was bound by all three VEGF switch devices (Figure 4.26). For the case of the designed toggle switch devices, folding analysis predicted that the toggle25t RNA-aptamer secondary structure would not form in these devices. Instead, nucleotide bases of the toggle25t RNA-aptamer were predicted to form more favourable interactions with nucleotide bases of the mRNA expression platform. Figure 4.33 presents the predicted secondary structure of the toggle switch device v02 mRNA as a representative example. These results were consistent with the experimental results from the EMSA experimet where none of the designed toggle switch devices was able to bind thrombin protein (Figure 4.30). Regarding the thrombin switch devices, *in silico* predicion showed that the anti-thrombin RNA-aptamer secondary structure would be present in devices v01, v03, v05 and v06. These predictions were in agreement with experimental data from the EMSA experiment (Figure 4.22). However, it was also predicted that the anti-thrombin RNA-aptamer structure is absent from devices v02, v04, v07, v08. These predictions were not in agreement with EMSA derived experimental results as the devices were able to bind thrombin protein. It was speculated that other conformormations of the thrombin switch devices mRNA can exist, that upon folding release free energy at levels very close to that of the minimum free energy (MFE) structure. These alternative conformation could be favoured in the specific buffer conditions in the EMSA system or in the presence of thrombin.

Table 4.4: Folding analysis of designed RNA-switch devices. Table of folding analysis for ~150 nucleotide long partial sequences of RNA-switch devices. Sequnces were folded to their minimun free energy (MFE) secondary structure conformation. The MFE conformation was examined for the presence of the RNA-aptamer secondary structure (column 2) and the associated Gibbs free energy (column 3). The RNA-switch devices that were found not to form the RNA-aptamer secondary structure were also folded with constrain parameters that forced formation of the RNA-aptamer secondary structure. The Gibbs free energy of the constrained MFE conformation was recorded (column 4) and compared to Gibbs free energy of the MFE conformation (columns 5 and 6). Folding analysis and MFE predictions were carried out using the Fold algorithm of the RNAstructure software tool (Mathews, 2004).

Designed RNA-switch devices	RNA-aptamer structure	MFE (kcal/mole)	Constrained MFE (kcal/mole)	Difference (kcal/mole)	Percentage difference (%)
Thrombin switch					
device v01	YES	-24.2			
device v02	NO	-20	-18.1	-1.9	9%
device v03	YES	-21.1			
device v04	NO	-20.3	-18	-2.3	11%
device v05	YES	-22.2			
device v06	YES	-23			
device v07	NO	-23.9	-17.6	-6.3	26%
device v08	NO	-20.4	-18	-2.4	12%
VEGF switch					
device v01	YES	-31.7			
device v02	YES	-37.8			
device v03	YES	-28.6			
Toggle switch					
device v01	NO	-25	-17.7	-7.3	29%
device v02	NO	-30.7	-27.2	-3.5	11%
device v03	NO	-31.4	-27.2	-4.2	13%
device v04	NO	-28.6	-23.6	-5	17%



ENERGY = -31.7 device v01

Figure 4.32: Minimum free energy structure of VEGF switch device v02. Image of the predicted minimum free energy structure prediction for VEGF switch device variants v02. Folding analysis used the 150 nucleotides long partial sequence of the device. The anti-VEGF RNA-aptamer primary sequence nucleotides are highlighted red. MFE prediction was carried out using the RNAstructure software tool and its Fold algorithm (Mathews, 2004).

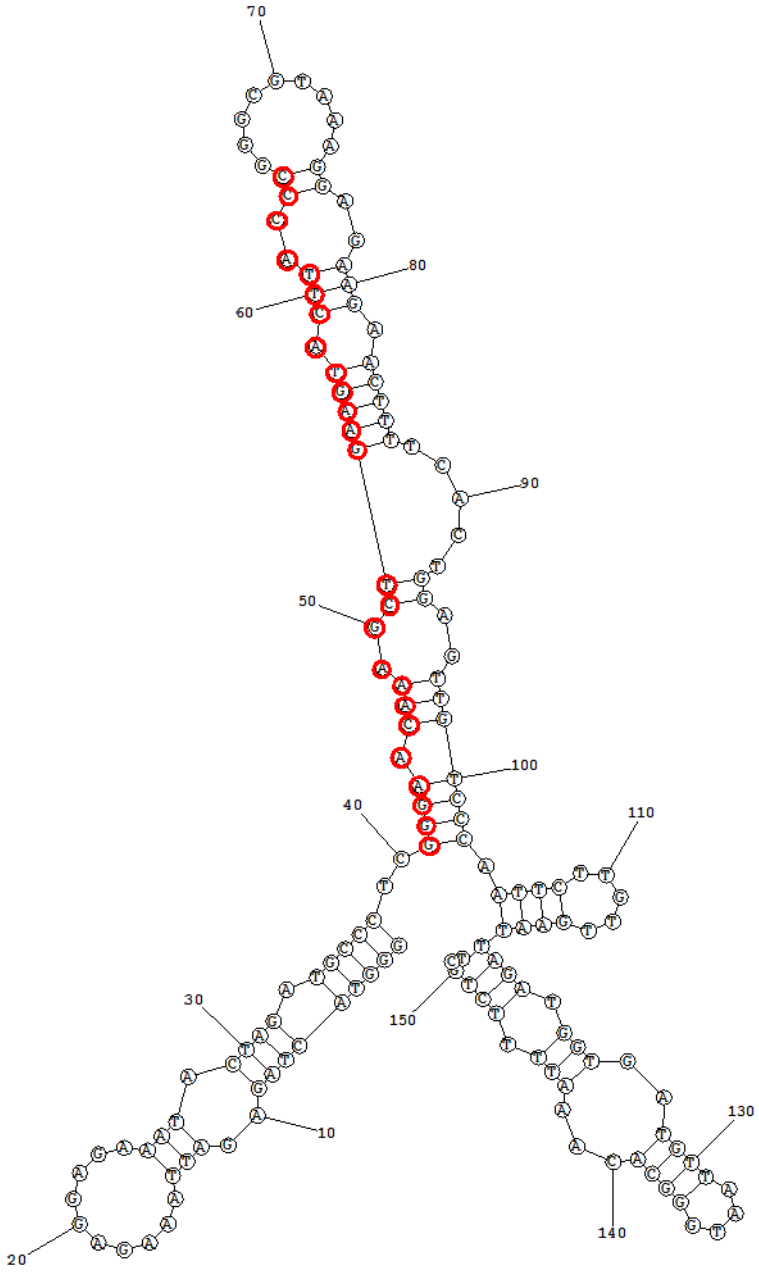


Figure 4.33: Minimum free energy structure of toggle switch device v02. Image of the predicted minimum free energy structure prediction for toggle switch device variants v02. Folding analysis used the 150 nucleotides long partial sequence of the device. The toggle-25t RNA-aptamer primary sequence nucleotides are highlighted red. MFE prediction was carried out using the RNAstructure software tool and its Fold algorithm (Mathews, 2004)

Further analysis was carried out for the RNA-switch devices that their predicted MFE conformations did not form the RNA-aptamer secondary structure. This analysis aimed to investigate the energy differences between the MFE conformation of a device and the device conformation that would facilate binding to the target protein through the formation of the RNA-aptamer secondary structure. Small differences between the two thermodynamic states could help explain the discrepancies between EMSA experimental data and predicted conformations for the thrombin switch devices in the previous experiment.

In this experiment, folding analysis was carried out as previously described but with the addition of constrain parameters that forced the formation of the RNA-secondary structure in the MFE conformation according to secondary structures in Figure 4.2. Then, the Gibbs free energy of the constrained MFE conformation was compared to that of the unconstrained MFE to determine the energy differerence between the two energy states.

The Gibbs free energy differerence between the MFE conformation and the constrained MFE conformation for each RNA-switch device are presented in Table 4.4. For thrombin switch devices, the energy differences for devices v02, v04 and v08 spanned the range of 1.9-2.4 kcal/mole. The percentage difference of the constrained MFE conformation of each device was less than 15% different from the MFE conformation. For energy differences of ~10% from the MFE conformation of the mRNA sequence, sub-optimal foldings can be deemed to be biochemically significant due to large uncertainties in the energy data and incompleteness of the models used in free energy minimisation calculations (Mathews, 2006; Zuker, 1989). Thus, the constrained MFE conformations of these devices were deemed to be biochemically relevant which resolved the discrepancy between experimentally derived EMSA data and prediction.

Nevertheless, the prediction for thrombin switch device v07 showed a large energy difference of 6.3 kcal/mole or 26% difference between the MFE conformation and the constrained MFE conformation. This energy increment was deemed to be too large to be attributed to the uncertainty of folding prediction algorithms and the discrepancy with experimental results remained. A possible explanation that would resolve the discrepancy can involve state transition from the MFE conformation to the constrained MFE conformation that is facilitated by the energetics of thrombin binding to the incorporated anti-thrombin RNA-

aptamer of the thrombin switch device v07, and this is discussed more extensively in the technical discussion section of this chapter.

For the case of toggle switch devices, energy difference between the MFE conformation and constrained MFE conformation was found to be within the uncertainty range for devices v02, v03 and v04, while the energy difference of 29% for device v01 was found to be outside the uncertainty range.

Summary

In summary, software tools that make use of free energy minimisation were used to predict the MFE conformation of the RNA-switch devices partial sequences (~150 nucleotides long models). The MFE conformations were then assessed for the presence of the secondary structure of the RNA-aptamer incorporated into the mRNA expression platform. All VEGF switch devices were predicted to facilitate the formation of the anti-VEGF RNA-aptamer structure. In all toggle switch devices, the nucleotide bases of the incorporated RNA-aptamer were predicted to form base pair interactions with bases of the mRNA expression platform instead of the toggle-25t RNA-aptamer secondary structure. For the thrombin switch devices, four devices facilitated the formation of the anti-thrombin RNA-aptamer secondary structure, and in the other four the RNA-aptamer nucleotides formed interactions with bases of the mRNA platform. Of the four latter devices, three devices were predicted to form sub-optimal conformations that included the anti-thrombin RNA-aptamer secondary structure. These conformations were deemed to be biochemically relevant as their associated Gibbs free energy fell within the acceptable uncertainty margin with respect to the MFE conformation energy state.

4.3 Technical Discussion

This chapter describes the development of a novel biosensor design for the detection of protein biomarkers with intended use in medical diagnostic applications. The design leverages on the high specificity and affinity of RNA-aptamers for their target molecule to act as the sensor elements of the biosensor. The RNA-aptamer is incorporated into a GFP encoding mRNA expression platform which acts as the transducer module of the device and reports target binding by the sensor element. The information of analyte binding to the sensor element is transduced through a mechanism of direct competition between the protein target and the ribosome molecules for binding the translation initiation region element of the mRNA molecule. The biosensor design also makes use of the of *E. coli* cell-extract technology that removes barriers and allows direct access of exogenously added protein analyte to the cellular molecular machinery. Additionally, the cell-extract expression system acts as the GFP reporter protein production factory.

Construction of designed RNA-switch devices for biosensor prototyping could easily be achieved by the incorporation of the RNA-aptamer primary sequence into the DNA plasmid encoded GFP expression device. The prototype thrombin switch device made use of the anti-thrombin RNA-aptamer (Section 4.2.2) of which primary sequence was retrieved from the literature. The RNA-aptamer sequence is 33 nucleotides long and its incorporation in the GFP device coding sequence was easily achieved by a single round of PCR followed by standard cloning procedures. This short turnaround cycle of prototype device construction should be applicable for the construction of RNA-switch biosensor devices for most target biomarkers as most RNA-aptamers reported in the literature have primary sequence 15-45 nucleotides long. The RNA-aptamer incorporation into the GFP expression device resulted in the DNA plasmid construct pNK402 genetically encoding the prototype thrombin switch device (Figure 4.7).

The successful construction of the plasmid system encoding the prototype thrombin switch device was followed by the production of the physical form of the biosensing device which is that of purified mRNA template. The mRNA template was produced by *in vitro* transcription using the pNK402 as the DNA template, and was followed by mRNA product

purification. The isolated thrombin switch device mRNA was analysed by RNA electrophoresis and sample migrated at the expected size as determined through the use of an RNA protein ladder standard (Figure 4.8). The sample appeared to contain two extra bands (Figure 4.8) which varied in intensity relative to the main product signal band between different production batches of the prototype device mRNA (data not shown). Possible origin of these signal bands could be degradation products of the main product due to RNase contamination, prematurely terminated or incomplete transcripts form the *in vitro* transcription reaction due to sequence specific properties or depletion of raw materials (eg. dNTPs) respectively, or RNA gel artefacts due to gel well overloading.

The prototype thrombin switch device was then tested for its functionality as a biosensing device as proposed by the biosensor design. Cell-extract reactions of the prototype thrombin switch device expressed lower levels of GFP reporter protein in the presence of thrombin protein as compared to reactions with no added thrombin (Figure 4.9). These results suggested that thrombin protein binding to the anti-thrombin RNA-aptamer of the device inhibited GFP reporter protein expression as proposed by the biosensor design. In addition, cell-extract reactions of prototype thrombin switch device in the presence of VEGF protein did not cause any decrease in the GFP expression levels (Figure 4.10). These results suggested that the biosensor device was specific for its thrombin protein target. Finally, cell-extract reactions of the prototype thrombin switch device with various concentrations of thrombin protein showed that the GFP expression repression was inversely proportional to the added thrombin protein concentrations (Figure 4.11). The relationship between the two variables was found to be linear over the range of thrombin protein concentrations tested. In conclusion, the experimental results demonstrated the potential of the prototype thrombin switch device as a biosensing assay for thrombin protein detection.

The design and set up of experiments that investigated the functionality of the thrombin switch device in cell-extract reactions posed some technical challenges. The cell-extract assays described in the chapter were carried out with the use of thrombin switch device in the form of purified mRNA template, despite initial attempts for testing the biosensor design were carried out with plasmid DNA template. The reasons for the latter choice was the stability of plasmid

DNA as an information template medium that allowed its easier implementation in the lab environment and also the procedures for its preparation were more economical. Overall, the cell-extract reaction set up was more practical and straightforward when plasmid DNA was used as the biosensor template medium in cell-extract reactions. Nevertheless, the use of plasmid DNA complicated downstream procedures in the biosensing assay implementation, including data analysis. Cell-extract reactions with plasmid DNA template transcribed unknown amounts of biosensor mRNA template molecules, and thus it was hard to determine the amount of added analyte needed to repress GFP expression. An attempt was made to tackle this technical difficulty by fusing the mRNA template to a genetically encoded fluorescence tag (Pothoulakis, Ceroni, Reeve, & Ellis, 2014), but its implementation in the cell-extract biosensing assay proved to be very demanding. Another complication of the use of DNA template was that the analyte mediated repression of GFP expression was very transient as new mRNA template molecules were constantly produced.

In contrast, the use of purified mRNA allowed for known amounts of biosensor device molecules to be added to the cell-extract reactions. However, GFP expression levels from the mRNA template were low and large amounts of mRNA template were required for fluorescent signal detection. In addition, the cell-extraction reaction setup when using the *E. coli* S30 Extract System from Promega Corporation allowed only for 18% volume/volume sample addition to the cell-extract reaction that necessitated production of mRNA template at high concentrations. Attempts to increase sample addition volume to more than 25% resulted in the inactivation of the cell-extract reactions. For repression assays, available sample volume in cell-extract reaction mix had to be divided between mRNA template and thrombin protein components. This necessitated the use of high concentrations of thrombin protein samples for assay setup. An adverse effect of the use of thrombin protein stock solutions of high concentrations was that pipetting inaccuracies were large due to the viscosity of protein samples. The aforementioned technical difficulties of the cell-extract expression and repression assays were tackled by the production of high concentration mRNA samples of thrombin switch device, careful consideration of reactants amounts and numerous experimental repeats.

Experimental results from cell-extract assays demonstrated the functionality of the prototype device as a biosensing device for thrombin detection. These experiments offered indirect evidence of thrombin binding to the anti-thrombin RNA-aptamer sensor element. A series of EMSA experiments were designed that allowed the direct observation of thrombin protein binding to the biosensor device, and thus further validated the device functionality as proposed for the biosensor design. Initially, an EMSA experiment was carried out with anti-thrombin RNA-aptamer and thrombin protein. EMSA analysis showed specific interaction between the two species and validated the EMSA setup for studying this system of interaction partners (Figure 4.12). The gel system used in this assay made use of HEPES/Imidazole buffer (pH 7.55). EMSA experiments carried out for this work that made use of the commonly used TBE-buffered (pH8.3) gel system (Hellman & Fried, 2007) showed no interaction between the anti-thrombin RNA-aptamer and thrombin protein (data not shown). This observation highlighted the importance of factors such as the pH and salts composition/concentration in maintaining a stable interaction between the RNA-aptamer and its target protein. As a consequence, caution should be taken in extrapolating experimental results from EMSA binding assays for interacting partners to biosensing assays in cell-extract systems that make use of a different buffer system to maintain desired pH.

Following the demonstration of the EMSA setup for studying the interaction between thrombin and anti-thrombin RNA-aptamer, EMSA experiments were carried out with the full length mRNA template of the prototype thrombin switch device and partial sequence models of the device. EMSA analysis of the whole prototype device with thrombin produced inconclusive results as the protein was shown to interact non-specifically to the mRNA molecules (Figure 4.13). In future experiments the assay could be optimised to tackle this limitation by the introduction of random RNA to reduce non-specific binding to the prototype device mRNA as demonstrated by others (Hellman & Fried, 2007). EMSA analysis of ~150 & ~250 nucleotides long partial sequences of the prototype device showed thrombin binding to the device mRNA sequences (Figure 4.14; Figure 4.15). These results provided further evidence of the functionality of prototype thrombin switch device according to the proposed model suggested by the biosensor design. The partial sequences served as model systems for the full-length

prototype thrombin switch device mRNA but further experiments could provide more insights for their suitability and limitations as model systems.

The successful prototyping of a thrombin switch device validated by EMSA analysis and cell-extract repression assays allowed for further investigations on the biosensor design. To reduce the cost of further experimentation with cell-extract assays, a coupled transcription/translation (TX-TL) cell-extract system derived from *E. coli* BL21(DE3) cell strain was prepared and optimised for high levels of protein expression (Figure 4.16; Figure 4.17). Experiments utilising this homemade cell-extract system were more cost effective than the use of the commercially available cell-extract system used in the proof of concept experiments. In addition, a thrombin biosensing assay utilising the prototype thrombin switch device demonstrated that the functionality of the device was retained in the newly prepared cell-extract system (Figure 4.18). In addition, the TX-TL cell-extract system allowed for 25% volume/volume sample addition to the cell-extract reactions which could be extended to up to 40% without inactivating the reaction.

Next, experiments attempted to investigate strategies for modulating characteristics of the developed thrombin biosensing assay with a focus on increasing output signal from devices. To this end, a library approach was used to identify device variants with a less thermodynamically stable structure at the lower stem helix of the RNA-aptamer, as this was thought to allow the ribosome easier access to the translation initiation region. The selected device variants exhibited higher levels of GFP expression in cell-extract expression assays in comparison to the prototype thrombin switch device (Figure 4.19), while at the same time they retained their ability for thrombin binding as demonstrated by EMSA analysis (Figure 4.22). One of the device variants, thrombin switch device v04, was demonstrated to be a functional thrombin biosensing device in cell-extract reactions (Figure 4.23). This demonstration hinted that the small library of thrombin switch devices includes more functional thrombin sensing devices. These variants may exhibit diverse analytical characteristics for biosensing assay development, but further cell-extract repression assay experiments are required to validate and thoroughly characterise the variants library.

An alternative approach to the construction of devices that exhibit improved signal output could have been the use of a different reporter protein for the RNA-switch design. For example, the use of enzymatic reporter systems, such as luciferase, can provide an amplification step to output signal generation or can allow for better signal to noise ratio with lower background signal emitted from the cell-extract mix. Nevertheless, the use of enzymatic reporter systems often requires the use of exogenously added substrate as opposed to the auto-fluorescent GFP reporter (Close, Ripp, & Sayler, 2009). In addition, the use of the anti-thrombin RNA-aptamer in a very different RNA context could have severely affected the sensor functionality. Another approach for achieving increased output signal from the thrombin switch device would have been the use of a cell-extract system with improved capacity for protein expression (Caschera & Noireaux, 2014). This approach would be less likely to impact the device functionality since no changes would be required to device design.

In search of a model that would allow the *in silico* design of thrombin devices for desired levels of output signal, the free energy of sub-sequences of thrombin switch devices v01-v08 (Table 4.1) were used to predict GFP expression of biosensing devices. The model was found to correlate poorly with experimentally derived data (Figure 4.19). Possible limitations to the predictive ability of the model could be due to the folding of the sub-sequence of the RNA-aptamer's local environment in isolation from the rest of the mRNA expression platform. Alternatively, the RBS calculator software tool was used for the prediction of GFP reporter expression from the thrombin switch devices v01-v08, but this model correlated poorly with experimental data as well (Figure 4.20). Possible limitations to the predictive ability of the RBS calculator model could be the use of a sub-sequence that stretches from the 5'-end of the mRNA sequence to the +35 position (+1 position assigned to the first nucleotide of the start codon) (Salis, 2011). The RNA-aptamer sub-sequences in thrombin switch devices reaches to +36 position on the GFP expression mRNA platform. As a result, the RBS calculator model is in effect "blind" to any nucleotide substitutions in position +36. This effect is obvious when considering thrombin switch devices v01 and v02. The mRNA sequence of the two devices is identical up to nucleotide position +35 (Table 4.1) and differ by a single nucleotide substitution (U→A) in position +36. As a result the RBS calculator predicts an identical TIR for devices v01

and v02 (Figure 4.20). In reality, the single nucleotide substitution in position +36 disrupts the continuous helical stem in the lower part of the RNA-aptamer that greatly enhances the TIR value of thrombin switch device v02 as reflected in the experimental data of GFP expression assays (Figure 4.19). In conclusion, further research could investigate other models that would enable prediction of GFP expression from the RNA-switch devices that would thus enable *in silico* design of desired output signal levels for devices.

Subsequently, the anti-VEGF RNA-aptamer and the toggle-25t RNA-aptamer were used to investigate the modularity of the biosensor design for sensor element incorporation. The anti-VEGF RNA-aptamer was utilised for the engineering of biosensor devices to detect VEGF protein and the toggle-25t RNA-aptamer was used for the engineering of biosensor devices to detect thrombin protein. The ability of these RNA-aptamers to specifically bind their target protein was demonstrated in EMSA experiments (Figure 4.24, Figure 4.28). A high percentage of the RNA-aptamers molecules did not seem to interact with their target protein and formed bright basal RNA bands on the EMSA gels (Figure 4.24, Figure 4.28). It is possible that the interaction between the RNA-aptamers and their target was unstable in the EMSA gel system and future experiments could explore alternative buffer systems that stabilise this interaction. The TBE-buffer system (pH 8.3) was used for EMSA experiments that utilised the anti-VEGF RNA-aptamer because a higher percentage of the RNA-aptamer molecules bound the target protein during the assay optimisation phase with the particular buffer system. This improved performance could be related to the isoelectric point of the VEGF protein at pH 8.25 value (Kaur & Yung, 2012). Interestingly, the isoelectric point of thrombin protein lies within the pH range of 6.35–7.6 (refer to product information) and the toggle-25t RNA-aptamer performed better in HEPES-Imidazole buffered (pH7.55) EMSA gel system.

Three different VEGF switch devices were engineered to investigate the modularity of the biosensor design (Table 4.2). EMSA analysis of partial sequences of the engineered devices showed a VEGF protein induced RNA band shift (Figure 4.26). This result suggested that the incorporation of the anti-VEGF RNA-aptamer into the GFP expressing mRNA platform did not interfere with the RNA-aptamer ability to specifically bind to VEGF protein. Subsequently, cell-extract expression assay with the VEGF switch devices showed that the engineered devices did

not express GFP protein in cell-extract reactions (Figure 4.27). This result suggested that the incorporation of the anti-VEGF RNA-aptamer hindered the ability of mRNA platform for GFP expression. Since the sensory module of the VEGF switch devices interfered with the functionality of the transducer module of the device, the biosensor design was determined not to be modular for anti-VEGF RNA-aptamer sensor element incorporation. It is believed that the limited modularity of the design is due to the strong secondary structure of the anti-VEGF RNA-aptamer that inhibits ribosome entry to the translation initiation region of the devices. It is possible that the use of an alternative version of the anti-VEGF RNA-aptamer that would include nucleotide substitutions that weaken its secondary structure stability without abolishing its binding functionality towards VEGF protein could produce a functional biosensing device.

In addition, four toggle switch devices were engineered to further investigate the modularity of the biosensor design (Table 4.3). In contrast to VEGF switch devices, the engineered toggle switch devices expressed GFP reporter in cell-extract reactions (Figure 4.31). This result suggested that the incorporation of the toggle-25t RNA-aptamer did not interfere with the ability of the mRNA platform to express GFP. However, EMSA analysis of partial sequences of the engineered devices did not show a thrombin protein induced RNA band shift (Figure 4.30), suggesting that the platform hindered the ability of the RNA-aptamer to bind its target molecule. Since the sensory element of the device lost its ability to carry out its function in the context of the mRNA platform, the biosensor design was determined not to be modular for the toggle-25t RNA-aptamer sensor element incorporation. Future designs of toggle switch devices could investigate structural RNA motifs that would strengthen the stability of the toggle-25t RNA aptamer through rational design or a library approach to produce functional biosensor devices.

To investigate the secondary structure of the engineered RNA-switch devices, the RNAstructure software tool was used. Folding analysis results showed that *in silico* predictions of the secondary structure RNA-switch devices (Table 4.4) agreed to a large extent to experimental data from EMSA experiments (Figure 4.22; Figure 4.26; Figure 4.30). This suggests

that the use of secondary structure prediction algorithms could be used in future experiments to aid the design of functional biosensing devices.

A small subset of the secondary structure predictions though did not agree with experimentally derived data and further research should be carried out to understand the limitations. One possibility is that the disagreement between predictions and EMSA experiments could arise from limitations of secondary structure algorithms in their prediction of the minimum energy structure (Mathews, 2006) but extra information can be obtained by examining sub-optimal structures of low energy (Table 4.4). Alternatively, the disagreement could arise due to differences in the environment in which the RNA molecule is folded (buffer system, salt concentrations, pH value, temperature etc) in comparison to models of secondary prediction algorithms. Lastly, it is possible that the presence of the target molecule of the RNA-switch devices impacts their secondary structure configuration. This parameter is used by the thermodynamic model of the Riboswitch Calculator software tool to predict state transitions of the RNA molecule upon ligand binding (Espah Borujeni, Mishler, Wang, Huso, & Salis, 2016). The use of the reported dissociation constant of the anti-thrombin RNA-aptamer for its target of 164nM (X. M. Liu et al., 2003) allows the calculation of the energy released upon binding of 9.6 kcal/mol (Espah Borujeni et al., 2016). This value is sufficiently large to cover the energy difference between the MFE configuration and the constrained MFE configuration (Table 4.4) for all thrombin switch devices that were predicted to not form the RNA-aptamer secondary structure. Thus, energy supplied by thrombin protein binding to the devices mRNA molecules could enable thermodynamically favourable state transition and resolve discrepancy between predicted MFE configurations and EMSA experimental results.

The use of *in silico* folding analysis tools for the prediction of RNA switch devices secondary structures is desirable as it avoids lengthy and costly laboratory procedures to experimentally test and validate designed devices. Given that secondary structure prediction strongly correlated with experimental data from EMSA experiments, this would allow for the design and testing an RNA-switch device *in silico* for thermodynamically favourable formation of the RNA-aptamer. Alternatively, the structure of the RNA-switch devices can be determining experimentally by using chemical probing (Kubota, Tran, & Spitale, 2015). For example, the

SHAPE method (Selective Hydroxyl Acylation analysed by Primer Extension) by the use of chemical modification of RNA bases is able to experimentally determine the structure of RNA molecules at single nucleotide resolution (Deigan, Li, Mathews, & Weeks, 2009; Wilkinson, Merino, & Weeks, 2006), but such methods are significant more time consuming and costly than *in silico* predictions.

Chapter 5

5 Conclusions

5.1 Discussion

This document describes the development of two biosensor platform technologies for the detection of protein analytes that were designed from a synthetic biology perspective. The biosensor platform technologies incorporate design specifications that could potentially make them suitable for the development of medical diagnostic devices. Both technologies were demonstrated to incorporate sensor elements of high affinity and high specificity for their intended target analyte. These sensor elements belong to classes of biomolecules for which there are established protocols for the development of sensors at will. Additionally, the design of both technologies addresses a limitation of synthetic biology biosensor devices for protein targets analysis and enables analyte transport to the sensor element of devices. Finally, both technologies were genetically encoded and their designs incorporated a strong element of modularity. Both of these characteristics are thought to be desirable for the development of low cost diagnostic tests as a result of lower production costs and more efficient prototyping cycles.

Biosensor design: Nanobody & whole cell chassis

Chapter 3 describes the design and prototyping of a biosensing platform technology in which nanobody sensor elements are surfaced displayed on *E. coli* cell surface (Figure 3.4). Experimental results showed that the *E. coli* cell can serve as a self-manufacturing platform of biosensor devices of this design (Figure 3.8), and that this platform can be used to specifically detect the protein analyte of interest in a plate agglutination assay format (Figure 3.13). These experiments provided proof of concept data for this novel biosensor technology.

The use of a nanobody sensor element is a prominent feature of this bacterial biosensor, especially because it incorporates this sensing element without the need for any purification, chemical treatment or immobilisation procedure as required for other biosensor technologies

(Sharma, Byrne, & O'Kennedy, 2016). Nanobodies and other members of the antibody family of biomolecules can exhibit very high specificity for the target, and for this reason they are frequently employed in analytical biosensors (Sharma et al., 2016) and especially for medical diagnostic applications (Mascini & Tombelli, 2008) where specificity is a necessary characteristic for reliable diagnostic testing.

Within the field of synthetic biology, there has been a great interest in developing synthetic biology design whole cell biosensors. The technology developed here allows for the detection of protein analytes that is a class of biomolecules for which technology tools were lacking. For example, synthetic biology biosensors have been constructed for the detection of small molecules analytes that are able to cross the bacterial cell membrane barrier through various mechanisms of transport. Such small molecule analytes include: heavy metal contaminants like arsenic (Aleksic et al., 2007), organic contaminants like benzene and toluene (Willardson et al., 1998), and drug molecules like theophylline (Lynch et al., 2007) amongst others. Also, synthetic biology biosensors have been constructed for detection of analytes at the bacterial cell surface such as: for metabolites like ribose (Reimer et al., 2014) and parasitic protease activity (Webb et al., 2016). The whole cell biosensor technology developed here is unique in detecting protein analyte at the bacterial cell surface in accordance to the extracellular detection paradigm of the latter examples. In accordance to the rest of synthetic biology design whole cell biosensors, the information of the biosensor design is genetically encoded and detection relies on recombinantly expressed genes in bacterial chassis. These features aim to take advantage of the bacterial cell as a self-manufacturing platform that its industrial production is thought to be of low cost.

In addition, the biosensor design was investigated from an engineering perspective for approaches that would enable the forward engineering and fast prototyping of further biosensing devices. Experimental results suggested that the use of a promoter library was a more robust tool for the forward engineering of biosensor devices with different dynamic ranges for GFP_2 analyte detection as compared to forward design using RBS design tool (Table 3.2; Table 3.3). The use of both tools, promoter library and RBS design, was aimed and facilitated the differential expression and cell surface display number of nanobody sensor

elements (Figure 3.14), and thus modulated the biosensor device analytical dynamic range (cross ref figure 3.11). The use of stronger/weaker promoter elements from the library resulted in biosensor devices that detected higher/lower analyte concentrations as predicted, as opposed with the same strategy with the use of RBS elements of varying strength. These results are in line with the current state of synthetic biology technology, where rules and techniques for the reliable use of promoter elements to fine tune gene expression are well established (Canton, Labno, & Endy, 2008; Davis, Rubin, & Sauer, 2011; Kelly et al., 2009), while similar rules for RBS design are in the refinement stage (Lou, Stanton, Chen, Munsky, & Voigt, 2012; Salis et al., 2009).

Furthermore, the modular structure of the biosensor platform display vector (Salema et al., 2013) allows the straightforward exchange of the sensor element to another with a different target specificity. Experimental results showed that the biosensor design could incorporate nanobodies specific for fibrinogen protein recognition and detect the presence of this analyte molecule in the plate agglutination assay format developed (Figure 3.15). The exchange of the nanobody sensor for a different sensor can be achieved by widely accessible molecular cloning technology within a working week timeframe. Given that protocols and procedures exist for the production and isolation of nanobodies to recognise any target analyte desirable (Pardon et al., 2014; Siontorou, 2013), the biosensor platform developed here could be potentially used to develop any diagnostic test desirable for the detection of disease markers, food and environmental contaminants, biological warfare agents and illicit drugs amongst others (Sharma et al., 2016). Moreover, the *in silico* model of the agglutination reaction (Section 2.2.8) could be used to inform the design of future agglutination tests to refine device's characteristics to suit particular applications without the need of extensive wet lab testing. Overall, the synthetic biology design of this biosensor platform -that includes all the aforementioned characteristics and approaches discussed- should enable shorter prototyping cycles and lower costs of development of future diagnostic tests that are based on the prototyped nanobody display whole cell biosensor design.

Biosensor design: RNA-aptamer & E. coli cell extract

Chapter 4 describes the design and prototyping of a second biosensing platform for the detection of protein analytes. In short, the design features a RNA-aptamer sensor element incorporated into a mRNA molecule that encodes a reporter protein. The translation of the reporter protein is allowed/disallowed depending on the absence/presence of the targeted analyte respectively. Experimental results demonstrated a functional prototype device of this design that utilised an anti-thrombin sensor element regulating the translation of GFP reporter protein (Figure 4.5). Importantly, this design was implemented in cell-extract reactions that allows the protein analyte for a direct route to sensing element.

Experimental results showed that the developed prototype device specifically detected thrombin protein analyte in cell-extract repression assays (Figure 4.9), mediated through the specific binding of the analyte molecules by the incorporated respective RNA-aptamer for both thrombin and VEGF protein analytes (Figure 4.14; Figure 4.24). RNA-aptamers can be selected *in vitro* to show very high specificity towards target recognition (Jenison et al., 1994), as their functional analogues antibodies, and thus show much potential as replacement or complimentary sensor elements for diagnostic test development (Jayasena, 1999). As previously discussed, specificity towards target recognition is a necessary characteristic for the development of reliable tests in medical diagnostics. Given the demonstrated ability of this biosensor design to incorporate RNA-aptamers as the sensing elements as well as the high selectivity of RNA-aptamers as evident from the literature, the RNA-switch in *E. coli* cell extract design could potentially constitute a novel biosensing platform for the development of a medical diagnostic tests.

Diagnostic point of care testing is an approach which in some instances is more preferable or appropriate to traditional centralised laboratory diagnostic testing (Yager, Domingo, & Gerdes, 2008). Pregnancy testing (Ehrenkranz, 2002) and measuring blood glucose levels (Turner, 2013) are two of well-known conditions for which point of care testing approach was successfully applied. The development of low-cost diagnostic biosensing devices would be well suited for such applications where the cost per test can be an important consideration (Yager et al., 2008). With cost per test in mind, the biosensing functionality of the developed RNA-switch

technology was demonstrated in a commercially available *E. coli* cell extract system (Figure 4.9) as well as in a different *E. coli* cell extract system that can be efficiently produced cheaply in-house (Figure 4.18). The latter cell extract has been reported to offer a 98% reduction in the cost of production as compared to similar commercially available systems (Sun et al., 2013). Similarly to this research, low cost diagnostic tests were developed utilising cell-extract technology in which mRNA sensors detected the Ebola virus (Pardee et al., 2014) and differentiated between Zika virus strains (Pardee, Green, et al., 2016), and these cell extract assays were incorporated into cheap paper matrix for distribution. Additionally, another study demonstrated that *E. coli* cell-extract systems can be used for *in-situ* production of biologics, including diagnostic reagents (Pardee, Slomovic, et al., 2016).

The biosensing platform technology developed incorporates a strong element of modularity to enable fast prototyping cycles and forward engineering. As described in (Section 4.2.1), and investigated (Section 4.2.7), the platform can potentially incorporate any RNA-aptamer sequence for the detection of the protein analyte of interest. RNA-aptamer sequences can be acquired from the academic literature (Yuce, Ullah, & Budak, 2015; Zimbres, Tarnok, Ulrich, & Wrenger, 2013), commercial databases (Aptagen) or developed using established SELEX protocols (Germer, Leonard, & Zhang, 2013). Then, the RNA-aptamer can be used to design novel RNA-switch devices, and prototype devices can be easily constructed (Section 4.2.2) by widely accessible molecular cloning technology within a working week timeframe. Currently in the synthetic biology field, the use of RNA and RNA-aptamers as modular functional domains for the design of synthetic sensing systems and logic computation systems is a popular frontier (Chappell, Watters, Takahashi, & Lucks, 2015; Saito & Inoue, 2007). Similar to the RNA-switch design presented here, a protein binding RNA-motif has been used to sterically hinder translation in human cells (Saito, Fujita, Kashida, Hayashi, & Inoue, 2011; Saito et al., 2010), while RNA/protein binding has been used for ligand induced intracellular logic through various signal relay mechanisms (Culler, Hoff, & Smolke, 2010; Kennedy et al., 2014; Qi, Lucks, Liu, Mutalik, & Arkin, 2012).

Despite our increasingly deep understanding of the interplay between RNA folding and its function (Chappell et al., 2015), researchers leverage both mutants library screening and

rational approaches aided by CAD tools for the engineering of synthetic RNA devices (Liang, Bloom, & Smolke, 2011). As an example of the former approach, a study used fluorescence-activated cell sorting for genetic library screening to identify ligand induced riboswitches that exhibited high-fold activation (Lynch et al., 2007). On the other hand, it has been demonstrated that riboswitches of desired characteristics can be designed *in silico* through a rational, forward engineering approach using a software tool (Espah Borujeni et al., 2016). Similarly, in this research a mutant library based approach has been demonstrated to allow for the modulation of output signal levels from thrombin switch devices (cross ref section. 3.3.6) that could be potentially useful for tailoring of biosensing assays with suitable characteristics for particular applications. Also, an *in silico* rational design approach was investigated to gain a greater understanding of the design rules for this technology (section 3.3.8) towards the goal of forward engineering biosensing RNA-switch devices by the use of a CAD tool.

5.2 Future Work

The sensitivity (limit of detection), specificity towards the target analyte and dynamic range of detection are the critical technical parameters for the successful application of analytical devices -such as the biosensor devices described in this work- in medical diagnostics (Courbet, Renard, & Molina, 2016). Future work could investigate strategies that would enable designers to achieve desirable characteristics as dictated by the specific application of the analytical device. The limit of detection of a biosensor device will be a function of the affinity of the sensor element toward its target and the ability of the transducer/output modules to relay and amplify the readout signal. Different reporter systems enable different sensitivity for the biosensor devices, and a comparative study between fluorescent protein and bioluminescence found the latter reporter system to enable faster and more sensitive detection (Sagi et al., 2003). Nevertheless, the absolute limit of sensitivity for a device will always be determined by the characteristics of its sensing element and thus the use of “high quality” sensing elements is of vital importance. In terms of modulation of the device dynamic range, other strategies than the ones presented here could be investigated. For example, a study used the architectural

design of the genetic circuit to achieve wide dynamic range of detection as a response to input analyte concentrations by splitting the biosensing circuit over a two plasmid system of different copy number origins of replication (Daniel, Rubens, Sarpeshkar, & Lu, 2013). An alternative strategy suggested is to leverage the product design (container of biosensor) to achieve traffic light sensors for the desired dynamic range of detection (van der Meer & Belkin, 2010).

In developing next-generation diagnostics, an important step is the ability to diagnose medical state through the interrogation of biomarker signatures that include multiple individual biomarkers (Oved et al., 2015). This is in contrast to the current state, especially for point-of-care testing, where diagnosis is based on a single biomarker analyte. For example, at the current state diagnosis of a medical condition would involve the detection of analyte A in a biological sample and determining whether the concentration of analyte A is above or below a critical threshold value. In contrast, diagnosis using biomarker signatures informed from medical knowledge could possibly involve three biomarkers; analyte A, analyte B and analyte C, where positive results could be indicated by logic such as: analyte A or B upregulated only when analyte C is downregulated. Synthetic biology offers a favourable framework through which such multiplexing, information processing diagnostic tests can be developed. Biochemical logic gates of all kinds - AND gate, NOT gate, NOR gate, XOR gate, NAND gate etc- have been demonstrated (Anderson et al., 2007; Tamsir et al., 2011; B. Wang, Kitney, Joly, & Buck, 2011; L. Wang et al., 2015) and integrated together to build complex biological computation (Moon, Lou, Tamsir, Stanton, & Voigt, 2012; Nielsen et al., 2016). Towards this goal, synthetic classifier genetic circuits have been demonstrated that interrogated a disease signature of six miRNA biomarkers to distinguish between HELA and non-HELA cells (Xie et al., 2011). Alternatively, multi-input diagnosis can be achieved separately from biological signal processing and executed by electronics components as demonstrated by printed arrays of whole cell bacterial sensors (Melamed et al., 2011). Such a transition to next generation diagnostics of biomarker signatures would be essential to tackle pressing global medical emergencies such as the rise of antimicrobial resistance and enable cancer prognosis (Manjili, Najarian, & Wang, 2012).

Future work in developing synthetic biology biosensors for medical diagnostics would need to address real world challenges to translate the technology into products that can be

used at the point of care. Proof of concept studies that demonstrate reliable detection of target analyte in purified form should be followed by studies that investigate the functionality of biosensor in the complex environment of biological samples and any possible signal interference. A recent study investigated the use of genetically engineered whole cell biosensors in blood and urine samples (Courbet et al., 2015). They found that although the synthetic gene networks remained functional in the clinical samples, signal digitization and amplification was required to counteract signal dampening effect from the biological samples and also that agents in biological samples were causing signal interferences. Further, another real world challenge that future research should address is the deploying of biosensor devices out of the lab environment. Towards this end, whole cell biosensor devices in *B. subtilis* chassis have been shown to be able to recover from lyophilisation and express biosensing modules from synthetic gene networks demonstrating the utility of the lyophilisation process for out of the lab deployment and cost effective transport (Webb et al., 2016). Additionally, cell-extract based glucose and Ebola virus diagnostic devices have been demonstrated to recover and maintain functionality following freeze-drying procedures onto inexpensive paper discs, and also the freeze-dried tests remained functional following storage for 1 year at room temperature (Pardee et al., 2014). Finally, future research should address the integration of biosensor devices to upstream sample processing procedures that would prepare the biological sample for analysis. In the case of the aforementioned Zika virus diagnostic test, the analyte molecules in the retrieved biological sample had to be amplified before analysis with the biosensor device (Pardee, Green, et al., 2016).

Future research should also address the social context to which the biosensor devices would be applied. Biosensor devices aimed for point of care testing and especially for cases where the end user is not a trained professional -e.g. home use, first responders at emergency sites among others- should require only elementary instructions to use and minimise possible sources of error. This would be facilitated by enabling technologies such as automated microfluidic systems (Gubala, Harris, Ricco, Tan, & Williams, 2012) that would be integrated to developed biosensor technologies. Additionally, product design considerations would be required for the different test settings - e.g. physician's office, resource limited settings etc.-

such as the use of chromoprotein reporter systems for visualisation by naked eye instead of fluorescence and bioluminescence reporter systems (French, Joshi, Elfick, Haseloff, & Ajioka, 2011). The widespread use of mobile phone technology and their connectivity to wireless networks would likely be heavily leveraged for test result quantification (Pardee et al., 2014), for data storage, monitoring and intervention (Mohan, Marin, Sultan, & Deen, 2008) towards the goal of delivering personalised medicine. Finally, as synthetic biology biosensing technology matures and approaches market entry there should be more active research in developing a concrete regulatory framework that guarantees the safe entry and release of synthetic biological devices into the environment. The use of abiotic devices utilising cell-extract technology, such as the one described here, the use of host chassis with a GRAS designation (Webb et al., 2016) and the use of technologies that disallow the unwanted horizontal gene transfer (HGT) of synthetic DNA (Wright, Delmans, Stan, & Ellis, 2015) are only a few examples of the approaches researched to address environmental concerns.

6 References

- Ajikumar, P. K., Xiao, W. H., Tyo, K. E., Wang, Y., Simeon, F., Leonard, E., . . . Stephanopoulos, G. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science, 330*(6000), 70-74. doi:10.1126/science.1191652
- Aleksic, J., Bizzari, F., Cai, Y., Davidson, B., Mora, K. d., Ivakhno, S., . . . Millar, A. (2007). Development of a novel biosensor for the detection of arsenic in drinking water. doi:doi:10.1049/iet-stb:20060002
- Anderson, J. C., Clarke, E. J., Arkin, A. P., & Voigt, C. A. (2006). Environmentally controlled invasion of cancer cells by engineered bacteria. *J Mol Biol, 355*(4), 619-627. doi:10.1016/j.jmb.2005.10.076
- Anderson, J. C., Voigt, C. A., & Arkin, A. P. (2007). Environmental signal integration by a modular AND gate. *Mol Syst Biol, 3*. doi:10.1038/msb4100173
- Annaluru, N., Muller, H., Mitchell, L. A., Ramalingam, S., Stracquadanio, G., Richardson, S. M., . . . Chandrasegaran, S. (2014). Total synthesis of a functional designer eukaryotic chromosome. *Science, 344*(6179), 55-58. doi:10.1126/science.1249252
- Aptagen. Apta Index™. Retrieved August 2016 https://www.aptagen.com/aptamer-index/aptamer-list.aspx
- Aquino-Jarquin, G., & Toscano-Garibay, J. D. (2011). RNA Aptamer Evolution: Two Decades of SELEction. *Int. J. Mol. Sci., 12*(12), 9155-9171. doi:10.3390/ijms12129155
- Arbabi Ghahroudi, M., Desmyter, A., Wyns, L., Hamers, R., & Muyldermans, S. (1997). Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett, 414*(3), 521-526.
- Atsumi, S., Hanai, T., & Liao, J. C. (2008). Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature, 451*(7174), 86-89. doi:10.1038/nature06450
- Atsumi, S., Higashide, W., & Liao, J. C. (2009). Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat Biotechnol, 27*(12), 1177-1180. doi:10.1038/nbt.1586
- Babitzke, P., Baker, C. S., & Romeo, T. (2009). Regulation of translation initiation by RNA binding proteins. *Annu Rev Microbiol, 63*, 27-44. doi:10.1146/annurev.micro.091208.073514
- Baek, J. H., Han, M. J., Lee, S. H., & Lee, S. Y. (2010). Enhanced Display of Lipase on the Escherichia coli Cell Surface, Based on Transcriptome Analysis▽†. In *Appl Environ Microbiol* (Vol. 76, pp. 971-973).
- Barabasi, A. L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat Rev Genet, 5*(2), 101-113. doi:10.1038/nrg1272
- Bassi, A. S., Ding, D. N., Gloor, G. B., & Margaritis, A. (2000). Expression of single chain antibodies (ScFvs) for c-myc oncoprotein in recombinant Escherichia coli membranes by using the ice-nucleation protein of Pseudomonas syringae. *Biotechnol Prog, 16*(4), 557-563. doi:10.1021/bp000053k
- Binder, U., Matschiner, G., Theobald, I., & Skerra, A. (2010). High-throughput sorting of an Anticalin library via EspP-mediated functional display on the Escherichia coli cell surface. *J Mol Biol, 400*(4), 783-802. doi:10.1016/j.jmb.2010.05.049

- Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., . . . Shao, Y. (1997). The complete genome sequence of Escherichia coli K-12. *Science, 277*(5331), 1453-1462.
- Borisov, S. M., & Wolfbeis, O. S. (2008). Optical biosensors. *Chem Rev, 108*(2), 423-461. doi:10.1021/cr068105t
- Bulich, A. (1979). *Use of Luminescent Bacteria for Determining Toxicity in Aquatic Environments.* Paper presented at the Second Annual Symposium on Aquatic Toxicology, Cleveland, Ohio.
- Bănică, F.-G. (2012). What are Chemical Sensors? In *Chemical Sensors and Biosensors* (pp. 1-20): John Wiley & Sons, Ltd.
- Cagatay, T., Turcotte, M., Elowitz, M. B., Garcia-Ojalvo, J., & Suel, G. M. (2009). Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell, 139*(3), 512-522. doi:10.1016/j.cell.2009.07.046
- Campuzano, S., Salema, V., Moreno-Guzman, M., Gamella, M., Yanez-Sedeno, P., Fernandez, L. A., & Pingarron, J. M. (2014). Disposable amperometric magnetoimmunosensors using nanobodies as biorecognition element. Determination of fibrinogen in plasma. *Biosens Bioelectron, 52*, 255-260. doi:10.1016/j.bios.2013.08.055
- Canton, B., Labno, A., & Endy, D. (2008). Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol, 26*(7), 787-793. doi:10.1038/nbt1413
- Caruthers, M. H. (1985). Gene synthesis machines: DNA chemistry and its uses. *Science, 230*(4723), 281-285.
- Caschera, F., & Noireaux, V. (2014). Synthesis of 2.3 mg/ml of protein with an all Escherichia coli cell-free transcription-translation system. *Biochimie, 99*, 162-168. doi:10.1016/j.biochi.2013.11.025
- Cass, A. E., Davis, G., Francis, G. D., Hill, H. A., Aston, W. J., Higgins, I. J., . . . Turner, A. P. (1984). Ferrocene-mediated enzyme electrode for amperometric determination of glucose. *Anal Chem, 56*(4), 667-671.
- Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., & Noller, H. F. (1999). X-ray crystal structures of 70S ribosome functional complexes. *Science, 285*(5436), 2095-2104.
- Celik, N., Webb, C. T., Leyton, D. L., Holt, K. E., Heinz, E., Gorrell, R., . . . Lithgow, T. (2012). A Bioinformatic Strategy for the Detection, Classification and Analysis of Bacterial Autotransporters. *PLOS ONE, 7*(8). doi:10.1371/journal.pone.0043245
- Ceroni, F., Algar, R., Stan, G. B., & Ellis, T. (2015). Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat Methods, 12*(5), 415-418. doi:10.1038/nmeth.3339
- Chappell, J., Jensen, K., & Freemont, P. S. (2013). Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res, 41*(5), 3471-3481. doi:10.1093/nar/gkt052
- Chappell, J., Watters, K. E., Takahashi, M. K., & Lucks, J. B. (2015). A renaissance in RNA synthetic biology: new mechanisms, applications and tools for the future. *Curr Opin Chem Biol, 28*, 47-56. doi:10.1016/j.cbpa.2015.05.018

- Chen, G., Cloud, J., Georgiou, G., & Iverson, B. L. (1996). A quantitative immunoassay utilizing Escherichia coli cells possessing surface-expressed single chain Fv molecules. *Biotechnol Prog, 12*(4), 572-574. doi:10.1021/bp960041s
- Chen, Y., Wei, L. N., & Muller, J. D. (2003). Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. *Proc Natl Acad Sci U S A, 100*(26), 15492-15497. doi:10.1073/pnas.2533045100
- Church, G. M., Elowitz, M. B., Smolke, C. D., Voigt, C. A., & Weiss, R. (2014). Realizing the potential of synthetic biology. *Nat Rev Mol Cell Biol, 15*(4), 289-294. doi:10.1038/nrm3767
- Clackson, T., Hoogenboom, H. R., Griffiths, A. D., & Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature, 352*(6336), 624-628. doi:10.1038/352624a0
- Clark, L. C., Jr., & Lyons, C. (1962). Electrode systems for continuous monitoring in cardiovascular surgery. *Ann N Y Acad Sci, 102*, 29-45.
- Close, D. M., Ripp, S., & Sayler, G. S. (2009). Reporter proteins in whole-cell optical bioreporter detection systems, biosensor integrations, and biosensing applications. *Sensors (Basel), 9*(11), 9147-9174. doi:10.3390/s91109147
- Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol, 13*(6), 343-359. doi:10.1038/nrmicro3456
- Coughlin, R. F., & Driscoll, F. F. (2001). *Operational Amplifiers and Linear Integrated Circuits* (6th Edition ed.). Prentice Hall, New Jersey: Pearson.
- Courbet, A., Endy, D., Renard, E., Molina, F., & Bonnet, J. (2015). Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates. *Sci Transl Med, 7*(289), 289ra283. doi:10.1126/scitranslmed.aaa3601
- Courbet, A., Renard, E., & Molina, F. (2016). Bringing next-generation diagnostics to the clinic through synthetic biology. *EMBO Mol Med, 8*(9), 987-991. doi:10.15252/emmm.201606541
- Culler, S. J., Hoff, K. G., & Smolke, C. D. (2010). Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. *Science, 330*(6008), 1251-1255. doi:10.1126/science.1192128
- Daniel, R., Rubens, J. R., Sarpeshkar, R., & Lu, T. K. (2013). Synthetic analog computation in living cells. *Nature, 497*(7451), 619-623. doi:10.1038/nature12148
- Danna, K., & Nathans, D. (1971). Specific cleavage of simian virus 40 DNA by restriction endonuclease of Hemophilus influenzae. *Proc Natl Acad Sci U S A, 68*(12), 2913-2917.
- Daringer, N. M., Dudek, R. M., Schwarz, K. A., & Leonard, J. N. (2014). Modular extracellular sensor architecture for engineering mammalian cell-based devices. *ACS Synth Biol, 3*(12), 892-902. doi:10.1021/sb400128g
- Daugherty, P. S., Chen, G., Olsen, M. J., Iverson, B. L., & Georgiou, G. (1998). Antibody affinity maturation using bacterial surface display. *Protein Eng, 11*(9), 825-832.
- Daunert, S., Barrett, G., Feliciano, J. S., Shetty, R. S., Shrestha, S., & Smith-Spencer, W. (2000). Genetically engineered whole-cell sensing systems: coupling biological recognition with reporter genes. *Chem Rev, 100*(7), 2705-2738.

- Davis, J. H., Rubin, A. J., & Sauer, R. T. (2011). Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res, 39*(3), 1131-1141. doi:10.1093/nar/gkq810
- de Wet, J. R., Wood, K. V., Helinski, D. R., & DeLuca, M. (1985). Cloning of firefly luciferase cDNA and the expression of active luciferase in Escherichia coli. *Proc Natl Acad Sci U S A, 82*(23), 7870-7873.
- Deigan, K. E., Li, T. W., Mathews, D. H., & Weeks, K. M. (2009). Accurate SHAPE-directed RNA structure determination. *Proc Natl Acad Sci U S A, 106*(1), 97-102. doi:10.1073/pnas.0806929106
- Deng, B., Lin, Y., Wang, C., Li, F., Wang, Z., Zhang, H., . . . Le, X. C. (2014). Aptamer binding assays for proteins: the thrombin example--a review. *Anal Chim Acta, 837*, 1-15. doi:10.1016/j.aca.2014.04.055
- Doherty, A. J., Connolly, B. A., & Worrall, A. F. (1993). Overproduction of the toxic protein, bovine pancreatic DNaseI, in Escherichia coli using a tightly controlled T7-promoter-based vector. *Gene, 136*(1-2), 337-340.
- Dolgosheina, E. B., Karulin, A., & Bobylev, A. V. (1992). A kinetic model of the agglutination process. *Math Biosci, 109*(1), 1-10.
- Dong, H., Nilsson, L., & Kurland, C. G. (1995). Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction. *J Bacteriol, 177*(6), 1497-1504.
- Ehrenkranz, J. R. L. (2002). Home and Point-of-Care Pregnancy Tests: A Review of the Technology. *Epidemiology, 13*(3 (Supplement)), S15-S18.
- Eisfeld, A. J., Neumann, G., & Kawaoka, Y. (2014). Influenza A virus isolation, culture and identification. *Nat Protoc, 9*(11), 2663-2681. doi:10.1038/nprot.2014.180
- Ellington, A. D., & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature, 346*(6287), 818-822. doi:10.1038/346818a0
- Ellington, A. D., & Szostak, J. W. (1992). Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature, 355*(6363), 850-852. doi:10.1038/355850a0
- Ellis, R. W., & Sobanski, M. A. (2000). Diagnostic particle agglutination using ultrasound: a new technology to rejuvenate old microbiological methods. *J Med Microbiol, 49*(10), 853-859. doi:10.1099/0022-1317-49-10-853
- Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature, 403*(6767), 335-338. doi:10.1038/35002125
- Endy, D. (2005). Foundations for engineering biology. *Nature, 438*(7067), 449-453. doi:10.1038/nature04342
- Engler, C., Kandzia, R., & Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. *PLOS ONE, 3*(11), e3647. doi:10.1371/journal.pone.0003647
- Espah Borujeni, A., Channarasappa, A. S., & Salis, H. M. (2014). Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res, 42*(4), 2646-2659. doi:10.1093/nar/gkt1139
- Espah Borujeni, A., Mishler, D. M., Wang, J., Huso, W., & Salis, H. M. (2016). Automated physics-based design of synthetic riboswitches from diverse RNA aptamers. *Nucleic Acids Res, 44*(1), 1-13. doi:10.1093/nar/gkv1289

- Eulberg, D., Buchner, K., Maasch, C., & Klussmann, S. (2005). Development of an automated in vitro selection protocol to obtain RNA-based aptamers: identification of a biostable substance P antagonist. *Nucleic Acids Res, 33*(4), e45. doi:10.1093/nar/gni044
- Everse, S. J., Spraggon, G., Veerapandian, L., Riley, M., & Doolittle, R. F. (1998). Crystal structure of fragment double-D from human fibrin with two different bound ligands. *Biochemistry, 37*(24), 8637-8642. doi:10.1021/bi9804129
- Ezure, T., Suzuki, T., Higashide, S., Shintani, E., Endo, K., Kobayashi, S., . . . Nishimura, O. (2006). Cell-free protein synthesis system prepared from insect cells by freeze-thawing. *Biotechnol Prog, 22*(6), 1570-1577. doi:10.1021/bp060110v
- Fairman, J. W., Dautin, N., Wojtowicz, D., Liu, W., Noinaj, N., Barnard, T. J., . . . Buchanan, S. K. (2012). Crystal structures of the outer membrane domain of intimin and invasin from enterohemorrhagic E. coli and enteropathogenic Y. pseudotuberculosis. *Structure, 20*(7), 1233-1243. doi:10.1016/j.str.2012.04.011
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., . . . et al. (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science, 269*(5223), 496-512.
- Franch, T., Gultyaev, A. P., & Gerdes, K. (1997). Programmed cell death by hok/sok of plasmid R1: processing at the hok mRNA 3'-end triggers structural rearrangements that allow translation and antisense RNA binding. *J Mol Biol, 273*(1), 38-51. doi:10.1006/jmbi.1997.1294
- Francisco, J. A., Campbell, R., Iverson, B. L., & Georgiou, G. (1993). Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface. *Proc Natl Acad Sci U S A, 90*(22), 10444-10448.
- French, C. E., Joshi, N., Elfick, A., Haseloff, J., & Ajioka, J. (2011). *Synthetic biology and the art of biosensor design*. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/
- Friedland, A. E., Lu, T. K., Wang, X., Shi, D., Church, G., & Collins, J. J. (2009). Synthetic gene networks that count. *Science, 324*(5931), 1199-1202. doi:10.1126/science.1172005
- Fuchs, P., Breitling, F., Dubel, S., Seehaus, T., & Little, M. (1991). Targeting recombinant antibodies to the surface of Escherichia coli: fusion to a peptidoglycan associated lipoprotein. *Biotechnology (N Y), 9*(12), 1369-1372.
- Gallagher, R. R., Li, Z., Lewis, A. O., & Isaacs, F. J. (2014). Rapid editing and evolution of bacterial genomes using libraries of synthetic DNA. *Nat Protoc, 9*(10), 2301-2316. doi:10.1038/nprot.2014.082
- Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature, 403*(6767), 339-342. doi:10.1038/35002131
- Garmendia, J., de las Heras, A., Galvao, T. C., & de Lorenzo, V. (2008). Tracing explosives in soil with transcriptional regulators of Pseudomonas putida evolved for responding to nitrotoluenes. *Microb Biotechnol, 1*(3), 236-246. doi:10.1111/j.1751-7915.2008.00027.x
- George, A. (2000). The Antibody Molecule. In A. George & C. Urch (Eds.), *Diagnostic and Therapeutic Antibodies* (1st ed., Vol. 40, pp. 1 -21): Humana Press.
- Georgiou, G., Stephens, D. L., Stathopoulos, C., Poetschke, H. L., Mendenhall, J., & Earhart, C. F. (1996). Display of beta-lactamase on the Escherichia coli surface: outer membrane phenotypes conferred by Lpp'-OmpA'-beta-lactamase fusions. *Protein Eng, 9*(2), 239-247.

- Germer, K., Leonard, M., & Zhang, X. (2013). RNA aptamers and their therapeutic and diagnostic applications. *Int J Biochem Mol Biol, 4*(1), 27-40.
- Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., . . . Venter, J. C. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science, 329*(5987), 52-56. doi:10.1126/science.1190719
- Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods, 6*(5), 343-345. doi:10.1038/nmeth.1318
- Glockshuber, R., Malia, M., Pfitzinger, I., & Pluckthun, A. (1990). A comparison of strategies to stabilize immunoglobulin Fv-fragments. *Biochemistry, 29*(6), 1362-1367.
- Goers, L., Kylilis, N., Tomazou, M., Wen, K., Freemont, P., & Polizzi, K. (2013). Engineering Microbial Biosensors. In H. Colin & W. Anil (Eds.), *Microbial Synthetic Biology* (1st ed., Vol. 40, pp. 119–156): Elsevier.
- Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., . . . Oliver, S. G. (1996). Life with 6000 genes. *Science, 274*(5287), 546, 563-547.
- Govaert, J., Pellis, M., Deschacht, N., Vincke, C., Conrath, K., Muyldermans, S., & Saerens, D. (2012). Dual beneficial effect of interloop disulfide bond for single domain antibody fragments. *J Biol Chem, 287*(3), 1970-1979. doi:10.1074/jbc.M111.242818
- Green, A. A., Silver, P. A., Collins, J. J., & Yin, P. (2014). Toehold switches: de-novo-designed regulators of gene expression. *Cell, 159*(4), 925-939. doi:10.1016/j.cell.2014.10.002
- Grieshaber, D., MacKenzie, R., Vörös, J., & Reimhult, E. (2008). Electrochemical Biosensors - Sensor Principles and Architectures. *Sensors (Basel), 8*(3), 1400-1458.
- Gubala, V., Harris, L. F., Ricco, A. J., Tan, M. X., & Williams, D. E. (2012). Point of care diagnostics: status and future. *Anal Chem, 84*(2), 487-515. doi:10.1021/ac2030199
- Guet, C. C., Elowitz, M. B., Hsing, W., & Leibler, S. (2002). Combinatorial synthesis of genetic networks. *Science, 296*(5572), 1466-1470. doi:10.1126/science.1067407
- Guilbault, G. G., & Lubrano, G. J. (1973). An enzyme electrode for the amperometric determination of glucose. *Anal Chim Acta, 64*(3), 439-455. doi:10.1016/s0003-2670(01)82476-4
- Gutknecht, J., & Walter, A. (1981). Histamine, theophylline and tryptamine transport through lipid bilayer membranes. *Biochim Biophys Acta, 649*(2), 149-154.
- Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E. B., . . . Hamers, R. (1993). Naturally occurring antibodies devoid of light chains. *Nature, 363*(6428), 446-448. doi:10.1038/363446a0
- Harmsen, M. M., Ruuls, R. C., Nijman, I. J., Niewold, T. A., Frenken, L. G., & de Geus, B. (2000). Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. *Mol Immunol, 37*(10), 579-590.
- Hartwell, L. H., Hopfield, J. J., Leibler, S., & Murray, A. W. (1999). From molecular to modular cell biology. *Nature, 402*(6761 Suppl), C47-52. doi:10.1038/35011540
- Harvey, B. R., Georgiou, G., Hayhurst, A., Jeong, K. J., Iverson, B. L., & Rogers, G. K. (2004). Anchored periplasmic expression, a versatile technology for the isolation of high-affinity

antibodies from Escherichia coli-expressed libraries. *Proc Natl Acad Sci U S A, 101*(25), 9193-9198. doi:10.1073/pnas.0400187101

- Hellman, L. M., & Fried, M. G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc, 2*(8), 1849-1861. doi:10.1038/nprot.2007.249
- Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C., & Ala'Aldeen, D. (2004). Type V Protein Secretion Pathway: the Autotransporter Story. *Microbiol Mol Biol Rev, 68*(4), 692-744. doi:10.1128/mmbr.68.4.692-744.2004
- Herrmann, I. K., Bertazzo, S., O'Callaghan, D. J., Schlegel, A. A., Kallepitis, C., Antcliffe, D. B., . . . Stevens, M. M. (2015). Differentiating sepsis from non-infectious systemic inflammation based on microvesicle-bacteria aggregation. *Nanoscale, 7*(32), 13511-13520. doi:10.1039/c5nr01851j
- Higashide, W., Li, Y., Yang, Y., & Liao, J. C. (2011). Metabolic engineering of Clostridium cellulolyticum for production of isobutanol from cellulose. *Appl Environ Microbiol, 77*(8), 2727-2733. doi:10.1128/aem.02454-10
- Holliger, P., & Hudson, P. J. (2005). Engineered antibody fragments and the rise of single domains. *Nat Biotechnol, 23*(9), 1126-1136. doi:10.1038/nbt1142
- Horvath, P., & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science, 327*(5962), 167-170. doi:10.1126/science.1179555
- Hulme, E. C., & Trevethick, M. A. (2010). Ligand binding assays at equilibrium: validation and interpretation. *Br J Pharmacol, 161*(6), 1219-1237. doi:10.1111/j.1476-5381.2009.00604.x
- Hung El, L. (2014). *Protein detection by whole-cell based biosensor device*. Thesis report for Systems and Synthetic Biology MRes course 2013-14, Imperial College London.
- Hüttenhofer, A., & Noller, H. F. (1994). Footprinting mRNA-ribosome complexes with chemical probes. *EMBO J, 13*(16), 3892-3901.
- Iizuka, N., Najita, L., Franzusoff, A., & Sarnow, P. (1994). Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from Saccharomyces cerevisiae. *Mol Cell Biol, 14*(11), 7322-7330.
- Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol, 3*, 318-356.
- Jayasena, S. D. (1999). Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clin Chem, 45*(9), 1628-1650.
- Jellinek, D., Green, L. S., Bell, C., & Janjic, N. (1994). Inhibition of receptor binding by high-affinity RNA ligands to vascular endothelial growth factor. *Biochemistry, 33*(34), 10450-10456.
- Jenison, R. D., Gill, S. C., Pardi, A., & Polisky, B. (1994). High-resolution molecular discrimination by RNA. *Science, 263*(5152), 1425-1429.
- Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S., & Winter, G. (1986). Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature, 321*(6069), 522-525. doi:10.1038/321522a0
- Juers, D. H., Matthews, B. W., & Huber, R. E. (2012). LacZ β-galactosidase: Structure and function of an enzyme of historical and molecular biological importance. *Protein Sci, 21*(12), 1792-1807. doi:10.1002/pro.2165

- Kabat, E. A., Wu, T. T., & Bilofsky, H. (1978). Variable region genes for the immunoglobulin framework are assembled from small segments of DNA--a hypothesis. *Proc Natl Acad Sci U S A, 75*(5), 2429-2433.
- Kannel, W. B., Wolf, P. A., Castelli, W. P., & D'Agostino, R. B. (1987). Fibrinogen and risk of cardiovascular disease. The Framingham Study. *Jama, 258*(9), 1183-1186.
- Karzbrun, E., Shin, J., Bar-Ziv, R. H., & Noireaux, V. (2011). Coarse-Grained Dynamics of Protein Synthesis in a Cell-Free System. *Phys. Rev. Lett., 106*(4), 048104. doi:10.1103/PhysRevLett.106.048104
- Kaur, H., & Yung, L. Y. (2012). Probing high affinity sequences of DNA aptamer against VEGF165. *PLOS ONE, 7*(2), e31196. doi:10.1371/journal.pone.0031196
- Kelly, J. R., Rubin, A. J., Davis, J. H., Ajo-Franklin, C. M., Cumbers, J., Czar, M. J., . . . Endy, D. (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. *J Biol Eng, 3*, 4. doi:10.1186/1754-1611-3-4
- Kennedy, A. B., Vowles, J. V., d'Espaux, L., & Smolke, C. D. (2014). Protein-responsive ribozyme switches in eukaryotic cells. *Nucleic Acids Res, 42*(19), 12306-12321. doi:10.1093/nar/gku875
- Khalil, A. S., & Collins, J. J. (2010). Synthetic biology: applications come of age. *Nat Rev Genet, 11*(5), 367-379. doi:10.1038/nrg2775
- Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Nakajima, R., Tanaka, A., & Yokoyama, S. (2004). Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. *J Struct Funct Genomics, 5*(1-2), 63-68. doi:10.1023/B:JSFG.0000029204.57846.7d
- Kim, D. M., & Swartz, J. R. (2001). Regeneration of adenosine triphosphate from glycolytic intermediates for cell-free protein synthesis. *Biotechnol Bioeng, 74*(4), 309-316.
- Kohler, G., & Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature, 256*(5517), 495-497.
- Kortmann, J., & Narberhaus, F. (2012). Bacterial RNA thermometers: molecular zippers and switches. *Nat Rev Microbiol, 10*(4), 255-265. doi:10.1038/nrmicro2730
- Kubota, M., Tran, C., & Spitale, R. C. (2015). Progress and challenges for chemical probing of RNA structure inside living cells. *Nat Chem Biol, 11*(12), 933-941. doi:10.1038/nchembio.1958
- Kwon, Y. C., & Jewett, M. C. (2015). High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci Rep, 5*, 8663. doi:10.1038/srep08663
- Laursen, B. S., Sorensen, H. P., Mortensen, K. K., & Sperling-Petersen, H. U. (2005). Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev, 69*(1), 101-123. doi:10.1128/mmbr.69.1.101-123.2005
- Leclerc, C., Charbit, A., Martineau, P., Deriaud, E., & Hofnung, M. (1991). The cellular location of a foreign B cell epitope expressed by recombinant bacteria determines its T cell-independent or T cell-dependent characteristics. *J Immunol, 147*(10), 3545-3552.
- Lee, M. H., Park, T. I., Park, Y. B., & Kwak, J. W. (2002). Bacterial expression and in vitro refolding of a single-chain fv antibody specific for human plasma apolipoprotein B-100. *Protein Expr Purif, 25*(1), 166-173. doi:10.1006/prep.2002.1623

- Leo, J. C., Grin, I., & Linke, D. (2012). Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos Trans R Soc Lond B Biol Sci, 367*(1592), 1088-1101. doi:10.1098/rstb.2011.0208
- Levskaya, A., Chevalier, A. A., Tabor, J. J., Simpson, Z. B., Lavery, L. A., Levy, M., . . . Voigt, C. A. (2005). Synthetic biology: engineering Escherichia coli to see light. *Nature, 438*(7067), 441-442. doi:10.1038/nature04405
- Leyton, D. L., Rossiter, A. E., & Henderson, I. R. (2012). From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol, 10*(3), 213-225. doi:10.1038/nrmicro2733
- Li, Y., Lee, H. J., & Corn, R. M. (2007). Detection of protein biomarkers using RNA aptamer microarrays and enzymatically amplified surface plasmon resonance imaging. *Analytical Chemistry, 79*(3), 1082-1088. doi:Doi 10.1021/Ac061849m
- Li, Z., Woo, C. J., Iglesias-Ussel, M. D., Ronai, D., & Scharff, M. D. (2004). The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev, 18*(1), 1-11. doi:10.1101/gad.1161904
- Liang, J. C., Bloom, R. J., & Smolke, C. D. (2011). Engineering biological systems with synthetic RNA molecules. *Mol Cell, 43*(6), 915-926. doi:10.1016/j.molcel.2011.08.023
- Liu, P. Y., Chin, L. K., Ser, W., Ayi, T. C., Yap, P. H., Bourouina, T., & Leprince-Wang, Y. (2014). An optofluidic imaging system to measure the biophysical signature of single waterborne bacteria. *Lab Chip, 14*(21), 4237-4243. doi:10.1039/c4lc00783b
- Liu, X. M., Cao, G. J., Ding, H. M., Zhang, D. J., Yang, G., Liu, N. L., . . . Shao, N. S. (2004). Screening of functional antidotes of RNA aptamers against bovine thrombin. *Febs Letters, 562*(1-3), 125-128. doi:Doi 10.1016/S0014-5793(04)00197-8
- Liu, X. M., Zhang, D. J., Cao, G. J., Yang, G., Ding, H. M., Liu, G., . . . Shao, N. S. (2003). RNA aptamers specific for bovine thrombin. *Journal of Molecular Recognition, 16*(1), 23-27. doi:10.1002/jmr.604
- Long, S. B., Long, M. B., White, R. R., & Sullenger, B. A. (2008). Crystal structure of an RNA aptamer bound to thrombin. *Rna, 14*(12), 2504-2512. doi:10.1261/rna.1239308
- Looger, L. L., Dwyer, M. A., Smith, J. J., & Hellinga, H. W. (2003). Computational design of receptor and sensor proteins with novel functions. *Nature, 423*(6936), 185-190. doi:10.1038/nature01556
- Lou, C., Stanton, B., Chen, Y. J., Munsky, B., & Voigt, C. A. (2012). Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat Biotechnol, 30*(11), 1137-1142. doi:10.1038/nbt.2401
- Lowe, C. R. (2008). Overview of Biosensor and Bioarray Technologies. In *Handbook of Biosensors and Biochips*: John Wiley & Sons, Ltd.
- Lowe, G. D., Rumley, A., & Mackie, I. J. (2004). Plasma fibrinogen. *Ann Clin Biochem, 41*(Pt 6), 430-440. doi:10.1258/0004563042466884
- Luo, Y., Frey, E. A., Pfuetzner, R. A., Creagh, A. L., Knoechel, D. G., Haynes, C. A., . . . Strynadka, N. C. (2000). Crystal structure of enteropathogenic Escherichia coli intimin-receptor complex. *Nature, 405*(6790), 1073-1077. doi:10.1038/35016618

- Lynch, S. A., Desai, S. K., Sajja, H. K., & Gallivan, J. P. (2007). A high-throughput screen for synthetic riboswitches reveals mechanistic insights into their function. *Chem Biol, 14*(2), 173-184. doi:10.1016/j.chembiol.2006.12.008
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., . . . Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science, 339*(6121), 823-826. doi:10.1126/science.1232033
- Mandal, M., & Breaker, R. R. (2004). Gene regulation by riboswitches. *Nat Rev Mol Cell Biol, 5*(6), 451-463. doi:10.1038/nrm1403
- Manjili, M. H., Najarian, K., & Wang, X. Y. (2012). Signatures of tumor–immune interactions as biomarkers for breast cancer prognosis. *Future Oncol, 8*(6), 703-711. doi:10.2217/fon.12.57
- Mascini, M., & Tombelli, S. (2008). Biosensors for biomarkers in medical diagnostics. *Biomarkers, 13*(7), 637-657. doi:10.1080/13547500802645905
- Mathews, D. H. (2004). Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *Rna, 10*(8), 1178-1190. doi:10.1261/rna.7650904
- Mathews, D. H. (2006). Revolutions in RNA secondary structure prediction. *J Mol Biol, 359*(3), 526-532. doi:10.1016/j.jmb.2006.01.067
- Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., & Turner, D. H. (2004). Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci U S A, 101*(19), 7287-7292. doi:10.1073/pnas.0401799101
- Mathews, D. H., & Turner, D. H. (2006). Prediction of RNA secondary structure by free energy minimization. *Curr Opin Struct Biol, 16*(3), 270-278. doi:10.1016/j.sbi.2006.05.010
- Matthews, D. R., Holman, R. R., Bown, E., Steemson, J., Watson, A., Hughes, S., & Scott, D. (1987). Pen-sized digital 30-second blood glucose meter. *Lancet, 1*(8536), 778-779.
- Meighen, E. A. (1991). Molecular biology of bacterial bioluminescence. *Microbiol Rev, 55*(1), 123-142.
- Melamed, S., Ceriotti, L., Weigel, W., Rossi, F., Colpo, P., & Belkin, S. (2011). A printed nanolitre-scale bacterial sensor array. *Lab Chip, 11*(1), 139-146. doi:10.1039/c0lc00243g
- Mikami, S., Kobayashi, T., & Imataka, H. (2010). Cell-free protein synthesis systems with extracts from cultured human cells. *Methods Mol Biol, 607*, 43-52. doi:10.1007/978-1-60327-331-2_5
- Mikami, S., Kobayashi, T., Masutani, M., Yokoyama, S., & Imataka, H. (2008). A human cell-derived in vitro coupled transcription/translation system optimized for production of recombinant proteins. *Protein Expr Purif, 62*(2), 190-198. doi:10.1016/j.pep.2008.09.002
- Mohan, P., Marin, D., Sultan, S., & Deen, A. (2008). MediNet: personalizing the self-care process for patients with diabetes and cardiovascular disease using mobile telephony. *Conf Proc IEEE Eng Med Biol Soc, 2008*, 755-758. doi:10.1109/iembs.2008.4649262
- Molina-Bolívar, J. A., & Galisteo-González, F. (2005). Latex Immunoagglutination Assays. *http://dx.doi.org/10.1081/MC-200045819, 45*(1). doi:Journal of Macromolecular Science Part C—Polymer Reviews, Vol. 45, No. 1, 2005: pp. 59–98

- Montange, R. K., & Batey, R. T. (2008). Riboswitches: emerging themes in RNA structure and function. *Annu Rev Biophys, 37*, 117-133. doi:10.1146/annurev.biophys.37.032807.130000
- Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C., & Voigt, C. A. (2012). Genetic programs constructed from layered logic gates in single cells. *Nature, 491*(7423), 249-253. doi:10.1038/nature11516
- Moran, U., Phillips, R., & Milo, R. (2010). SnapShot: key numbers in biology. *Cell, 141*(7), 1262-1262.e1261. doi:10.1016/j.cell.2010.06.019
- Morrison, S. L., Johnson, M. J., Herzenberg, L. A., & Oi, V. T. (1984). Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proc Natl Acad Sci U S A, 81*(21), 6851-6855.
- Mujahid, A., & Dickert, F. L. (2015). Blood Group Typing: From Classical Strategies to the Application of Synthetic Antibodies Generated by Molecular Imprinting. *Sensors (Basel), 16*(1). doi:10.3390/s16010051
- Mullis, K. B., Erlich, H. A., Arnheim, N., Horn, G. T., Saiki, R. K., Scharf, S. J., & Corporation, C. (1987).
- Na, D., & Lee, D. (2010). RBSDesigner: software for designing synthetic ribosome binding sites that yields a desired level of protein expression. *Bioinformatics, 26*(20), 2633-2634. doi:10.1093/bioinformatics/btq458
- Natale, P., Bruser, T., & Driessen, A. J. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim Biophys Acta, 1778*(9), 1735-1756. doi:10.1016/j.bbamem.2007.07.015
- Newman, D. J., Henneberry, H., & Price, C. P. (1992). Particle enhanced light scattering immunoassay. *Ann Clin Biochem, 29 (Pt 1)*, 22-42.
- Nguyen, V. K., Hamers, R., Wyns, L., & Muyldermans, S. (2000). Camel heavy-chain antibodies: diverse germline VHH and specific mechanisms enlarge the antigen-binding repertoire. *EMBO J, 19*(5), 921-930. doi:10.1093/emboj/19.5.921
- Nielsen, A. A., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., . . . Voigt, C. A. (2016). Genetic circuit design automation. *Science, 352*(6281), aac7341. doi:10.1126/science.aac7341
- Oberhettinger, P., Schutz, M., Leo, J. C., Heinz, N., Berger, J., Autenrieth, I. B., & Linke, D. (2012). Intimin and invasin export their C-terminus to the bacterial cell surface using an inverse mechanism compared to classical autotransport. *PLoS One, 7*(10), e47069. doi:10.1371/journal.pone.0047069
- Oda, Y., Nakamura, S., Oki, I., Kato, T., & Shinagawa, H. (1985). Evaluation of the new system (umu-test) for the detection of environmental mutagens and carcinogens. *Mutat Res, 147*(5), 219-229.
- Ogawa, A. (2011). Rational design of artificial riboswitches based on ligand-dependent modulation of internal ribosome entry in wheat germ extract and their applications as label-free biosensors. *RNA, 17*(3), 478-488. doi:10.1261/rna.2433111
- Orlandi, R., Gussow, D. H., Jones, P. T., & Winter, G. (1989). Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. *Proc Natl Acad Sci U S A, 86*(10), 3833-3837.

- Ortega-Vinuesa, J. L., & Bastos-Gonzalez, D. (2001). A review of factors affecting the performances of latex agglutination tests. *J Biomater Sci Polym Ed, 12*(4), 379-408.
- Oved, K., Cohen, A., Boico, O., Navon, R., Friedman, T., Etshtein, L., . . . Eden, E. (2015). A novel host-proteome signature for distinguishing between acute bacterial and viral infections. *PLoS One, 10*(3), e0120012. doi:10.1371/journal.pone.0120012
- Padlan, E. A., Silverton, E. W., Sheriff, S., Cohen, G. H., Smith-Gill, S. J., & Davies, D. R. (1989). Structure of an antibody-antigen complex: crystal structure of the HyHEL-10 Fab-lysozyme complex. *Proc Natl Acad Sci U S A, 86*(15), 5938-5942.
- Palumbo, J. S., Kombrinck, K. W., Drew, A. F., Grimes, T. S., Kiser, J. H., Degen, J. L., & Bugge, T. H. (2000). Fibrinogen is an important determinant of the metastatic potential of circulating tumor cells. *Blood, 96*(10), 3302-3309.
- Pardee, K., Green, A. A., Ferrante, T., Cameron, D. E., DaleyKeyser, A., Yin, P., & Collins, J. J. (2014). Paper-based synthetic gene networks. *Cell, 159*(4), 940-954. doi:10.1016/j.cell.2014.10.004
- Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., . . . Collins, J. J. (2016). Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell, 165*(5), 1255-1266. doi:10.1016/j.cell.2016.04.059
- Pardee, K., Slomovic, S., Nguyen, P. Q., Lee, J. W., Donghia, N., Burrill, D., . . . Collins, J. J. (2016). Portable, On-Demand Biomolecular Manufacturing. *Cell, 167*(1), 248-259.e212. doi:10.1016/j.cell.2016.09.013
- Pardon, E., Laeremans, T., Triest, S., Rasmussen, S. G. F., Wohlkönig, A., Ruf, A., . . . Steyaert, J. (2014). A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc, 9*(3), 674-693. doi:10.1038/nprot.2014.039
- Peabody, D. S. (1990). Translational repression by bacteriophage MS2 coat protein expressed from a plasmid. A system for genetic analysis of a protein-RNA interaction. *J Biol Chem, 265*(10), 5684-5689.
- Pelham, H. R., & Jackson, R. J. (1976). An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur J Biochem, 67*(1), 247-256.
- Pelton, J. T., & McLean, L. R. (2000). Spectroscopic methods for analysis of protein secondary structure. *Anal Biochem, 277*(2), 167-176. doi:10.1006/abio.1999.4320
- Pinero-Lambea, C., Bodelon, G., Fernandez-Perianez, R., Cuesta, A. M., Alvarez-Vallina, L., & Fernandez, L. A. (2015). Programming controlled adhesion of E. coli to target surfaces, cells, and tumors with synthetic adhesins. *ACS Synth Biol, 4*(4), 463-473. doi:10.1021/sb500252a
- Pothoulakis, G., Ceroni, F., Reeve, B., & Ellis, T. (2014). The spinach RNA aptamer as a characterization tool for synthetic biology. *ACS Synth Biol, 3*(3), 182-187. doi:10.1021/sb400089c
- Price, C. P. (2001). Agglutination Techniques for Detecting Antigen–Antibody Reactions. *eLS*. doi:10.1038/npg.els.0001114
- Qi, L., Lucks, J. B., Liu, C. C., Mutalik, V. K., & Arkin, A. P. (2012). Engineering naturally occurring trans-acting non-coding RNAs to sense molecular signals. *Nucleic Acids Res, 40*(12), 5775-5786. doi:10.1093/nar/gks168

- Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. *Cell, 108*(4), 557-572.
- Ramanathan, S., Shi, W., Rosen, B. P., & Daunert, S. (1997). Sensing antimonite and arsenite at the subattomole level with genetically engineered bioluminescent bacteria. *Anal Chem, 69*(16), 3380-3384.
- Raut, N., O'Connor, G., Pasini, P., & Daunert, S. (2012). Engineered cells as biosensing systems in biomedical analysis. *Anal Bioanal Chem, 402*(10), 3147-3159. doi:10.1007/s00216-012-5756-6
- Registry_of_Standard_Biological_Parts. (2016). Promoters/Catalog/Anderson - parts.igem.org. Retrieved from http://parts.igem.org/Promoters/Catalog/Anderson
- Reimer, A., Yagur-Kroll, S., Belkin, S., Roy, S., & van der Meer, J. R. (2014). Escherchia coli ribose binding protein based bioreporters revisited. *Sci Rep, 4*. doi:10.1038/srep05626
- Ritter, M. (2000). Polyclonal and Monoclonal Antibodies. In A. George & C. Urch (Eds.), *Diagnostic and Therapeutic Antibodies* (1st ed., pp. 23-34): Humana Press.
- Ro, D. K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., . . . Keasling, J. D. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature, 440*(7086), 940-943. doi:10.1038/nature04640
- Roberts, R. J. (2005). How restriction enzymes became the workhorses of molecular biology. *Proc Natl Acad Sci U S A, 102*(17), 5905-5908. doi:10.1073/pnas.0500923102
- Rosa, M. D. (1979). Four T7 RNA polymerase promoters contain an identical 23 bp sequence. *Cell, 16*(4), 815-825.
- Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol, 5*, 172. doi:10.3389/fmicb.2014.00172
- Rothbauer, U., Zolghadr, K., Tillib, S., Nowak, D., Schermelleh, L., Gahl, A., . . . Leonhardt, H. (2006). Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat Methods, 3*(11), 887-889. doi:10.1038/nmeth953
- Ruiz, N., Kahne, D., & Silhavy, T. J. (2006). Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Microbiol, 4*(1), 57-66. doi:10.1038/nrmicro1322
- Rutherford, N., & Mourez, M. (2006). Surface display of proteins by gram-negative bacterial autotransporters. *Microb Cell Fact, 5*, 22. doi:10.1186/1475-2859-5-22
- Sagi, E., Hever, N., Rosen, R., Bartolome, A. J., Rajan Premkumar, J., Ulber, R., . . . Belkin, S. (2003). Fluorescence and bioluminescence reporter *Sensors and Actuator B: Chemical, 90*(1-3), 2-8. doi:http://dx.doi.org/10.1016/S0925-4005(03)00014-5
- Saito, H., Fujita, Y., Kashida, S., Hayashi, K., & Inoue, T. (2011). Synthetic human cell fate regulation by protein-driven RNA switches. *Nat Commun, 2*, 160. doi:10.1038/ncomms1157
- Saito, H., & Inoue, T. (2007). RNA and RNP as new molecular parts in synthetic biology. *J Biotechnol, 132*(1), 1-7. doi:10.1016/j.jbiotec.2007.07.952
- Saito, H., Kobayashi, T., Hara, T., Fujita, Y., Hayashi, K., Furushima, R., & Inoue, T. (2010). Synthetic translational regulation by an L7Ae-kink-turn RNP switch. *Nat Chem Biol, 6*(1), 71-78. doi:10.1038/nchembio.273

- Salema, V., & Fernandez, L. A. (2013). High yield purification of nanobodies from the periplasm of E. coli as fusions with the maltose binding protein. *Protein Expr Purif, 91*(1), 42-48. doi:10.1016/j.pep.2013.07.001
- Salema, V., Marin, E., Martinez-Arteaga, R., Ruano-Gallego, D., Fraile, S., Margolles, Y., . . . Fernandez, L. A. (2013). Selection of single domain antibodies from immune libraries displayed on the surface of E. coli cells with two beta-domains of opposite topologies. *PLoS One, 8*(9), e75126. doi:10.1371/journal.pone.0075126
- Salis, H. M. (2011). The ribosome binding site calculator. *Methods Enzymol, 498*, 19-42. doi:10.1016/b978-0-12-385120-8.00002-4
- Salis, H. M., Mirsky, E. A., & Voigt, C. A. (2009). Automated Design of Synthetic Ribosome Binding Sites to Precisely Control Protein Expression. *Nat Biotechnol, 27*(10), 946-950. doi:10.1038/nbt.1568
- Sassanfar, M., & Szostak, J. W. (1993). An RNA motif that binds ATP. *Nature, 364*(6437), 550-553. doi:10.1038/364550a0
- Scott, W. G., & Klug, A. (1996). Ribozymes: structure and mechanism in RNA catalysis. *Trends Biochem Sci, 21*(6), 220-224.
- Seo, S. W., Yang, J. S., Cho, H. S., Yang, J., Kim, S. C., Park, J. M., . . . Jung, G. Y. (2014). Predictive combinatorial design of mRNA translation initiation regions for systematic optimization of gene expression levels. In *Sci Rep* (Vol. 4).
- Serganov, A., & Nudler, E. (2013). A decade of riboswitches. *Cell, 152*(1-2), 17-24. doi:10.1016/j.cell.2012.12.024
- Shaner, N. C., Steinbach, P. A., & Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nat Methods, 2*(12), 905-909. doi:10.1038/nmeth819
- Sharma, S., Byrne, H., & O'Kennedy, R. J. (2016). Antibodies and antibody-derived analytical biosensors. *Essays Biochem, 60*(1), 9-18. doi:10.1042/ebc20150002
- Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., & Ueda, T. (2001). Cell-free translation reconstituted with purified components. *Nat Biotechnol, 19*(8), 751-755. doi:10.1038/90802
- Shin, J., & Noireaux, V. (2010). Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70. *J Biol Eng, 4*, 8. doi:10.1186/1754-1611-4-8
- Shine, J., & Dalgarno, L. (1974). The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci U S A, 71*(4), 1342-1346.
- Siddiqui, M. Z. (2010). Monoclonal antibodies as diagnostics; an appraisal. *Indian J Pharm Sci, 72*(1), 12-17. doi:10.4103/0250-474x.62229
- Siegal-Gaskins, D., Noireaux, V., & Murray, R. M. (2013). *Biomolecular resource utilization in elementary cell-free gene circuits*. PeerReviewed retrieved from http://authors.library.caltech.edu/39352/
- Sinha, J., Reyes, S. J., & Gallivan, J. P. (2010). Reprogramming bacteria to seek and destroy an herbicide. *Nat Chem Biol, 6*(6), 464-470. doi:10.1038/nchembio.369

- Siontorou, C. G. (2013). Nanobodies as novel agents for disease diagnosis and therapy. *Int J Nanomedicine, 8*, 4215-4227. doi:10.2147/ijn.s39428
- Smith, H. O., & Wilcox, K. W. (1970). A restriction enzyme from Hemophilus influenzae. I. Purification and general properties. *J Mol Biol, 51*(2), 379-391.
- Somleva, M. N., Snell, K. D., Beaulieu, J. J., Peoples, O. P., Garrison, B. R., & Patterson, N. A. (2008). Production of polyhydroxybutyrate in switchgrass, a value-added co-product in an important lignocellulosic biomass crop. *Plant Biotechnol J, 6*(7), 663-678. doi:10.1111/j.1467-7652.2008.00350.x
- Stathopoulos, C., Georgiou, G., & Earhart, C. F. (1996). Characterization of Escherichia coli expressing an Lpp'OmpA(46-159)-PhoA fusion protein localized in the outer membrane. *Appl Microbiol Biotechnol, 45*(1-2), 112-119.
- Steitz, J. A. (1969). Polypeptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. *Nature, 224*(5223), 957-964.
- Sun, Z. Z., Hayes, C. A., Shin, J., Caschera, F., Murray, R. M., & Noireaux, V. (2013). Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. *J Vis Exp*(79), e50762. doi:10.3791/50762
- Sun, Z. Z., Yeung, E., Hayes, C. A., Noireaux, V., & Murray, R. M. (2014). Linear DNA for rapid prototyping of synthetic biological circuits in an Escherichia coli based TX-TL cell-free system. *ACS Synth Biol, 3*(6), 387-397. doi:10.1021/sb400131a
- Synthetic_Biology_Leadership_Council. (2016). *Biodesign for the Bioeconomy - UK Synthetic Biology Strategic Plan 2016*. Retrieved from https://connect.innovateuk.org/web/synthetic-biology-special-interest-group/roadmap-for-synthetic-biology
- Tabor, S., & Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci U S A, 82*(4), 1074-1078.
- Takai, K., & Endo, Y. (2010). The cell-free protein synthesis system from wheat germ. In *Methods Mol Biol* (2010/03/06 ed., Vol. 607, pp. 23-30).
- Takyar, S., Hickerson, R. P., & Noller, H. F. (2005). mRNA helicase activity of the ribosome. *Cell, 120*(1), 49-58. doi:10.1016/j.cell.2004.11.042
- Tamsir, A., Tabor, J. J., & Voigt, C. A. (2011). Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature, 469*(7329), 212-215. doi:10.1038/nature09565
- Tang, H. L., Tang, H. M., Fung, M. C., & Hardwick, J. M. (2015). In vivo CaspaseTracker biosensor system for detecting anastasis and non-apoptotic caspase activity. *Sci Rep, 5*, 9015. doi:10.1038/srep09015
- Thevenot, D. R., Toth, K., Durst, R. A., & Wilson, G. S. (2001). Electrochemical biosensors: recommended definitions and classification. *Biosens Bioelectron, 16*(1-2), 121-131.
- Tortora, G. J., Funke, B. R., & Case, C. L. (2012). *Microbiology: An Introduction* (11th ed.). London: Pearson.
- Trang, P. T., Berg, M., Viet, P. H., Van Mui, N., & Van Der Meer, J. R. (2005). Bacterial bioassay for rapid and accurate analysis of arsenic in highly variable groundwater samples. *Environ Sci Technol, 39*(19), 7625-7630.

- Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science, 249*(4968), 505-510.
- Turner, A. P. (2013). Biosensors: sense and sensibility. *Chem Soc Rev, 42*(8), 3184-3196. doi:10.1039/c3cs35528d
- van Bloois, E., Winter, R. T., Kolmar, H., & Fraaije, M. W. (2011). Decorating microbes: surface display of proteins on Escherichia coli. *Trends Biotechnol, 29*(2), 79-86. doi:10.1016/j.tibtech.2010.11.003
- van den Berg, B. (2010). Crystal structure of a full-length autotransporter. *J Mol Biol, 396*(3), 627-633. doi:10.1016/j.jmb.2009.12.061
- van der Meer, J. R., & Belkin, S. (2010). Where microbiology meets microengineering: design and applications of reporter bacteria. *Nat Rev Microbiol, 8*(7), 511-522. doi:10.1038/nrmicro2392
- Veiga, E., de Lorenzo, V., & Fernandez, L. A. (2004). Structural tolerance of bacterial autotransporters for folded passenger protein domains. *Mol Microbiol, 52*(4), 1069-1080. doi:10.1111/j.1365-2958.2004.04014.x
- Verma, R., & Boleti, E. (2000). Engineering Antibody Molecules. In A. George & C. Urch (Eds.), *Diagnostic and Therapeutic Antibodies* (1 ed., pp. 35-52): Humana Press.
- Vo-Dinh, T., & Cullum, B. (2000). Biosensors and biochips: advances in biological and medical diagnostics. *Fresenius J Anal Chem, 366*(6-7), 540-551.
- Wang, B., Barahona, M., & Buck, M. (2014). Engineering modular and tunable genetic amplifiers for scaling transcriptional signals in cascaded gene networks. *Nucleic Acids Res, 42*(14), 9484-9492. doi:10.1093/nar/gku593
- Wang, B., Kitney, R. I., Joly, N., & Buck, M. (2011). Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nat Commun, 2*, 508. doi:10.1038/ncomms1516
- Wang, J. (2008). Electrochemical glucose biosensors. *Chem Rev, 108*(2), 814-825. doi:10.1021/cr068123a
- Wang, L., Qian, K., Huang, Y., Jin, N., Lai, H., Zhang, T., . . . Wang, D. (2015). SynBioLGDB: a resource for experimentally validated logic gates in synthetic biology. *Sci Rep, 5*, 8090. doi:10.1038/srep08090
- Wang, Y., & Rando, R. R. (1995). Specific binding of aminoglycoside antibiotics to RNA. *Chem Biol, 2*(5), 281-290.
- Webb, A. J., Kelwick, R., Doenhoff, M. J., Kylilis, N., MacDonald, J. T., Wen, K. Y., . . . Freemont, P. S. (2016). A protease-based biosensor for the detection of schistosome cercariae. *Sci Rep, 6*, 24725. doi:10.1038/srep24725
- Weber, W., Schoenmakers, R., Keller, B., Gitzinger, M., Grau, T., Daoud-El Baba, M., . . . Fussenegger, M. (2008). A synthetic mammalian gene circuit reveals antituberculosis compounds. *Proc Natl Acad Sci U S A, 105*(29), 9994-9998. doi:10.1073/pnas.0800663105
- Wendel, S., Fischer, E. C., Martínez, V., Seppälä, S., & Nørholm, M. H. H. (2016). A nanobody:GFP bacterial platform that enables functional enzyme display and easy quantification of display capacity. *Microb Cell Fact, 15*. doi:10.1186/s12934-016-0474-y

- White, R., Rusconi, C., Scardino, E., Wolberg, A., Lawson, J., Hoffman, M., & Sullenger, B. (2001). Generation of species cross-reactive aptamers using "toggle" SELEX. *Mol Ther, 4*(6), 567-573. doi:10.1006/mthe.2001.0495
- Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., & Hochstrasser, D. F. (1999). Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol, 112*, 531-552.
- Wilkinson, K. A., Merino, E. J., & Weeks, K. M. (2006). Selective 2 '-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nature Protocols, 1*(3), 1610-1616. doi:10.1038/nprot.2006.249
- Willardson, B. M., Wilkins, J. F., Rand, T. A., Schupp, J. M., Hill, K. K., Keim, P., & Jackson, P. J. (1998). Development and Testing of a Bacterial Biosensor for Toluene-Based Environmental Contaminants. In *Appl Environ Microbiol* (Vol. 64, pp. 1006-1012).
- Winkler, W., Nahvi, A., & Breaker, R. R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature, 419*(6910), 952-956. doi:10.1038/nature01145
- Wise, A. A., & Kuske, C. R. (2000). Generation of Novel Bacterial Regulatory Proteins That Detect Priority Pollutant Phenols. *Appl Environ Microbiol, 66*(1), 163-169.
- Wright, O., Delmans, M., Stan, G. B., & Ellis, T. (2015). GeneGuard: A modular plasmid system designed for biosafety. *ACS Synth Biol, 4*(3), 307-316. doi:10.1021/sb500234s
- Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., . . . Turner, D. H. (1998). Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry, 37*(42), 14719-14735. doi:10.1021/bi9809425
- Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., & Benenson, Y. (2011). Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science, 333*(6047), 1307-1311. doi:10.1126/science.1205527
- Yager, P., Domingo, G. J., & Gerdes, J. (2008). Point-of-care diagnostics for global health. *Annu Rev Biomed Eng, 10*, 107-144. doi:10.1146/annurev.bioeng.10.061807.160524
- Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E., & Famulok, M. (1996). Structural basis of ligand discrimination by two related RNA aptamers resolved by NMR spectroscopy. *Science, 272*(5266), 1343-1347.
- Yuce, M., Ullah, N., & Budak, H. (2015). Trends in aptamer selection methods and applications. *Analyst, 140*(16), 5379-5399. doi:10.1039/c5an00954e
- Yusupova, G. Z., Yusupov, M. M., Cate, J. H., & Noller, H. F. (2001). The path of messenger RNA through the ribosome. *Cell, 106*(2), 233-241.
- Zangi, R., & Filella, M. (2012). Transport routes of metalloids into and out of the cell: a review of the current knowledge. *Chem Biol Interact, 197*(1), 47-57. doi:10.1016/j.cbi.2012.02.001
- Zhang, J., Lan, W., Qiao, C., Jiang, H., Mulchandani, A., & Chen, W. (2004). Bioremediation of organophosphorus pesticides by surface-expressed carboxylesterase from mosquito on Escherichia coli. *Biotechnol Prog, 20*(5), 1567-1571. doi:10.1021/bp049903c

- Zimbres, F. M., Tarnok, A., Ulrich, H., & Wrenger, C. (2013). Aptamers: novel molecules as diagnostic markers in bacterial and viral infections? *Biomed Res Int, 2013*, 731516. doi:10.1155/2013/731516
- Zimmermann, G. R., Jenison, R. D., Wick, C. L., Simorre, J. P., & Pardi, A. (1997). Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nat Struct Biol, 4*(8), 644-649.
- Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science, 244*(4900), 48-52.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res, 31*(13), 3406-3415.

7 Appendix A – Extended materials and methods

7.1 Materials

The materials have been grouped in a series of tables where each table includes materials that fall under a generalised reagent category described by the caption of the table.

Table 7.1: Buffer recipes

Name	Ingredients
Protein Purification Lysis Buffer	50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole (pH8.0)
Protein Purification Wash Buffer	50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole (pH8.0)
Protein Purification Elution Buffer	50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole (pH8.0)
Protein Purification Storage Buffer	50 mM NaH_2PO_4, 150 mM NaCl, (pH8.0)
PBS	50 mM potassium phosphate, pH 7.2; 150 mM NaCl
HEPES buffer (25x)	500mM HEPES pH7.4, 3M Nacl, 125mM KCl, 25mM$CaCl_2$, 12.5% Tween
DNA sample loading buffer (6x)	30% (v/v) glycerol, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol
RNA gel system	5% (w/v) acrylamide, 8M urea, 1x TBE-buffer, 0.17% (w/v) APS, 0.1% (v/v) TEMED
RNA gel running buffer system	1x TBE-buffer
TBE-buffer (5x)	0.45 M Tris-base, 0.45 M Boric Acid, 5 mM Na_2EDTA (pH 8.3)
RNA gel sample loading buffer	3.5 M Urea,0.5x TBE-buffer, 6% (w/v) Ficoll Type 400, 0.005% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol
Laemmli sample loading buffer (2x)	62.5 mM Tris-HCl (pH 6.8), 2% (w/v) SDS, 25% (v/v) glycerol, 0.01% (w/v) bromophenol blue

TGS running buffer	25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3
EMSA gel system	(5 – 12) % (w/v) acrylamide, 1x Buffer system, 0.17% (w/v) APS, 0.1% (v/v) TEMED
HEPES-Imidazole buffer system (5x)	175 mM HEPES, 215 mM Imidazole (pH7.55)
Destain solution	10% (v/v) methanol, 7% (v/v) acetic acid in ddH_2O

Table 7.2: Restriction Enzymes

Material	Source	Catalog Number (where applicable)
RE XbaI RE	NEB	R0145
RE HindIII RE	NEB	R3104
RE DpnI	NEB	R0176
CutSmart Buffer	NEB	B7204S
NEBuffer2.1	NEB	B7202

Table 7.3: Proteins

Material	Source	Catalog Number (where applicable)
Human Fibrinogen (Factor I)	Enzyme Research Laboratories	Fib3
Thrombin from human plasma	Sigma-Aldrich	T6884 SIGMA
VEGF165 human	Sigma-Aldrich	SRP4363 SIGMA

Table 7.4: Chemicals

Material	Source	Catalog Number (where applicable)
1kb Plus DNA ladder	ThermoFischer Scientific	10787018
4–20% Mini-PROTEAN® TGX™ Precast Protein Gels	BioRad	4561095
Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-3 membrane	Merck Millipore	UFC903008
Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-30 membrane	Merck Millipore	UFC900308
Ampicillin sodium salt	Sigma-Aldrich	A0166 SIGMA
APS (Ammonium persulphate)	Sigma-Aldrich	A3678
Benzonase® Nuclease	Sigma-Aldrich	E1014
Boric Acid	Sigma-Aldrich	B7901
Breathe-Easy® sealing membrane	Sigma-Aldrich	Z380059
Chroamphenicol	Sigma-Aldrich	C3175 SIGMA
Corning® 96 well plates polystyrene round bottom, non-treated, well volume 330 µl, sterile	Sigma-Aldrich	CLS3795
Corning® Costar® cell culture plates 96 well, flat bottom, tissue-culture treated	Sigma-Aldrich	CLS3595
GelRed	Cambridge Bioscience	BT41003
IPTG	Sigma-Aldrich	I6758

Lysozyme Solution	ThermoFischer Scientific	90082
Microplate, 384 Well, Ps, F-Bottom, µclear®, Black, Med. Binding	Greiner bio-one	781096
Myc-Tag (9B11) Mouse mAb (PE Conjugate)	NEB	3739
NTA-Ni agarose	Qiagen	30230
PageRuler™ Plus Prestained Protein Ladder	ThermoFischer Scientific	26619
Phusion® High-Fidelity PCR Master Mix with HF Buffer	NEB	M0531
SDS (Sodium dodecyl sulfate)	Sigma-Aldrich	L3771
SimplyBlue™ SafeStain	ThermoFischer Scientific	LC6060
SybrSafe	ThermoFischer Scientific	S33102
SYPRO® Ruby Protein Gel Stain	ThermoFischer Scientific	S12000
T4 DNA Ligase	NEB	M0202
T4 DNA Ligase Reaction Buffer	NEB	B0202
T4 Polynucleotide Kinase	NEB	M0201
TEMED	Sigma-Aldrich	T9281

Table 7.5: Molecular biology kits

Name	Distributor	Catalog Number (where applicable)
MEGAscript® T7 Transcription Kit	ThermoFischer Scientific	AM1333
MEGAclear™ Transcription Clean-Up Kit	ThermoFischer Scientific	AM1908
QIAprep Spin Miniprep Kit	Qiagen	27104
QIAquick Gel Extraction Kit	Qiagen	28704
MinElute Gel Extraction Kit	Qiagen	28604
QIAquick PCR Purification Kit	Qiagen	28104
S30 T7 High-Yield Protein Expression System	Promega	L1110

Table 7.6: Instrumentation

Name	Distributor	Catalog Number (when applicable)
Attune NxT acoustic focusing cytometer	Life Technologies	
BD FACScan flow cytometer	Becton Dickinson	
Dark Reader blue LED transilluminator	Clare Chemical	DR89
Gel Doc XR	BioRad	
POLARstar microplate reader	BMG Labtech	
Vibra Cell VCX 130 model	Sonics & Materials	

Nanodrop 1000 spectrophotomer	Thermo Scientific	

Table 7.7: Software tools

Name	Distributor
Prism 7	GraphPad
FlowJo Vx0.	FlowJo
MATLAB R2016a	MathWorks
ImageJ	National Institute of Health

Table 7.8: Biological parts DNA sequences

Name	Sequence	Source
GFP gene part: E0040	http://parts.igem.org/Part:BBa_E0040	Registry of Standard Biological Parts
Plasmid Backbone: pSB1A2	http://parts.igem.org/Part:pSB1A2	Registry of Standard Biological Parts
Promoter part J23105	tttacggctagctcagtcctaggtactatgctagc	Registry of Standard Biological Parts
Promoter part J23106	tttacggctagctcagtcctaggtatagtgctagc	Registry of Standard Biological Parts
Promoter part: J23101	tttacagctagctcagtcctaggtattatgctagc	Registry of Standard Biological Parts
T7 promoter	taatacgactcactataggg	Registry of Standard Biological Parts
RBS part 12K	aaaggaaatcta	Designed using RBScalculator

RBS part 25K	taggagcgcaaa	Designed using RBScalculator
RBS part 50K	aggagctctcaa	Designed using RBScalculator
RBS part: B0030	attaaagaggagaaa	Registry of Standard Biological Parts
Terminator part: B0015	http://parts.igem.org/Part:BBa_B0015	Registry of Standard Biological Parts

7.2 Plasmid construct sequences

>His-L_GFP /pPROEX-Htb

LacI AmpR ColE1 IPTG inducible expression His-GFP

GTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATAAT
TCGTGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTAATCATCCG
GTCCGTATAATCTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGACCATGTCGTACTACCATCACCATCACCATCACGATTACGATATCCCAACGA
CCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCATGCGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATG
GGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAA
CACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTG
ATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTT
AACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAAC
CATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATG
GATGAACTATACAAATAAAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGA
ATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAG
AGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACA
AATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAGAAGGCC
ATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTC
AATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTG
GTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAAC
GTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGA
ATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCG
GCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCT
GAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTACAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCC
GGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGT
GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAAC
GAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTA
ATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAA
AGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACT
CTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCA
GCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCAC
GCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT
ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTA
CGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCG
CCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGC
ATAATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA
TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACT
ACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAG
CCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCG
CCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTACC

AGTCACGTAGCGATATCGGAGTGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGG
CTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCAG
ATCAATTCGCGCGCGAAGGCGAAGCGGCATGCATTTACGTTGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCA
ATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACG
TTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGA
TTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATG
GTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATG
CCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCG
ACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATA
AATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATC
GTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTGG
GATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAA
CTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCACCCAATACGCAAACCGCCTCTCCCCGCGC
GTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCGCGAATTGATCTG

>12L-GFP2

To assemble plasmid 12L-GFP2, replace His-GFP in His-L_GFP /pPROEX-Htb construct with the sequence given below

His-GFP$_2$12L

ATGTCGTACTACCATCACCATCACCATCACGATTACGATATCCCAACGACCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCATGCGTAAAGGAGAAGAACTT
TTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACT
TACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCAGATCATA
TGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTC
AAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCA
CACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTA
TCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCA
CATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAAGGCCATGGCACCGGCAGCACCGGCAGCGGCAGCAGCATG
CGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGA
TGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTG
CGAGATACCCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTAC
AAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATT
GGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCG
TTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCC
CAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATAAGAATTC

>pNVgfp

LacI CamR pBR322 Plac promoter NVgfp VhhGFP myc-tag E-tag

ACCCGACACCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATAC
GATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGA
TGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCG
TCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCG
GTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGC
GTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCA
GCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGG
AACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGAT
GGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGACATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATC
CCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCT
GTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGG
TTTCCCGACTGGAAAGCGGGCAGTGAGCGGTACCCGATAAAAGCGGCTTCCTGACAGGAGGCCGTTTTGTTTTGCAGCCCACCTCAACGCAATTAATGTGAGTT
AGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACC
ATGATTACGAATTTCTAGATAACGAGGGCAAATCATGATTACTCATGGTTGTTATACCCGGACCCGGCACAAGCATAAGCTAAAAAAAACATTGATTATGCTTAG
TGCTGGTTTAGGATTGTTTTTTTATGTTAATCAGAATTCATTTGCAAATGGTGAAAATTATTTTAAATTGGGTTCGGATTCAAAACTGTTAACTCATGATAGCTATC
AGAATCGCCTTTTTTATACGTTGAAAACTGGTGAAACTGTTGCCGATCTTTCTAAATCGCAAGATATTAATTTATCGACGATTTGGTCGTTGAATAAGCATTTATA
CAGTTCTGAAAGCGAAATGATGAAGGCCGCGCCTGGTCAGCAGATCATTTTGCCACTCAAAAAACTTCCCTTTGAATACAGTGCACTACCACTTTTAGGTTCGGC
ACCTCTTGTTGCTGCAGGTGGTGTTGCTGGTCACACGAATAAACTGACTAAAATGTCCCCGGACGTGACCAAAAGCAACATGACCGATGACAAGGCATTAAATT
ATGCGGCACAACAGGCGGCGAGTCTCGGTAGCCAGCTTCAGTCGCGATCTCTGAACGGCGATTACGCGAAAGATACCGCTCTTGGTATCGCTGGTAACCAGGC
TTCGTCACAGTTGCAGGCCTGGTTACAACATTATGGAACGGCAGAGGTTAATCTGCAGAGTGGTAATAACTTTGACGGTAGTTCACTGGACTTCTTATTACCGTT
CTATGATTCCGAAAAAATGCTGGCATTTGGTCAGGTCGGAGCGCGTTACATTGACTCCCGCTTTACGGCAAATTTAGGTGCGGGTCAGCGTTTTTTCCTTCCTGC
AAACATGTTGGGCTATAACGTCTTCATTGATCAGGATTTTTCTGGTGATAATACCCGTTTAGGTATTGGTGGCGAATACTGGCGAGACTATTTCAAAAGTAGCGT
TAACGGCTATTTCCGCATGAGCGGCTGGCATGAGTCATACAATAAGAAAGACTATGATGAGCGCCCAGCAAATGGCTTCGATATCCGTTTTAATGGCTATCTACC
GTCATATCCGGCATTAGGCGCCAAGCTGATATATGAGCAGTATTATGGTGATAATGTTGCTTTGTTTAATTCTGATAAGCTGCAGTCGAATCCTGGTGCGGCGAC
CGTTGGTGTAAACTATACTCCGATTCCTCTGGTGACGATGGGGATCGATTACCGTCATGGTACGGGTAATGAAAATGATCTCCTTTACTCAATGCAGTTCCGTTA
TCAGTTTGATAAATCGTGGTCTCAGCAAATTGAACCACAGTATGTTAACGAGTTAAGAACATTATCAGGCAGCCGTTACGATCTGGTTCAGCGTAATAACAATAT
TATTCTGGAGTACAAGAAGCAGGATATTCTTTCTCTGAATATTCCGCATGATATTAATGGTACTGAACACAGTACGCAGAAGATTCAGTTGATCGTTAAGAGCAA
ATACGGTCTGGATCGTATCGTCTGGGATGATAGTGCATTACGCAGTCAGGGCGGTCAGATTCAGCATAGCGGAAGCCAAAGCGCACAAGACTACCAGGCTATT
TTGCCTGCTTATGTGCAAGGTGGCAGCAATATTTATAAAGTGACGGCTCGCGCCTATGACCGTAATGGCAATAGCTCTAACAATGTACAGCTTACTATTACCGTT
CTGTCGAATGGTCAAGTTGTCGACCAGGTTGGGGTAACGGACTTTACGGCGGATAAGACTTCGGCTAAAGCGGATAACGCCGATACCATTACTTATACCGCGAC
GGTGAAAAAGAATGGGGTAGCTCAGGCTAATGTCCCTGTTTCATTTAATATTGTTTCAGGAACTGCAACTCTTGGGGCAAATAGTGCCAAAACGGATGCTAACG
GTAAGGCAACCGTAACGTTGAAGTCGAGTACGCCAGGACAGGTCGTCGTGTCTGCTAAAACCGCGGAGATGACTTCAGCACTTAATGCCAGTGCGGTTATATTT
TTTGATGGTGCGCCGGTGCCGTATCCGGATCCGCTGGAACCGGCCCAGCCGGCCATGGCTCAGGTGCAGCTGGTGGAGTCTGGGGGAGCCTTGGTGCAGCCG
GGGGGGTCTCTGAGACTCTCCTGTGCAGCCTCTGGATTCCCCGTCAATCGCTATAGTATGAGGTGGTACCGCCAGGCTCCAGGGAAGGAGCGCGAGTGGGTCG
CGGGTATGAGTAGTGCTGGTGATCGTTCAAGTTATGAAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACGACGCCAGGAATACGGTGTATCTGCAAAT
GAACAGCCTGAAACCTGAGGACACGGCCGTGTATTACTGTAATGTCAATGTGGGCTTTGAGTACTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCAGCGGCC
GCCGAACAAAAACTCATCTCAGAAGAGGATGCAGCTGCATAAAAGCTTGACCTGTGAAGTGAAAAATGGCGCACATTGTGCGACATTTTTTTTGTCTGCCGTTTA
CCGCTACTGCGTCACGGATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTA
GCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCGATTTAGTGCTTTA
CGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTT

AATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATG
AGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATT
TTTCTAAATACATTCAAATATGTATCCGCTCATGTCGAGACGTTGGGTGAGGTTCCAACTTTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGAGTT
ATCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCAT
TTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTAT
TCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAGTTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTT
CCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAA
CCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAAC
TTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTCTGTGATGGCTTCCAT
GTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCCTGGTGCTAC
GCCTGAATAAGTGATAATAAGCGGATGAATGGCAGAAATTCGAAAGCAAATTCGACCCGGTCGTCGGTTCAGGGCAGGGTCGTTAAATAGCCGCTTATGTCTA
TTGCTGGTTTACCGGTTTATTGACTACCGGAAGCAGTGTGACCGTGTGCTTCTCAAATGCCTGAGGCCAGTTTGCTCAGGCTCTCCCCGTGGAGGTAATAATTGC
TCGACATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAAT
CTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCA
GATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCT
GCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAG
CCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCG
GTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGC
GTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATG

>pNVFib1

To assemble plasmid pNVfib1, replace VhhGFP in pNVgfp construct with sequence below

Vhhfib1

ATGGCTCAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTCGGTGCAGGCTGGAGGGTCTCTGAGACTCTCCTGTGCAGCCTCTGGATACACCTACAATAACAACT
GCATGGGCTGGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAGGGGGTCGCAGTTATAAATCGTGGTGGTGGTAGGACATACTACACCGACTCCGTGAAGGGCC
GATTCACCATCTCCCAAGACAACGCCAACAACACGGTGTATCTGACAATGAACAGCCTGAAACCTGAGGACACGGCCATCTATTACTGTGCGGCAAGATGGGG
GTGGGCGTCCAGTAGTAACTGGTACGATATGGGAAAGTATAACTACTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCA

>pNVfib2

To assemble plasmid pNVfib2, replace VhhGFP in pNVgfp construct with sequence below

VhhFib2

CGGCCATGGCTCAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTCGGTGCAGGCTGGGGGGTCTCTGAGACTCTCCTGCACAGCCCCTGGATTCACCCCCAATAC
GTGCGGCGTAGTCTGGAGCCGCCAGGCTGCAGGGAAGCAGCGCGAGTGGGTCTCAAGTATTAAGTCTGATGGCAGCACAACCTATGCAGACTCCGTGAAGGG
CCGATTCACCATTTCCGAAGACAAAGTTAAGGACGTGGTGTATCTGCAAATGAACGGCCTGACACCTGACGACACGGCCATGTATTACTGTTGGAGACGCTGTG
CACCCGAGTCCTGGGGCCGGGGGACCCAGGTCACCGTCTCCTCAGC

>pNK302

To assemble plasmid pNK303-310, replace Bba_J23105 RBS12K in pNK302 construct with sequences from Table 7.8 using Table 2.2 for guidance.

LacI CamR pBR322 Bba_J23105 RBS12K NVgfp VhhGFP myc-tag E-tag

TTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCAAAGGAAATCTAATGATTACTCATGGTTGTTATACCCGGACCCGGCACAAGCATAAGCTAAAAAAAACAT
TGATTATGCTTAGTGCTGGTTTAGGATTGTTTTTTTATGTTAATCAGAATTCATTTGCAAATGGTGAAAATTATTTTAAATTGGGTTCGGATTCAAAACTGTTAACT
CATGATAGCTATCAGAATCGCCTTTTTTATACGTTGAAAACTGGTGAAACTGTTGCCGATCTTTCTAAATCGCAAGATATTAATTTATCGACGATTTGGTCGTTGA
ATAAGCATTTATACAGTTCTGAAAGCGAAATGATGAAGGCCGCGCCTGGTCAGCAGATCATTTTGCCACTCAAAAAACTTCCCTTTGAATACAGTGCACTACCAC
TTTTAGGTTCGGCACCTCTTGTTGCTGCAGGTGGTGTTGCTGGTCACACGAATAAACTGACTAAAATGTCCCCGGACGTGACCAAAAGCAACATGACCGATGAC
AAGGCATTAAATTATGCGGCACAACAGGCGGCGAGTCTCGGTAGCCAGCTTCAGTCGCGATCTCTGAACGGCGATTACGCGAAAGATACCGCTCTTGGTATCGC
TGGTAACCAGGCTTCGTCACAGTTGCAGGCCTGGTTACAACATTATGGAACGGCAGAGGTTAATCTGCAGAGTGGTAATAACTTTGACGGTAGTTCACTGGACT
TCTTATTACCGTTCTATGATTCCGAAAAAATGCTGGCATTTGGTCAGGTCGGAGCGCGTTACATTGACTCCCGCTTTACGGCAAATTTAGGTGCGGGTCAGCGTT
TTTTCCTTCCTGCAAACATGTTGGGCTATAACGTCTTCATTGATCAGGATTTTTCTGGTGATAATACCCGTTTAGGTATTGGTGGCGAATACTGGCGAGACTATTT
CAAAAGTAGCGTTAACGGCTATTTCCGCATGAGCGGCTGGCATGAGTCATACAATAAGAAAGACTATGATGAGCGCCCAGCAAATGGCTTCGATATCCGTTTTA
ATGGCTATCTACCGTCATATCCGGCATTAGGCGCCAAGCTGATATATGAGCAGTATTATGGTGATAATGTTGCTTTGTTTAATTCTGATAAGCTGCAGTCGAATCC
TGGTGCGGCGACCGTTGGTGTAAACTATACTCCGATTCCTCTGGTGACGATGGGGATCGATTACCGTCATGGTACGGGTAATGAAAATGATCTCCTTTACTCAAT
GCAGTTCCGTTATCAGTTTGATAAATCGTGGTCTCAGCAAATTGAACCACAGTATGTTAACGAGTTAAGAACATTATCAGGCAGCCGTTACGATCTGGTTCAGCG
TAATAACAATATTATTCTGGAGTACAAGAAGCAGGATATTCTTTCTCTGAATATTCCGCATGATATTAATGGTACTGAACACAGTACGCAGAAGATTCAGTTGAT
CGTTAAGAGCAAATACGGTCTGGATCGTATCGTCTGGGATGATAGTGCATTACGCAGTCAGGGCGGTCAGATTCAGCATAGCGGAAGCCAAAGCGCACAAGAC
TACCAGGCTATTTTGCCTGCTTATGTGCAAGGTGGCAGCAATATTTATAAAGTGACGGCTCGCGCCTATGACCGTAATGGCAATAGCTCTAACAATGTACAGCTT
ACTATTACCGTTCTGTCGAATGGTCAAGTTGTCGACCAGGTTGGGGTAACGGACTTTACGGCGGATAAGACTTCGGCTAAAGCGGATAACGCCGATACCATTAC
TTATACCGCGACGGTGAAAAAGAATGGGGTAGCTCAGGCTAATGTCCCTGTTTCATTTAATATTGTTTCAGGAACTGCAACTCTTGGGGCAAATAGTGCCAAAA
CGGATGCTAACGGTAAGGCAACCGTAACGTTGAAGTCGAGTACGCCAGGACAGGTCGTCGTGTCTGCTAAAACCGCGGAGATGACTTCAGCACTTAATGCCAG
TGCGGTTATATTTTTGATGGTGCGCCGGTGCCGTATCCGGATCCGCTGGAACCGGCCCAGCCGGCCATGGCTCAGGTGCAGCTGGTGGAGTCTGGGGGAGCC
TTGGTGCAGCCGGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTCTGGATTCCCCGTCAATCGCTATAGTATGAGGTGGTACCGCCAGGCTCCAGGGAAGGAGC
GCGAGTGGGTCGCGGGTATGAGTAGTGCTGGTGATCGTTCAAGTTATGAAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACGACGCCAGGAATACGGT
GTATCTGCAAATGAACAGCCTGAAACCTGAGGACACGGCCGTGTATTACTGTAATGTCAATGTGGGCTTTGAGTACTGGGGCCAGGGGACCCAGGTCACCGTC
TCCTCAGCGGCCGCCGAACAAAAACTCATCTCAGAAGAGGATGCAGCTGCATAAAAGCTTGACCTGTGAAGTGAAAAATGGCGCACATTGTGCGACATTTTTTT
TGTCTGCCGTTTACCGCTACTGCGTCACGGATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTT
GCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTCCG
ATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGA
GTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATT
GGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCC
CTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGTCGAGACGTTGGGTGAGGTTCCAACTTTCACCATAATGAAATAAGATCACTACCGGGCG
TATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAAC
ATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTA
TCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAGTTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTG
TTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGT
TACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTGGCCA
ATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTCTGTG
ATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAAACG

CCTGGTGCTACGCCTGAATAAGTGATAATAAGCGGATGAATGGCAGAAATTCGAAAGCAAATTCGACCCGGTCGTCGGTTCAGGGCAGGGTCGTTAAATAGCC
GCTTATGTCTATTGCTGGTTTACCGGTTTATTGACTACCGGAAGCAGTGTGACCGTGTGCTTCTCAAATGCCTGAGGCCAGTTTGCTCAGGCTCTCCCCGTGGAG
GTAATAATTGCTCGACATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTT
CTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTC
AGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTG
TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTT
CGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGG
ACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACC
TCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTT
TTGCTCACATG
ACCCGACACCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATAC
GATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGA
TGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCG
TCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCG
GTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGC
GTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCA
GCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGG
AACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGAT
GGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGACATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATC
CCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCT
GTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGG
TTTCCCGACTGGAAAGCGGGCAGTGAGCGGTACCCGATAAAAGCGGCTTCCTGACAGGAGGCCGTTTTGTTTTGCAGCCCACCTCAACGCAATTAATGTGAGTT
AGCTCACTCATTAGGCACCCCAGGC

>pNK400

To assemble pNK400 construct, insert sequence below in vector backbone pSB1C3 (Table 7.8).

BBa_J23101 BBa_B0034 BBa_E0040 BBa_B0015

TTTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCTACTAGAGAAAGAGGAGAAATACTAGATGCGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTC
TTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTA
CTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCAGATCATATGAAACAGCATGACTTTTTCAA
GAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTG
TTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAG
ACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGC
GATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAA
CAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTC
GTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA

>pNK401

To assemble pNK401 construct, insert sequence below in vector backbone pSB1A2 (Table 7.8).

T7 promoter BBa_B0030 BBa_E0040 BBa_B0015

TAATACGACTCACTATAGGGTACTAGAGATTAAAGAGGAGAAATACTAGATGCGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAG
ATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTA
CCTGTTCCATGGCCAACACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCATGC
CCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATC
GAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAA
GAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTG
TCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGG
GATTACACATGGCATGGATGAACTATACAAATAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGT
TGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA

>pNK402

To assemble pNK402 construct, insert sequence below in vector backbone pSB1A2 (Table 7.8).

T7 promoter BBa_B0030 BBa_E0040 BBa_B0015 anti-thrombin RNA-aptamer

TAATACGACTCACTATAGGGTACTAGAGATTAAAGAGGAGAAATACTAGATGAATTTGGAAGATAGCTGGAGAACTAACCAAATTCGTAAAGGAGAAGAACTT
TTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACT
TACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCAGATCATA
TGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTC
AAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCA
CACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTA
TCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCA
CATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATAATAATACTAGAGCCAGGCATCAAATAAAACGAAAGGCT
CAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA

> pNK403-416

To assemble pNK403-416 constructs replace ATGAAanti-thrombin RNA-aptamerTTCGT in pNK402 with sequences in Table 4.1, Table 4.2 and Table 4.3 according to guidance from Table 2.3.

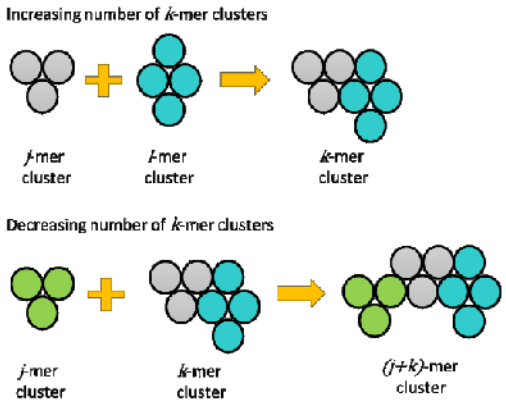
8 Appendix B – Agglutination reaction mathematical model

This section provides relevant parts from (Hung El, 2014) as these are related to section 3.2.5 of this document. More specifically, this section provides the derivation of the bacterial agglutination reaction mathematical model, the model simplification, the parameter values used and the Matlab script for simulating the agglutination reaction. The script includes input parameter values from this study where these are different from (Hung El, 2014).

Model derivation

Equation 1: Cell clusters aggregation dynamics

The time-dependent amount of certain clusters of *k* cells (or *k*-mer cluster) ($n_k(t)$), where *k* is a positive whole integer, is affected by several parameters. Since the interaction between bound ligand and free receptor is a prerequisite of cell aggregation, the dynamics of cell aggregation is dependent on the rate constants associated with aggregation, bound ligand density, free receptor density and the total number of ways involved in constructing and deconstructing *k*-mer clusters.



Considering the event of increasing number of *k*-mer clusters by coalescence of two clusters of size *j*-mer and *l*-mer, and the event of decreasing number of *k*-mer clusters by coalescence of two clusters of size *k*-mer and *j*-mer to form a (*j*+*k*)-mer cluster, as shown in **Figure 3.1**.

Figure 3.1: Schematic diagram describing the construction and deconstruction of a k-mer cluster. Each circle is a cell, and all cells are identical; color code is just for easier visualization of the changes of number of cells in a cluster.

The time-dependent evolution of *k*-mer clusters can be described in a Smoluchowski kinetic equation[39,40] (that describes kinetics of collision between colloidal particles) as shown by Equation (1) below.

$$\frac{dn_k(t)}{dt} = \begin{cases} C_1(a(t))[1-g(a(t))] \times \frac{1}{2}\sum_{j+l=k}^{k-1} K(n_j(t), n_l(t)) - n_k(t)\sum_{j=1}^{\infty} K(n_j(t)), & if\ g(a(t)) < a(t) < 1-g(a(t)) \\ 0, & otherwise \end{cases} \quad (1)$$

A threshold value, *g*, is included such that if the bound ligand density $a(t)$ and the free receptor density $1-a(t)$ falls below the threshold, i.e. $a(t) < g$ and $1-a(t) < g$ respectively, the tethering forces between the existing bound ligand and free receptor will not be sufficient to maintain a stable aggregate[27,41]. Under this condition, the $\frac{dn_k(t)}{dt}$ will simply be 0, since no new k-

mer clusters will be formed, for both the construction and the deconstruction of k-mer clusters. However in our model this threshold is not considered (i.e. setting $g = 0$) due to difficulty in experimentally characterizing the aggregation behaviour.

The first part of the non-zero Equation (1)

$$C_1(a_t)[1 - a_t]$$

describes the rate of aggregation between cells with bound ligand density of (a_t) and cells with free receptor density of ($1 - a_t$) (unbound ligand = free receptor). This can be viewed as a second-order reaction with rate constant C_1 and the two interacting species. The value of a_t ranges from 0 to 1 as it can be described as the ratio of number of bound ligand to total number of receptor on cell surface. The product $a_t[1 - a_t]$ is proportional to the number of mutual pairs of bound ligand (a_t) and free receptor ($1 - a_t$) on cell surface, since the aggregation event requires the binding interaction between a bound ligand and a free receptor on cell surface. C_1 is a kinetic constant related to the aggregation dynamics of two clusters of cells. Currently this value cannot be estimated from experimental characterization as it is difficult to have a quantitative measurement of aggregation.

The second part of Equation (1)

$$\frac{1}{2}\sum_{j=1}^{k-1} K_{j,l}(n_j, t)(n_l, t)$$

describes all possible events that generate clusters of *k* cells (*k*-mer cluster) by the collision of two clusters consisting *j* and *l* cells where *k=j+l*. Since the largest value of *j* is *k-1* (therefore *l=1*), the sum upper limit is *k-1* rather than *k*, as *l=0* is meaningless (minimum of *l* is 1).

A halving factor, ½, is introduced to avoid double-counting of ways constructing *k*-mer clusters, as the sum has already considered all possible ways to construct the *k*-mer cluster twice[42]. For example, assuming that *k*=5, there are 2 possible ways to form a 5-mer cluster: the coalescence of a 1-cell and a 4-cell clusters, and that of a 2-cell and a 3-cell clusters. If we follow the sum *j=1* to *k-1* strictly, it would give 4 possible ways, as in

$$\sum_{j=1}^{4} K_{j,l}(n_j, t)(n_l, t) = (n_j(1)\,n_l(4) + n_j(2)\,n_l(3) + n_j(3)\,n_l(2) + n_j(4)\,n_l(1))$$

4 apparent ways to construct a 5-mer cluster

However, it is important to note that there is no difference between the coalescence of a 1-cell and a 4-cell cluster and that of a 4-cell and a 1-cell cluster. Similarly a 2-cell coalescing with a 3-cell cluster is the same if the order is reversed. In other words, the order of coalescing cluster in an aggregation event is not important. Therefore, the actual number of ways to construct a 5-mer cluster is 2, instead of 4. So to correct for this double counting inherent in the sum, a halving factor must be included.

The third part of Equation (1)

$$-n(k,t)\sum_{j=1}^{\infty}[K_{jk}n(j,t)]$$

describes events that decrease *k*-mer clusters by the collision of with other clusters of any size (represented by sum of *j*=1 to infinity). For example, when a 5-mer (*k*) cluster coalesce with a 3-mer (*j*) cluster, this produces a (5+3=8)-mer cluster but reducing the number of 5-mer clusters in the system. Since this decreases the amount of *k*-mer clusters in the system, a negative sign is required.

The kernel of Equation (1)

$$K_{jl}=\frac{(D_j+D_l)(R_j+R_l)}{4D_1R_1} \quad ; \quad K_{jk}=\frac{(D_j+D_k)(R_j+R_k)}{4D_1R_1}$$

Represents the probability of aggregation in the event of collision between clusters of size *j* & *l*, and *j* & *k* respectively. To simplify the model, it can be assumed that the two clusters can always aggregate when they collide (otherwise the model will be overly complicated), and the rate of collision is only weakly dependent on the aggregate size, such that $K_{jl} \approx K_{jk} \approx 1$[27,43].

Equation 2: Free ligand dynamics

$$\frac{dn_l(t)}{dt}=-C_2n_l(t)[1-a(t)]f\sum_{k=1}^{\infty}[n(k,t)k]+C_3a(t)f\sum_{k=1}^{\infty}[n(k,t)k]$$

The first part of Equation (2)

$$-C_2n_l(t)[1-a(t)]f\sum_{k=1}^{\infty}[n(k,t)k]$$

describes events that decrease free ligand in the system via binding to receptor on cell surface. This can be viewed as a second-order rate reaction with rate constant C_2 and the two interacting

species, free ligand ($n_l(t)$) and total number of free receptors on surface of all cells. The latter is a product of three components: free receptor density, $1 - a(t)$, average number of receptor per cell, f, and the total number of cells, $\sum_{k=1}^{\infty}[n(k,t)k]$, which is equivalent to the total sum of all products between k and number of k-mer clusters ($n(k,t)$), for which k=1 to infinity. Since the binding reaction decreases the amount of free ligand in the system, a negative sign is required.

The second part of Equation (2)

$$+C_3 a(t) f \sum_{k=1}^{\infty}[n(k,t)k]$$

describes events that increase free ligand in the system via dissociation from bound receptor on cell surface. This can be viewed as a first-order rate reaction with rate constant C_3 and the only species, bound ligand, which is the product of bound ligand density, $a(t)$, average number of receptor per cell, f, and the total number of cells, $\sum_{k=1}^{\infty}[n(k,t)k]$. Since the dissociation increases the amount of free ligand in the system, a positive sign is included.

Equation 3: Bound ligand dynamics (on cell surface)

$$\frac{da(t)}{dt} = C_2 n_l(t)[1 - a(t)] - C_3 a(t)$$

The first part of Equation (3)

$$C_2 n_l(t)[1 - a(t)]$$

describes events that increase bound ligand in the system via free ligand binding to receptor on cell surface. This can be viewed as a second-order reaction with rate constant C_2 and the two interacting species, free ligand ($n_l(t)$) and free receptor density, $1 - a(t)$. Note that this only considers events that occur on surface of individual cells rather than the whole system, therefore the total number of free receptor is not required here.

The second part of Equation (3)

$$-C_3 a(t)$$

describes events that decrease bound ligand in the system via dissociation from receptor on cell surface. This is simply a first-order reaction with rate constant C_3 and only one species, bound ligand density $a(t)$. Since this decreases the amount of bound ligand, a negative sign is required.

Description of parameters

A list of parameters and their corresponding descriptions are as shown below in **Table 4**. Typical value of certain parameters are also included where appropriate, with references indicated in square brackets.

Table 4: List of parameters used in model derivation.

Parameter	Description	Typical value [Ref]
$n(k,t)$	Time-dependent amount of k-mer cluster, where j + l = k	N/A
$n(j,t)$	Time-dependent amount of j-mer cluster, where j + l = k	N/A
$n(l,t)$	Time-dependent amount of l-mer cluster, where j + l = k	N/A
C_1	Kinetic constant related to aggregation interaction between aggregating clusters	1 [a]
C_2	Kinetic constant related to binding interaction between free ligand and receptor, equivalent to the on-rate of nanobody-GFP interaction	7.68 x 10^5 M^{-1} s^{-1} [29]
C_3	Kinetic constant related to dissociation of bound ligand from receptor, equivalent to the off-rate of nanobody-GFP interaction	1.74 x 10^{-4} s^{-1} [29]
$a(t)$	Time-dependent bound ligand density (per cell), where $a(t) \in [0\ 1]$	N/A
$1 - a(t)$	Time-dependent free receptor density (per cell)	N/A
K_{jl}	Probability of aggregation in the event of collision between a *j*-mer cluster and a *l*-mer cluster	1 [27]
K_{jk}	Probability of aggregation in the event of collision between a *j*-mer cluster and a *k*-mer cluster	1 [27]
D_j	Diffusion coefficient of a *j*-mer cluster, where $\mathrm{D_j} = D_1 R_1 / \mathrm{R_j}$	N/A
D_l	Diffusion coefficient of a *l*-mer cluster, where $\mathrm{D_l} = D_1 R_1 / \mathrm{R_l}$	N/A
D_k	Diffusion coefficient of a *k*-mer cluster, where $\mathrm{D_k} = D_1 R_1 / \mathrm{R_k}$	N/A
D_1	Diffusion coefficient of one cell	4 x 10^{-10} m^2/s [44]
R_j	Radius of a *j*-mer cluster, where $R_j = R_1 \sqrt[3]{j}$	N/A
R_l	Radius of a *l*-mer cluster, where $R_l = R_1 \sqrt[3]{l}$	N/A
R_k	Radius of a *k*-mer cluster, where $R_k = R_1 \sqrt[3]{k}$	N/A
R_1	Radius of one cell	5 x 10^{-7} m [45]
$n_l(t)$	Time-dependent amount of free ligand in the system	N/A
f	Average number of receptors per cell on cell surface	10^4 [b]
N_0	Total number of cells in system	10^6 to 10^9 [c]
N_l	Total number of free ligand in system	10^{10} to 10^{13} [d]
$\mathrm{N}(t)$	Normalized number of aggregated particles	N/A

Notes:

[a]	C_1 is set to 1 as the default value, as there is insufficient experimental characterization data for estimating the kinetic constant of aggregation reaction.
[b]	Determined from nanobody quantification assays performed in this project.
[c]	This range is selected to correspond to the typical value of number of cells that allows visualization of cell pellets in round-bottom wells in agglutination assays
[d]	This range is selected to correspond to the typical conditions tested in agglutination assays

Model simplification

The section aims to simplify the three fundamental equations described above into two, the first one governs time-dependent evolution of normalized number of aggregating particles, while the second governs time-dependent evolution of bound ligand density. This system of two ordinary differential equations is then solved and analyzed with MATLAB® to determine the aggregation output given the input variables (as elaborated in Results and Discussion).

Time-dependent evolution of normalized number of aggregating particles, $\frac{dN(t)}{dt}$

Considering the non-zero function of Equation (1) only:

$$\frac{dn(k,t)}{dt} = C_1 a(t)[1-a(t)] \times \left\{ \frac{1}{2} \sum_{\substack{j=1 \\ j+l=k}}^{k-1} [K_{jl} n(j,t) n(l,t)] - n(k,t) \sum_{j=1}^{\infty} [K_{jk} n(j,t)] \right\}$$

Since it can be assumed that $K_{jl} \approx K_{jk} \approx 1$

$$\frac{dn(k,t)}{dt} = C_1 a(t)[1-a(t)] \times \left\{ \frac{1}{2} \sum_{\substack{j=1 \\ j+l=k}}^{k-1} [n(j,t) n(l,t)] - n(k,t) \sum_{j=1}^{\infty} [n(j,t)] \right\}$$

Equation (1.1)

Let $N(t)$ be the normalized number of aggregating particles such that

$$N(t) = \sum_{k=1}^{\infty} n(k,t)$$

Equation (1.2)

Summing both sides of Equation (1.1) from k=1 to ∞,

$$\sum_{k=1}^{\infty} \frac{dn(k,t)}{dt} = \sum_{k=1}^{\infty} \left[C_1 a(t)[1-a(t)] \times \left\{ \frac{1}{2} \sum_{\substack{j=1 \\ j+l=k}}^{k-1} [n(j,t) n(l,t)] - n(k,t) \sum_{j=1}^{\infty} [n(j,t)] \right\} \right]$$

Rearranging the sum notations to their relevant summands,

$$\sum_{k=1}^{\infty}\frac{dn(k,t)}{dt}=C_1a(t)[1-a(t)]\times\left\{\frac{1}{2}\sum_{k=1}^{\infty}\sum_{\substack{j=1\\ j+l=k}}^{k-1}[n(j,t)n(l,t)]-\sum_{k=1}^{\infty}n(k,t)\sum_{j=1}^{\infty}n(j,t)\right\}$$

Equation (1.3)

On the LHS, summing the differential terms give

$$\sum_{k=1}^{\infty}\frac{dn(k,t)}{dt}=\frac{dN(t)}{dt}$$

On the RHS, looking at the double summation, since $l=k-j$,

$$\sum_{k=1}^{\infty}\sum_{\substack{j=1\\ j+l=k}}^{k-1}[n(j,t)n(l,t)]=\sum_{k=1}^{\infty}\sum_{\substack{j=1\\ j+l=k}}^{k-1}[n(j,t)n(k-j,t)]$$

Expanding the inner summation series over j=1 to j=k-1 (omitting *t*),

$$\sum_{k=1}^{\infty}\sum_{\substack{j=1\\ j+l=k}}^{k-1}[n(j)n(k-j)]=\sum_{k=1}^{\infty}[n(1)n(k-1)+n(2)n(k-2)+\cdots+n(k-2)n(2)+n(k-1)n(1)]$$

Note that, since k only takes positive integers, therefore $n(k)=0$ for $k\leq 0$

With this in mind, we can now expand the outer summation series over k=1 to infinity (omitting *t*),

$$\begin{aligned}
&\sum_{k=1}^{\infty}[n(1)n(k-1)+n(2)n(k-2)+\cdots+n(k-2)n(2)+n(k-1)n(1)]\\
&=\ n(1)n(0)+n(2)n(-1)+n(3)n(-2)+\cdots\\
&\quad +n(1)n(1)+n(2)n(0)+n(3)n(-1)+\cdots\\
&\quad +n(1)n(2)+n(2)n(1)+n(3)n(0)+\cdots\\
&\quad +n(1)n(3)+n(2)n(2)+n(3)n(1)+\cdots\\
&\quad +\cdots\\
&=\ n(1)\cdot 0+n(2)\cdot 0+n(3)\cdot 0+\cdots\\
&\quad +n(1)n(1)+n(2)\cdot 0+n(3)\cdot 0+\cdots\\
&\quad +n(1)n(2)+n(2)n(1)+n(3)\cdot 0+\cdots\\
&\quad +n(1)n(3)+n(2)n(2)+n(3)n(1)+\cdots\\
&\quad +\cdots
\end{aligned}$$

$$= n(1)[n(1)+n(2)+n(3)+\cdots]+n(2)[n(1)+n(2)+n(3)+\cdots]$$
$$+n(3)[n(1)+n(2)+n(3)+\cdots]+\cdots$$
$$= [n(1)+n(2)+n(3)+\cdots]\times[n(1)+n(2)+n(3)+\cdots]$$
$$=\sum_{j=1}^{\infty} n(j)\cdot\sum_{k-j=1}^{\infty} n(k-j)$$
$$=\sum_{j=1}^{\infty} n(j)\cdot\sum_{l=1}^{\infty} n(l)$$

Therefore from Equation (1.3),

$$\frac{dN(t)}{dt}=C_1a(t)[1-a(t)]\times\left[\frac{1}{2}\sum_{j=1}^{\infty} n(j,t)\sum_{l=1}^{\infty} n(l,t)-\sum_{k=1}^{\infty} n(k,t)\sum_{j=1}^{\infty} n(j,t)\right]$$

Applying Equation (1.2),

$$\frac{dN(t)}{dt}=C_1a(t)[1-a(t)]\times\left[\frac{1}{2}N(t)\cdot N(t)-N(t)\cdot N(t)\right]$$

$$\therefore\frac{dN(t)}{dt}=C_1a(t)[1-a(t)]\left(-\frac{1}{2}N^2(t)\right)$$

Equation (4)

Time-dependent evolution of bound ligand density, $\frac{da(t)}{dt}$

From Equation (2),

$$\frac{dn_l(t)}{dt}=-C_2n_l(t)[1-a(t)]f\sum_{k=1}^{\infty}[n(k,t)k]+C_3a(t)f\sum_{k=1}^{\infty}[n(k,t)k]$$

Conservation law of cells apply, where it is assumed that in this system cells do not divide nor die, which means that the total number of cells, either freely diffusing or in aggregating particles, remain the same.

Conservation law of cells can be described by:

$$\sum_{k=1}^{\infty}[n(k,t)k]=N_0\ ,\qquad \forall t\geq 0$$

Equation (2.1)

N_0 is the total number of cells. This summation means that the sum of the products of the number of cells in a k-mer cluster and the number of k-mer clusters, would be equal to the total number of cells.

Substituting Equation (2.1) into Equation (2), and extracting the common factor $N_0 f$,

$$\frac{dn_l(t)}{dt} = N_0 f[-C_2 n_l(t)[1 - a(t)] + C_3 a(t)]$$

Equation (2.2)

However, the terms in the outermost square bracket in Equation (2.2) is just the negative of Equation (3).

$$-\frac{da(t)}{dt} = -C_2 n_l(t)[1 - a(t)] + C_3 a(t)$$

Equation (2.3)

Substituting Equation (2.3) into Equation (2.2)

$$\therefore \frac{dn_l(t)}{dt} = -N_0 f \frac{da(t)}{dt}$$

Equation (2.4)

It would be helpful to express $n_l(t)$ in terms of $a(t)$ to combine Equations (2) and (3) into one.

Let

$$s = \frac{N_0 f}{N_l}$$

Equation (2.5)

Substituting Equation (2.5) into Equation (2.4),

$$\frac{dn_l(t)}{dt} = -sN_l \frac{da(t)}{dt}$$

Equation (2.6)

Integrate Equation (2.6) both sides with respect to dt:

$$\int \frac{dn_l(t)}{dt} dt = -sN_l \int \frac{da(t)}{dt} dt$$

$$n_l(t) = -sN_l\,[a(t) + C]$$

Equation (2.7)

Normalizing $n_l(t)$:

$$\tilde{n}_l(t) = \frac{n_l(t)}{N_l}$$

Equation (2.8)

Such that Equation (2.7) becomes

$$\tilde{n}_l(t) = -s\,a(t) + C$$

Equation (2.9)

To calculate C, let $t = 0$.

Since initially, number of free ligand is equal to total number of ligand, thus

$$\tilde{n}_l(0) = \frac{n_l(0)}{N_l} = \frac{N_l}{N_l} = 1$$

Equation (2.10)

Since initially, there is no ligand bound on cell surface, thus

$$a(0) = 0$$

Equation (2.11)

Substituting Equation (2.10) and Equation (2.11) into Equation (2.9)

$$\tilde{n}_l(0) = -sa(0) + C$$

$$1 = -s(0) + C$$

$$C = 1$$

Equation (2.12)

Substituting Equation (2.12) into Equation (2.9)

$$\tilde{n}_l(t) = 1 - sa(t)$$

Equation (2.13)

Substituting Equation (2.8) into Equation (2.13)

$$n_l(t) = N_l[1 - sa(t)]$$

Equation (2.14)

Now we can use this to combine Equation (2) and (3) into one.

From Equation (3),

$$\frac{da(t)}{dt} = C_2 n_l(t)[1 - a(t)] - C_3 a(t)$$

Substituting Equation (2.14) into Equation (3),

$$\frac{da(t)}{dt} = C_2 N_l[1 - sa(t)][1 - a(t)] - C_3 a(t)$$

Equation (3.1)

Expanding the factors and extracting the common factors

$$\frac{da(t)}{dt} = C_2 N_l \left[sa^2(t) - \left(1 + s + \frac{C_3}{C_2 N_l}\right) a(t) + 1 \right]$$

Equation (3.2)

Let

$$\alpha = \frac{C_3}{C_2 N_l}$$

Equation (3.3)

Substituting Equation (3.3) into Equation (3.2)

$$\therefore \frac{da(t)}{dt} = C_2 N_l[sa^2(t) - (1 + s + \alpha)a(t) + 1]$$

Equation (5)

Matlab Script

```
%defining parameters
%C1=aggregation rate constant
%C2=association rate constant
%C3=dissociation rate constant
%N0=Initial number of cells
%Nl=Initial number of free ligands
%f=average number of receptors/nanobody per cell
%g=threshold for binding

C1=1;
C2=3600*7.68*10^5;
C3=3600*1.74*10^-4;
N0=0.210^9;
iteration=1000;
Nl=10^10*(1:iteration);
Nl_nM=Nl.*10^9./(6.023*10^23*20*10^-6);
f=6750;
Tspan=[0 18];

s=f.*N0./Nl;
a=C3./(C2.*Nl);
Nmat=zeros(length(s),1);
constants=[C1,C2,N0,f];
opt=odeset('AbsTol',1e-6,'RelTol',1e-3,'MaxStep',0.5);

parfor i=1:length(s)
    s_i=s(i);
    a_i=a(i);
    Nl_i=Nl(i);
    wrapper=@(t,y)smo(t,y,s_i,C1,C2,a_i,Nl_i);
[t,y]=ode23s(wrapper,Tspan,[1,0]',opt);

Nmat(i,1)=y(end,1);
end
warning('off','all');
semilogx(Nl,1-Nmat(:,1))
title('Cell Aggregation Model');
xlabel('Number of GFP dimer');
ylabel('Normalised aggregation, N');
end

function dydt=smo(t,y,s_i,C1,C2,a_i,Nl_i)
    dydt=[C1*y(2)*(1-y(2))*(-0.5*(y(1).^2));
    C2*Nl_i*(s_i*y(2)^2-(1+s_i+a_i)*y(2)+1)];
end
```

9 Appendix C – RNA sequences folding analysis

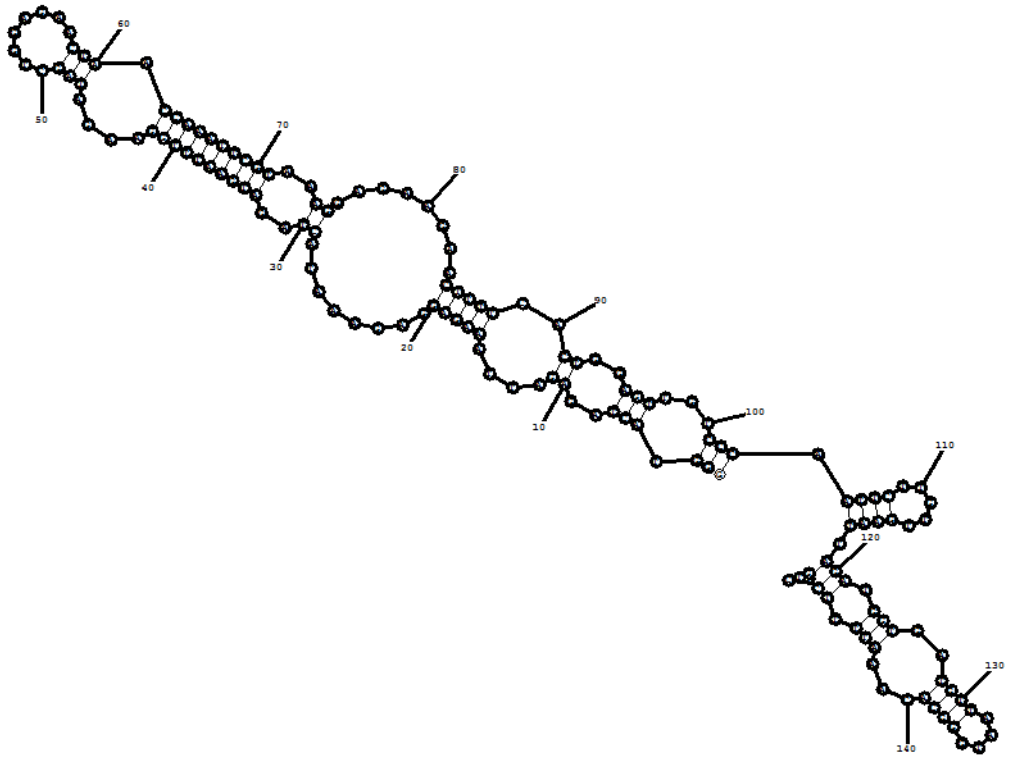
This section provides extended information on the experimental results of section 4.2.8, which analysed the secondary structure folding of designed RNA-switch devices (~150 nucleotides partial sequences). The secondary structure analysis was carried out with the Fold algorithm of the RNAstructure software package. The information provided in this section for each designed RNA-switch device includes: (A) the input sequence for secondary structure folding, (B) the output image of the minimum free energy structure for the folded sequence and (C) the constraints used (where applicable) to force formation the RNA-aptamer secondary structure (Figure 4.2) when this was not present in the minimum free energy structure.

>Thrombin switch device v01|150 nts partial sequence|

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGAATTTGGAAGATAGCTGGAGAACTAACCAAATTCGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



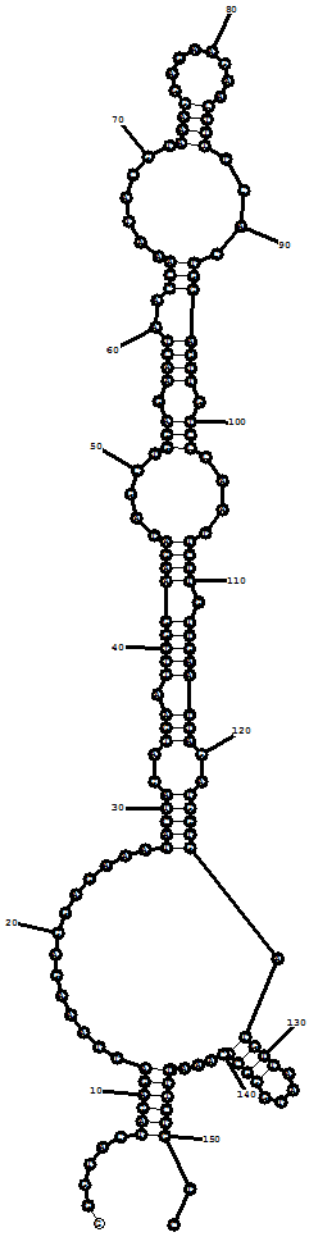
ENERGY = -24.2 device v01

>Thrombin switch device v02|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGAATTTGGAAGATAGCTGGAGAACTAACCAAATACGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



ENERGY = -20.0 device v02

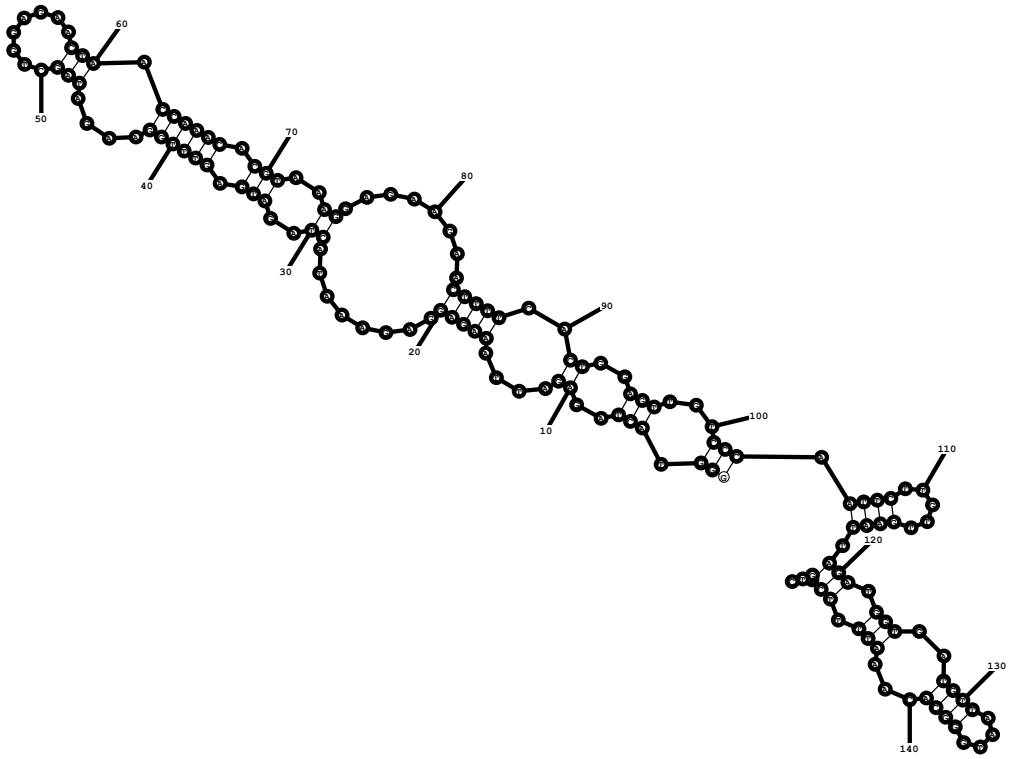
Folding constraints: 38:66, 39:65; 40:64, 41:63, 42:62, 47:60, 48:59, 39:58

>Thrombin switch device v03|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGAGTTTGGAAGATAGCTGGAGAACTAACCAAACACGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



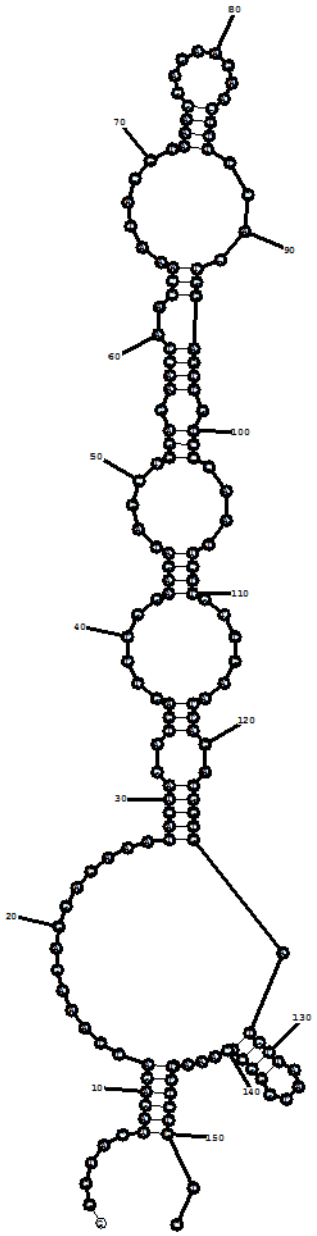
ENERGY = -21.1 device v03

>Thrombin switch device v04|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGAGTTTGGAAGATAGCTGGAGAACTAACCAAATACGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



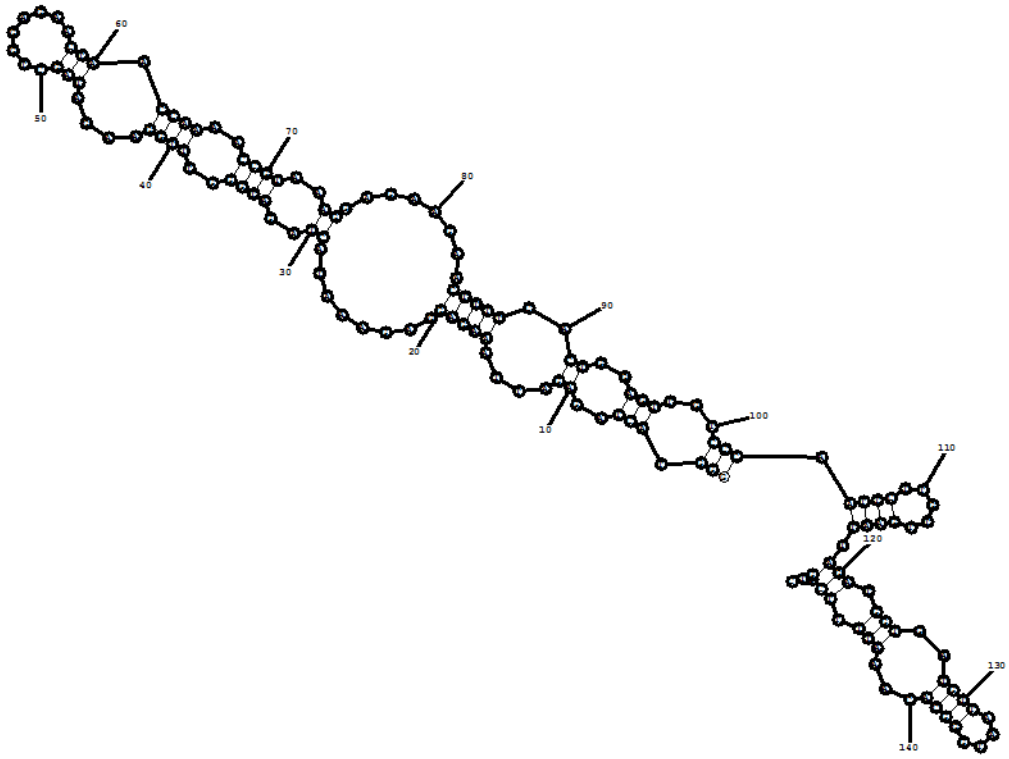
Folding constraints: 38:66, 39:65; 40:64, 41:63, 42:62, 47:60, 48:59, 39:58

>Thrombin switch device v05|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGGGTTTGGAAGATAGCTGGAGAACTAACCAAAACCGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



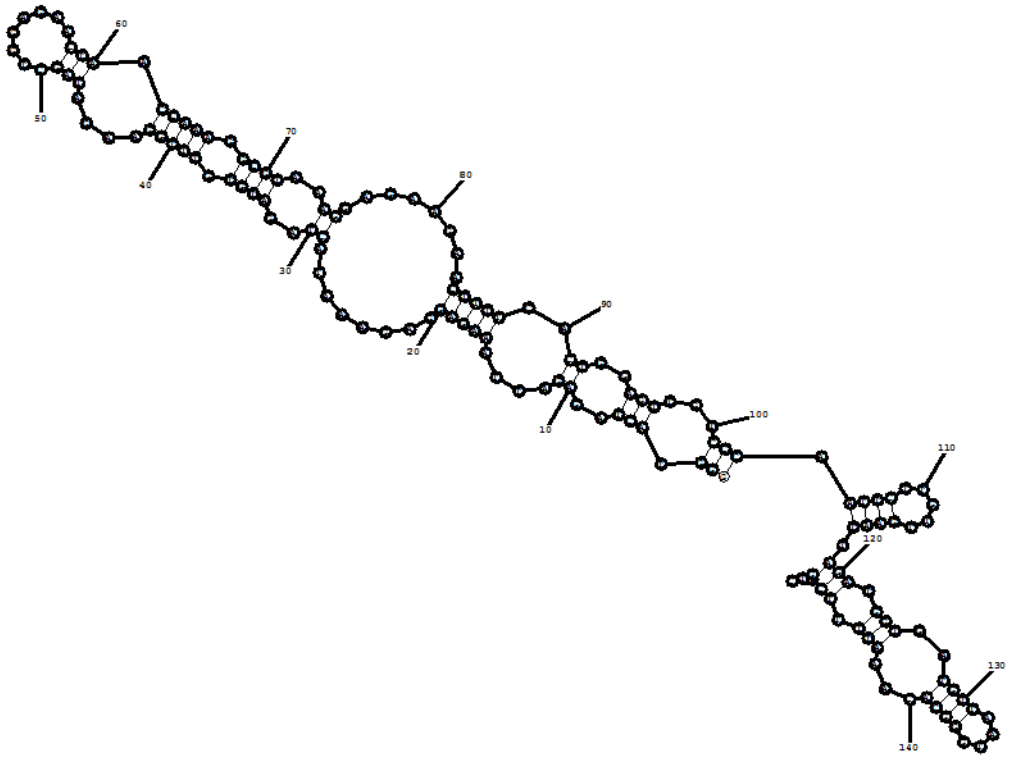
ENERGY = -22.2 device v05

>Thrombin switch device v06|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGTGTTTGGAAGATAGCTGGAGAACTAACCAAAGGCGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



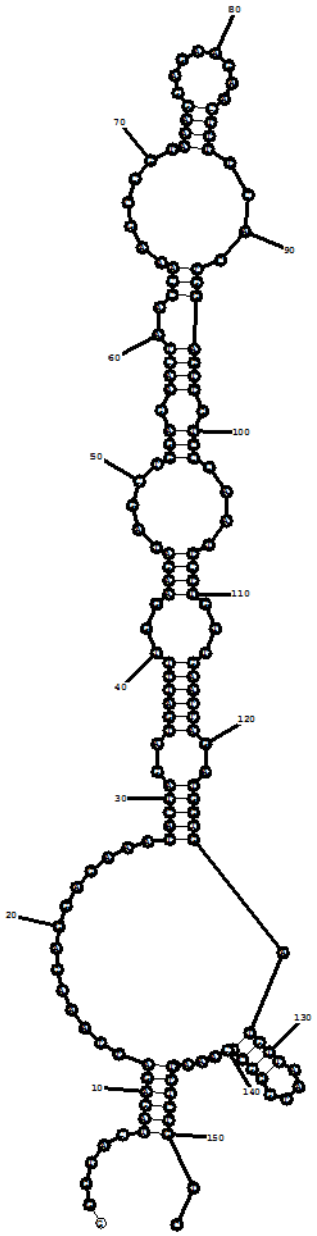
ENERGY = -23.0 device v06

>Thrombin switch device v07|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGATTTTGGAAGATAGCTGGAGAACTAACCAAAGCCGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



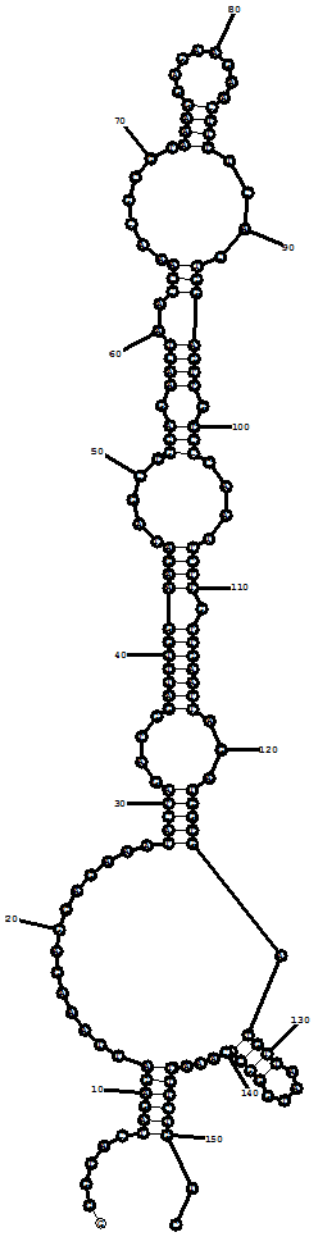
Folding constraints: 38:66, 39:65; 40:64, 41:63, 42:62, 47:60, 48:59, 39:58

>Thrombin switch device v08|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGGATTTGGAAGATAGCTGGAGAACTAACCAAACGCGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



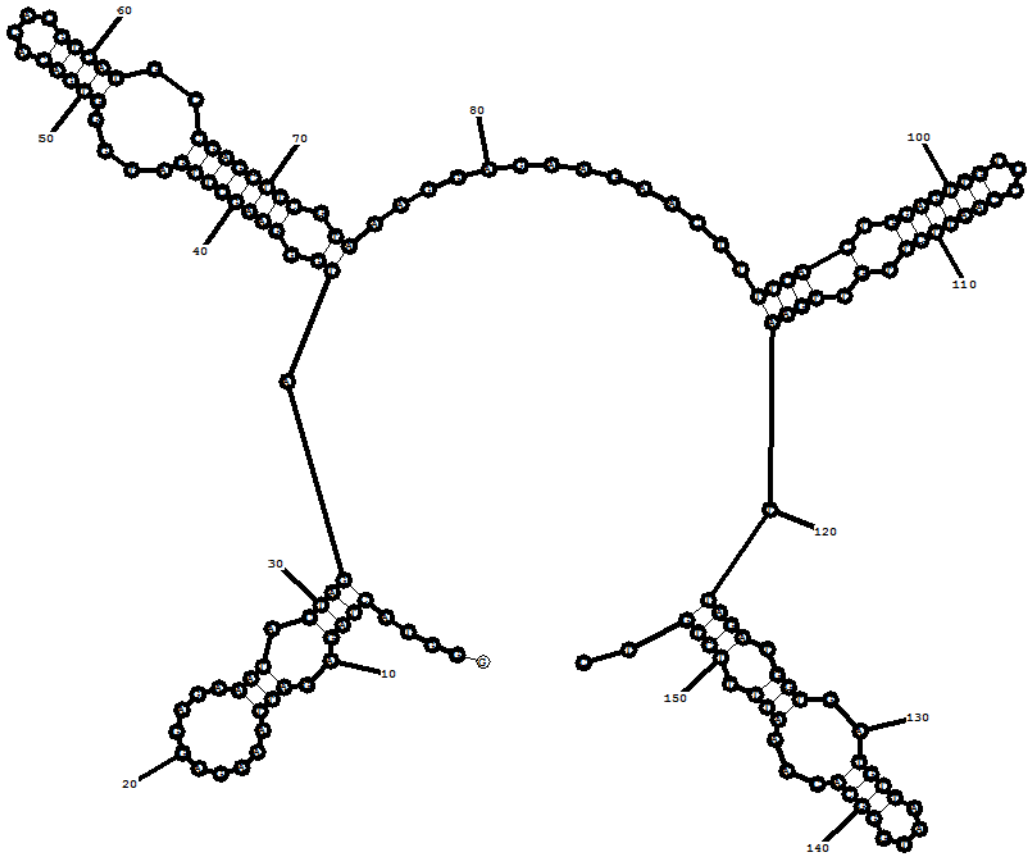
Folding constraints: 38:66, 39:65; 40:64, 41:63, 42:62, 47:60, 48:59, 39:58

>VEGF switch device v01|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGGGAAGCTTGATGGGTGACACACGTCATGCCGAGCTTCGT

AAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTT

CTGTC



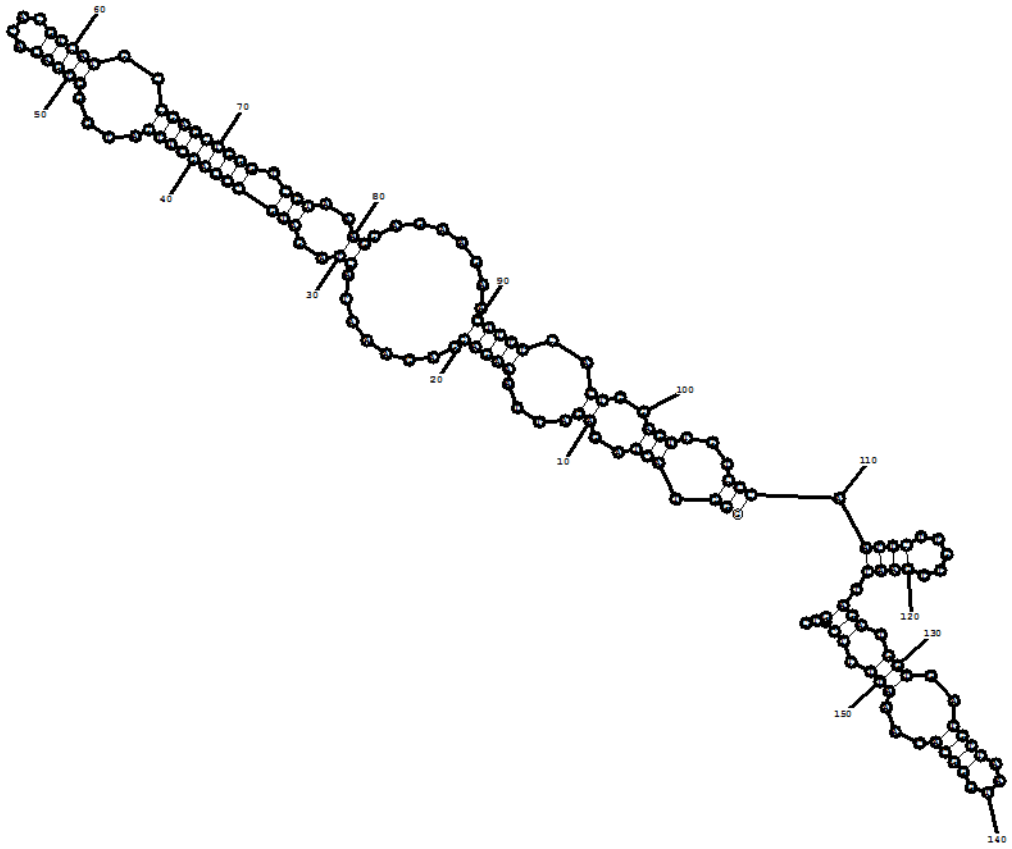
ENERGY = -31.7 device v01

>VEGF switch device v02|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGCCCAGCTTGATGGGTGACACACGTCATGCCGAGCTGGGT

CGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAAT

TTTCTGTC



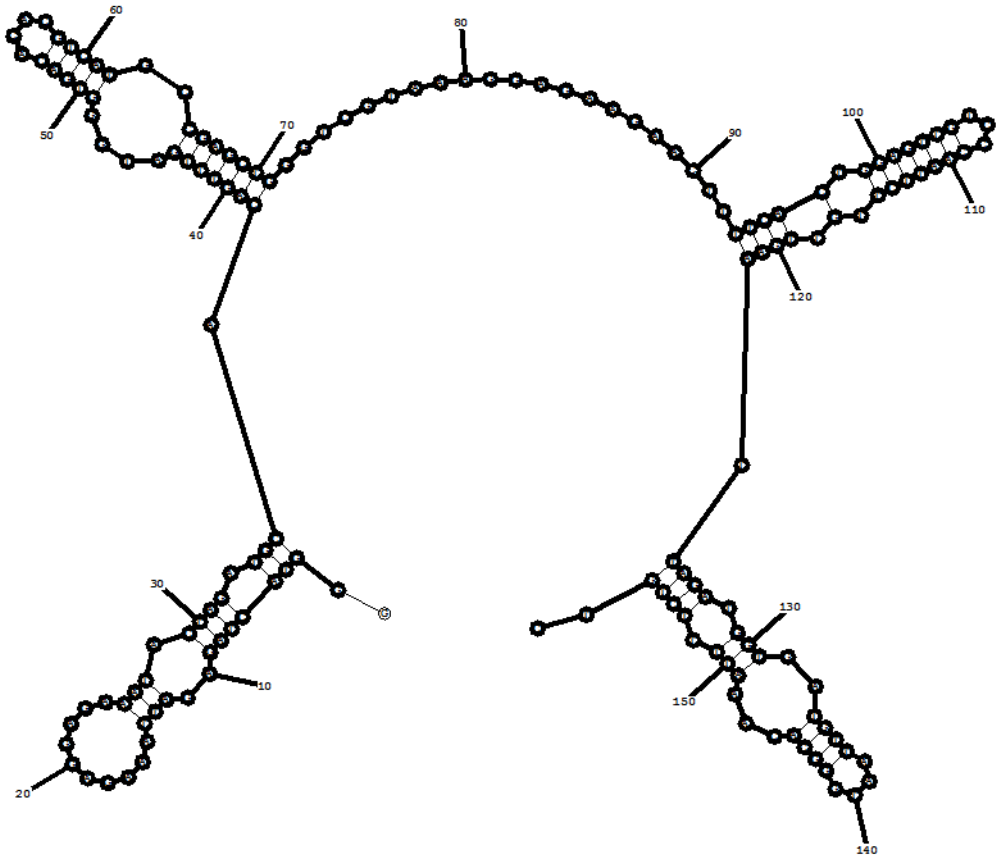
ENERGY = -37.8 device v02

>VEGF switch device v03|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGCACAGTTTGATGGGTGACACACGTCATGCCGAGCTGGGT

CGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAAT

TTTCTGTC

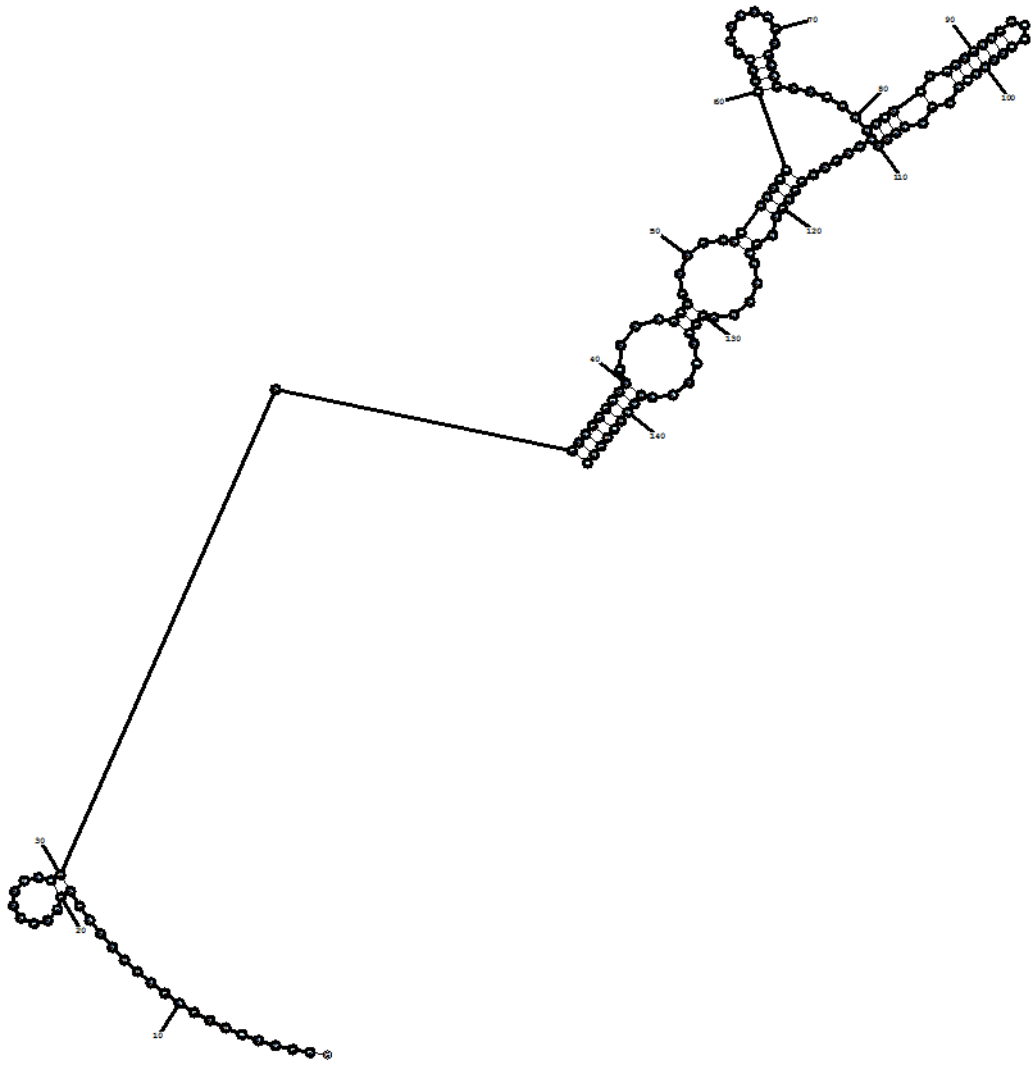


ENERGY = -28.6 device v03

>Toggle switch device v01|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGGGGAACAAAGCTGAAGTACTTACCCTTCGTAAAGGAGAA

GAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTC



ENERGY = -25.0 device v01

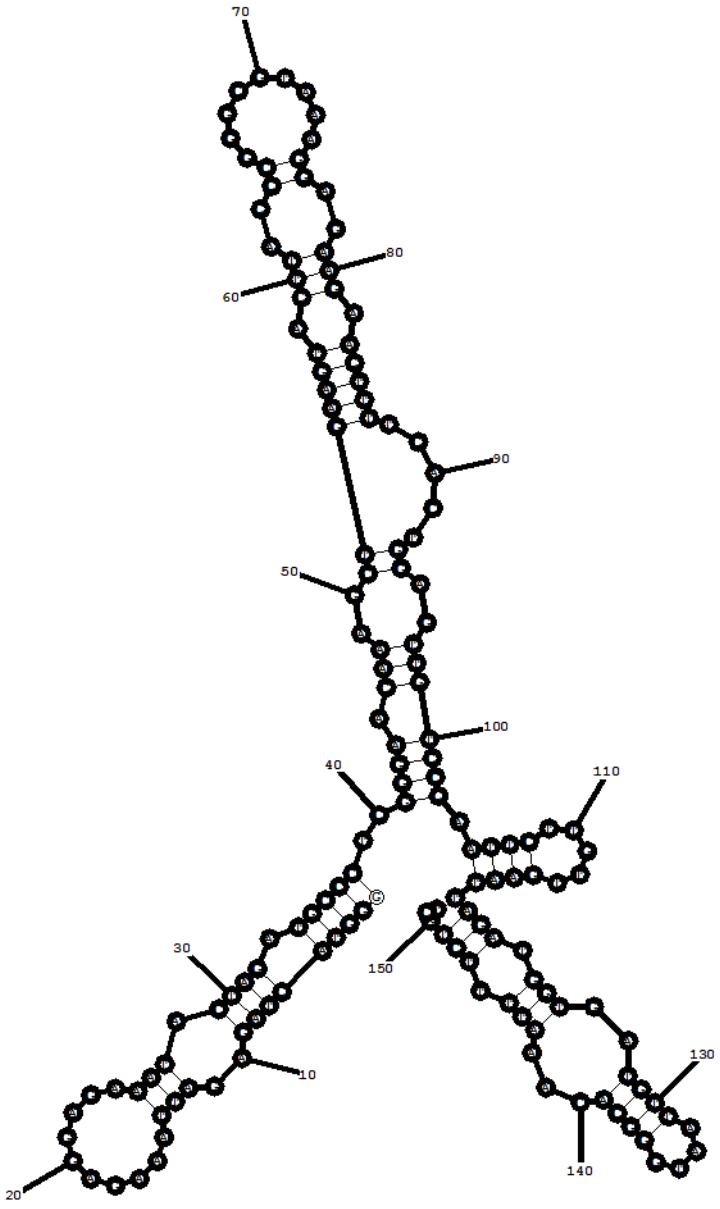
Folding constraints: 36:60, 37:59, 38:58, 43:56, 44:55, 45:54

>Toglgle switch device v02|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGCCCTCGGGAACAAAGCTGAAGTACTTACCCGGGCGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



ENERGY = -30.7 device v02

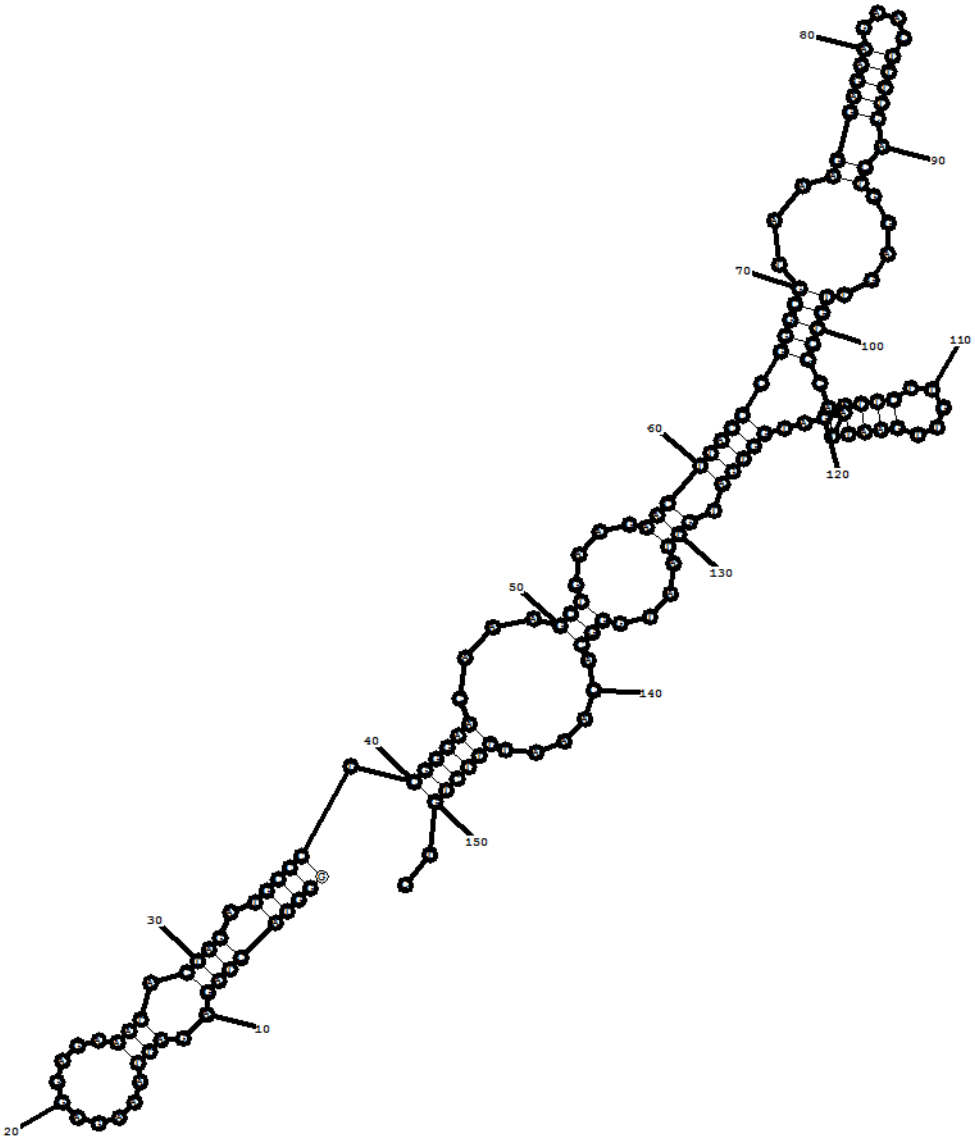
Folding constraints: 41:65, 42:64, 43:63, 48:61, 49:60, 50:59

>Toggle switch device v03|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGCCCTCGGGAACAAAGCTGAAGAACTTACCCGGGCGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



ENERGY = -31.4 device v03

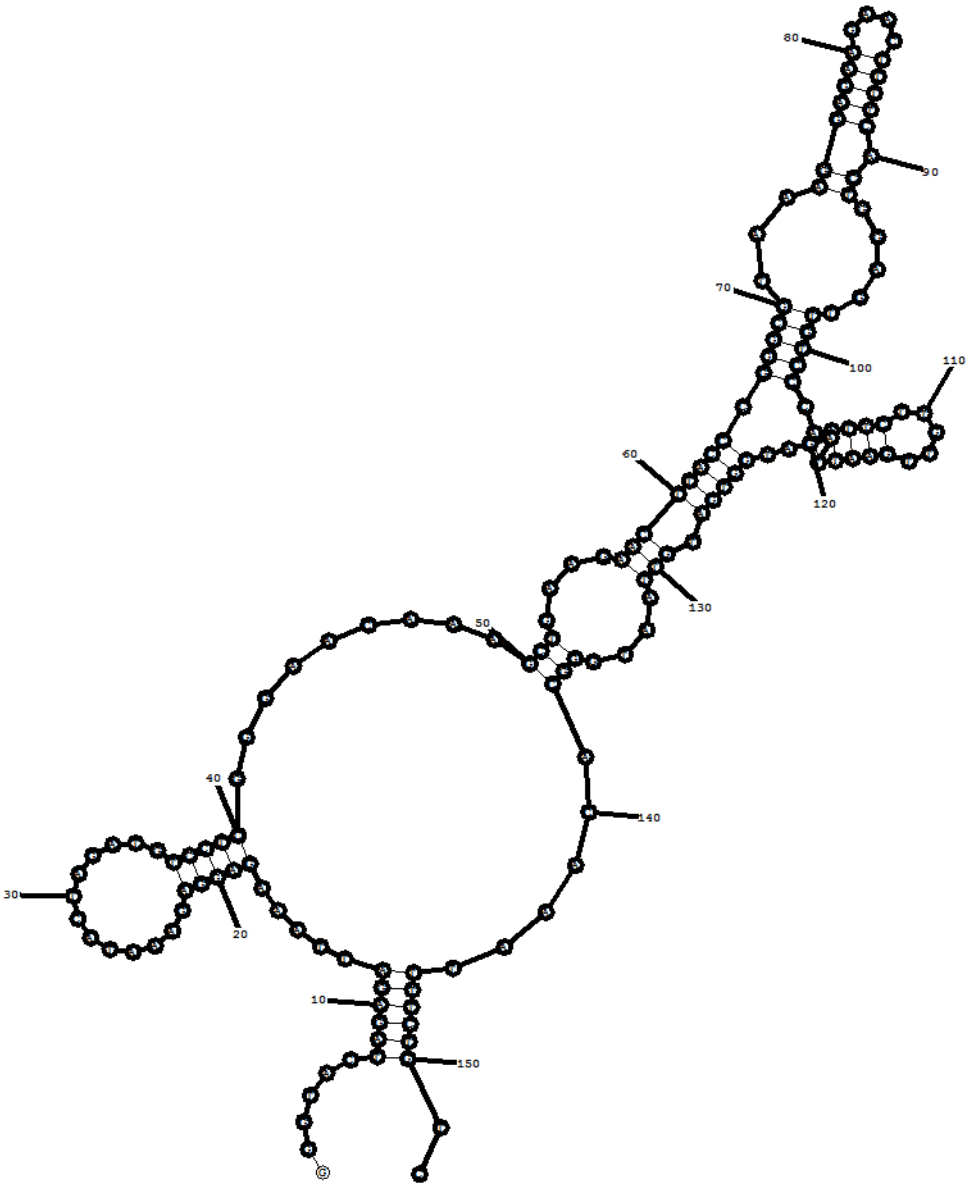
Folding constraints: 41:65, 42:64, 43:63, 48:61, 49:60, 50:59

>Toggle switch device v04|150 nts partial sequence

GGGTACTAGAGATTAAAGAGGAGAAATACTAGATGTCCTCGGGAACAAAGCTGAAGAACTTACCCGGGCGTAAA

GGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGT

C



ENERGY = -26.0 device v04

Folding constraints: 41:65, 42:64, 43:63, 48:61, 49:60, 50:59

10 Appendix D – Reprint permissions

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Jan 03, 2017

This Agreement between Nicolas Kylilis ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

License Number	4021251509173
License date	Jan 03, 2017
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Microbiology
Licensed Content Title	From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis
Licensed Content Author	Denisse L. Leyton, Amanda E. Rossiter and Ian R. Henderson
Licensed Content Date	Mar 1, 2012
Licensed Content Volume Number	10
Licensed Content Issue Number	3
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 1.2: Classical (type Va) autotrasporter protein biogenesis pathway
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Synthetic biology biosensor design for medical diagnostics
Expected completion date	Jan 2017
Estimated size (number of pages)	250
Requestor Location	Nicolas Kylilis B610 Bessemer Building Imperial College London South Kensington Campus London, SW7 2AZ United Kingdom Attn: Nicolas Kylilis
Billing Type	Invoice
Billing Address	Nicolas Kylilis B610 Bessemer Building Imperial College London

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

Jan 04, 2017

This Agreement between Nicolas Kylilis ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

License Number	4021820995678
License date	Jan 04, 2017
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Protocols
Licensed Content Title	Influenza A virus isolation, culture and identification
Licensed Content Author	Amie J Eisfeld, Gabriele Neumann, Yoshihiro Kawaoka
Licensed Content Date	Oct 16, 2014
Licensed Content Volume Number	9
Licensed Content Issue Number	11
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 3.3: Agglutination assay.
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Synthetic biology biosensor design for medical diagnostics
Expected completion date	Jan 2017
Estimated size (number of pages)	250
Requestor Location	Nicolas Kylilis B610 Bessemer Building Imperial College London South Kensington Campus London, SW7 2AZ United Kingdom Attn: Nicolas Kylilis
Billing Type	Invoice
Billing Address	Nicolas Kylilis B610 Bessemer Building Imperial College London South Kensington Campus London, United Kingdom SW7 2AZ



RightsLink®

Home | Account Info | Help



AMERICAN SOCIETY FOR MICROBIOLOGY

Title:	Initiation of Protein Synthesis in Bacteria
Author:	Brian Søgaard Laursen, Hans Peter Sørensen, Kim Kusk Mortensen et al.
Publication:	Microbiology and Molecular Biology Reviews
Publisher:	American Society for Microbiology
Date:	Mar 1, 2005

Copyright © 2005, American Society for Microbiology

Logged in as:
Nicolas Kylilis
Account #:
3001097211
LOGOUT

Permissions Request

ASM authorizes an advanced degree candidate to republish the requested material in his/her doctoral thesis or dissertation. If your thesis, or dissertation, is to be published commercially, then you must reapply for permission.

BACK CLOSE WINDOW

Copyright © 2017 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jan 21, 2017

This Agreement between Nicolas Kylilis ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4025381474627
License date	Jan 10, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	Elsevier Books
Licensed Content Title	Methods in Enzymology
Licensed Content Author	Howard M. Salis
Licensed Content Date	2011
Licensed Content Pages	24
Start Page	19
End Page	42
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier chapter?	No
Will you be translating?	No
Order reference number	
Original figure numbers	Figure 2.1
Title of your thesis/dissertation	Synthetic biology biosensor design for medical diagnostics
Expected completion date	Jan 2017
Estimated size (number of pages)	250
Elsevier VAT number	GB 494 6272 12
Requestor Location	Nicolas Kylilis B610 Bessemer Building Imperial College London South Kensington Campus London, SW7 2AZ United Kingdom Attn: Nicolas Kylilis
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jan 21, 2017

This Agreement between Nicolas Kylilis ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4026100108430
License date	Jan 11, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	Chemistry & Biology
Licensed Content Title	A High-Throughput Screen for Synthetic Riboswitches Reveals Mechanistic Insights into Their Function
Licensed Content Author	Sean A. Lynch,Shawn K. Desai,Hari Krishna Sajja,Justin P. Gallivan
Licensed Content Date	February 2007
Licensed Content Volume Number	14
Licensed Content Issue Number	2
Licensed Content Pages	12
Start Page	173
End Page	184
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	
Original figure numbers	Figure 4
Title of your thesis/dissertation	Synthetic biology biosensor design for medical diagnostics
Expected completion date	Jan 2017
Estimated size (number of pages)	250
Elsevier VAT number	GB 494 6272 12
Requestor Location	Nicolas Kylilis B610 Bessemer Building Imperial College London South Kensington Campus