University of Alberta

Development and optimization of a high through-put screening methodology for rapid dynamic range improvement of FRET-based biosensors

by

Ahmed Abdel Mohsen Abdel Latif Ibraheem

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry

Spring 2012
Edmonton, Alberta

Abstract

Fluorescent protein (FP)-based biosensors based on the principle of intramolecular Förster resonance energy transfer (FRET) enable the visualization of a variety of biochemical events in living cells. The sensitivity of such biosensors is proportional to the change in ratiometric emission, and so there is a pressing need for methods to maximize the ratiometric change of existing biosensor constructs in order to increase the breadth of their utility. To accelerate the development and optimization of improved FRET-based biosensors, we have developed a method for function-based high-throughput screening of biosensor variants in colonies of *Escherichia coli*. We have demonstrated this technology by undertaking the optimization of a biosensor for detection of methylation of lysine 27 of histone H3 (H3K27). Application of this screening methodology led to the identification of an optimized H3K27-trimethylation biosensor with a two-fold improved emission ratio change relative to that of the initially constructed biosensor.

Acknowledgement

I find myself at a loss for words that can genuinely capture the essence of the gratitude that I harbour for my supervisor, Dr. Robert E. Campbell, who unfailingly provides his students with the academic and moral support, encouragement, and professional camaraderie that makes working with him a memorable experience.

I am also grateful to Dr. Anna D. Jordan, Dr. Mariusz Klobukowski, Dr. Todd L. Lowary, and Mrs Karen McKinley for their moral support. I extend my thanks to the University of Alberta and its working staff that gave me the opportunity to work and study in such a supportive environment where I have met the remarkable people I mentioned, along with many others that I dare not start naming lest I accidentally forget to mention someone.

I am forever grateful and deeply indebted to my mother.

Table of contents

Chapter 1: Introduction 1

1.1 *Aequorea victoria* green fluorescent protein 5

1.2 FP- based biosensors 13

1.2.2 Group I: FRET-based biosensors 15

1.2.2.1 Introduction to the concept of FRET 15

1.2.2.2 Group Ia: intramolecular FRET-based biosensors 25

1.2.2.3 Group Ib: intermolecular FRET-based biosensors 28

1.2.3 Group II: BiFC-based biosensor 29

1.2.4 Group III: Single FP-based biosensors 31

1.2.4.1 Group IIIa: single FP-based biosensors with exogenous MREs 31

1.2.4.2 Group IIIb: single FP-based biosensors with endogenous MREs 31

1.3 Objectives of the thesis 33

Chapter 2: Development and optimization of the dual expression system pUADE 36

2.1 Introduction 37

2.2 Materials and methods 43

2.2.1 General methods and materials 43

2.2.2 H3K27-MetBio1 and H3K9-MetBio1 series construction 45

2.2.3 Protein purification 47

2.2.4 *In vitro* PTM enzymatic activity detection of vSET 48

2.2.5 Construction of the dual expression plasmid (pUADE) 49

2.2.5.1 pET22b vector modification 49

2.2.5.2 pUADE assembly 50

2.2.3 Domain library (lib1) construction 51

2.2.6 Optimization of plates composition 52

2.2.6.1 Western blot 52

2.2.6.2 Fluorescence intensity analysis 53

2.2.7 Screening (protocol-1) 53

2.3 Results and discussion 55

2.4 Conclusion 69

Chapter 3: Dynamic range optimization of H3K27-MetBio 70

3.1 Introduction 71

3.2 Materials and methods 79

3.2.1 General material and methods 79

3.2.2 Construction of lib2 and lib3 linker libraries 79

3.2.3 Screening of linker libraries 83

3.2.4 Construction of lib4 rearranged-domains library 83

3.2.4.1 Construction of the plasmids 83

3.2.4.2 PCRs and substrate assembly 84

3.2.4.3 Assembly of the libraries 85

3.2.4 Recombinant protein production and *in vitro* FRET assay 85

3.2.5 Alternative screening procedure (protocol-2) 86

3.3 Results and discussion 87

3.4 Conclusion 101

Chapter 4: Conclusions and future work 102

4.1 Reoptimization of the screening conditions 105

4.1.1 Temporal control over the PTM enzymatic activity 106

4.1.1.1 Controlling the on-set of vSET expression 106

4.1.1.2 Controlling the on-set of the enzymatic reaction 107

4.1.2 Obtaining a vSET variant with faster kinetics 109

4.2 Screening the domain-shuffled libraries 110

Bibliography 111

Appendix 126

List of figures

Figure 1 Simplified Jabłoński diagram demonstrating the phenomenon of fluorescence. 3

Figure 2 Mechanism of avGFP chromophore formation as proposed by Tsien and coworkers 7

Figure 3 Mechanism of avGFP chromophore formation as proposed by Wachter and coworkers 9

Figure 4 Branched chromophore formation mechanism as proposed by Wachter and coworkers 10

Figure 5 Cartoon representation of avGFP 12

Figure 6 Jabłoński diagram representation of the concept of FRET 16

Figure 7 The relative orientation between a FRET pair 18

Figure 8 FRET efficiency profile at $\kappa 2$ value of 2/3 for three chromophore pairs with different Förster ranges 19

Figure 9 Schematic representation of the membrane potential biosensor design reported by Tsien and coworkers 22

Figure 10 Profile of FRET efficiency as a function of the relative orientation factor 24

Figure 11 Intramolecular FRET-based biosensor designs (group Ia) 26

Figure 12 Schematic model of a generic intermolecular FRET-based biosensor (group Ib) 29

Figure 13 Schematic model of a generic BiFC biosensor design (group II) 30

Figure 14 Schematic representation of generic single FP-based biosensor designs (group III)..... 33
Figure 15 Choosing a proper MRE for the construction of a given biosensor..... 41
Figure 16 Representation of the H3K27-trimethylation biosensor (H3K27-MetBio) 57
Figure 18 Nucleotide and amino acid sequences of H3K27-MetBio1 60
Figure 19 Overview of the colony-based screening strategy..... 62
Figure 20 Induction and repression profile of pUADE..... 64
Figure 21 Screening plates composition optimization 67
Figure 22 Library screening methodology..... 68
Figure 23 Arranged domain Lib4-(A-P) 85
Figure 24 Data histogram of FRET ratio changes of all colonies screened from Lib2 88
Figure 25 Linker combinations identified during screening of lib2..... 89
Figure 26 Data histogram of FRET ratio changes of colonies screened from lib391
Figure 27 Linker combinations identified during screening of lib3..... 92
Figure 28 Normalized emission scans of H3K27-MetBio2 (A) and H3K27-MetBio3 (B)..... 93
Figure 29 Comparison between the best identified variants from lib2 (H3K27-MetBio2) and lib3 (H3K27-MetBio3) and the original H3K27-MetBio1 94
Figure 30 Linker combinations of additional clones isolated from lib2 and lib3 . 95
Figure 31 Schematic representation of screening procedures using protocol-2.... 98
Figure 32 Data histogram of screening results using protocol-2..... 99

List of abbreviations

°C	Degree Celsius
ADP	Adenosine diphosphate
AKAR	A kinase activity reporter
ATP	Adenosine triphosphate
avGFP	*Aequorea victoria* green fluorescent protein
BCA	Bicinchoninic acid
BiFC	Bimolecular fluorescence complementation
bp	Base pairs
BSA	Bovine serum albumin
Cbx	Chromobox domain
CCD	Charged coupled devise
CFU	Colony forming unit
cGMP	Cyclic guanosine monophosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double stranded deoxyribonucleic acid
DsRFP	*Discosoma* red fluorescent protein
E	FRET efficiency
E coli	*Escherichia coli*
ECFP	Enhanced cyan fluorescent protein
EDTA	Ethylenediamine tetraacetic
EGFP	Enhanced green fluorescent protein
EYFP	Enhanced yellow fluorescent protein
FACS	Fluorescence activated cell sorting
FHA1	Forkhead associated domain 1
FLIM	Fluorescence lifetime imaging
FLIP	Fluorescent indicator protein
FP	Fluorescent protein

FRET	Förster resonance energy transfer
Glx1	Glutaredoxin-1
GPCR	G-protein coupled receptor
GSK3(BETA)	Glycogen synthase kinase 3 β
H3K27-MetBio	Histone H3 Lys27 methylation biosensor
hCNT3	Human concentrative nucleoside transporter 3
HCV	Hepatitis C virus
HEV	Human enterovirus
HIV-1	Human immunodeficiency virus type 1
HKMT	Histone lysine methyltransferase
HP1	Heterochromatin associated protein
HPLC	High performance liquid chromatography
hr	Hour
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Infrared light
K_d	dissociation constant
KIE	kinetic isotope effect
LB	Lauria-Bertini medium
MAPK	Mitogen activated protein kinase
MBSU	Molecular biology services unit at the University of Alberta
MCS	Multiple cloning site
min	Minute
mM	Millimolar
MRE	Molecular recognition element
mTFP1	Monomeric teal fluorescent protein 1
NADH	Reduced nicotinamide adenine dinucleotide
Ni-NTA	Ni^{2+} nitrilotriacetate
nm	Nanometer
O.N.	Over night
PBP	Periplasmic binding protein

PCR	Polymerase chain reaction
PDB	Protein data base
PKA	Protein kinase A
PTM	Posttranslational modification
PVDF	Polyvinylidene fluoride
R_0	Förster radius
roGFP	Redox-sensitive GFP
rpm	Revolution per minute
rrGFP	*Renilla reniformis* green fluorescent protein
RT	Room temperature
SAM	S-adenosylmethionine
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNP	Single nucleotide polymorphism
ssDNA	Single stranded deoxyribonucleic acid
ssRNA	Single stranded ribonucleic acid
TAE	Tris-acetate EDTA
uM	Micromolar
UV	Ultraviolet light
VR	Vibrational relaxation
vSET	Viral suppressor of variance enhancer of Zeste-trithorax
WGA	Wheat germ agglutinin
YC3.60	Yellow cameleon 3.60
κ^2	The orientation factor

Chapter 1:

Introduction

Fluorescence is the phenomenon of photonic emission exhibited by a class of molecules known as fluorophores upon being subjected to a source of electromagnetic radiations of a given wavelength suitable to cause electronic excitation in these molecules. The process of fluorescence can be demonstrated by the simplified Jabłoński diagram in Figure 1. In the ground state, the molecules of a given fluorophore populate the lowest vibrational energy level of the electronic ground state (S_0). Upon irradiation of these molecules with a source of electromagnetic radiations of sufficient energy, the molecules absorb photons and are promoted to occupy the different vibrational energy levels of the electronic excited state (S_1) where they undergo vibrational relaxation (VR) to occupy the lowest vibrational energy level of this excited state. Finally, the excited molecules dissipate the absorbed excitation energy by emitting photons (fluorescing) and repopulate the electronic ground state.

Owing to the sensitivity of fluorescence detection (1) and the possibility of using multiple fluorophores of different colours in a single assay, fluorescent probes have become widely popular in different research, forensics, and routine analytical applications. For example, in the field of nucleic acid analysis, detection of the different forms of these polymers can be performed using a number of fluorescent dyes which preferentially bind to different forms of nucleic acids (*e.g.*, single stranded deoxyribonucleic acid (ssDNA), double stranded deoxyribonucleic acid (dsDNA), or single stranded ribonucleic acid (ssRNA)) (2).

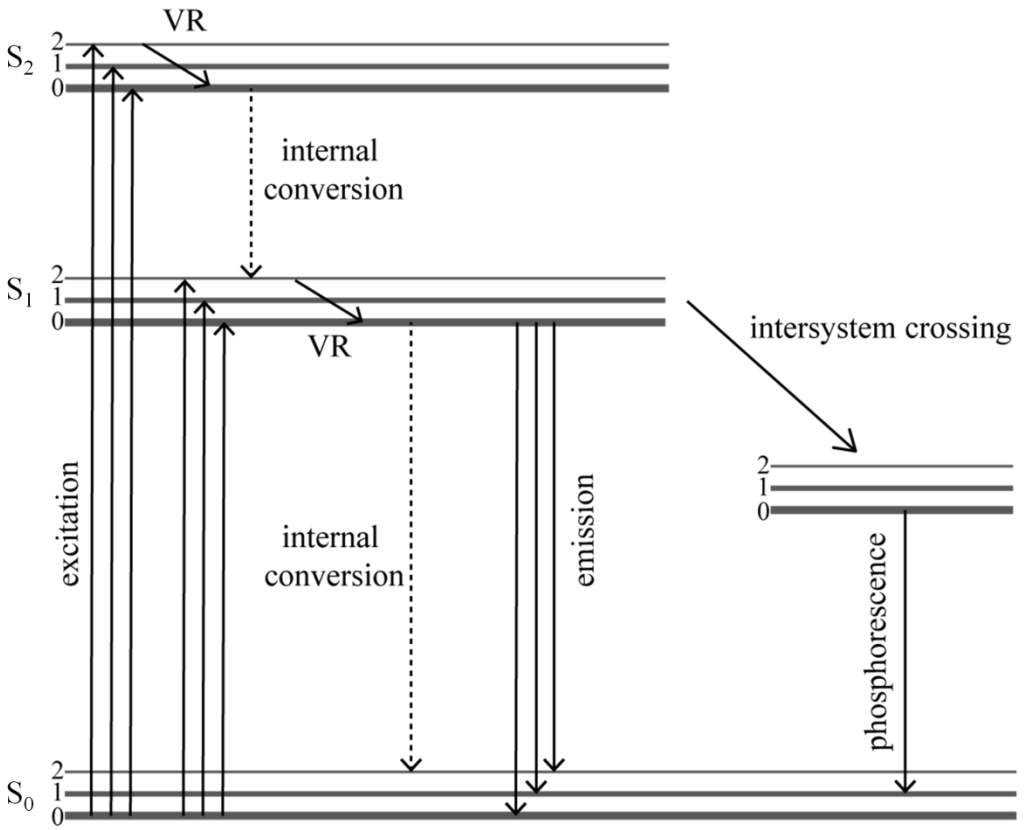

Figure 1 Simplified Jabłoński diagram demonstrating the phenomenon of fluorescence. In addition to fluorescing, excited molecules have other de-excitation routes to dissipate the absorbed energy such as internal conversion, Förster resonance energy transfer (FRET), or intersystem crossing which might culminate in phosphorescence.

Other applications of fluorescence include: DNA sequencing using fluorescently labelled chain terminators (3); single nucleotide polymorphism (SNP) detection using molecular beacons which has been applied to detect SNP associated with antibiotic resistance in *Mycobacterium tuberculosis* (4); protein detection using molecular aptamer beacons, a method that has been applied for the detection of a number of human immunodeficiency virus type 1 (HIV-1) proteins (5, 6); immuno-staining using fluorescently-labelled antibodies which are used in

Western blotting, and fluorescence activated cell sorting (FACS) to name but few applications.

Fluorescent probes are generally classified into intrinsic and extrinsic groups, where the former encompasses compounds naturally occurring in the system under investigation and that can be fluorescently detected without the aid of any additional reagent (7). A few examples of this class include aromatic amino acids (tryptophan, phenylalanine, and tyrosine), and enzyme cofactors (*e.g.*, reduced nicotinamide adenine dinucleotide (NADH)) (7). Whereas extrinsic fluorophore group extends to include fluorescent, and fluorogenic, compounds added heterologously to a given system to obtain information about this system through their spectral properties.

The extrinsic fluorophores class can be further subdivided into synthetic and genetically encoded families of fluorescent molecules (7). One group of molecules that belong to the genetically encoded family is the fluorescent proteins (FPs) group, which includes many variants that have been laboratory-evolved from many marine organisms (8).

Herein, a brief account on the early discovery of FPs, followed by a discussion of employing the laboratory-evolved FP variants in the construction of FP-based biosensors, as an example of the many applications of this class of fluorescent probes (8, 9), shall be presented.

1.1 *Aequorea victoria* green fluorescent protein[1]

Aequorea victoria green fluorescent protein (avGFP) was discovered by Shimumora *et al.* in 1962 during their study of the photoprotein aequorin in the jelly fish *Aequorea victoria* (10). AvGFP was proposed to be the light-emitting species in the jelly fish through a Förster-type mechanism that involved aequorin as the donor luminophore (11). Contemporaneous studies of other bioluminescent organisms, *e.g.*, the sea pansy *Renilla reniformis* (12), lead to the discovery of other naturally occurring GFPs which were also proposed to play a similar role as their avGFP counterpart (13).

Despite the distinct excitation spectra exhibited by avGFP and *Renilla reniformis* GFP (rrGFP) in their native forms, these proteins showed identical spectral properties after denaturation or enzymatic digestion (14), an observation which raised the speculation that these FPs might possess similar chromophore structures, and that the spectral properties of each protein was a function of the micro-environment created by the unique amino acid sequence of each protein (14).

The chemical structure of avGFP chromophore was studied by Shimumora in 1979 who, using the protocol for the isolation and purification of avGFP reported by Morise *et al.* (15), subjected the isolated protein to steps of thermal denaturation, limited enzymatic digestion, and chromatographic separation of the

[1] Parts of this chapter were published by Ibraheem *et al.* in 2010 (186)

yielded peptides (16). After analyzing the amino acid composition of the peptide fraction that showed absorption above 300 nm, and comparing its absorption spectrum to that of a synthetic analogue, Shimumora proposed that the avGFP chromophore structure was based on 4-(*p*-hydroxybenzylidene)-5-imidazolone moiety that was integrated into the amino acid sequence of the protein (16). The chromophore structure of avGFP was later restudied by Cody *et al.* in 1993 who confirmed the core structure proposed by Shimumora but corrected the substitution pattern and the amino acid sequence in which the chromophore exists (17). Interestingly, the first four amino acids in the chromogenic peptides derived from enzymatic digestion of avGFP and rrGFP were identical (FSYGDR in avGFP and FSYGVQ in rrGFP) which further supported the hypothesis that the distinct spectral properties of these proteins, despite possessing the same chromophore structure, may be attributed to the different micro-environments created by their sequences (18). The availability of avGFP coding DNA sequence (19) allowed different research groups to demonstrate, through mutagenesis experiments, that FP spectral properties were indeed influenced by the amino acid composition of both the chromophore (20, 21) as well as the sequence of the non-chromogenic part of the protein (22, 23).

The formation of the avGFP chromophore proceeds via an autocatalyzed posttranslational modification (PTM) reaction that involves the three residues Ser65, Tyr66, and Gly67 (23, 24). The mechanism of this process has been proposed by Tsien and coworkers to proceed via three consecutive steps beginning with nucleophilic attack of the electron lone pair of the nitrogen atom

of Gly67 on the carbonyl group of Ser65 to form hydroxylated pentacyclic ring of imidazolidinone moiety (Figure 2, a1), which subsequently undergoes a dehydration step to form the imidazolone moiety (Figure 2, a2). Finally an oxidation step that involves the C_α-C_β of Tyr66 by a molecule of oxygen takes place leading to the formation of the mature chromophore (Figure 2, a4) (23, 24).

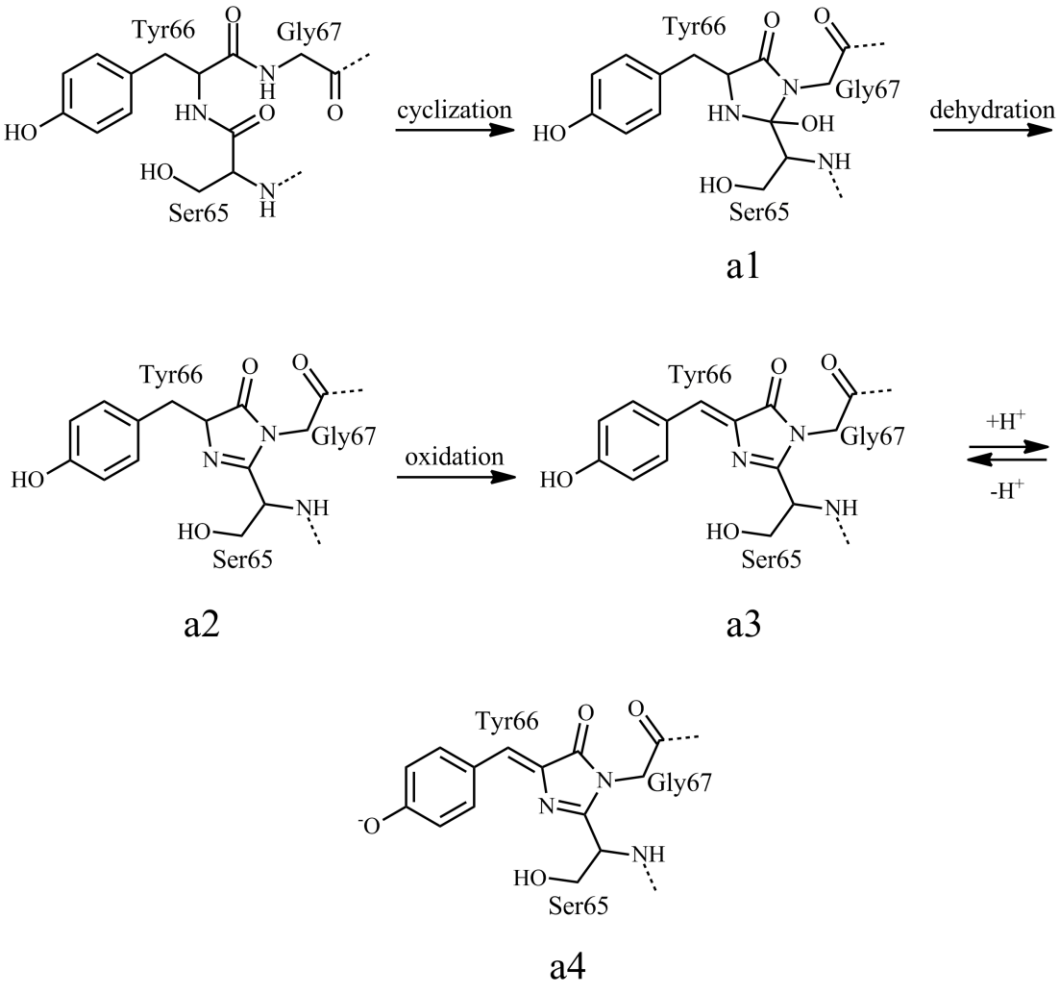

Figure 2 Mechanism of avGFP chromophore formation as proposed by Tsien and coworkers (23, 24).

However, crystallographic study of the non-fluorescent GFP Y66L mutant revealed that the C_α of Leu66, involved in the heterocyclic ring, adopted a trigonal planar geometry (indicating an unsaturated bond with either the carbonyl carbon, C_β or the nitrogen atom of Leu66 residue) while the hydroxyl group, formed after the nucleophilic attack of the initial cyclization step, was still attached to the ring (25). Findings of this study led the authors to propose a new mechanism of the green fluorophore formation (Figure 3) in which the oxidation step follows the initial cyclization and involves the conversion of the imidazolidinone ring to a hydroxylated cyclic imine moiety (Figure 3, b1). Ring oxidation is then succeeded by a dehydration step that involves a proton extraction from the C_α of Tyr66 culminating in the formation of the mature chromophore (25). This sequence of events was consistent with the results reported later by Zhang *et al.* who observed that the maturation of the green fluorophore succeeds the production of hydrogen peroxide (26). Further studies reported by Pouwels *et al.* demonstrated, through kinetic isotope effect (KIE) studies, that, indeed, peroxide production does not demonstrate any KIE during the chromophore formation of a GFP that contained dideuterated Tyr66 C_α, which repudiates the hypothesis of the involvement of its hydrogen in the process of molecular oxygen reduction (27). Therefore, the green chromophore formation is likely to proceed through cyclization, oxidation, and dehydration steps consecutively as demonstrated in Figure 3 (28, 29).

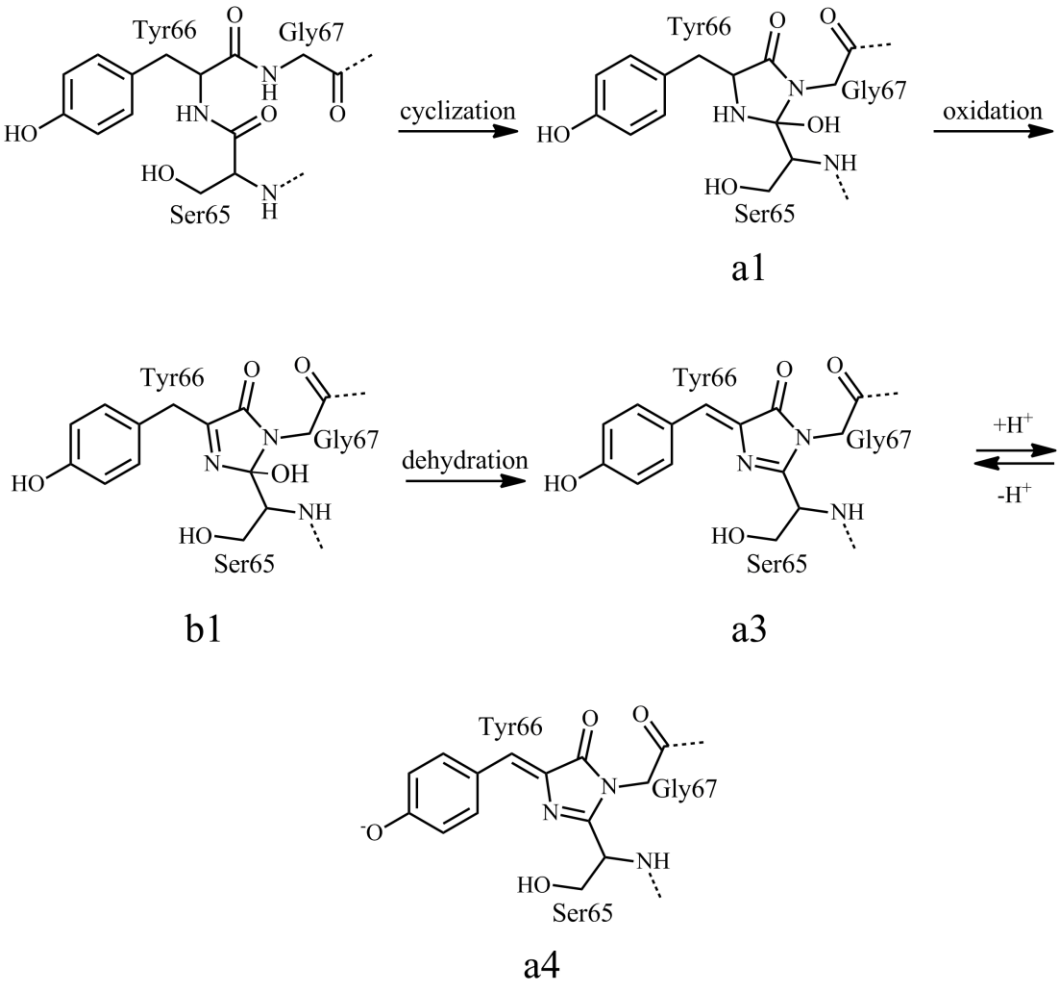

Figure 3 Mechanism of avGFP chromophore formation as proposed by Wachter and coworkers (25-29).

Not only do the non-chromogenic amino acids influence the spectral properties of the chromophore, but also they influence the course of its development. The red FP isolated from *Discosoma* (DsRFP) (30) is a homotetramer that contains both green and red chromophores (31). In this protein, the nature of non-chromogenic residue at position 83 (DsRed numbering) influence the formation of the red

chromophore (32). For example, introducing K83R abolishes the formation of red chromophore but not the green (32).

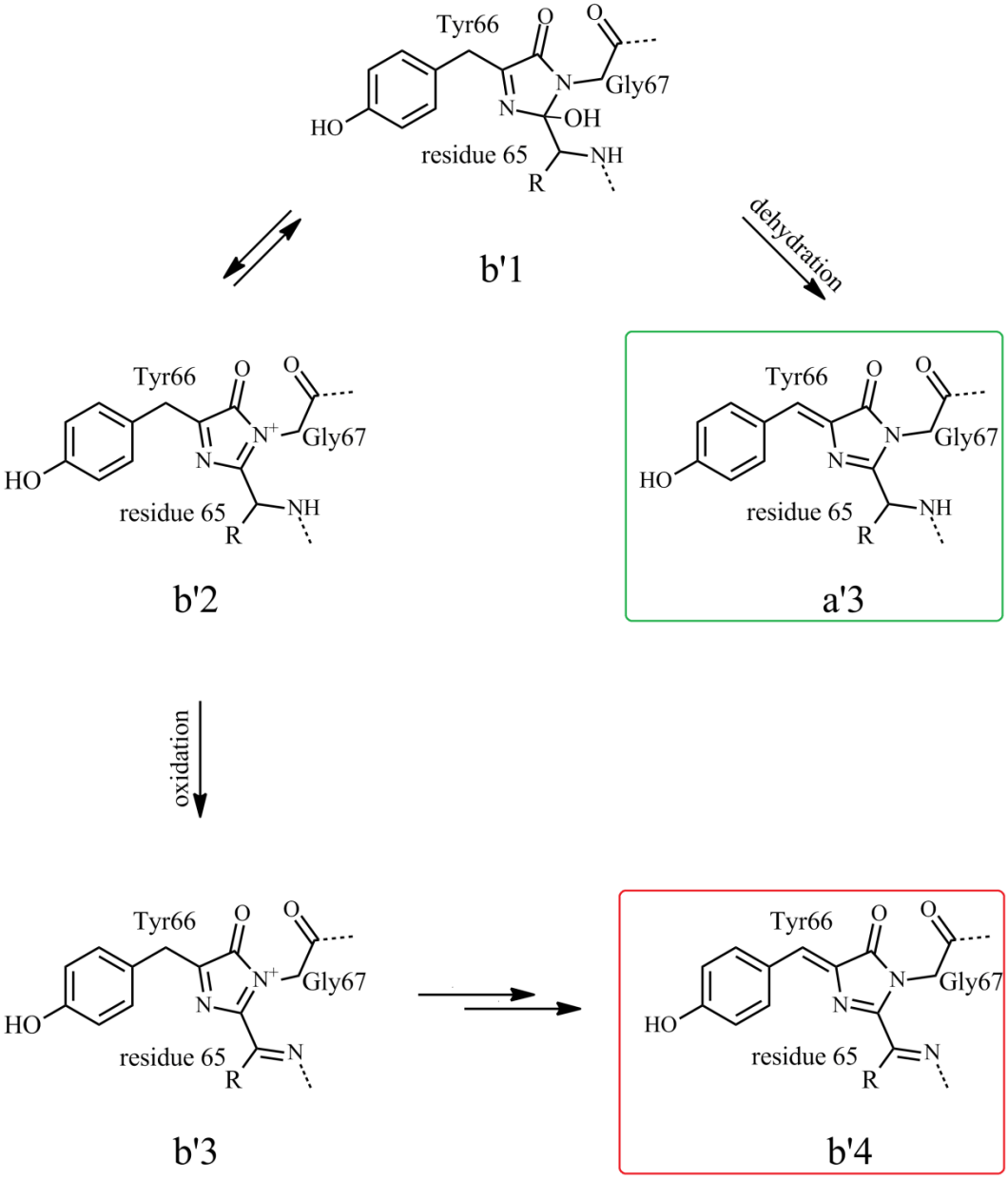

Figure 4 Branched chromophore formation mechanism as proposed by Wachter and coworkers. Residue numbering is based on avGFP. Structures a'3 and b'4 exist in an equilibrium with their deprotonated forms that are responsible for the green and red fluorescence, respectively.

An avGFP-like green chromophore was thought to be the precursor for the formation of the red chromophore through an additional oxidation step involving the C_α-N bond of Gln66 residue of DsRFP (position 65 in avGFP numbering) that results in an extended π system (Figure 4, b'4) (32). However, Strack *et al.* have brought this mechanism proposal into question in light of their experimental findings (28). They convincingly demonstrate that the hydroxylated cyclic imine (Figure 3, b1; Figure 4, b'1), formed as a result of the pentacyclic ring oxidation, may be present in an equilibrium with dehydroxylated form of this moiety (Figure 4, b'2) (28). And based on the amino acid composition of the protein, this equilibrium step may constitute a branching point that can lead to the formation of either the green (Figure 4, 3a) or the red (Figure 4, b'4) chromophores (28).

Structural information about the micro-environment created by amino acids surrounding the chromophore have aided the researchers in better accounting for the altered spectral properties that stems from introducing different mutations in the sequence of FPs (33, 34) and have guided the directed evolution of FP variants with desirable qualities (34, 35). The crystal structure of avGFP was solved in 1996 by two research groups independently (36, 37). Briefly, it consists of eleven β strands that form the unique β-barrel fold which is spanned by an α-helix that connects the third and fourth β strands (36, 37). This α-helix strand bears the chromogenic sequence Ser65-Tyr66-Gly67 which is encapsulated in the cavity of the β-barrel (Figure 5) (36, 37).

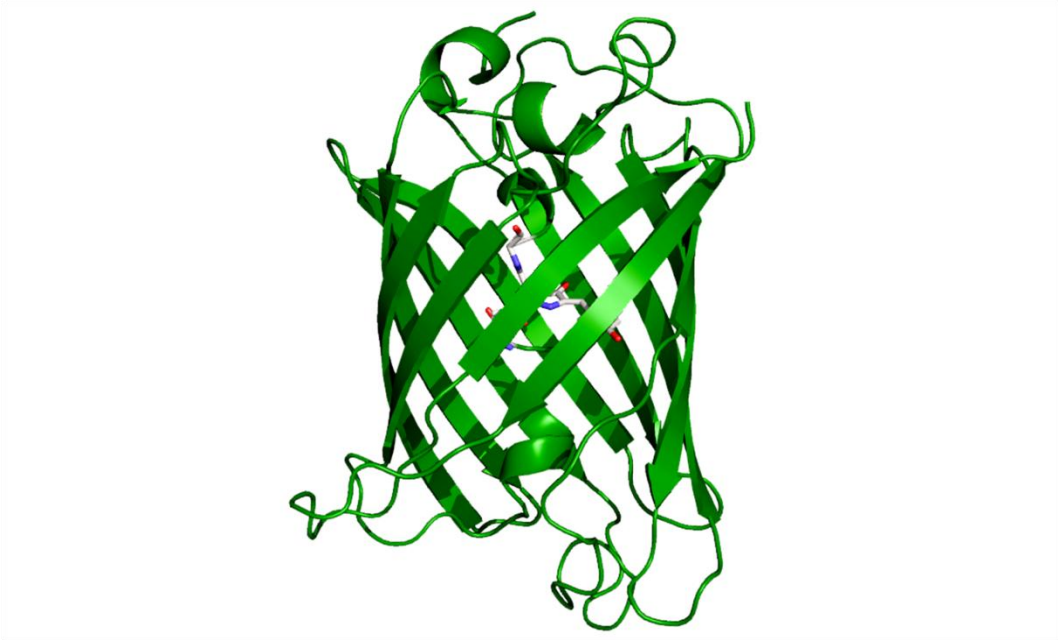

Figure 5 Cartoon representation of avGFP (PDB: 1EMA) (36). Figure was rendered in PyMOL.

Owing to the fact that the formation of avGFP does not require any additional enzymes or cofactors and that it was successfully expressed ectopically in *prokaryota* and *eukaryota* (38), avGFP – and later, engineered FP variants derived from the naturally occurring FPs obtained from *Aequorea victoria* and other marine organisms – has claimed a major role in the study of cell biology and other research areas (8, 9). Directed evolution of avGFP and other naturally occurring FPs became the research area of a number of research groups but was pioneered by Roger Y. Tsien who was awarded the Nobel prize in Chemistry jointly with Martin Chalfie and Osamu Shimumora in 2008 for "the discovery and development of the green fluorescent protein, GFP". Currently, FP engineered variants, derived from many natural sources, span the entire visible region and

even extend into the near IR (8, 9). A detailed discussion of these variants is beyond the scope of this dissertation and interested readers are advised to consult the recent comprehensive reviews by Chudakov *et al.* (8) and Newman *et al.* (9).

FPs have been employed in a many applications (8, 9). One of the applications of FPs is their use in the construction of FP-based biosensors. In the following section, different classes of such biosensors along with their applications shall be discussed.

1.2 FP- based biosensors

Biosensing encompasses a diverse array of techniques for the generation of an experimentally accessible 'read-out' of a molecular interaction between a biologically-derived molecular recognition element (MRE) (*e.g.*, a protein domain) and an analyte of interest (*e.g.*, a small molecule, another protein, or an enzymatic activity). Molecular entities or devices capable of translating the nanometer-scale biorecognition event into an instrumentally comprehensible 'readout' are referred to as biosensors (39). One of the changes that accompany biorecognition events is the change in molecular conformation of the MRE. This change could be a distance change between the MRE and its analyte as in the case of protein-protein interaction, or a distance change between N- and C- termini of the MRE triggered by conformational changes following recognition and binding of the MRE to its cognate analyte as in the case of allosteric proteins. As shall be discussed below, researchers have devised a variety of strategies by which conformational changes

of an MRE can modulate the fluorescence colour or intensity of FP(s) genetically tethered to the MRE.

FP-based biosensors have numerous advantages over alternative technologies (8, 9, 40). For example, FP-based biosensors are relatively easy to construct using standard molecular biology techniques; can be non-invasively introduced into living cells where they are produced using the cellular transcriptional and translational machinery; can be designed to yield information about a given biorecognition process in the natural habitat of the protein of interest thus preserving spatial and temporal components of the interaction event; and can be targeted to most cellular compartments using specific signal sequence tags. Moreover, FPs autonomously develop their fluorophores and thus do not require the addition of any external catalyst.

Practically all genetically encoded FP-based biosensors can be classified into three groups depending on their mode of signal transduction. Group I includes biosensors that utilize the phenomenon of FRET as the signal transducer. This group can be further subcategorized into two subgroups: Group Ia) Intramolecular FRET-based biosensors which encompasses biosensors that have all of their components in a single polypeptide chain (Figure 11); and Group Ib) Intermolecular FRET-based in which, in contrast to Group Ia, the two FPs are in two different polypeptide chains and are brought into close proximity by a protein-protein interaction or any other biorecognition event (Figure 12). Group II includes those biosensors based on bimolecular fluorescence complementation (BiFC). In this biosensing strategy, a biorecognition event is used to bring two

fragments of a split FP to suitable spatial proximity that allows for the reconstitution of an intact (and fluorescent) FP (Figure 13).

Groups III includes biosensors based on single FPs encoded by a single polypeptide chain (Figure 14). Likewise, this group can be subcategorized into Group IIIa and Group IIIb where the difference between these two subgroups is whether or not the MRE element of the biosensor is exogenous (Group IIIa) or endogenous (Group IIIb) with respect to the FP. In the case of an exogenous MRE, the binding of the analyte causes conformational changes that are relayed to the chromophore environment and alter its spectral properties. In the case of an endogenous MRE, the FP plays a dual role: it is responsible for both the molecular recognition (that is, without the aid of any additional entity fused to its sequence) and the fluorescence readout.

Below, we will provide examples of the different designs of genetically-encoded FP-based biosensors belonging to the aforementioned groups and describe recent progress in their development and application.

1.2.2 Group I: FRET-based biosensors

1.2.2.1 Introduction to the concept of FRET

FRET was first described by the German scientist Theodore Förster in 1946 (41). This phenomenon is defined as the non-radiative energy transfer that is observed between an excited blue-shifted fluorophore (the donor) and a red-shifted chromophore (the acceptor), whose absorption spectrum overlaps with the

emission spectrum of the donor, when the two entities exist within a suitable spatial distance defined by their spectral properties. Figure 6 provides schematic representation of FRET using a simplified Jabłoński diagram. The spatial range over which FRET can be observed is referred to as the Förster range and shall be discussed in more details below. Within this range, the FRET efficiency (*E*) exhibits a sixth power dependence on the distance separating the two chromophores and can be described by the following equation:

$$E = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6} \tag{1}$$

where r is the distance between the two fluorophores and R_0 is the Förster radius, which is defined as the distance at which 50% of the maximum FRET efficiency is observed at a given value of relative orientation between the two fluorophores.

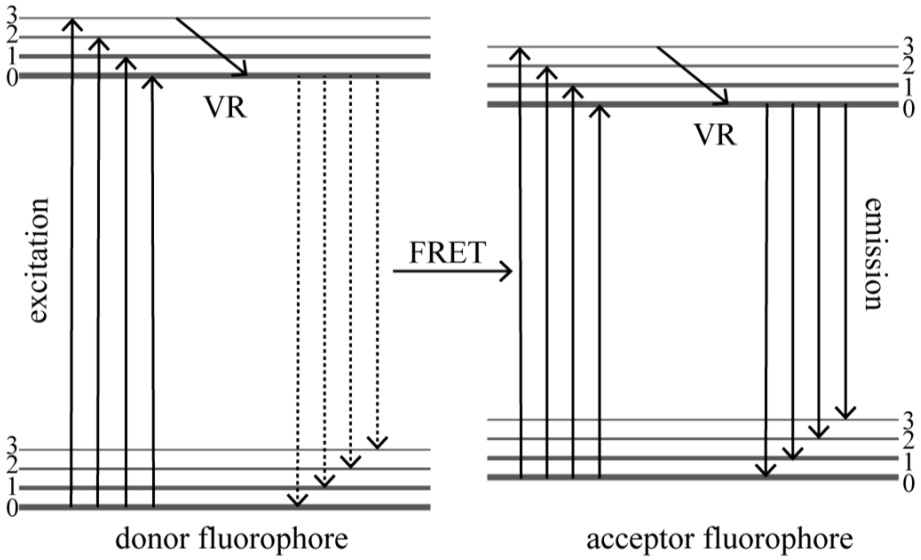

Figure 6 Jabłoński diagram representation of the concept of FRET.

Förster radius is largely dependent on the spectral properties of the two fluorophores and is calculated using the following equation:

$$R_0^6 = 8.79\ x\ 10^{23}[\kappa^2 \phi_D^0 n^{-4} J(\lambda)] \quad (2)$$

where κ^2 is the relative orientation of the two fluorophores, which can assume any value ranging between 0 to 4 for perpendicular or co-linear relative orientations respectively (Figure 7); ϕ_D^0 is the quantum yield of the donor in the absence of acceptor; n is the refractive index of the medium, which is usually assumed to have the value of 1.3 for aqueous systems; and $J(\lambda)$ is the spectral overlap integral of the two chromophores which is calculated using the following formula:

$$J(\lambda) = \int_0^\infty I_D(\lambda)\, \varepsilon_A(\lambda) \lambda^4\, d\lambda \quad (3)$$

where $I_D(\lambda)$ and $\varepsilon_A(\lambda)$ are the fluorescent intensity of the donor and the molecular extinction coefficient of the acceptor, respectively, at a given wavelength (λ).

From equation 2, it is readily realized that the value of Förster radius at a given orientation factor is a characteristic constant for a given fluorophore pair. For simplicity, equation 2 can be written as:

$$R_0^6 = (C\kappa^2)^6 \quad (4)$$

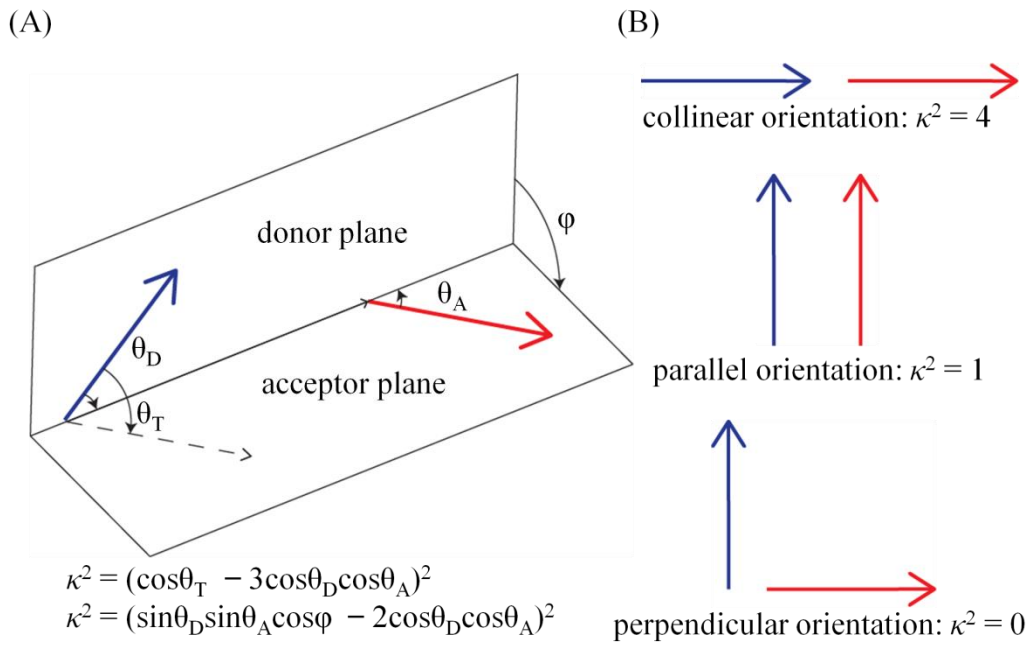

Figure 7 The relative orientation between a FRET pair. The blue vector represents the donor fluorophore excitation dipole while the red vector represents that of the acceptor. (A) Calculation of the orientation factor value (κ^2); where φ is the angle between the donor and acceptor planes as defined by a hypothetical vector connecting the transition dipoles of both vectors; θ_D and θ_A are the angles formed by the vector connecting the transition dipoles of the vectors with the donor and acceptor transition dipoles, respectively; and θ_T is the angle between the transition dipoles of the FRET pair. (B) Values of κ^2 at collinear, parallel, and perpendicular relative orientations.

Since the orientation factor can assume any value between 0 and 4 (Figure 7), the corresponding values of Förster radii, calculated over the entire range of values that can be assumed by the orientation factor, define what is referred to as the Förster range which is the spatial distance range over which FRET can be observed for a given chromophore pair. According to the extent of spectral overlap between the two chromophores in addition to the other parameters defined by equation 2, different chromophore combinations provide researchers with

different Förster working ranges (42, 43). It is customary to calculate R_0 at κ^2 = 2/3, which is the average value of orientation factor calculated for all possible relative orientations. Figure 8 provides three examples of chromophore combinations with R_0 values (at κ^2 = 2/3) of 1 nm (2.3-diazabicyclo[2.2.2]oct-2-ene/tryptophan pair (44)); 3.5 nm (EDANS/Dabcyl pair (45)); and 10.4 nm (terbium complex/(CdSe/ZnS) pair (46)).

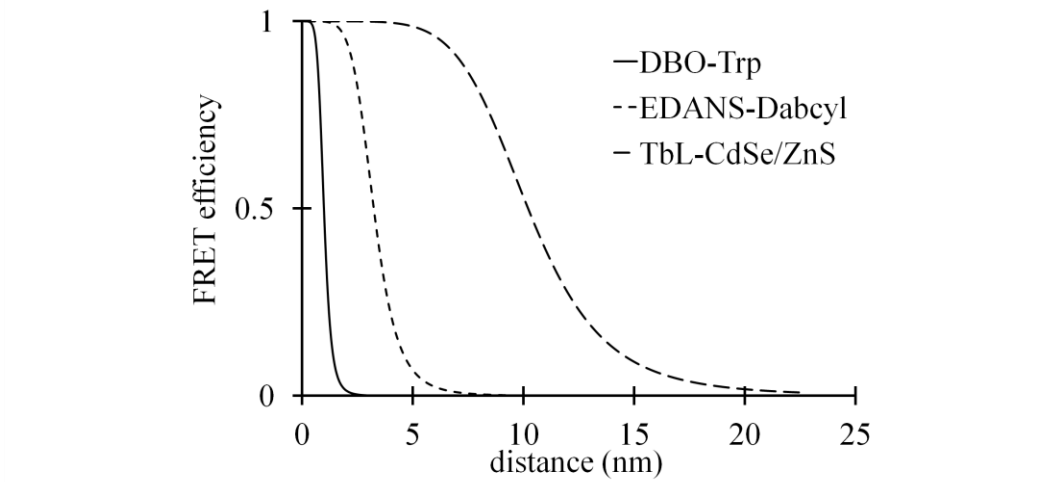

Figure 8 FRET efficiency profile at κ^2 value of 2/3 for three chromophore pairs with different Förster ranges. The diagram demonstrates the different spatial ranges that can be probed by three chromophore pairs with Förster radii (κ^2 value of 2/3) ranging from; small: 1 nm (2.3-diazabicyclo[2.2.2]oct-2-ene/tryptophan pair (44)); intermediate: 3.5 nm (EDANS/Dabcyl pair (45)); and large: 10.4 nm (terbium complex/(CdSe/ZnS) pair (46)).

Generally, the spatial range over which FRET is observed between most chromophore pairs is within 1-10 nm, a range which overlaps with the dimensions of most biological molecules and the distance changes that accompany biorecognition events. Therefore, information about a given biological event can

be gathered by allosterically coupling a chromophore pair to the event under investigation whereby distance changes are revealed by changes in the FRET efficiency (ΔE) between the chromophores used. This technique of 'translating' distance changes in biological events into detectable changes in FRET efficiency constitutes the essence of the FRET-based biosensors.

One of the considerations that face researchers during the construction of a FRET-based biosensor is the choice of a suitable FRET pair. A judicious choice of a chromophore pair the Förster radius of which (that is, the FRET working range afforded by the chromophore pair at the most probable relative orientation value, which is 2/3 in most cases when the chromophores are not spatially restricted) overlaps with the anticipated dimensions of the biological event under investigation is crucial for maximizing the response dynamic range of the biosensor. In other words, the inter-chromophore distance change, coupled to the process under investigation, should be maximized relative to the Förster radius value of the chosen FRET pair.

The rationale behind choosing a different donor fluorophore during the optimization of the initial design of the cell membrane potential biosensor that was reported by Gonzalez *et al.* in 1995 (47) can serve as an illustrative example (48). This biosensor was designed to monitor changes in membrane potential by monitoring the changes in FRET between a donor fluorophore that was anchored to the extracellular side of the membrane (fluorescein-labeled wheat germ agglutinin (WGA) – a lectin), and a mobile hydrophobic anion acceptor fluorophore (an oxonol dye) which, in the resting state of the membrane, is

concentrated on the positively charged extracellular leaflet of the lipid bilayer. Upon membrane depolarization, the acceptor fluorophore permeates the lipid bilayer and translocates to the inner leaflet of the membrane, bringing about a decrease in FRET (Figure 9) (47). One of features of the biosensor that was targeted for optimization was the moiety to which the donor fluorophore was attached (48) which was substituted for a phospholipid. Substituting the lectin moiety for a phospholipid provided two advantages: more uniform loading in different cell lines, and a decrease in the initial distance between the two fluorophores thus maximizing the initial FRET (48). However, the FRET pair used in the initial construct had a R_0 value of 5 nm which is larger than the distance change between the two fluorophores brought about by membrane depolarization (48). To maximize the FRET efficiency change, a coumarin fluorophore with lower spectral overlap with the acceptor fluorophore was chosen as the donor fluorophore (48). This pair had a R_0 value of 4.6 nm, therefore, the same distance traveled by the acceptor fluorophore would produce a larger change in FRET efficiency (48).

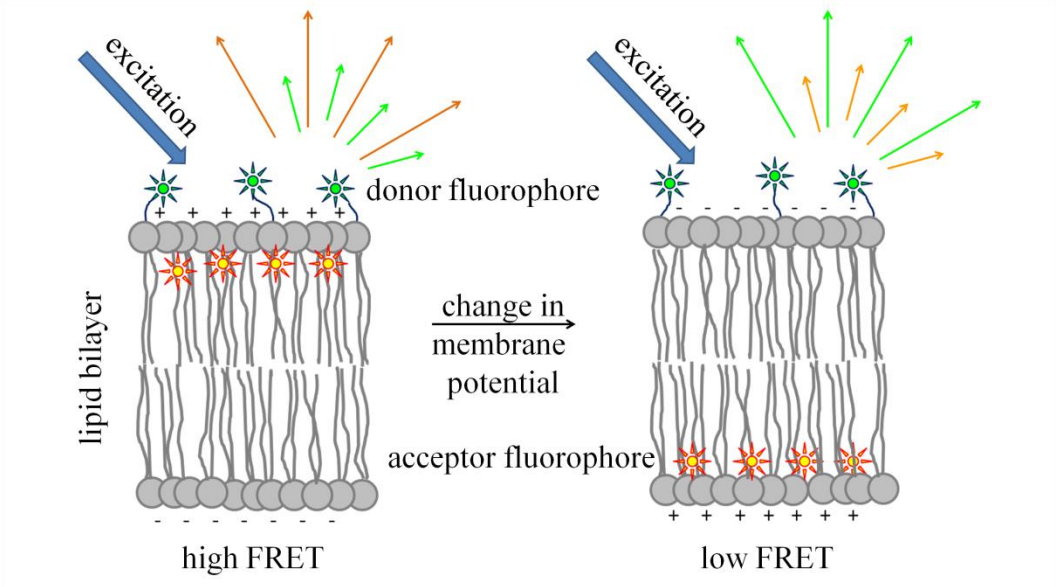

Figure 9 Schematic representation of the membrane potential biosensor design reported by Tsien and coworkers (47, 48). In this illustration, the donor fluorophore is immobilized on the outer surface of the lipid bilayer while the mobility of the acceptor fluorophore (a hydrophobic fluorescent anion) is determined by the polarization of the membrane. Under resting membrane potential, the acceptor fluorophore is concentrated on the outer leaflet of the membrane in close spatial proximity to the donor fluorophore. Excitation of the donor fluorophore (blue arrow) results in high sensitized emission from the acceptor fluorophore. Upon membrane depolarization and the subsequent migration of the acceptor fluorophore towards the inner leaflet of the membrane, excitation of the donor fluorophore results in lower sensitized emission (relative intensities of donor and acceptor fluorophore emissions are represented by different lengths of the green and orange arrows, respectively)

After choosing the proper FRET pair, a suitable change in FRET efficiency should be achieved. Typically, a minimum change in FRET efficiency of 0.1 is required to qualify the biosensor as an appropriate candidate for obtaining measurements in live cells imaging (49). An expression that describes FRET efficiency change, stating explicitly the variables upon which ΔE depends can be written as follows, after substituting equation 4 into equation 1:

$$\Delta E = E_{final} - E_{initial} = \frac{1}{1 + \left(\frac{r_{final}}{C\kappa^2_{final}}\right)^6} - \frac{1}{1 + \left(\frac{r_{initial}}{C\kappa^2_{initial}}\right)^6} \quad (5)$$

Given that FRET efficiency is a function of both the distance and relative orientation, $E\ (r, \kappa^2)$ (equation 5), a suitable ΔE can be achieved by a change in the distance separating the two chromophores, a change in their relative orientation, or a combination thereof.

In practice, the general approach of designing a FRET-based biosensor is based primarily on the anticipated inter-chromophore distance changes obtained by coupling the chromophore pair to the MRE with little, if any, attention paid to achieving changes in orientation factor (49). For example, in the case of FP FRET-based biosensors, FPs are genetically fused to the MRE via linkers that contain amino acids encoded by the built-in N- or C- terminal linkers of the FPs, amino acids encoded by restriction palindromes, linkers in the plasmid used, linkers added by the researcher, unstructured ends of the MRE, or a combination thereof giving rise to what is referred to as the 'composite linkers' (50). These unstructured linkers may provide spatial freedom to the fluorophores allowing them to acquire all possible spatial relative orientations in both states of the biosensor, thus effectively limiting the useful Förster spatial range to its value at the statistical average of the orientation factor (Figure 10).

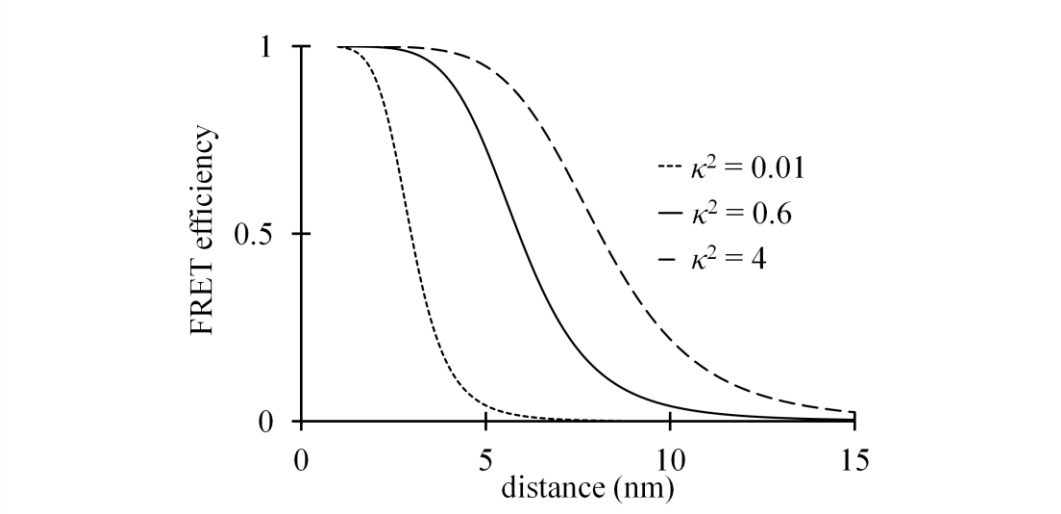

Figure 10 Profile of FRET efficiency as a function of the relative orientation factor. The graph illustrates FRET efficiency at the indicated values of relative orientation for a hypothetical FRET pair with an R_0 value of 6 nm at κ^2 value of 2/3. A biosensor design that allows for unrestricted spatial freedom of the chromophore pair in this example limits the working Forster range to 3.5-9.5 nm, whereas a design that allows for achieving maximal changes may afford a Forster working range 1.5-11.5 nm.

In addition to the aforementioned numerous advantages of using FPs in biosensor construction, the amenability of FPs for sequence insertion (51, 52), circular permutations (53, 54) and their tolerance of N- and C- terminal linkers truncation (55) have been used as strategies to improve the response dynamic range of a number of FRET-based biosensors. Indeed, and as shall be discussed in more details in Chapter 3, results reported by Frommer and coworkers for the dynamic range optimization of different FP FRET-based biosensors strongly suggest that the improved dynamic ranges achieved in some of their biosensors may be attributed to the optimization of the impact of the relative orientation factor on FRET efficiency changes (50, 56-58).

1.2.2.2 Group Ia: intramolecular FRET-based biosensors

The canonical structure of biosensors belonging to this group consists of two FPs flanking an MRE (Figure 11A). Changes in the MRE conformation alter the distance and/or the relative orientation between the two FPs producing changes in FRET efficiency. This basic design of FP-based biosensors has been successfully applied to detect proteolytic activities, PTM enzymes activities, and small molecules. In the following paragraphs, different designs of MREs will be briefly discussed.

FP FRET-based biosensors designed to detect proteolytic activity consist of a chromophore pair tethered by a polypeptidic MRE that owes its sequence to the substrate motif cognate to the protease under investigation. In the absence of the enzyme, the two FPs are in close spatial proximity that allows for FRET to occur. Enzymatic activity severs the MRE resulting in an infinite increase in the distance between the FRET pair and the subsequent loss of FRET (Figure 11B). Biosensors with this design have been developed for the detection of a large number of proteases *e.g.*, viral proteases belonging to human enterovirus (HEV) (59) and hepatitis C virus (HCV) (60).

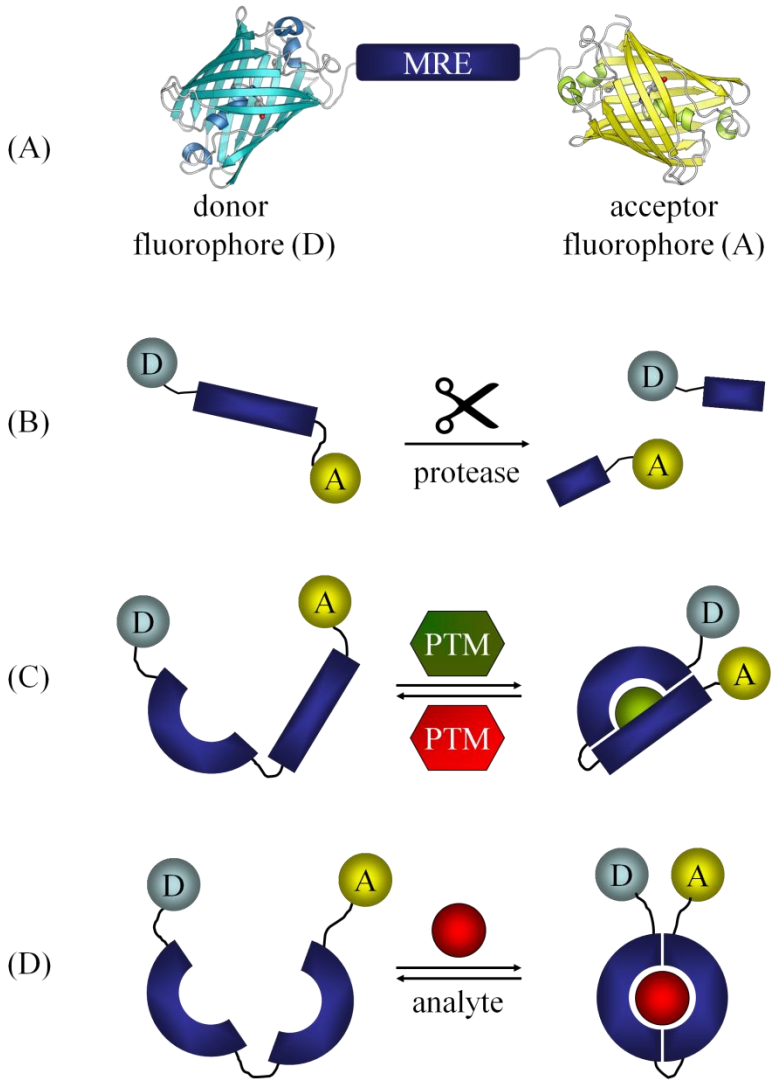

Figure 11 Intramolecular FRET-based biosensor designs (group Ia). (A) Schematic model of a generic intramolecular FRET-based biosensor. A FP FRET pair flanks an MRE that undergoes a conformational change that alters the distance and/or orientation of the FPs relative to each other. (B) An MRE suitable for the detection of protease activity. (C) An MRE for the detection of PTM enzymatic activities where the modification of the peptide substrate creates a binding dock for the binding domain resulting in a FRET change. (D) An MRE in which the conformational change is triggered by the presence of its analyte.

PTM enzymes catalyze the covalent modification (*e.g.*, phosphorylation, acetylation, methylation) of the side chain of a given amino acid residue when present in a specific amino acid context recognizable by a given enzyme. A common strategy for the detection of PTM enzymatic activity is the use of a two-part MRE chimera composed of: a specific substrate motif recognizable by the PTM enzyme of interest; and a binding domain that preferentially binds the modified substrate with a high affinity, whereby conformational changes in the MRE that ensues the enzymatic modification of the substrate motif and the subsequent intramolecular domain-substrate binding can be detected by changes in FRET efficiency of a given chromophore pair tethered to the MRE (Figure 11C). This design has been successfully used to construct several FRET-based biosensors for the detection of diverse PTM enzymatic activities (9, 61) *e.g.*, kinases (62, 63), histone lysine methyltransferase (64), histone acetyltransferase (65), O-GlcNAc transferase (66) activities.

In addition to the aforementioned biosensors designed to detected enzymatic activities, FRET has also been used to construct numerous biosensors for small molecules detections (67, 68). Intuitively, a suitable MRE for such biosensing application is a protein domain that undergoes conformational changes upon binding to its cognate small molecule analyte (Figure 11D). A celebrated family of proteins that exhibit this behaviour is the bacterial periplasmic binding proteins (PBP) (69, 70) the motion of which has been described as a Venus-trap or a clam-shell like movement. This family provides researchers with a repertoire of MREs with varying affinities and specificities towards diverse molecules (70) and,

indeed, PBPs have been used to construct various FRET-based biosensors for numerous analytes such as sugars (*e.g.*, ribose (71), glucose (56), and sucrose (57)), amino acids (*e.g.*, glutamate (72), and tryptophan (73)), and ions (*e.g.*, phosphate (74)). Another example of an MRE that undergoes conformational changes upon binding to its analyte is the chimera of calmodulin and Ca^{2+}/calmodulin-binding peptide which served as the sensory element in the family of Ca^{2+} FRET-based biosensors known as cameleons (75, 76).

1.2.2.3 Group Ib: intermolecular FRET-based biosensors

Biosensors belonging to this group are necessarily split constructs, in which the MRE is fused to one of the FPs and the analyte protein is fused to the other (Figure 12). This design of biosensors is particularly useful for the study of protein-protein interactions. Intermolecular FRET has been applied to study the oligomerization state of different member of the G-protein coupled receptor (GPCR) superfamily (77, 78). However, the versatility of this design of biosensors does not end at merely deducing the oligomerization state of receptors. Conformational changes of the activated receptors, read-out by changes in FRET, can be used to determine the kinetic parameters of the receptor activation as was demonstrated in a recent study for mGluR1, a member of the GPCR superfamily (79).

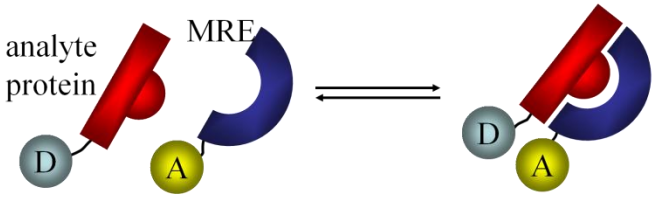

Figure 12 Schematic model of a generic intermolecular FRET-based biosensor (group Ib) for the detection of protein–protein interaction. Protein-protein interaction can be visualized in live cells by tagging each one of the proteins to one member of a FP FRET pair and observing the changes in donor/acceptor intensities.

Intermolecular FRET-based biosensor design has also been used for the detection of PTM (80). In this study, Chun *et al.* studied the consequences of activating the enzyme glycogen synthase kinase 3 β (GSK3β) on intermolecular association between tau protein which is involved in the development of Alzheimer's disease (80).

One of the problems that can hinder accurate and reproducible intermolecular FRET measurements is the variation in the expression level of the two biosensor halves. This could be a major concern when ratiometric measurements are employed. To circumvent the variations in concentration, fluorescence lifetime imaging (FLIM) can be employed (81). Another caveat to the use of intermolecular FRET measurements is that caution must be exercised in the interpretation of results, since FRET also can sometimes occur between two FP-fused proteins that do not directly interact (82).

1.2.3 Group II: BiFC-based biosensor

BiFC is dependent on the intrinsic ability of some FP variants, when expressed in a split form tagged to a pair of interacting proteins, to refold properly into the β-barrel structure and thus reconstitute the fluorescent form of the protein. BiFC-based biosensors are necessarily split constructs in which the MRE is genetically fused to one fragment of the FP and the analyte protein is fused to the other (Figure 13). Several reviews provide a thorough treatment of the practical aspects of BiFC and how to correctly implement this technique (83, 84).

BiFC-based biosensors have been employed to visualize a variety of protein-protein interactions (85) and PTMs in live cells. For example, BiFC was used to reveal the recruitment of members of the chromobox domain (Cbx) family to different parts of the chromatin through their interaction with histone H3 (86).

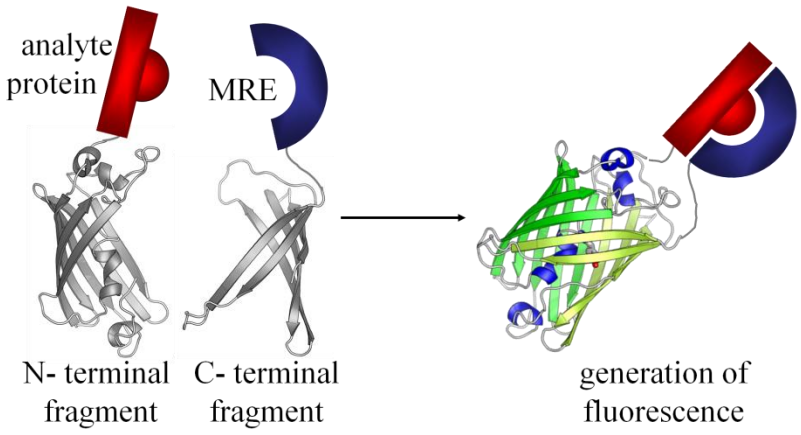

Figure 13 Schematic model of a generic BiFC biosensor design (group II). The figure illustrates protein–protein interaction detection via BiFC.

Interestingly, fluorescent reconstitution is possible between fragments belonging to different FPs, creating chimeras with a variety of fluorescent hues (87). This

allows for simultaneous imaging of more than one event in live cells. Utilizing the multicolour BiFC, Vidi *et al.* studied the homodimeric and heterodimeric oligomerization patterns between adenosine A_{2A} and dopamine D_2 receptors (88).

1.2.4 Group III: Single FP-based biosensors

1.2.4.1 Group IIIa: single FP-based biosensors with exogenous MREs

This class of genetically-encoded single FP based biosensors depends on the ability of some of the variants of FPs to tolerate protein insertion and circular permutations at certain locations. This property has allowed researchers to construct ligand sensitive single FP-based biosensors. The biorecognition event is carried out by an exogenous MRE and information about this event is relayed to the chromophore changing its spectral properties (Figure 14A). This design of biosensor has been applied for the detection of Ca^{2+} (53, 89), Zn^{2+} (90), cyclic guanosine monophosphate (cGMP) (91), and adenosine triphosphate to adenosine diphosphate (ATP:ADP) ratio (33).

1.2.4.2 Group IIIb: single FP-based biosensors with endogenous MREs

In this group of biosensors, changes in the analyte concentration affect the fluorophore microenvironment causing changes in its fluorescence (Figure 14B). For example, some FP variants show pH-dependent spectral changes (92) *e.g.*, the engineered avGFP variants known as enhanced green FP (EGFP), enhanced cyan FP (ECFP), and enhanced yellow FP (EYFP) have pKa's for fluorescence

quenching of 6.15, 6.4, and 7.1, respectively (93). An engineered variant of *Discosoma* red FP (DsRFP), known as mNectarine, was shown to exhibit a useful pH-dependency (94). To demonstrate its potential, the authors fused mNectarine to the cytoplasmic amino acid terminus of human concentrative nucleoside transporter (hCNT3). The readout of the mNectarine and the other pH- sensitive FPs mentioned above is a change in their fluorescence intensity. These intensity-based measurements have the disadvantages of not being easily calibrated and large cell-to-cell variation. To overcome the concentration dependence, and other limitations that are inherent in intensiometric measurements (1), the pH-dependent changes in EGFP were measured using fluorescence lifetime rather than intensity (95). Fluorescence lifetime is a characteristic parameter of a given fluorophore that does not depend on the fluorophore concentration and is not affected by the fluorophore photobleaching. Another way to overcome the intensity measurement shortcomings is to utilize an FP variant that shows ratiometric changes of its spectral properties such as ratiometric pHluorin reported by Miesenböck *et al.* (96), and E^1GFP reported by Serresi *et al.* (97) Other FP-based pH indicators have been reviewed by Bizzarri *et al.* (98).

Another example of a biosensor with an endogenous MRE is redox-sensitive GFP (roGFP). The substitution of two surface amino acid residues of a GFP variant with a cysteine pair at an appropriate distance from each other - to facilitate a disulfide bond formation - rendered this GFP variant sensitive to the redox state of its environment. This roGFP allows for ratiometric measurement of the cell redox status (99). An improved redox biosensor was created by fusing roGFP to human

glutaredoxin-1 (Glx1) which catalyzes rapid equilibration between roGFP and glutathione, thus improving the response rate of roGFP (100).

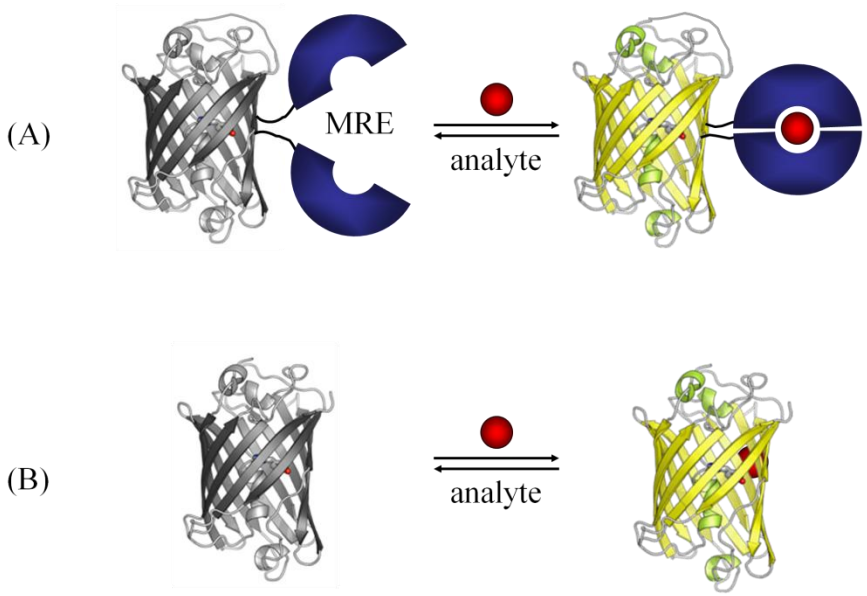

Figure 14 Schematic representation of generic single FP-based biosensor designs (group III). (A) Single FP biosensors with an exogenous MRE (group IIIa). (B) Single FP biosensors with an endogenous MRE (group IIIb).

1.3 Objectives of the thesis

Despite the widespread use of FPs in biosensor construction, most FP FRET-based biosensors suffer from a small response dynamic range (101). To overcome this problem, different approaches aimed at improving the dynamic range of these biosensors have been reported in the literature. These approaches include: trimming the N- or C- FP terminal linkers (58), modulating the degree of dimerization tendency among FPs in a given FRET pair (102, 103) using

circularly permuted variants (104), inserting the FPs in permissive surface sites of the MRE used (50), using rigid or structured linkers (105), or a combination thereof. However, it should be stressed that optimizing the response dynamic range of a given biosensor should be considered on a case-to-case basis; a strategy successfully applied to optimize the dynamic range of a given biosensor may not yield satisfactory results when applied to another. For example, using minimalized linker regions to decrease the degrees of freedom available for the employed FRET pair may have a favourable impact on the FRET efficiency in a given biosensor, however, the same strategy may hinder the proper folding of a given MRE in another, in which case using longer flexible linkers and a FRET pair with a moderate tendency towards dimerization may be a more favourable option (106).

Optimizing the response dynamic range of a given biosensor may be a cumbersome process. For example, linker optimization of a biosensor employing a single domain MRE would require the separate cloning and testing of a number of variants that is proportional to the number of the linkers tested. In one of the most dedicated studies that used linker region optimization to improve the response dynamic range of a FRET-based glutamate biosensor entailed the cloning of 176 variants and their subsequent testing (72).

Recently, a method of optimizing the dynamic range of FRET-based biosensors designed for calcium detection has been reported (107). However, to our knowledge, no equivalent methodology has been reported for optimizing PTM FRET-based biosensors. Therefore, we sought to provide the research community

with the means by which such optimization can be quickly and conveniently achieved. Our method depends on the differential expression of the PTM enzyme in bacterial colonies that constitutively produce the FRET-based biosensor. Chapter 2 will address the development and optimization of this methodology and will demonstrate through a proof-of-concept experiment the potential of applying this system for screening libraries of variants of a given FRET-based biosensor. In Chapter 3, we will discuss the screening of two libraries of variants of the prototype biosensor and address the shortcomings and limitations of this system. Suggested optimization of the reported methodology along with other potential applications shall be described in Chapter 4.

Chapter 2:
Development and optimization of the dual expression system pUADE

2.1 Introduction[2]

Protein-protein interaction is the basis of communication in all living systems. Enzymatic modification of proteins by the addition (or removal) of certain chemical groups (*i.e.*, PTM) have several biological consequences; for example, these modifications can change the affinity between different proteins, inhibit the catalytic activity of an enzyme, or activate the catalytic activities of others (108). Using these modifications, living systems have evolved intricate pathways consisting of cascades of such modifications by means of which information from external or internal stimuli might be remitted to the proper cellular destination responsible to assume a proper response. For example, the mitogen-activated protein kinase (MAPK) signalling pathways are composed of a number of tiers of protein kinases (*e.g.*, an MAPKKK phosphorylating a certain MAPKK which in turn phosphrylates a downstream MAPK) (109). These signalling pathways, depending on the cell type and the specific signalling pathway, culminate in controlling a number of transcription factors that are involved in cell proliferation, differentiation, or death (110).

As was discussed in the introductory chapter, FP-based biosensors are powerful tools for the study of protein-protein interactions in living cells. Being genetically encoded, these nano-scaled biosensors are produced intracellularly rendering unnecessary the use of invasive techniques such as cell membrane permeablization. In addition, they can be easily targeted to a specific cellular

[2] Parts of this chapter were published by Ibraheem *et al.* 2011 (118)

compartment using the appropriate localization signals and thus used to gain information about a given biological event in the context of its 'natural habitat'. Furthermore, with the constant additions of engineered FP variants with improved properties, along with the improvements in microscopic techniques and detection methods, the sensitivity of these biosensors continues to improve.

To construct PTM FP-based biosensors, the two strategies of BiFC and FRET could be employed (for the sake of this discussion, we will exclude the proteolytic PTM activity and will only deal with the enzymes that modify their substrates by the addition or deletion of certain chemical groups). The BiFC strategy (reviewed in (83, 84, 111)) depends on increasing the effective concentration of two fragments of a split FP variant brought about by the interaction of the, necessarily, split MRE tethered to each fragment, which results in the complementation process giving rise to a fluorescence signal. BiFC has been used to study a number of post-translational modifications (83, 84, 111). For instance, it has been used to study different patterns of acetylated histones recognition by bromodomain-containing proteins (112). FRET strategy is more versatile for biosensor construction (49). It was successfully used to construct biosensors capable of detecting a number of PTM enzymatic activities (61).

Although the sensitivity of BiFC module is widely professed (since the signal change goes, ideally, from essentially zero to very high fluorescence intensities), FRET module, despite its inherently lower sensitivity compared to BiFC, is more appealing since it: a) provides a wavelength-ratiometric measurement, thus overcoming limitations imposed by using fluorescence intensities as the signal

output; and b) is reversible and therefore could be used to study the dynamics of a given biological process of interest. Most PTMs are reversible processes, where one specific enzyme introduces the modification and a second specific enzyme removes it. One illustrative example is the PTM of phosphate group acquisition, typical of many cell signalling pathways. In the presence of a certain stimulus, a cascade of protein-protein interactions is initiated through the activity of families of kinases and/or phosphatases. That is, the acquisition or loss of phosphate groups by protein kinases and phosphatases respectively, results in increasing or decreasing the binding affinities between a hierarchy of proteins involved in a specific signalling pathway. But once the stimulus is removed, the initial state must be restored. Therefore, depending on the intended application, an ideal biosensor is the one that not only specifically detects the PTM activity of a given enzyme (for example a kinase activated at a certain step in a cascade of events), but also detects the reverse enzymatic modification of other enzymes (in this example, a phosphatase) that exert their enzymatic modifications on the same substrates of the enzyme of interest (Figure 15). This dynamic monitoring could be achieved using biosensors based on FRET modulation while it would not be feasible in those based on BiFC since FP complementation is an irreversible process.

The development and optimization of biosensors for the detection of protein kinase A (PKA) catalytic activity is an illustrative example of how to develop and optimize a given FRET-based biosensor. In 2001, Zhang *et al.* reported a biosensor capable of detecting the PTM activity of PKA. They named this

biosensor A-kinase activity reporter (AKAR1) (113). This biosensor was constructed using what became the canonical PTM biosensor design: a FRET pair (ECFP and mCitrine) sandwiching an MRE consisting of the phosphoamino acid binding domain (14-3-3 tau) and a substrate motif (113). This biosensor exhibited a FRET ratio change of 30% (113). However, the very high affinity of the employed binding domain to the phosphorylated substrate decreased its sensitivity to detect phosphatases which limited its use to study the signalling pathway they took an interest in (114).

To increase the utility of this kinase biosensor, the binding domain was substituted for another phosphoamino acid binding domain, forkhead associated domain 1 (FHA1), which had a lower affinity towards the phosphorylated substrate modified by PKA (114). The optimized biosensor, AKAR2, demonstrated the specificity and the reversibility required of a suitable biosensor that could be used to study a given biological event that depend on reversible modifications of proteins involved in it. However, optimization of the MRE had an impact on overall FRET ratio change which decreased to be about 17% (114, 115). To improve the dynamic range of AKAR2, Allen *et al.* constructed a series of biosensors employing the same MRE, and the donor FP, but using different variants of the circularly permuted Venus (cpVenus) (115). By tweaking the relative orientations of the FP pair, they obtained AKAR3 with FRET ratio change of about 34% (115). This example could serve as a suitable guide for researchers attempting to develop a PTM activity reporter.

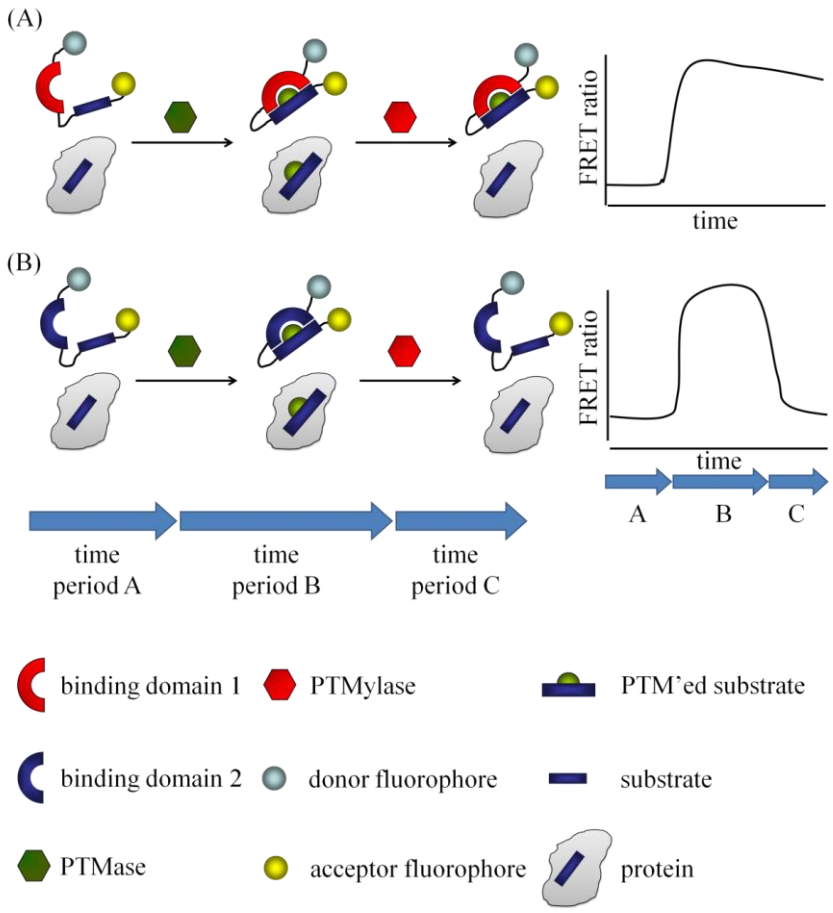

Figure 15 Choosing a proper MRE for the construction of a given biosensor. This diagram illustrates two hypothetical FRET-based biosensors designed to detect the PTM activity of a given enzyme. The MRE of both biosensors borrow their substrate motif from the target protein of the PTMase but utilize two different binding domains. Upon the activation of the PTMase (beginning of time period B), both biosensors are modified along with the target protein and an increase in FRET is recorded in both cases. After the removal of the stimulus and the restoration of the initial state by the PTMylase (beginning of time period C), the biosensor (A) fails to respond to the latter enzyme because of the high affinity of its binding domain to the modified substrate, and therefore fails to faithfully represent the system under investigation, whereas the accessibility of biosensor (B) to the enzymatic activity of PTMylase and the subsequent drop in the recorded FRET qualifies the latter biosensor as a suitable candidate to study the system. The figure is inspired by AKAR1 and AKAR2 (113, 114).

As will be demonstrated in the next chapter, there are a number of approaches by means of which the dynamic ranges of different FRET-based biosensors have been increased. One of these approaches depends on the optimization of linkers tethering the constitutive domains of a given biosensor. Largely, this has been achieved by cloning a number of variants of a given biosensor followed by testing each one of them individually. As an example, the 6.2-fold dynamic range improvement in the glutamate biosensor SuperGluSnFR relative to the initial GluSnFR was achieved by constructing 176 biosensor variants which had systematically truncated N- and C- termini of the MRE used followed by testing each one of them *in vitro* (72).

In this chapter, we will discuss the development and optimization of an '*in colony*' screening methodology which could be used to rapidly and conveniently screen libraries of a PTM biosensor variants for a compositionally optimized variant with a higher signal change. Our method consisted of using the dual expression system 'pUADE' (that has been developed in our laboratory) which employs two differentially induced promoters, namely P_{BAD} and P_{tac} to differentially control the production of two recombinant proteins. We show that by using pUADE according to the provided protocol, we were able to select H3K27-trimethylation biosensors (hereafter abbreviated as H3K27-MetBio) capable of detecting methyltransferase activity of vSET from those incapable of its detection.

2.2 Materials and methods

2.2.1 General methods and materials

All primers and polynucleotides were purchased from Integrated DNA Technologies (IDT Inc). OligoAnalyzer (SciTools, IDT website) was used to calculate melting temperatures and to predict other properties of the designed polynucleotides. Polymerase chain reactions (PCRs) were performed using either *Taq* DNA polymerase (Invitrogen Inc) or *Pfu* DNA polymerase (Fermentas Inc) according to the provided manufacturers' protocols. Deoxyribonucleotide triphosphate (dNTP) solutions used in PCRs were purchased from Invitrogen. DNA restriction at the desired palindromic sequences was achieved using cognate restriction endonucleases (New England Biolabs Inc) following the manufacturer's recommendations. For routine analytical DNA digestion, FastDigest® restriction endonucleases (Fermentas Inc) were used. DNA agarose gel electrophoresis was routinely performed in tris-acetate-EDTA buffer (TAE) using electrophoresis equipment purchased from Bio-Rad Inc. Subsequently, DNA bands were visualized by UV excitation of the intercalating fluorescent dye ethidium bromide (Bio-Rad Inc) that was added to agarose gels prior to congealing. QIAquick Gel-Extraction kit (Qiagen Inc) was used for DNA extraction and purification. DNA ligation was carried out using T4 DNA ligase (Invitrogen Inc) in reaction mixtures that contained a vector-to-insert molar ratio of 1: 4 unless otherwise indicated.

Electrocompetent *E. coli* strain ElectroMAX DH10B™ (Invitrogen Inc) was used for routine plasmid propagation, library construction and screening, and recombinant protein production. Electroporation was performed using 0.2 cm MicroPulser cuvettes (Bio-Rad Inc) and MicroPulser electroporator (Bio-Rad Inc).

Plasmid DNA was isolated and purified using Fermentas GeneJET™- plasmid miniprep kit following the manufacturer's protocol. DNA sequencing reactions were performed using either DYEnamic™ ET (Amersham) or BigDye® (Applied Biosciences) terminator cycle sequencing kits. Ethanol precipitation was performed to clean up the sequencing reactions before they were submitted to the Molecular Biology Services Unit at the University of Alberta (MBSU) for analysis. Sequencing chromatograms were visualized using FinchTV 1.4.0 (developed by Geospiza Inc) and sequence-alignment was performed using ClustalW (European Bioinformatics Institute).

Bicinchoninic acid (BCA) protein assay (Pierce) was routinely used to measure protein concentration following the manufacturer's microplate format procedures in 96-well plates (Corning). Absorbance values of developed assays were recorded on TECAN Safire^2 microplate reader. To determine the concentration of chimeric proteins that contained mTFP1 (116), absorbance at 462 nm was recorded on Beckman-Coulter DU-800 spectrophotometer and concentration was calculated using Beer-Lambert's law.

For purity assessment and/or rough estimation of the extent of proteolytic degradation of recombinant proteins, glycine-SDS-PAGE or tricine-SDS-PAGE –

the choice of which was based on the estimated molecular weight of the protein of interest (117) – was performed using 2 µg protein/lane. PageRuler™ prestained protein ladder (Fermentas) was used as a protein-size reference guide. Procedures and solutions compositions are described in details somewhere else (117). Electrophoresis equipment was purchased from Bio-Rad.

Unless otherwise indicated, all chemicals and reagents were purchased from Fisher Scientific. CDYL1a chromodomain, C20orf140 and JMJD2A tudor domains in pGEX-6P-1 vector (GE Life Sciences) were valuable gifts from Dr Mark T. Bedford at the University of Texas, Austin. Cbx7 A91K variant, JMJD2A D945K variant, and D945R variant were engineered for better selectivity towards methylated H3K27 (by Dr. Hongkin Yap (118). Viral histone H3 Lys27 methyltransferase vSET in pET22b vector (Novagen) was a valuable gift from Dr Alice Y. Ting at MIT.

GenBank accession numbers of the used protein domains and their positions in the respective proteins are as follows: CDYL1a: AAD22735 (amino acids 56-118); C20orf140: NP_057520 (amino acids 77-128); JMJD2A: NP_055478 (amino acids 890-1011); and Cbx7: EDL04620 (amino acids 36-97).

2.2.2 H3K27-MetBio1 and H3K9-MetBio1 series construction

The gene encoding mTFP1 was subjected to three rounds of successive PCRs using the primers combinations: ‘Trugon1 Forward1 (Seq)’ and ‘Trugon1 Backward1 (Compl)’; ‘Trugon2 Forward2 (Seq)’ and ‘Trugon2 Backward2 (Compl)’; and ‘Trugon3 Forward3 (Seq)’ and ‘Trugon3 Backward3 (Compl).’

The amplified product was digested with XhoI and EcoRI and ligated into pBAD/His B vector that was digested with the same enzymes. The resulting vector (referred to as 'Trugon-TFP') had the five unique restriction sites *XhoI*, *XbaI*, *AvrII*, *PstI* and *BclI* upstream of *mtfp1* sequence; and *BglII*, *EcoRI* and *HindIII* downstream of it.

A separate PCR was performed to amplify *mtfp1* with 'TFP-XhoI-FD-SEQ' and 'TFP-XbaI-BK-CPL' primers that introduced *XhoI* and *XbaI* restriction sites at the 5' and 3' ends of mTFP1. The PCR-product was subsequently digested with XhoI and XbaI enzymes (New England Biolabs) and ligated into pBAD/His B based 'Trugon-TFP' that was digested with the same restriction enzymes. The resulting vector was digested with XbaI and PstI restriction enzymes to be used in a subsequent ligation reaction.

The binding domains were each subcloned into a pBAD/His B-derived plasmid that had the following nucleotide scheme: *AvrII-protein domain-218 linker-XbaI-PstI*. The '218 linker' is a previously reported sequence with good proteolytic stability (119). Plasmids containing these protein domains were each digested with AvrII and PstI and the obtained inserts were ligated with the previously mentioned XbaI and PstI digested plasmid. The plasmid products were digested with XbaI and HindIII enzymes and used in subsequent ligation reactions.

Nucleotide sequence encoding the two histone substrates were added to the 5' end of *mcitrine* through two rounds of successive PCRs. H3K27-mCitrine was obtained using the primers: '1-XbaI-H3K27-SEQ' and 'YFP-BK-CPL'; and '2-

XbaI-H3K27-SEQ' and '2-YF-BK-CPL.' Whereas H3K9 was obtained using: '1-XbaI-H3K9-SEQ' and 'YFP-BK-CPL'; and '2-XbaI-H3K9-SEQ' and '2-YF-BK-CPL.' The obtained products were digested with XbaI and HindIII enzymes and ligated with the previously mentioned XbaI and HindIII digested plasmid. This procedure produced plasmids encoding the 14 distinct potential biosensor constructs.

2.2.3 Protein purification

Unless otherwise indicated, pBAD/His B (Invitrogen Inc) was the plasmid of choice for recombinant protein production where nucleotide sequences encoding proteins of interest were placed downstream of P_{BAD} promoter and in-frame with the 5' nucleotide sequence encoding a hexa histidine (6xHis) tag. Over-expression of the 6xHis-tagged protein of interest was achieved by L-arabinose induction of transformed ElectroMAX DH10B™. Briefly, a transformant colony from a freshly plated bacterial culture was used to inoculate 500 mL LB-Lennox media supplemented with ampicillin and L-arabinose (Sigma) to final concentrations of 0.04% and 0.2% respectively. The culture was incubated for 22 hr at 37 °C and 240 rpm shaking speed (Innova 4330 shaker, New Brunswick Scientific). Bacterial cells were collected by centrifugation at 5000 rpm on Beckman rotor after which they were suspended in 30 mL lysis buffer (50 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazol, pH 7.5). The suspended cells were lysed using French pressure cell press. To clear the lysate, centrifugation at 14,000 rpm was performed using Beckmann Rotor centrifuge. Ni-NTA beads (Qiagen) were added to the cleared cell lysate to capture the 6xHis-tagged

proteins. After 30 min of incubation at 4 ℃, the beads were collected in disposable plastic Polyprep columns (BioRad) using a vacuum manifold. Beads were extensively washed with wash buffer (50 mM sodium phosphate, 300 mM sodium chloride, 20 mM imidazol, pH 7.5) and the recombinant protein was subsequently eluted with two column volumes of elution buffer (50 mM sodium phosphate, 300 mM sodium chloride, 500 mM imidazol, pH 7.5). To minimize proteolytic degradation of recombinant proteins, Complete Protease inhibitor cocktail tablets (Roche) were added to all buffers used at the manufacturer's recommended dose and all steps were performed at 4 °C.

H3K27-MetBio1 protein chimeras were subjected to gel-filtration to isolate intact protein from degraded forms on ÄKTA™-design HPLC system operated by UNICORN™ software using HiLoad™ 16/60 Superdex™ 75 pg XK column (GE Life Sciences). Flow rate was set to 1 mL/min and the buffer used was 50 mM sodium phosphate pH 8.0, 300 mM sodium chloride. Chromatographic separation was monitored by absorbance at 280 nm, 462 nm and 513 nm wavelengths using Monitor UV-900 (GE Life Sciences).

2.2.4 *In vitro* PTM enzymatic activity detection of vSET

The *in vitro* FRET assay was performed in corning 96-well plates. Each reaction contained 200 µM S-adenosylmethionine (SAM), 1 µM vSET enzyme, 1 µM biosensor protein, and 2 mg/mL BSA. Reactions were buffered with 50 mM sodium phosphate, 300 mM sodium chloride, pH 8.0. Similar reaction mixtures that lacked SAM were simultaneously assembled. Excitation wavelength was set

to 420 nm and fluorescence emission was collected at 494 nm and 526 nm wavelengths. Measurements were collected at a time interval of 2 min for 6 hr using TECAN Safire^2 microplate reader. Ratio of fluorescence intensity at 526 nm to that at 494 nm were calculated for each measurement. Ratios calculated for the wells that lacked SAM were averaged and the average values were used as the baseline ratios or R_{min} for each biosensor construct. Changes in FRET was calculated as $\Delta R/R_{min}$ % and was plotted against time (Figure 17).

2.2.5 Construction of the dual expression plasmid (pUADE)

Construction of pUADE consisted of modifying pET22b vector (Novagen) as shall be described below. A step that included changing P_{T7} to P_{tac} promoter; modifying the nucleotide sequence downstream of the promoter to encode a 6xHis tag at the 5' end of the multiple cloning site (MCS), which was modified to introduce *XhoI* and *HindIII* palindromes in the same reading frame to their counterparts in pBAD/His B vector; and adding a transcription terminator downstream of the MCS. Required sequence of the modified pET22b-based vector was PCR-amplified and inserted into the unique restriction site *BsaAI* in pBAD/His B vector whose MCS has been modified by eliminating *XhoI* and *HindIII* palindromes and introducing *EcoRI* and *BglII* restriction sites.

2.2.5.1 pET22b vector modification

The two primers 'pGEX-For-TAC (Seq)' and 'pGEX-Back-TAC (Comp)' were used to amplify a 264 bp fragment that contained P_{tac} from pGEX-6P-1 vector (GE life sciences) and to introduce *BamHI* and *NcoI* restriction sites at 5' and 3'

ends, respectively, of the amplified product. The obtained DNA fragment was digested with BamHI and NcoI restriction enzymes and was subsequently ligated into pET22b vector (Novagen) that was digested with BglII and NcoI enzymes. Next, 'Forward-BAD-transfer' and 'Backward-BAD-transfer' were used to amplify the part of pBAD/His B extending from nucleotide 333 to 765 (original plasmid numbering) and to introduce *PciI* and *BglII* restriction sites at the 5' and 3' ends respectively. The obtained DNA was digested with PciI and BglII enzymes and was ligated into NcoI and BamHI digested pET22b vector that was obtained from the previous step.

The obtained vector was used as a template for a PCR using the primers 'POLY-Prmr1 (pet)- Frwd' and 'POLY-Prmr1-(pet)- Backwrd' to amplify the part of the plasmid that encoded the LacI, the modified promoter and its downstream MCS and transcription terminator.

2.2.5.2 pUADE assembly

A dsDNA fragment with 5' and 3' overhangs compatible for ligation with XhoI and HindIII restricted DNA respectively was obtained by slowly cooling a buffered solution of 'MCS-SEQUENCE' and 'MCS Complementary' primers, at a concentration of 0.5 mM each, from 95 °C to RT at a temperature drop rate of 1 °C /3 min. The obtained dsDNA was ligated into XhoI-HindIII (New England Biolabs) digested pBAD/His B vector (Invitrogen). Owing to the nucleotide sequence of the ligated insert, *XhoI* and *HindIII* palindromes were lost and two new restriction sites, *EcoRI* and *BglII*, were introduced downstream of P_{BAD}. The

modified pBAD/His B vector was then digested with EcoRI and BglII and used to ligate the similarly restricted gene encoding vSET that was PCR-amplified using the two primers 'vSET-EcoRI-Frwd (seq)' and 'vSET-BglII-Bckd (Compl).' The resulting vector, named 'New MCS-pBAD,' was digested with BsaAI restriction enzyme and was ligated with the PCR-amplified product of the modified pET22b vector which was digested with PmlI restriction enzyme (obtained from the last step discussed under 'pET22b vector modification.' When the recombinant proteins encoded by pUADE need to be explicitly stated, we will use the notion pUADE[P_{BAD}(*recombinant protein 1*)/P_{tac}(*recombinant protein 2*)], which specifies the transcriptional control of each of the recombinant proteins as well as their identity.

2.2.3 Domain library (lib1) construction

pBAD/His B vectors containing the aforementioned seven H3K27-MetBio1 constructs were pooled together in equimolar amounts and subsequently digested with XhoI and HindIII restriction enzymes. DNA electrophoresis followed by gel extraction were performed to purify the nucleotide inserts encoding the biosensors which were subsequently ligated into XhoI-HindIII restricted pUADE.

2.2.6 Optimization of plates composition

2.2.6.1 Western blot

To prevent interference from the 6xHis-tagged H3K27-MetBio1, pBAD/His B vector containing the vSET encoding nucleotide sequence (under the control of P_{BAD}) was used to assess the expression pattern of the 6xHis-tagged enzyme as a function of plate composition. ElectroMAX DH10B™ cells (Invitrogen) were transformed with pBAD/His B vector encoding vSET, spread on plates with three different compositions; A) 1 mM IPTG/13 mM L-arabinose; B) 1 mM IPTG/11 mM glucose; and C) 1 mM IPTG. All plates contained ampicillin to a final concentration of 0.04%. Some of the 1mM IPTG/11 mM glucose plates were sprayed with a 1 M L-arabinose solution and incubated at RT for 3 hr to assess the possibility of 'turning on' the expression at request. To detect the 6xHis-tagged vSET, Western blot experiment was performed as described by Li *et al.* (120). Briefly, colonies from the four sets of plates were gently scraped and collected in 2 mL Eppendorf tubes. B-PER II (Pierce) was used to extract the soluble proteins from the cell paste according to the manufacturer's protocol. Protein concentration was then determined using BCA assay kit (Pierce). 2 µg and 20 µg (in equal volumes) were subjected to polyacrylamide gel electrophoresis using tricine-SDS-PAGE as described by Schägger *et al.* (117). Proteins were then electroblotted on PVDF membrane (Millipore). Tris-buffered saline pH 7.4 containing 0.1% Tween-20 and 5% non-fat milk was used to block the membrane overnight at 4 °C. The membrane was subsequently washed, incubated with anti-

His antibody conjugated with horseradish peroxidase (Roche), and rewashed. Locations of the 6xHis-tagged proteins were then detected using ECL chemiluminescence substrate (Pierce) and BioMax light film (Kodak).

2.2.6.2 Fluorescence intensity analysis

Four types of LB-agar plates that differed in their L-arabinose and glucose concentrations were prepared: A) contained 20 mM glucose; B) 10 mM glucose; C) 10 mM glucose/20 mM L-arabinose; and D) 10 mM glucose/10 mM L-arabinose. All plates contained IPTG and ampicillin to final concentrations of 1 mM and 0.04% respectively.

ElectroMAX DH10B™ cells (Invitrogen) transformed with pUADE[P_{BAD}(*vset*)/ P_{tac}(*h3k27-metbio1-cbx7*)] were spread on the prepared LB-agar plates. The plates were incubated at 37 °C overnight after which they were used to acquire a set of three images for each plate as shall be described in details under 2.2.7 Screening (protocol-1). FRET/mTFP1 ratios were calculated for each plate and data histogram was obtained using Microsoft Excel.

2.2.7 Screening (protocol-1)

For *in colony* screening of H3K27-MetBio1 constructs, pUADE[P_{BAD}(*vset*)/P_{tac}(*lib1*)] was used to transform ElectroMAX DH10B™ cells (Invitrogen) and transformants were grown on LB-agar supplemented with ampicillin to a final concentration of 0.04% in disposable polystyrene Petri dishes (100 mm x 15 mm). Sterilized toothpicks were used to stab transformant colonies

to be spotted on two LB-agar plates at the same relative position for comparison. One of the plates (referred to as 'arabinose plate') contained 0.04% ampicillin, 10 mM L-arabinose, 10 mM glucose and 1 mM IPTG (to induce the expression of the FRET-based biosensor and vSET enzyme). Whereas the other plate (referred to as 'glucose plate') contained 0.04% ampicillin, 20 mM glucose and 1 mM IPTG (to completely suppress vSET enzyme expression). Images of fluorescent colonies were captured using custom-made imaging system that was described elsewhere (121). Briefly, the system employed a 175W xenon-arc lamp (Sutter Instrument Company, Novato, CA) as the light source for fluorescence excitation. Wavelength selection was achieved by passing the light through one of a set of band-pass filters housed in a filter-wheel (Sutter): A 426 nm to 446 nm band-pass filter (Chroma Technology Corp.), or 490 nm to 510 nm for excitation of mTFP1 or mCitrine respectively. Filtered light was channelled to a black-fabric shrouded cubicle through a bifurcated fibre optic bundle (Newport Corporation, Stratford, CT) that was positioned to deliver the light to a holder with circular recession designed to house a 10 cm diameter dish. Another filter-wheel housing a number of band-pass filters was placed vertically in close proximity to the illuminated plate within the shrouded cubicle. The band-pass filters used in experiments discussed herein were the 460 nm to 500 nm and the 520 nm to 550 nm band-pass filters for mTFP1 and mCitrine emissions respectively. A Retiga 1300i 12-bit CCD camera (QImaging, Burnaby, BC), placed immediately behind the emission filter, was used to capture images of the fluorescent intensities of the screened colonies. Positioning of excitation and emission filter-wheels, image capturing

and processing were automated by custom macros run using Image Pro Plus (Media Cybernetics Inc., Silver Spring, MD). Three images using the following combination of excitation and emission filters were captured for each plate: 'donor image' using 426-446 nm excitation and 460-500 nm emission filters; 'FRET image' using 426-446 nm excitation and 520-550 nm emission filters; and 'acceptor image' using 490-510 nm excitation and 520-550 nm emission filters. Values of fluorescent intensities of the three images for both 'arabinose' and 'glucose' plates that had the same clones were exported to Microsoft Excel program. Ratios of sensitized emission of mCitrine (FRET image intensity values) to mTFP1 emission (donor image intensity values) were calculated for both plates. FRET change was calculated using the equation $\Delta R/R_{min}$% (49) using the formula provided in Figure 22. Colonies that showed the highest FRET change were propagated and their plasmid DNA was extracted and sequenced.

2.3 Results and discussion

Eukaryotic DNA is organized into macromolecular complex known as chromatin which consists of building blocks referred to as nucleosomes (122, 123). Each of these building blocks consists of a hetero octameric protein assembly of two units of histone H2A, H2B, H3, and H4 proteins twined by two turns of dsDNA of around 145-147 base pairs (bp) in length (122, 123). PTMs of the histone proteins – in addition to other mechanisms collectively known as epigenetic mechanisms (124) – result in the recruitment of various proteins (or protein complexes) that are involved in executing different biological functions (125). One of these PTMs

is the addition of methyl group(s) to the epsilon nitrogen of histone lysine residues, which is catalyzed by histone lysine methyltransferase enzymes (HKMTs) (126).

HKMTs are involved in a number of biological functions such as heterochromatin formation, X-chromosome inactivation, and transciptional regulation (126). Dysregulation of HKMTs activities has been linked to various diseases and disorders such as psychiatric disorders (127), and cancer (128).

Despite the numerous advantages offered by FRET-based biosensors in studying PTM enzymatic activities dynamically *in vivo*, there have been only one report describing the development of a FRET-based biosensor for the detection of HKMT activity, namely histone H3 Lys27 methyltransferase activity (64). However, this biosensor suffered from a modest dynamic change of 30% (64). Using the methodology described herein, we used vSET, a viral HKMT that contains the SET domain, as a model HKMT enzyme with histone H3 Lys27 methyltransferase activity to optimize the dynamic range of H3K27-MetBio.

Lin *et al.* reported the construction of two FRET-based biosensors to detect histone H3 Lys9 (H3K9) and Lys27 (H3K27) targeted methyltransferase activities (64). The binding domain of H3K9-targeted methyltransferase activity was the HP1 chromodomain that was previously shown to recognize the trimethylated H3K9 sequence context (129, 130). Whereas H3K27-targeted methyltransferase biosensor depended on polycomb chromodomain which is capable of recognizing the trimethylated Lys27 in the context its histone H3 sequence (131, 132).

Following the same design adopted by Lin *et al.* (64), we constructed two series of first generation H3K9-MetBio1 and H3K27-MetBio1 to detect enzymatic methylation of H3K9 (not shown) and H3K27 (Figure 16) respectively). The binding domains we used were the tudor domain of C20orf140 and the tandem hybrid tudor domain of JMJD2A which were reported by Kim *et al.* to bind dimethylated and trimethylated H3K9 respectively (133). Also, we used CDYL1a chromodomain (reported by Kim *et al.* to bind di- and trimethylated H3K9 (133)), and Cbx7 chromodomain which binds both H3K9me3 and H3K27me3 (134-136). We also used two mutated variants of JMJD2A domain and one mutated variant of Cbx7 that were engineered in our laboratory to increase their affinity and binding specificity towards H3K27me3 (by Dr. Hongkin Yap (118)).

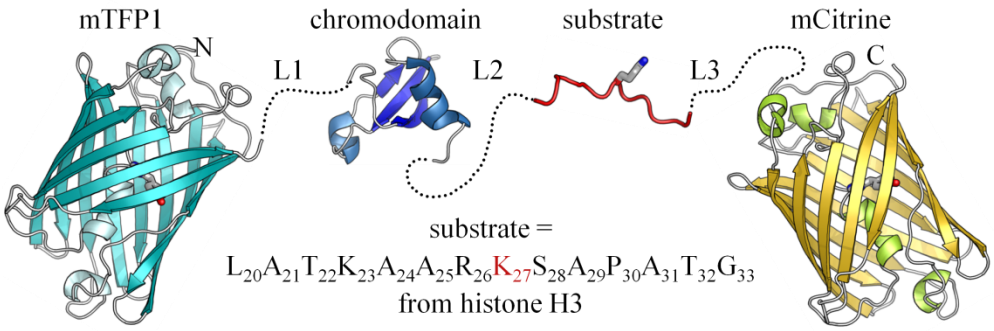

Figure 16 Representation of the H3K27-trimethylation biosensor (H3K27-MetBio). The biosensor is a four-part genetic chimera composed of the FRET donor mTFP1 (116), a domain that binds to methylated lysine residues, residues 20 to 33 of histone H3 (accession CAA40407.1), and the FRET acceptor mCitrine (137). The first (L1), second (L2), and third (L3) linkers are represented with dotted lines. This biosensor design has been previously reported (64).

The *in vitro* methylation assay showed that H3K27-MetBio1-(Cbx7) and H3K27-MetBio1-(Cbx7 A91K) biosensors were the only constructs that displayed FRET-changes in the presence of vSET enzyme and SAM (Figure 17). Cbx7 can bind the trimethylated states of both H3K9 and H3K27 (134-136) with yet better affinity towards trimethylated H3K9 (134, 135), and therefore was expected to be a suitable binding domain to detect both H3K9- and H3K27- targeted methyltransferase activities. However, we did not detect any FRET change in H3K9-MetBio1-(Cbx7) as opposed to its H3K27-MetBio1-(Cbx7) counterpart, suggesting a substrate specificity of the vSET enzyme – a finding that is in accordance with the results obtained by Zhou and coworkers (138-140). The chromodomain of CDYL1a provided by Dr Bedford and verified by sequencing – GenBank accession number AAD22735, Supplementary information (133) – was reported by Kim *et al.* to bind both dimethylated and trimethylated H3K9 (133). However later reports have demonstrated that this chromodomain showed no affinity towards any methylation state of either H3K9 or H3K27 (141, 142), which may explain our results for the lack of response of biosensors that were based on this binding domain.

Experimental results reported by Lin *et al.* which showed that vSET can indiscriminately recognize and methylate both H3K9 and H3K27 (64) seem to be in conflict with our observation of substrate specificity demonstrated by the enzyme, which corroborate the same observations reported earlier by Zhou and coworkers (138-140). However, we do not have enough evidence to assert the substrate specificity of vSET.

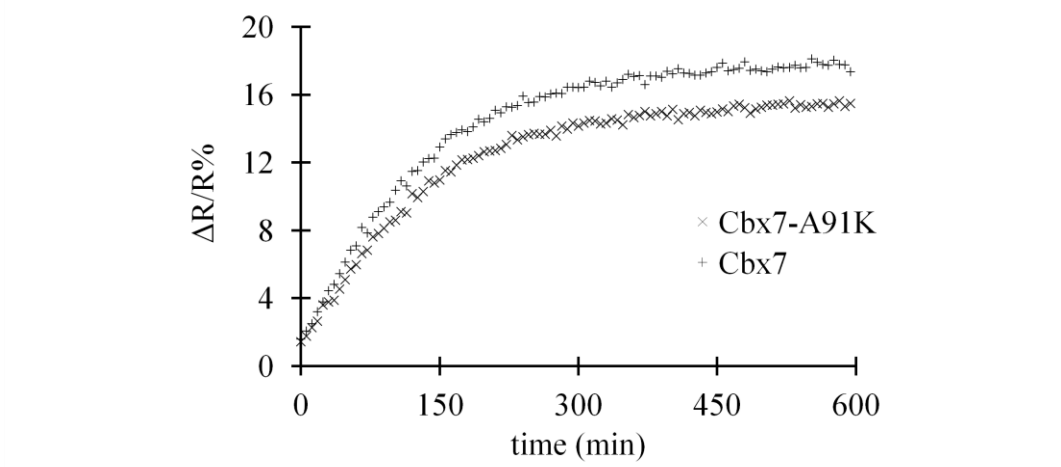

Figure 17 Time-course graph demonstrating the calculated ΔR/R% of H3K27-MetBio1 based on Cbx7 and Cbx7 A91K (work of Dr. Yap (118)) in the presence of vSET and SAM (emission intensities used to calculate ratios were acquired at 494 nm and 526 nm, after excitation at 420 nm).

The lack of response of our H3K9 methylation biosensors may have different origins: substrate specificity of vSET; binding affinities of the chosen protein domains; hindered intramolecular binding arising from improper positioning of the substrate motif with respect to the binding domain in the biosensor. However, we did not investigate this point any further and instead, we decided to devote our attention to the optimization of H3K27-MetBio1-(Cbx7) biosensor.

H3K27-MetBio1-(Cbx7) (Figure 18), hereafter simply referred to as H3K27-MetBio1, displayed a FRET ratio change of 29% (Figure 17and Figure 21B) – which was very similar to the ratio change of the H3K27 methylation biosensor reported by Lin *et al.* (64). Although a FRET ratio change of about 30% may qualify the biosensor to be used successfully in cell imaging or other applications (49), we aspired to develop a methodology by which the FRET signal output of a

given biosensor may conveniently be optimized; and, as a model system, we decided to undertake the task of optimizing H3K27-MetBio1 by changing the composition and the number of amino acids linking the four constitutive domains of this protein chimera.

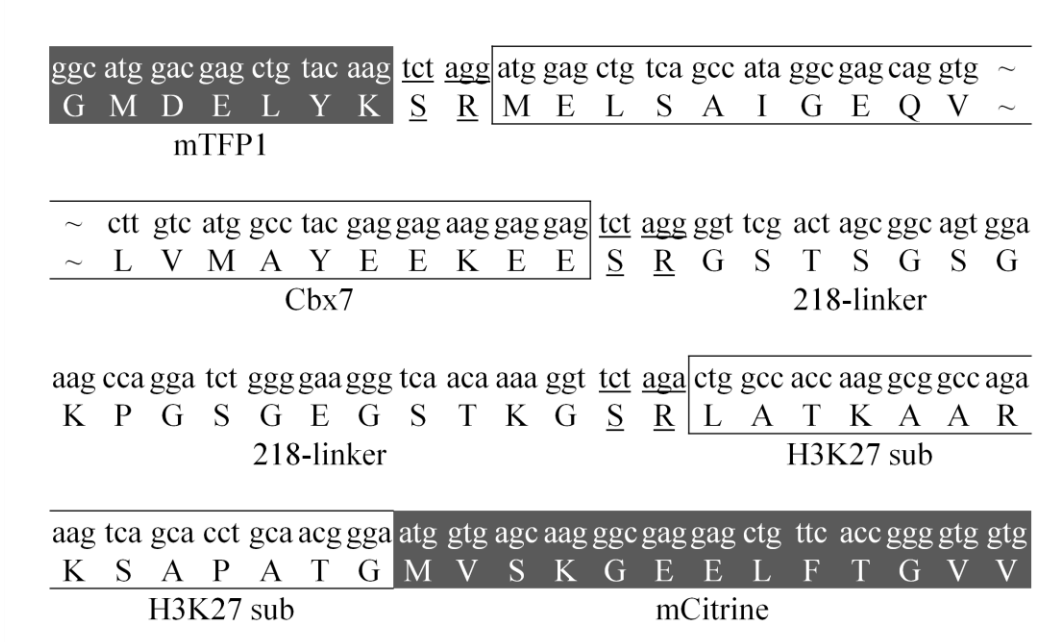

Figure 18 Nucleotide and amino acid sequences of H3K27-MetBio1. The three bolded sequences indicate positions of the restriction sites used. Note that the first and second sites originate from XbaI/AvrII compatible ligation. Amino acid spacer between Cbx7 binding domain and the substrate motif is based on the previously reported 218-Linker (119).

The key for creating a successful biosensor is fine tuning of initial and final distances and orientations of the FP barrels (the FRET pair), coupled to the biological event under investigation, to obtain a suitable FRET ratio change. This may be achieved empirically by different means which will be discussed in the following chapter.

The remainder of this chapter will cover the development of the screening system, and the optimization of the procedures. It will then conclude with a proof-of-concept experiment that demonstrates the functionality of this screening methodology. The application of this system to screen libraries of biosensor constructs for improved signal change shall be dealt with in the proceeding chapter.

Inspired by the fact that fluorescence screening in bacterial colonies has been the technology of choice for the directed evolution of improved FPs, we sought to extend this methodology to the screening of biosensors. However, unlike individual FPs that have a static and unchanging fluorescence, biosensors have a dynamic fluorescence emission that must be imaged in both its initial (*i.e.,* unstimulated) and final (*i.e.,* stimulated) states. Accordingly, the primary challenge of screening biosensors in bacterial colonies is how to induce the enzymatic activity (or other biochemical change) that the biosensor is designed to sense. To address this challenge we have developed a screening system in which the response of a FRET-based biosensor for a PTM can be artificially executed in live bacterial colonies. Our strategy was influenced by the work reported by Kimura *et al.* in which they demonstrated the possibility of using the bacterial cell as a 'test tube' to monitor the signal change of a FRET-based biosensor designed to detect the enzymatic activity of a certain protease (143). This was achieved by placing the DNA encoding the biosensor and the enzyme under two differentially induced promoters in an experimental set-up which allowed for the controlled and sequential expression of the two recombinant proteins (143).

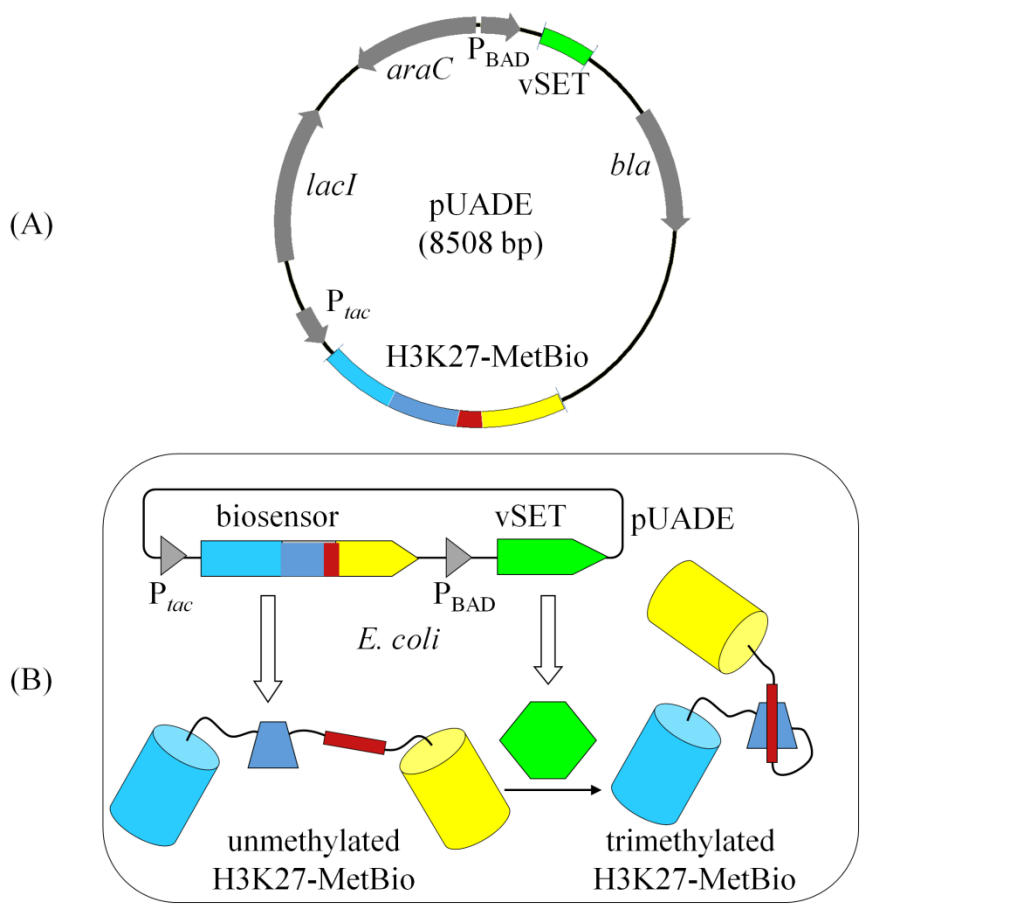

Figure 19 Overview of the colony-based screening strategy. (A) Plasmid map of pUADE. P_{tac} is an IPTG-inducible promoter, P_{BAD} is an L-arabinose-inducible promoter, and the enzyme is vSET. (B) Schematic representation of protein expression from pUADE in *E. coli* in the presence of both IPTG and L-arabinose. In the absence of L-arabinose, the P_{BAD} promoter is inactive and no vSET is produced.

Following a similar approach, we decided to use the two promoters P_{tac} (144) and P_{BAD} (145) to control the expression of the H3K27-MetBio and the vSET enzyme respectively. However, to facilitate the experimental procedures and eliminate the need for more than one resistance marker, we incorporated the two promoters into one plasmid which we decided to refer to as 'pUADE' (Figure 19).

As a first step, we set to determine if the promoters in pUADE retained their functionality and were amenable to artificial induction. Towards that goal, Ni-NTA purified protein preparations from differentially induced cultures of pUADE-transformed *E. coli* were analyzed using SDS-PAGE. Coomassie blue staining of the polyacrylamide gel revealed the presence of two major bands at positions that agreed with the molecular weights of the two encoded proteins: a band of about 13 kDa (vSET enzyme); and another of about 44 kDa (a protein fusion with mCitrine). The appearance of the protein band which corresponded to the P_{BAD}-controlled vSET matched the induction profile and was only visible in protein preparations obtained from L-arabinose induced cultures. However, the P_{tac} controlled protein gave a visible band in all protein preparations even in those obtained from uninduced cultures (Figure 20A).

Although bacterial liquid culture was not the intended experimental setup for library screening, we reasoned that it would be a more convenient way to obtain a reasonable amount of recombinant proteins to assess the functionality of the plasmid. With that point established, further experiments conducted to determine plate composition suitable for screening purposes (that is the concentration of L-arabinose and glucose to turn on or off the expression of the vSET enzyme) were performed using ElectroMAX DH10B™ cells grown on LB-agar plates; and the expression level of vSET was assayed with the more sensitive method of Western blot.

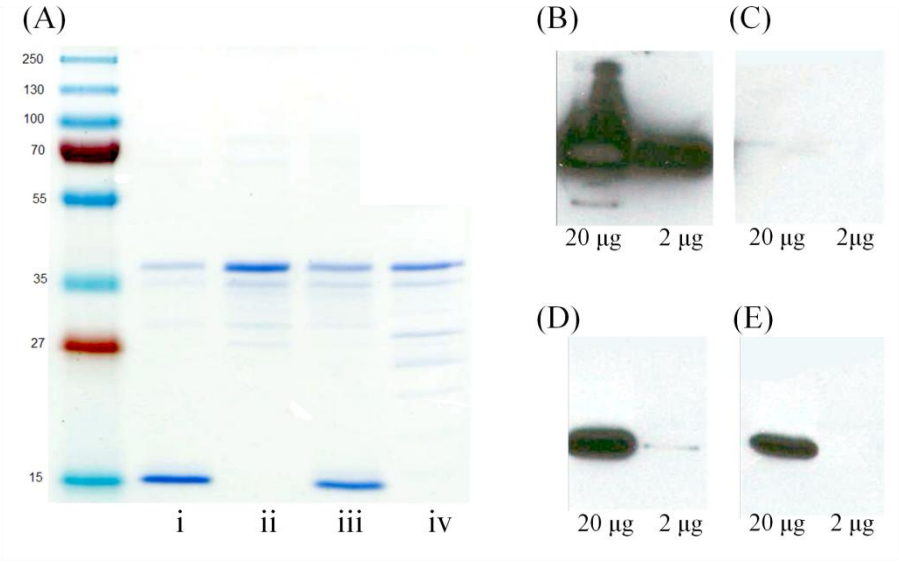

Figure 20 Induction and repression profile of pUADE. (A) Coomassie blue stained polyacrylamide gel of recombinant proteins produced by different induction conditions of pUADE. Each lane contains 2 microgram of Ni-NTA purified protein from: i) L-arabinose induced; ii) IPTG induced; iii) L-arabinose/IPTG induced; and iv) non induced cultures. (B-E) Western blots showing expression of vSET for colonies of *E. coli* grown in the presence or absence of inducer. For each condition, 20 µg was loaded in the left lane and 2 µg was loaded in the right lane. (B) Media containing IPTG (1 mM) and L-arabinose (13 mM). (C) Media containing IPTG (1 mM) and D-glucose (11 mM). (D) Media conditions identical to (C), 3 hr after being sprayed with a concentrated solution of L-arabinose. (E) Media containing IPTG (1 mM).

Although vSET was not visually detected on the Coomassie blue stained gel of the protein preparation from cultures that lacked L-arabinose, low level of transcription 'leakiness' of uninduced P_{BAD} (145) was confirmed by Western blot analysis (Figure 20E**)**. However, unlike P_{tac}, P_{BAD} is subject to catabolite repression (146) and it provided a tight control of vSET expression in the presence of glucose. Based on Western Blot analysis, the expression level of vSET in colonies grown on 11 mM glucose/1 mM IPTG plates is estimated to be less than 0.02 µg of the enzyme per 20 µg total soluble protein extract (Figure

20C). Induction of vSET expression by L-arabinose spraying of *E. coli* colonies grown on plates of similar composition (11 mM glucose/1 mM IPTG) resulted in the accumulation of vSET enzyme to a concentration at least an order of magnitude higher than its level obtained from uninduced plates, over a time period of 4 hr (Figure 20D). Colonies grown on 13 mM L-arabinose/1 mM IPTG seem to have accumulated a large amount of the enzyme, roughly estimated to be more than 10% of the total bacterial protein (Figure 20B). However, the size of the colonies was very small suggesting a metabolic burden on the hosting cells (147).

Results obtained from Western blot analysis were not enough to make a decision regarding the plate composition for screening experiments, and different L-arabinose and glucose concentrations had to be tested. However, further analysis of the vSET expression level was not determined by Western blot technique for two reasons: we were not able to obtain any signal of the 6xHis-tagged enzyme when the concentration of the standard vSET enzyme was dropped to 2 ng in the standards lanes (not shown); and the signal of the band obtained from colonies grown on 11 mM glucose/1 mM IPTG seems to occupy a position that suggests a slightly bigger protein size than vSET.

To prevent any misinterpretation due to ambiguities, we resorted to fluorescence analysis of the FRET signal of the H3K27-MetBio1. This step was based on comparing FRET/mTFP1 ratio averages calculated for colonies grown on plates of different compositions (*i.e.*, different concentrations of glucose and/or L-

arabinose) to the *in vitro* FRET/mTFP1 ratio change of the same biosensor in the presence and absence of vSET.

Our experimental results suggested that the biosensor behaved similarly *in vivo* and *in vitro*; and colonies grown on glucose plates (suppressed vSET expression) had a low FRET/mTFP1 ratio averages while those grown on L-arabinose plates (expressed vSET) showed higher ratio averages (correlating with a low FRET/mTFP1 ratio which increases by 29% in the presence of vSET when the methylation assay is performed *in vitro*). FRET/mTFP1 ratios of colonies grown on 20 mM glucose showed a lower average than those grown on 10 mM glucose plates, therefore the former concentration was decided on for suppressing vSET expression. On the other hand, colonies grown on 10 mM L-arabinose/10 mM glucose showed a higher FRET/mTFP1 ratio average than those grown on 20 mM L-arabinose/10 mM glucose, and likewise the former composition was chosen to turn on vSET expression (Figure 21A). Unlike colonies grown on plates containing L-arabinose and IPTG only, those spread on L-arabinose plates that also contained glucose grew to a bigger size (expected of this strain of *E. coli*). Probably adding glucose 'tuned down' the foreign protein expression to a level metabolically less burdensome.

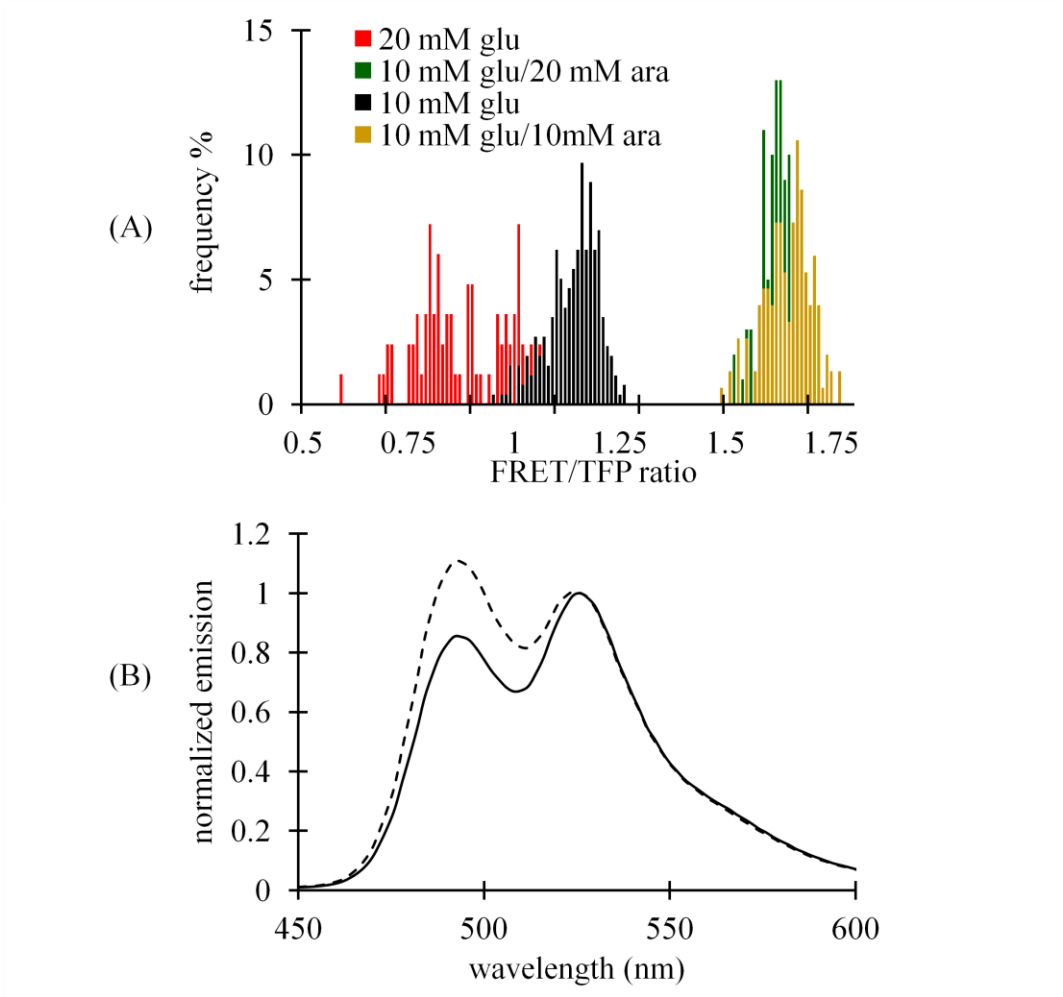

Figure 21 Screening plates composition optimization. (**A**) Image-based emission ratios for H3K27-MetBio1 in colonies of *E. coli* that were grown on different plate compositions. (B) Emission spectra of the H3K27-MetBio1 protein in both the methylated (solid line) and unmethylated state (dashed line).

Given that *in vitro* FRET change could be calculated as $(R_{Final} - R_{Initial})/R_{Initial}$, it seemed reasonable to substitute this formula for $(R_{arabinose} - R_{glucose})/R_{glucose}$ to calculate the FRET changes *in vivo* (Figure 21). To test this hypothesis, we transformed *E. coli* ElectroMax DH10B cells with pUADE[P_{BAD}(*vset*)/P_{tac}(*lib1*)] and proceeded as described under 2.2.7 Screening (protocol-1). To obtain $R_{arabinose}$

($I_{A,ara/IPTG}/I_{D,araIPTG}$) and $R_{glucose}$ ($I_{A,glu/IPTG}/I_{D,glu/IPTG}$) for screened cloned, we manually spotted the bacterial colonies on the same relative position of two plates 'arabinose' and 'glucose' plates as described in details in the experimental section (Figure 22). After data processing, ten clones that showed the highest ΔR/R% were picked and submitted to MBSU for identification. Consistent with our *in vitro* observations, sequencing results showed that 9/10 of the picked clones had Cbx7 as the binding domain.

This proof-of-concept experiment demonstrated the utility of our dual expression vector pUADE, used along the proposed experimental procedures, to be used for screening libraries of FRET-based biosensors.

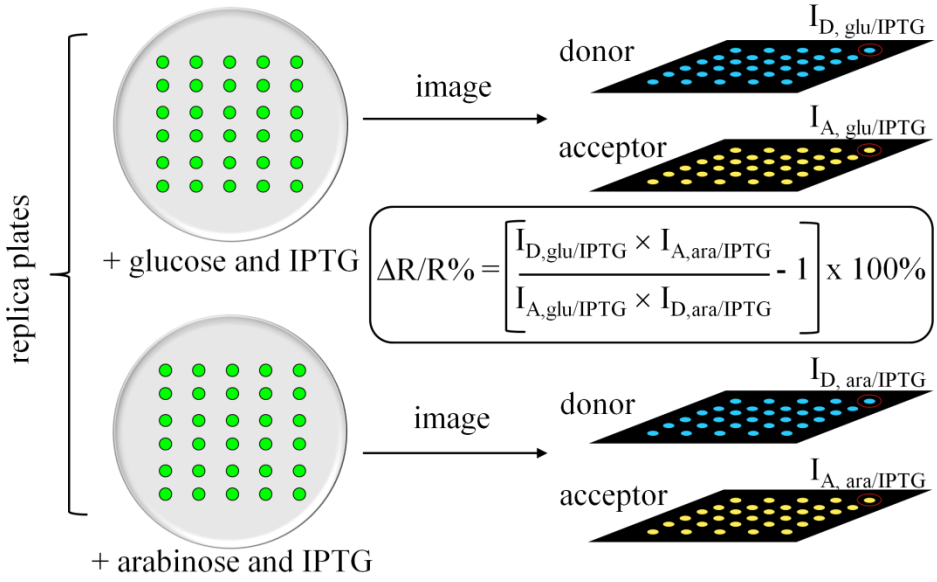

Figure 22 Library screening methodology. To screen libraries of biosensor variants for clones exhibiting improved ratio changes, an image-based method was used to determine FRET ratio changes in individual colonies of *E. coli* previously transformed with the plasmid library.

2.4 Conclusion

In this chapter, we have demonstrated that our methodology which is based on the differential expression of two recombinant proteins *in colony* can potentially be used to screen libraries of variants of a given PTM FRET based biosensor (as shall be seen in the proceeding chapter). Our proof-of-concept experiment confirmed the ability of this screening methodology to differentiate between biosensors designed with the appropriate MRE to detect the methyltransferase activity of vSET from others that cannot detect it. Our methodology can be applied to design and optimize FRET-based biosensors capable of detecting other PTM enzymes.

Another potential application of this system is to be used to screen libraries of MREs to find either the cognate substrate motif or the binding motif for a given PTM enzyme, provided that the enzyme is produced in its active form in *E. coli*. However, it should be noted that false negatives in this case might limit this potential application. These false negatives can arise from either very small FRET ratio change, that might not be detectable using this screening protocol, or from the lack thereof due to unsuitable design that hinders intramolecular recognition upon substrate motif modification.

Chapter 3:
Dynamic range optimization of H3K27-MetBio

3.1 Introduction[3]

Owing to the several advantages genetically-encoded FRET-based biosensors enjoy, the FRET modulation design strategy has become a very popular approach for constructing a wide array of biosensors to gain information about diverse biological events. However, a frequently encountered disadvantage of this group of biosensors is the small dynamic range of their FRET response which limits their sensitivity (101). Optimization of the dynamic range and other properties of the biosensor are, therefore, crucial developmental steps to obtain a reliable biosensor capable of faithfully conveying an accurate picture of the biological event under investigation however small or transient in nature it might be. There are several goals that should be sought, but never independently from one another, during the optimization of a given biosensor: improving the dynamic range of the biosensor to increase its sensitivity to detect small concentration changes; tuning its detection range (*i.e.*, apparent K_d') to overlap with the physiologically relevant concentrations; optimizing its response kinetics to allow for the detection of rapid concentration fluctuations; increasing its selectivity for the cognate analyte over potentially interfering molecules present in the same milieu; decreasing its sensitivity to changes in environmental factors; and increasing its autonomy to circumvent diminished response signals arising from untoward interference from the native biological molecules with the MRE.

[3] Parts of this chapter were published by Ibraheem *et al.* 2011 (118)

In this introductory section, our main focus will be the discussion of a number of approaches and methodologies reported in the literature that have been used to optimize the response dynamic range of different FRET-based biosensors. However, it should be noted that optimization of biosensors should be considered on a case-to-case basis. That is, optimum biosensor elements benefiting the response of a given biosensor may not be beneficial to others and may even result in inferior responses, especially if they were applied without further 'fine tweaking'. For example, introducing the A206K mutation, which induced monomerization of avGFP derived variants that otherwise naturally dimerize at high concentrations (148), in the FRET pair used to construct a Src kinase FRET based biosensor improved its dynamic range from 25% to 43% (149). However, when this mutation was introduced into the FRET pair of previously reported biosensors (150, 151) constructed to detect protein kinase C activity, their dynamic ranges were severely diminished (152). We contend that neither of these two seemingly contradictory findings should be used to postulate any universally applicable rule about the dynamic range impact that might result from employing these monomerized FP variants. Moreover, the latter biosensors, with the monomerizing mutations, might not be after all impervious to dynamic range 're'-optimization. It would be interesting to see the outcome of subjecting these biosensors to other optimization strategies (*e.g.*, linker length and composition optimization or topological redesigning). The dynamic range of the originally reported biosensors – both very modest – might have benefited from the natural propensity of avGFP derived variants to dimerize at high effective concentrations

in one state or the other of such biosensors (152), but the conformational changes of the MRE in their design might have not been optimal to cause adequate FRET efficiency changes in the biosensor constructs using the monomerized FP variants, a problem which is very likely to be amenable for improvement by compositional re-optimization of these biosensors.

The problem of sub-optimal tuning of FRET efficiency to the conformational change of the MRE becomes very clear in biosensors employing PBPs as the MRE (49). The PBP family provides researchers with a wealth of binding domains of varying affinities and specificities to a wide range of analytes. Generally these proteins consist of two lobes connected by a hinge region (70). Upon binding their cognate analyte, they undergo a hinge-twist movement to engulf the analyte (described as Venus fly-trap or clam shell movements). When these PBPs are employed to construct FRET-based biosensors following the general design of sandwiching the MRE between the FRET pair, such constructs show modest dynamic range, especially if the linker regions are not properly designed (50). These 'composite linkers' (50) may cause rotational averaging of the FPs dipoles. This potential problem may be alleviated by tethering the FPs more tightly to the biosensor thus reducing the degrees of freedom available to the FPs. Indeed, it has been shown that systematic truncations of the aforementioned composite linker regions has been very useful to improve the dynamic range of FRET-based biosensors employing PBPs as the MRE (50, 57, 58, 72). These truncations might have afforded tighter allosteric coupling of the FPs to the motion of the PBPs used or reduced rotational averaging thus improving the

signal changes (50). A similar trend of dynamic range improvement observed with reducing the degrees of freedom available for the FPs has also been reported for other biosensors (105, 153, 154).

Another way to improve the dynamic range of FRET-based biosensors employing PBPs as their MRE, possibly through tuning of the relative orientation, was demonstrated by Deuschle *et al.* in 2005 where they changed the topology of a previously reported glucose biosensor FLIPglu-600μ (-12% emission ratio change) (155) by examining the internal insertion of the donor FP in a number of different locations in the sequence of the employed PBP where it was 'displayed' on its surface. The best variant they identified was given the name FLII^{12}Pglu-600u and had an emission ratio change of 58% (50). Similarly they improved the dynamic range of a previously reported glutamate biosensor FLIPEglu-600n (-11%) (156) to FLII^{81}PE-1μ (-48%) (50). Furthermore, the problem of sub-optimal tuning of FRET efficiency to the conformational change of the MRE was addressed by Okada *et al.* using a different approach (157). By circularly permuting members of PBP family, they were able to successfully construct a number of FRET-based biosensors with improved dynamic ranges, even using members which when fused to a FRET pair in their wild-type nature failed to give any FRET change (157).

The aforementioned approaches aimed at increasing the dynamic range of biosensors by optimizing the relative orientation and/or by tight allosteric coupling of the employed FPs to the conformational change of the MRE by truncating the unnecessary composite linkers, and/or internal fluorophore

insertion. Another way of tuning the relative orientation of the FPs in FRET-based biosensors to maximize the dynamic range is to use circularly permuted variants of the FPs. Circular permutation of the FPs was first reported in 1999 (53, 54). Using circularly permuted variants of FPs has resulted in dramatic improvement of the dynamic range of a number of biosensors (104, 154, 158).

However, recent findings suggest that dynamic range improvement achieved by employing circularly permuted variants of the FPs may not be entirely attributed to the relative orientation of the FRET pair and could also be attributed to anti-parallel interaction of the FPs (102). In their study, Kotera *et al.* examined the impact on response dynamic range different dimerization propensities of employed FPs of a given biosensor could cause. They systematically modified the dimerization interface of avGFP derived variants employed by a number of previously reported biosensor (*e.g.*, the calcium biosensors TN-XL (158), and YC3.60 (104) both of which had circularly permuted variants of mCitrine and Venus respectively and the avGFP wild type dimerization interface, in addition to other biosensors (102)) to the wt dimerization interface of avGFP (prone to dimerization at high concentrations), monomerized variants interface (A206K mutation (148)), and dimerization-enhanced variants (containing the mutations S208F/V224L (103, 159)). Both TN-XL and YC3.60 responded in the same pattern to these mutations: the dynamic ranges were reduced with both monomer-associated and dimerization-enhancing mutations, while they were optimum with the natural avGFP interface (102). These observation in addition to others reported in this study led the authors to conclude that: according to the mode of

analyte recognition of the biosensor (remaining intact as in calcium biosensors or getting cleaved as in protease reporters), different degrees of dimerization propensities could be favourable (102).

Although Kotera *et al.* have provided the scientific community with compelling evidence that reversible FRET-based biosensors (those that do not get cleaved) may benefit from the weak natural tendency of avGFP variants to dimerize (102) (on par with results reported by Jost *et al.* (152)), others have reported that the dynamic ranges of their biosensors were improved by using dimerization-enhancing mutations (160), while others reported dynamic range improvements attributed to the monomerized versions (149). Based on the observation reported by Hires *et al.* (72) that changing the FRET pair in their linker-optimized glutamate biosensor $GluSnFR_{8N5C}$ (which they obtained after systematic truncations of the N- and C- termini of the binding domain used) that employed ECFP-mCitrine to other variants led to decreased dynamic range (72), we contend that changing the FRET pair to another may bring the optimization process to 'square one', unless the motion of the MRE was by mere chance tuned to produce high FRET change in the new adopted pair. However, that may not preclude the possibility that this biosensor could be re-optimized with the new FPs.

Until now, we have only discussed different reported approaches that aim to optimize the dynamic range of some biosensors without considering the potential impact on the detection range and other properties of biosensors these modifications may incur. It could hardly be surprising to know that subjecting the MRE to operations such as FP insertion or circular permutation could alter its

analyte binding affinity (50, 157). But even when the identity of the MRE is not modified, different linkers or properties of the FPs can impact the detection range of the biosensor (102, 161, 162). For example, Kotera *et al.* observed different apparent K_d' of the calcium biosensors they tested by changing the propensity of the FRET pairs to dimerize (102). While this might be a side effect of dynamic range optimization, the impact on the apparent K_d' brought about by different linker lengths, separating two parts of a given MRE, was used to generate a series of FRET-based biosensors capable of detecting zinc over a wide range of concentrations – which has been explained in terms of the varied effective concentrations of the two halves of the MRE (161, 162). Different linker lengths have been also observed to affect the response kinetics of biosensors(143).

A general outline of designing a useful biosensor could be as follows: The first step of building a biosensor is the judicious choice of an MRE if there exists a number of potential candidates, or subjecting a given MRE to a protein engineering method to tune its detection range to overlap with the expected physiological concentrations it is intended to monitor, and to eliminate undesirable qualities such as promiscuous analyte binding, slow response kinetics, or insufficient autonomy, by which we refer to the untoward interference of biological molecules with the action of the employed MRE as was the case in the calcium biosensors employing calmodulin and M13 (163). After the proper MRE is obtained, response dynamic range should be maximized. There are several approaches by which this could be achieved such as changing the lengths and/or composition of the linker regions tethering the constitutive domains, using

circularly permuted FP variants, changing the degree of adhesiveness of the FPs used, and changing the topology of the biosensor. All of these approaches share the same principle of finely tuning the relative distance and orientation changes of a FRET pair to the biological event under investigation to maximize FRET efficiency changes and consequently increase the sensitivity of this biosensor. During the process of dynamic range optimization, other properties of the biosensor, such as detection range or response kinetics, should also be monitored, especially if the MRE is subjected to excessive manipulation.

Although computational methods for prediction of FRET changes has witnessed some progress (164) and has been successfully used (105, 165), empirical approaches remain by far the most reliable and successful means to achieve optimization. In a recent report, Piljic *et al.* describe a methodology developed to rapidly obtain a biosensor with optimized properties (107). Briefly, their technique depends on the cloning of a given MRE in a number of plasmids containing different FRET pairs with linkers of different lengths, then spotting these clones on a special chamber glass where they are used to reverse transfect mammalian cells, followed by examining the behaviour of the *in vivo* expressed biosensors using fluorescent microscopy (107). However, to our knowledge there has not been any report describing an *in vivo* methodology for rapid optimization of PTM FRET-based biosensors.

In this chapter, we will describe using the dual expression plasmid pUADE that has been developed in our laboratory and discussed in the previous chapter to screen libraries of hundreds of H3K27-MetBio clones with different linker lengths

for structurally optimized variants with improved FRET emission ratio changes. Provided that a given PTM enzyme could be produced in an active form in *E. coli*, our screening methodology could be used to rapidly obtain a cognate biosensor with improved dynamic range. We will also describe a different screening methodology which could be used, once optimized, to screen libraries on a scale of thousands of variants.

3.2 Materials and methods

3.2.1 General material and methods

Laboratory equipments and supplies, chemical compounds, enzymes, general procedures and protocols were described in details in the preceding chapter.

3.2.2 Construction of lib2 and lib3 linker libraries

To facilitate cloning of linker libraries, we incorporated new restriction sites between the constitutive parts of the H3K27-MetBio to have the following scheme: *XhoI-mtfp1-KpnI-cbx7-SalI-h3k27-EagI-mcitrine-HindIII*. To introduce the new restriction sites in the MCS downstream of P_{tac} in pUADE, a stepwise assembly of an initial construct with these new sites was performed followed by its introduction in pUADE. Briefly, double stranded DNA encoding the H3K27 substrate was obtained by slowly cooling a buffered solution that contained both coding polynucleotide strand 'SalI(RF2)-H3K27-Sense' and its reverse compliment 'EagI(RF3)-H3K27-Antisense' (at 0.5 µM each) from 95 °C to 10 °C at temperature drop rate of -1 °C every 3 min. The product was purified using

QIAquick Nucleotide Removal Kit after which it was digested with SalI and re-purified. The obtained product was ligated with SalI digested *cbx7* PCR product that had been previously amplified using 'FrWd-KpnI(RF1)-Cbx7' and '0-BkWd-SalI(RF2)-Cbx7' primers. After four hours of ligation, an aliquot of the reaction mixture was used as the template in a PCR reaction with the primers 'LigCbx7-H3K27-forward' and 'LigCbx7-H3K27-Backward'.

Purified product from the previous step was digested with EagI and ligated with similarly digested *mcitrine* PCR product that had been previously amplified using '0-FrWd-EagI(RF3)-YFP' and 'YFP-Primer 4' primers. The ligation reaction was allowed to proceed for four hours before an aliquot was removed and used as the template in a PCR with the 'FrWd-KpnI(RF1)-Cbx7' and 'YFP-Primer 4' primers. The product of this PCR reaction was a 3-part chimera encoding Cbx7-H3K27-mCitrine. This product was digested with KpnI and ligated with KpnI-digested *mtfp1* PCR product that had been previously amplified using the 'TFP-Primer 1' and '0-BkWd-KpnI(RF1)-TFP' primers. The ligated construct was subsequently subjected to an amplification step using 'TFP-Primer 1' and 'YFP-Primer 4' primers to produce the full 4-part chimera encoding H3K27-MetBio with the new restriction sites. This PCR product was digested with XhoI and HindIII and ligated with the pUADE[P_{BAD}(*vset*)/P_{tac}(*h3k27-metbio1*)] plasmid that had been digested with the same enzymes. This plasmid was then used for the construction of the linker lengths.

The second library (lib2) consisted of 392 variants of possible combinations of: seven different representatives of L1; eight different linkers for L2; and seven

different linkers for L3. Diversification at L1 region was introduced into the library through PCR using *mtfp1* as the template where residue Gly225 was chosen as the reference point (116); and the linker which omitted all proceeding mTFP1 residues was designated as L1 = 0. Other linkers of this region had 2, 4, 6, 8, 10, and 14 amino acid additions relative to Gly225 of mTFP1. The 3' end of linker-1 region had the restriction site *KpnI* which connected *mtfp1* with the *cbx7*, however amino acid residues encoded by this palindrome (glycine, threonine) were not counted in the linker lengths. Likewise, Cbx7 encoding DNA served as the template in a series of PCRs to incorporate the second linker region variations into the library construction. The first linker of this series, L2 = 0, had no amino acid addition between the last residue of the chromodomain and the nucleotide sequence 'ggg tcg aca' which encoded *SalI* restriction site (underlined). Other linkers designed for this region had 2, 4, 6, 8, 10, 14, and 20 amino acid spacers upstream of the aforementioned nucleotide sequence. The third region targeted for variation was L3 which connected the substrate to mCitrine. The last codon of *h3k27* substrate motif was immediately proceeded by the nucleotide sequence 'acg gcc ggc' which encoded *EagI* restriction site (underlined); and the first linker of this series, L3 = 0, had no further amino acid additions between the latter nucleotide sequence and that encoding Glu6 of *mcitrine* (137). Other linkers of this region had 2, 4, 6, 8, 10, and 14 amino acid spacers relative to Glu6 of mCitrine which was used as the template in the PCRs that introduced these different linkers into the library. Each linker region in this library was thus represented with different linker lengths which spanned from no inter-domain

amino acid spacers up to fourteen amino acids (L1 and L3) and 20 amino acids in case of L2.

The third library (lib3) was constructed with further trimming of the C- terminal of mTFP1 (that is, beyond Gly225) and N- terminal of mCitrine (after Glu6). This library consisted of 640 variants arising from the possible combination of ten linkers from L1 region; eight linkers from L2 region; and eight members from L3 region. L1 was given one of the following values: -4, -3, -2, -1, 0, 1, 2, 3, 4, or 5 which indicates the number of added (positive value) or truncated (negative value) amino acids relative to residue Gly225 of mTFP1 . L2 region encoded 6, 8, 10, 14, 15, 20, 25, or 30 amino acid spacers between the last codon of Cbx7 chromodomain and the nucleotide sequence containing *SalI* restriction site (ggg tcg aca). Similarly, L3 acquired one of the values: 3, -2, -1, 0, 1, 2, 3, or 4 which indicates the number of added or truncated amino acids relative to residue Glu6 of mCitrine.

PCR amplifications with each pair of primers were done individually, and the resulting PCR products pooled prior to digesting and ligation into pUADE. Following ligation, 3 µL of the reaction mixture were used to transform 60 µL of electrocompetent *E. coli* which were then added to 3 mL of fresh LB medium containing ampicillin in a 10 mL culture tube. After 16 h at 37 °C with shaking at 220 rpm, the culture was spun down and plasmid DNA was extracted. The obtained plasmid DNA was then digested with EagI and HindIII enzymes to be used in ligation reaction with mCitrine library. Ligation, transformation and plasmid DNA preparation proceeded as described above. Similarly, the KpnI and

SalI digested Cbx7 library was ligated into the pUADE plasmid that contained mTFP1 and mCitrine libraries.

PCR mixtures were assembled to have the following composition: 20 mM Tris-HCl (pH 8.8), 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 0.1 mg/ml BSA, 0.1% (v/v) Triton X-100, 2 mM $MgSO_4$, 0.2 mM of each dNTP, 0.5 µM of each forward and backward primers, 0.1 µg DNA template and 1.25 U *Pfu* DNA polymerase. Thermal cycling parameters used for mTFP1 and mCitrine amplification were: (1) 95 °C: 2 min; (2) 35 cycles of 95 °C: 45 sec, 64 °C: 30 sec and 72 °C: 1.5 min; (3) 72 °C: 10 min. Whereas those used to amplify Cbx7 were: (1) 95 °C: 2 min; (2) 35 cycles of 95 °C: 45 sec, 65 °C: 30 sec and 72 °C: 1 min; (3) 72 °C: 10 min.

3.2.3 Screening of linker libraries

Screening of the linker libraries proceeded as described in the previous chapter.

3.2.4 Construction of lib4 rearranged-domains library

To further examine the impact of the biosensor design on its response dynamic range, we constructed sixteen different libraries (lib4-A - P), each with a unique arrangement of the four domains constituting the biosensor. For simplicity, we arbitrarily assigned the letters ‘A’, ‘B’, ‘C’, and ‘D’ to refer to mTFP1, Cbx7, the substrate motif, and mCitrine, respectively. Figure 23 demonstrates the different designs of these libraries.

3.2.4.1 Construction of the plasmids

To facilitate cloning, sixteen different MCSs, encoding the chosen palindromes, were first introduced into pUADE downstream of P_{tac} between the restriction sites *XhoI* and *HindIII*. This step was performed by primer annealing reaction of primers designed to include 5' and 3' overhangs followed by ligation into the restricted pUADE. Briefly, dsDNA fragments with 5' and 3' overhangs compatible for ligation with XhoI and HindIII restricted DNA respectively was obtained by slowly cooling a buffered solution of a given 'sense' primer with its respective 'antisense' complementary primer (*e.g.*, 'ABCD-sense' with 'ABCD-antisense' to yield the MCS for lib4-E), at a concentration of 0.5 mM each, from 95 °C to RT at a temperature drop rate of 1 °C/3 min. The obtained dsDNA was ligated into XhoI-HindIII (New England Biolabs) digested pUADE.

3.2.4.2 PCRs and substrate assembly

PCRs were carried out as described above. For mTFP1, 84 reactions were separately assembled using the 'n-FrWd-XhoI(RF1)-TFP' and 'n-BkWd-KpnI(RF1)-TFP' primer series, where 'n' denotes added or truncated amino acids. Similarly, 72 and 33 PCRs were performed for mCitrine, using the primer series 'n-FrWd-EagI(RF3)-YFP' and 'n-BkWd-HindIII(RF1)-YFP', and Cbx7, using 'n-FrWd-AvrII(RF1)-Cbx7' and 'n-BkWd-SalI(RF2)-Cbx7', respectively.

Four substrate motifs were assembled separately using the four sense/antisense pairs of L-S-Sub, L-L-Sub, S-L-Sub, and S-S-Sub primers, where 'S' and 'L' stand for short and long respectively (check Appendix for sequences).

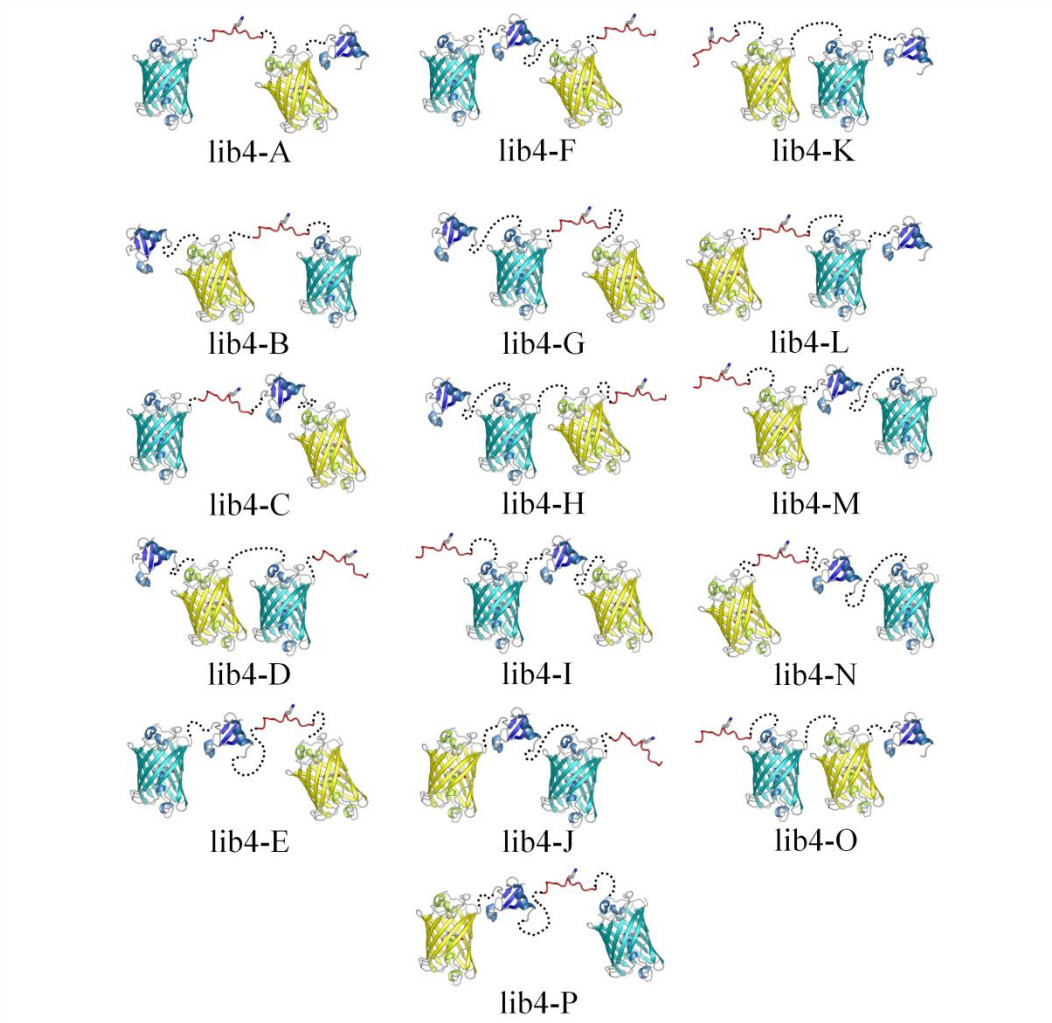

Figure 23 Arranged domain Lib4-(A-P).

3.2.4.3 Assembly of the libraries

Assembly o the lib4-(A-P) proceeded essentially as described for lib2 and lib3.

3.2.4 Recombinant protein production and *in vitro* FRET assay

Selected FRET-based biosensor clones were subcloned into pBAD/His B vector between *XhoI* and *HindIII* restriction sites before they were expressed. Recombinant protein production and purification was performed as described in the previous chapter.

The *in vitro* FRET assay was performed in Corning 96-well plates. Each reaction contained 200 µM SAM, 1 µM vSET enzyme, 1 µM FRET-based biosensor protein, and 2 mg/mL BSA. Reactions were buffered with 50 mM sodium phosphate, 300 mM sodium chloride, pH 8.0. Excitation wavelength was set to 420 nm and fluorescence emission was collected at 494 nm and 526 nm wavelengths. Measurements were collected at a time interval of 2 min for 6 hr using TECAN Safire^2 microplate reader.

3.2.5 Alternative screening procedure (protocol-2)

Transformation efficiency of electrocompetent *E. coli* strain ElectroMAX DH10B™ (Invitrogen Inc) was determined for pUADE and pBAD/His B (Invitrogen) (166). The same bacterial strain was transformed with pBAD/His B encoding mCitrine, and the two plasmids pUADE[P_{BAD}(*vset*)/P_{tac}(*h3k27-metbio3*)] and pUADE[P_{BAD}(*nohkmt*)/P_{tac}(*h3k27-metbio3*)] (where *h3k27-metbio3* is the DNA encoding the optimized H3K27-MetBio obtained after screening lib3, and *nohkmt* is a non histone lysine methyltransferase encoding DNA). The amounts of transformed *E. coli* expected to contain 600-800 CFU were spread on LB-agar plates (containing 10 mM glucose, 1 mM IPTG and 0.04% ampicillin). Plates

were incubated at 37 °C overnight and were subsequently transferred to a fridge where they were incubated for 16 hr at 4 °C.

After incubation, three images (Ex-donor/Em-donor), (Ex-donor/Em-acceptor), (Ex-acceptor/Em-acceptor) were acquired for each plate. After that, plates were sprayed with 200-250 µL of 1 M L-arabinose solution using a plastic atomizer spray bottle, incubated at RT for 4-6 hr and were used to acquire another three similar images. Using ImagePro the six images of each plate were processed and the values of fluorescence intensity and colony size were exported to a Microsoft Excel spreadsheet where data was analyzed (Figure 31).

3.3 Results and discussion

Based on the results obtained from screening the first library, our screening methodology seemed capable of winnowing out clones which did not detect the methyltransferase activity of vSET from those capable of its detection. And having identified the preferred binding domain for H3K27-MetBio, our next goal was to find the combination of linkers that provided the optimal FRET ratio change. Accordingly, we constructed a second library (lib2) that was designed to contain 392 members that each had one of seven different lengths of L1 (ranging from 0 to 14 residues), eight different lengths of L2 (ranging from 0 to 20 residues), and seven different lengths for L3 (ranging from 0 to 14 residues). All linkers were composed primarily of glycine, alanine and serine, likely rendering them unstructured and highly flexible. By design, lib2 was relatively 'coarse' as

our primary goal was to reveal the general trends in preferred linker length, such that we could eventually design a more refined second linker library.

To identify the best linker length combinations from lib2, six pairs of 'glucose' and 'arabinose' plates, each with 45 re-spotted colonies, were screened as shown in Figure 22. The average ratio change for all colonies was determined to be 14% with a standard deviation of 4% (Figure 24).

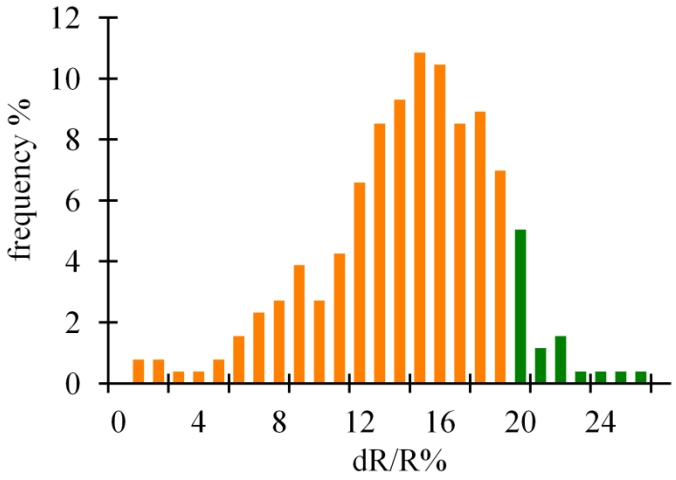

Figure 24 Data histogram of FRET ratio changes of all colonies screened from Lib2. Average value was about 14% with a standard deviation of about 4%.

The three colonies which showed the highest FRET ratio changes; one which showed the lowest; and other 9 clones which showed average values were picked, cultured, and the H3K27-MetBio genes were sequenced. Sequencing results revealed that colonies which showed the highest ratio changes had L1 and L3 linker lengths of 0, and L2 linker lengths of 10, 14, or 20 (Figure 25). The variant with L2 = 20 residues was designated as H3K27-MetBio2 (Figure 28A and Figure 29).

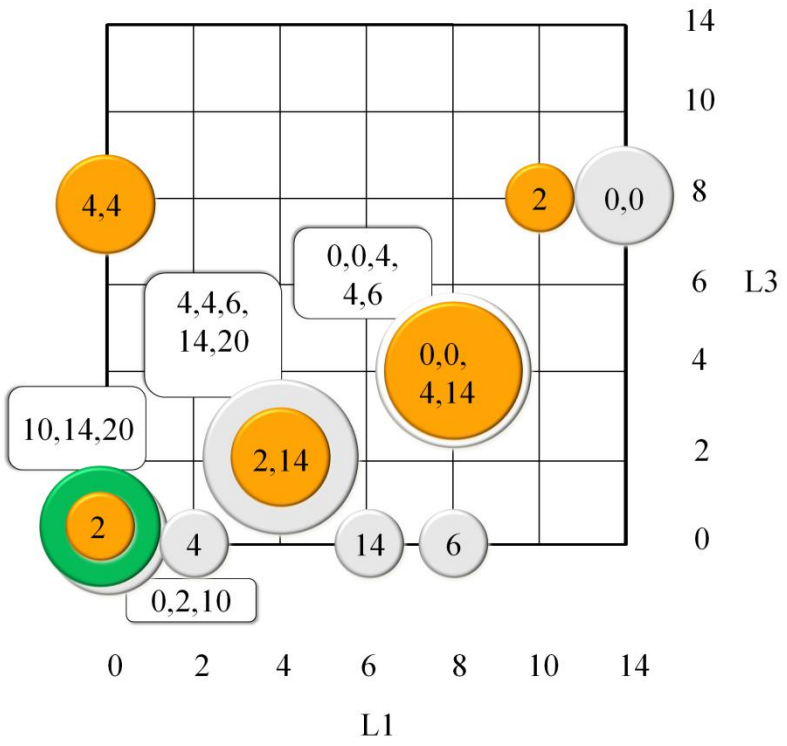

Figure 25 Linker combinations identified during screening of lib2. The horizontal and the vertical axes represent the number of amino acids constituting the first linker (L1) and the third linker (L3), respectively. The inputs for the second linker (L2) are sequences of 0, 2, 4, 6, 8, 10, 14, 16, and 20 residues in length. Identified L1 and L3 combinations are indicated on the grid by coloured circles the size of which represents the number of sequenced clones with a given L1-L3 combination, while the number of amino acids constituting L2 of these clones is indicated by the numerical values assigned to the circles. Linker combinations exhibiting the highest FRET ratio changes (>20%) are represented by the green circle; orange circles represent clones with average-to-low ratio changes (<20%); and grey circles represent clones picked at random.

The results of lib2 screening suggested to us that the combination of a short L1, a long L2, and a short L3 linker would provide H3K27-MetBio variants with the highest FRET changes. This observed linker length pattern can be rationalized on the basis of the solution NMR structure of the Cbx7 chromodomain in complex

with a trimethylated H3K27-containing peptide (135). The structure reveals that the N- terminus of Cbx7 (to which mTFP1 is fused in H3K27-MetBio) is in close proximity to the C- terminus of the H3K27 peptide (to which mCitrine is fused in H3K27-MetBio). Accordingly, we expect that shorter L1 and L3 linkers, located at the N- terminus of Cbx7 and the C- terminus of the substrate peptide, respectively, could serve to hold the mTFP1-mCitrine FRET pair in closer proximity in the compacted (*i.e.*, trimethylated) state of the biosensor and provide higher FRET efficiency than longer linkers would. In the extended (*i.e.*, unmethylated) state of the biosensor, a longer L2 linker could serve to increase the overall distance between the FPs and provide a lower FRET efficiency than shorter linkers would. So overall, the observed linker combination is a reasonable solution to providing the maximal change in FRET efficiency upon methylation of H3K27-MetBio.

Guided by the results of the lib2 library screening, we attempted to further refine the H3K27-MetBio linkers by construction and screening of a second linker library designated lib3. The lib3 library consisted of 640 variants arising from the possible combination of ten L1 linkers (ranging from -4 to +5 residues), eight L2 linkers (ranging from 6 to 30 residues), and eight L3 linkers (ranging from -3 to +4 residues). Negative linkers indicate truncations relative to either Gly225 of mTFP1 in the case of L1 or relative to Glu6 of mCitrine in the case of L3.

For screening of lib3, twelve pairs of 'glucose' and 'arabinose' plates were prepared, and the FRET ratio changes of the colonies were calculated. The

average ratio change was found to be 38% with a standard deviation of 16% (Figure 26).

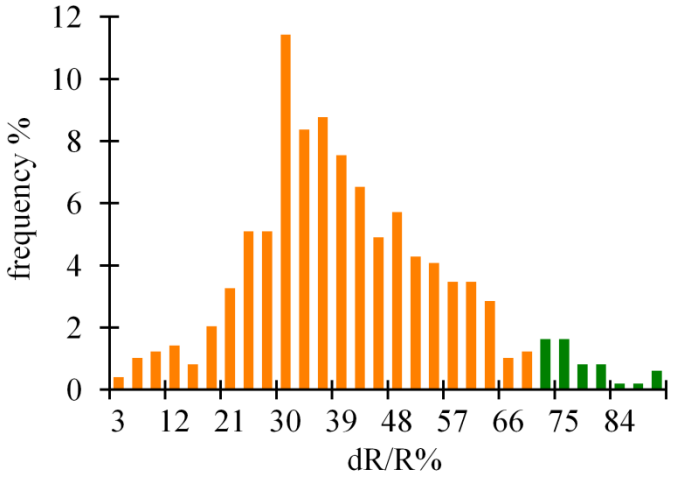

Figure 26 Data histogram of FRET ratio changes of colonies screened from lib3.

Nine clones exhibiting high ratio changes (>70%), and twenty-two with average-to-low ratio changes (<70%) were picked for clonal expansion and further characterization. Sequencing results revealed that the L1 linker length in the high ratio change variants ranged from -4 to +3, indicating that there was no strong preference for any particular linker length within this range (Figure 27). In contrast, 7 of 9 high ratio change variants had an L3 of -2, indicating that there is a much stronger dependence on this particular linker for achieving high ratio change. Similarly, 8 of the 9 clones with the highest ratio changes had L2 of 14 residues or longer. The variant with L1 = -2, L2 = 20, and L3 = -2 residues was designated as H3K27-MetBio3 (Figure 28B,and Figure 29)

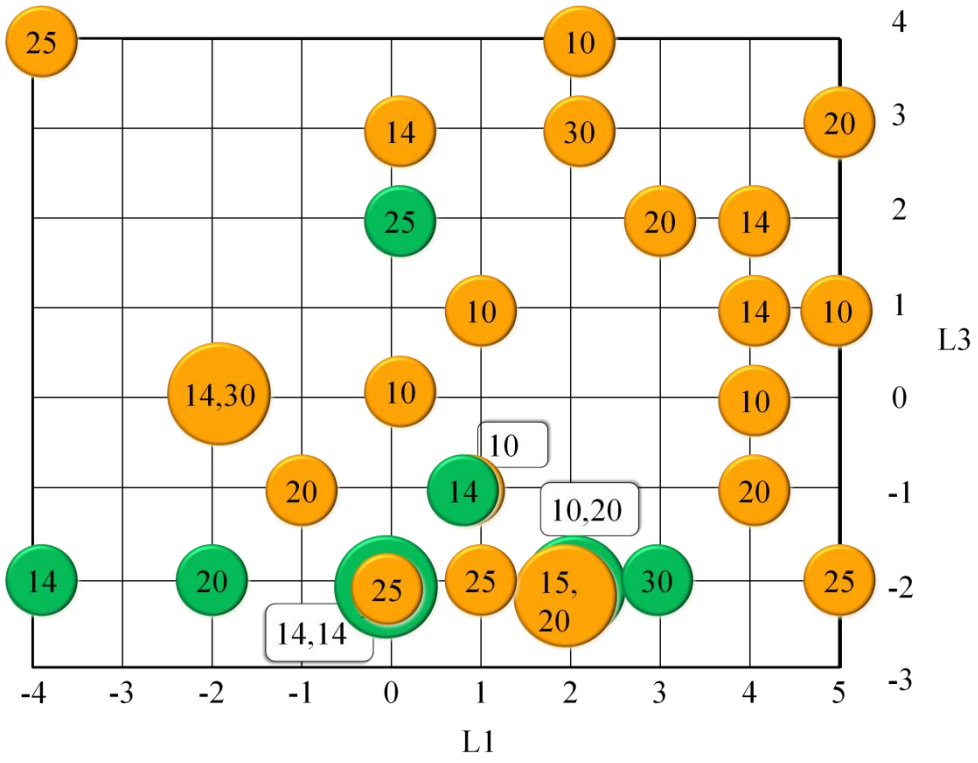

Figure 27 Linker combinations identified during screening of lib3. Input for L2 are linkers with 6, 8, 10, 14, 15, 20, 25, or 30 residues in length. Linker combinations exhibiting high FRET ratio changes (>70%) are represented with green circles; those showing average-to-low ratio changes (<70%) are represented by orange circles. These thresholds were based on the calculated average of ratio changes of compiled data from this experiment which was about 38% with a standard deviation of 16%.

The best two variants identified from lib2 and lib3, H3K27-MetBio2 and H3K27-MetBio3, respectively, were purified in their unmodified states and a portion of the protein was subjected to an *in vitro* methylation assay. Fluorescence emission scans were acquired using an excitation wavelength of 420 nm (Figure 28) and ΔR/R% was calculated for each biosensor. H3K27-MetBio2 demonstrated only a modest improvement in signal change (37%) relative to the H3K27-MetBio1

(29%). However, H3K27-MetBio3 exhibited a substantially improved ΔR/R% of 66%, which is 2.3× higher than H3K27-MetBio1.

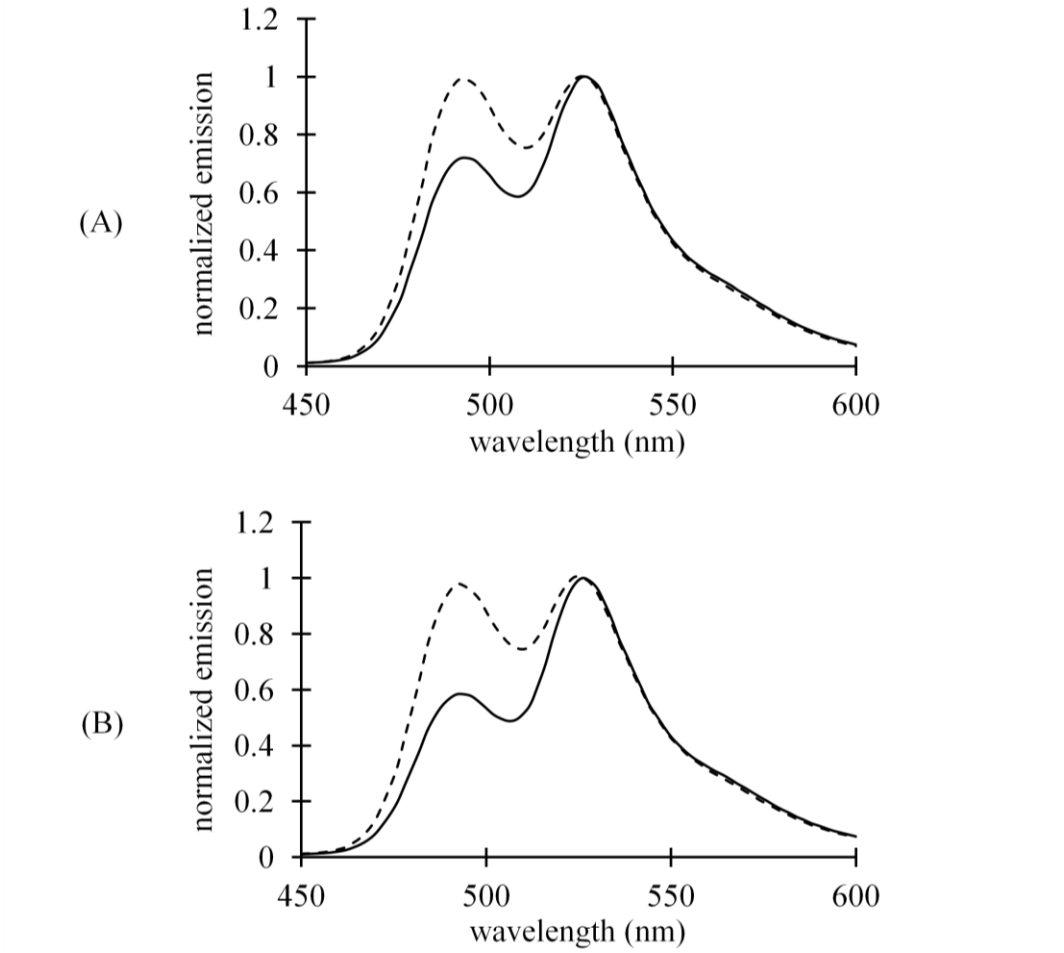

Figure 28 Normalized emission scans of H3K27-MetBio2 (A) and H3K27-MetBio3 (B) after prolonged treatment with vSET in the presence (solid line) or absence (dashed line) of SAM.

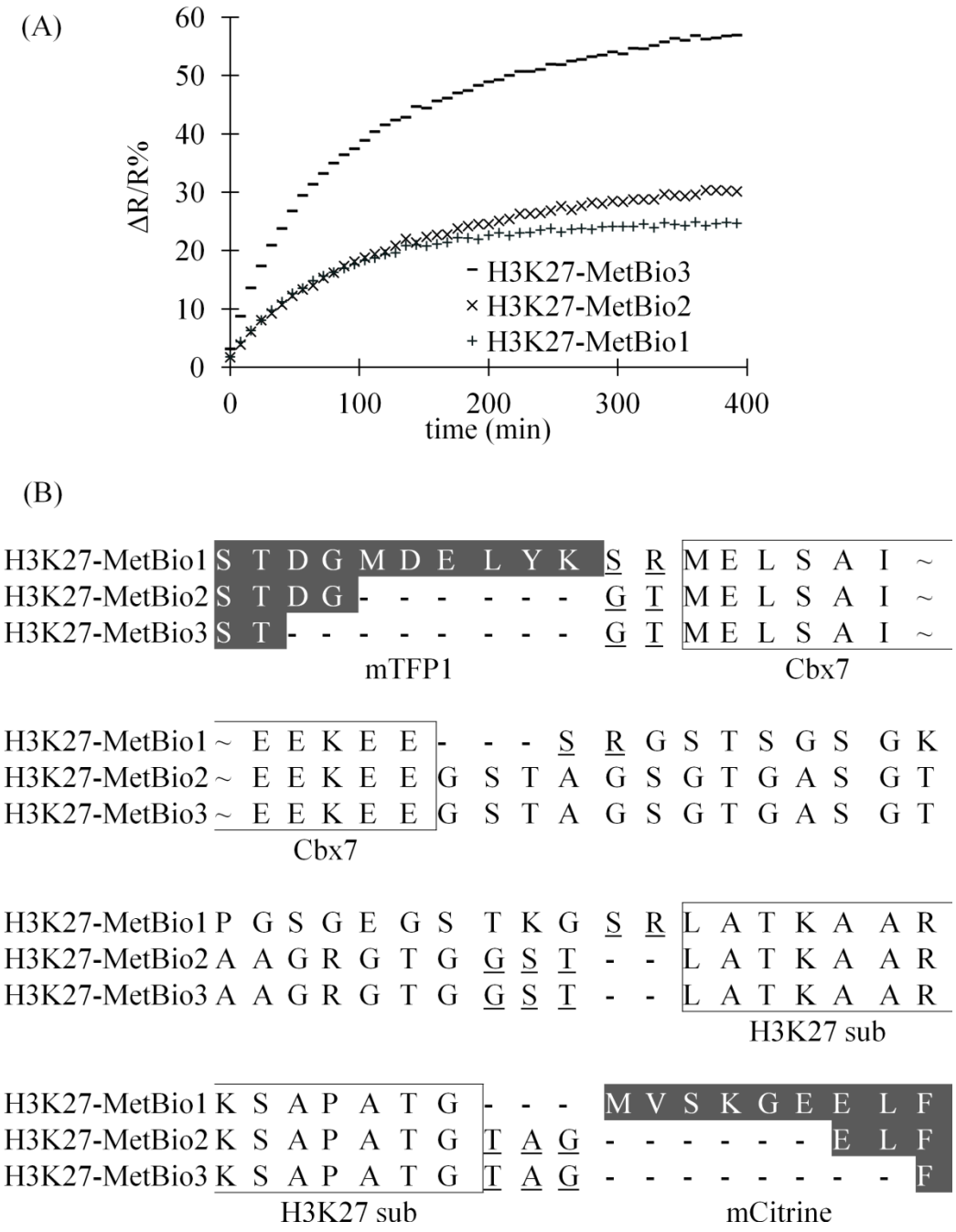

Figure 29 Comparison between the best identified variants from lib2 (H3K27-MetBio2) and lib3 (H3K27-MetBio3) and the original H3K27-MetBio1. (A) FRET ratio changes of H3K27-MetBio1-3 calculated using fluorescence intensities acquired at 494 and 526 nm at time intervals of 4 min for a period of 6 hr 40 min. (B) Sequence alignment of the three constructs. Underlined residues denote the amino acids encoded by restriction sites.

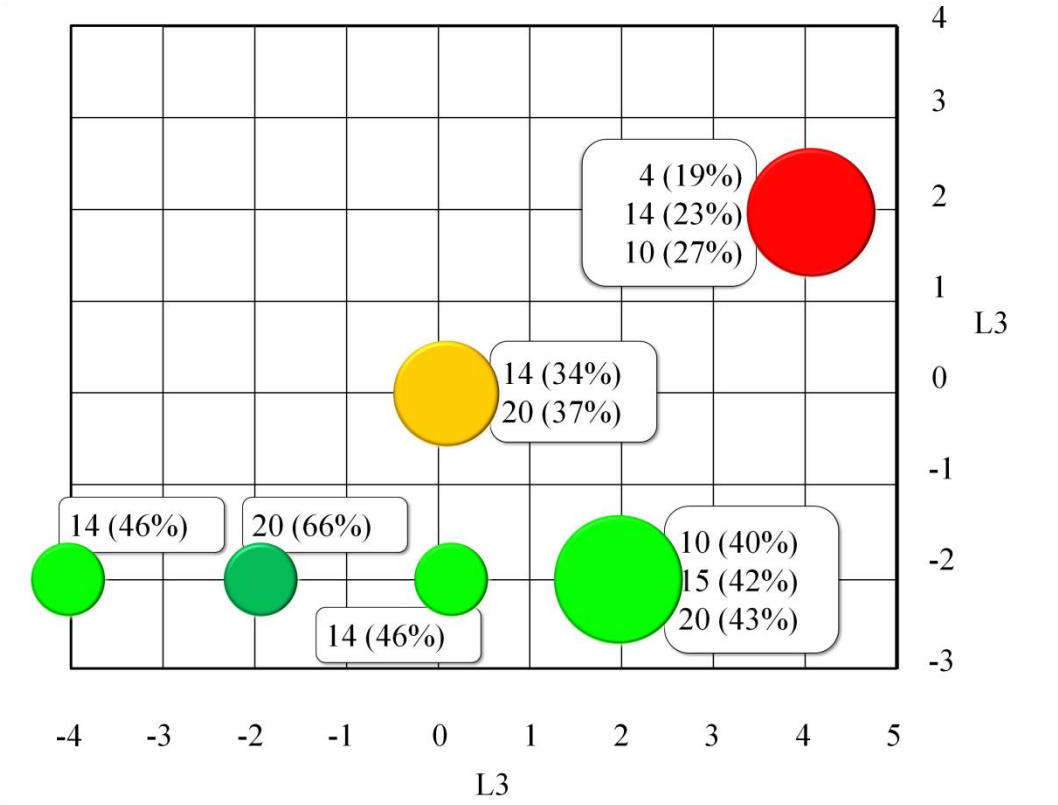

Figure 30 Linker combinations of additional clones isolated from lib2 and lib3, that were recombinantly produced and subsequently subjected to prolonged treatment with vSET. The dark green circle represents H3K27-MetBio3 which was the best variant we identified. Light green circles represent clones with intermediate ratio changes, yellow circle represent clones with modest changes, while the red circle represent clones with poor ratio changes. Values of ΔR/R% are indicated for each clone in brackets next to its L2 length.

The linkers used were mostly composed of glycine and serine, potentially imparting flexibility to the linkers. Long L2 were prevalent in the clones that showed high FRET ratio changes, but after exceeding a certain minimum, further increments to this linker region seemed to have a very little effect on the final ratio change. This becomes apparent when we compare the ratio changes of the three variants with the linker compositions (2, 10, -2), (2, 15, -2) and (2, 20, -2).

These variants had a final ratio change of 40%, 42%, and 43% respectively. Interestingly, however, L2 = 14 seems to be a disfavoured linker; and the two variants with the linker combinations (4, 10, 2) and (4, 14, 2) changed in ratio by 27% and 23% respectively (Figure 30).

It should be mentioned that during the construction of our libraries, we only investigated the number of amino acids constituting the linkers while we paid very little attention to the potential impact on the final ratio change which could be achieved by the identity of amino acids used. As we have mentioned above, we used amino acids that imparted flexibility to the linker. Investigating both the number and identity of amino acids constituting the linker regions could potentially achieve better improvements in terms of ratio changes of screened biosensors (105, 167). Changing the orientation of the FP by employing the circularly permuted variants could also have a very favourable outcome on the FRET ratio change (104, 115).

With the successful identification of linker combinations that provide improved ratio changes, we have demonstrated that this dual expression library screening strategy, in its current form, is of practical utility for optimizing FRET-based biosensors designed to respond to PTM. We expect that this system could be used to optimize a wide variety of biosensors for PTM, provided that the a constitutively active enzyme with the activity of interest can be expressed in its functional form in *E. coli*. However, the current implementation of the screening system does have some drawbacks that will hopefully be addressed in future versions. Specifically, our reliance on manual spotting for the replication of

bacterial colonies severely limited the throughput of the assay and increased the likelihood of human error. An improved method for library screening would involve the on-plate induction of enzyme expression. With this approach, images of the same plate, before and after enzyme induction, could be acquired and processed to identify the colonies with the largest changes in emission ratio.

With the improved H3K27-MetBio3 in hand, we sought to develop an on-plate induction methodology to render pUADE more useful for screening bigger libraries. The first step of developing the new on-plate induction protocol was to find the proper glucose concentration that: a) is sufficient to inhibit vSET expression; and b) is not too high to impede induction of transcription upon spraying with L-arabinose. Towards that goal, we transformed DH10B cells with pBAD/His B encoding mCitrine which were used to seed plates with different concentrations of glucose (20 mM, 15 mM, or 10 mM) in addition to 1 mM IPTG and 0.04% ampicillin. After overnight incubation, each plate was used to acquire an image using the cognate excitation and emission filter wheels for mCitrine as described in Chapter 2. Next, these plates were sprayed with a 1 M solution of L-arabinose, and incubated at room temperature for four 4 to 6 hours before they were used to acquire new images. Initial intensities of colonies grown on plates with 20 mM or 15 mM glucose were essentially indistinguishable from background intensities whereas those grown on 10 mM glucose showed slightly higher intensities than the background and therefore plates with the latter concentration were omitted from the experiment. As opposed to colonies grown on plates that contained 20 mM glucose, those grown on 15 mM glucose plates

corresponded to L-arabinose induction as reveal by increased fluorescence intensities, therefore we decided to use the latter concentration (data not shown).

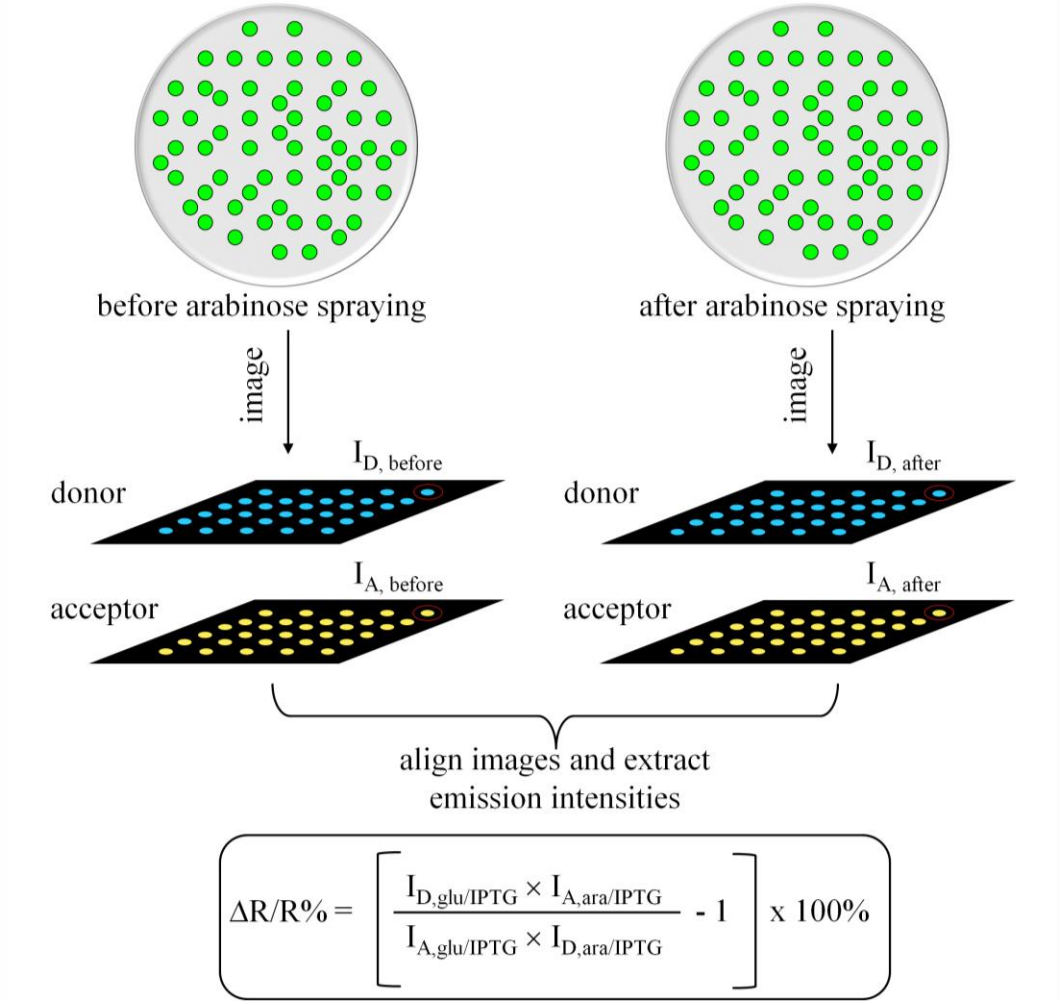

Figure 31 Schematic representation of screening procedures using protocol-2. Briefly, colonies grown on P_{BAD} repressive media are imaged prior to L-arabinose spraying. After spraying, plates are incubated at RT for few hours and subsequently used to acquire a new set of images, which are then aligned to the formerly acquired set and used to extract emission intensities to calculated ΔR/R% for each colony.

After deciding on the glucose concentration, we tested the consequences of on-plate induction of vSET expression on the FRET emission ratio change for each

colony as shall be briefly discussed in the following lines. The two plasmids pUADE[P_{BAD}(*vset*)/P_{tac}(*h3k27-metbio3*)] and pUADE[P_{BAD}(*nohkmt*)/P_{tac}(*h3k27-metbio3*)] were used to transform DH10B cells which were subsequently used to seed plates that that contained 15 mM glucose, 1 mM IPTG, and 0.04% ampicillin. Two sets of images were acquired for each plate before and after L-arabinose spraying as described above. ImagePro was used to align the two sets of images after which fluorescence intensities of colonies were extracted and transferred to Microsoft Excel for data processing. FRET change was calculated as ΔR/R% and data histogram showing the percentage of colonies with a given ΔR/R% value was constructed (Figure 31 dashed line).

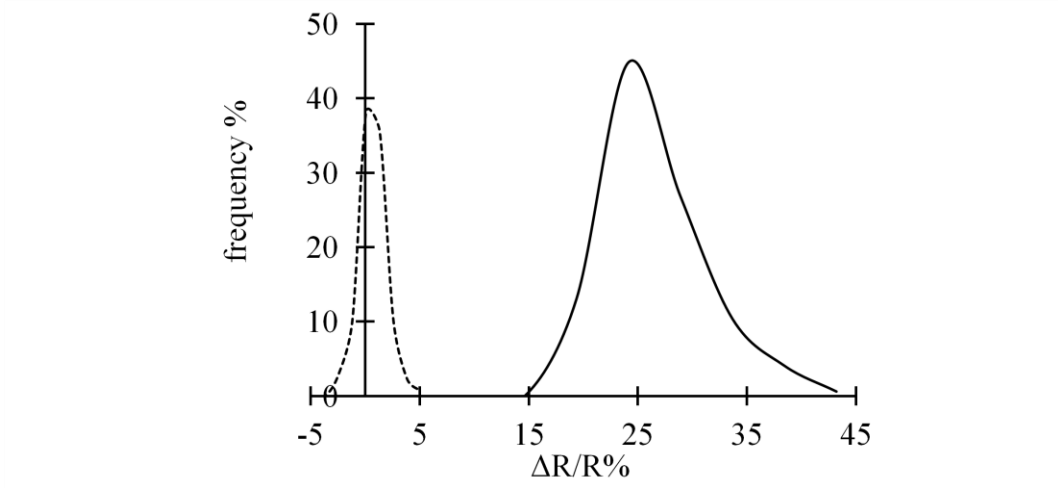

Figure 32 Data histogram of screening results using protocol-2. The diagram shows the frequency of ΔR/R% of two bacterial populations each transformed with the indicate plasmid and subjected to protocol-2. Population transformed with pUADE[P_{BAD}(*vset*)/P_{tac}(*h3k27-metbio3*)] show an average ΔR/R% of 26% (solid line) while colonies transformed with pUADE[P_{BAD}(*nohkmt*)/P_{tac}(*h3k27-metbio3*)] did not demonstrate any significant FRET ratio change (dashed line).

As expected, colonies transformed with pUADE[P_{BAD}(*nohkmt*)/P_{tac}(*h3k27-metbio3*)] did not show any significant change in ΔR/R% and the population ΔR/R% was centered around 0% change. On the other hand, colonies transformed with pUADE[P_{BAD}(*vset*)/P_{tac}(*h3k27-metbio3*)] centered around 26% change (Figure 32 solid line).

However, unfortunately, on-plate induction methodology did not provide satisfactory results due to the heterogeneity of the application. We expect that this limitation could be overcome through the use of a promoter that could be induced at the same level across the whole plate. One potential solution is to use a cold-inducible promoter (168), which could presumably be induced at the same level in all colonies by simply changing the incubation conditions of the plate. However, it is unclear whether an alternate promoter could provide an equivalent level of repression to P_{BAD} in the uninduced state.

With improvements in this library screening strategy, it should be possible to substantially increase the screening throughput. In the current implementation, we were limited to screening hundreds of variants. Accordingly, we only investigated the effect of linker length on ratio change and did not attempt to adjust linker composition. We strongly suspect that further improvements in ratio change could be achieved by exploring both altered linker lengths and compositions (105, 167). We also suspect that biosensors with further improved ratio changes could be identified by altering the orientation of the donor and acceptor FPs by employing circularly permuted variants (104) or using ‘sticky’ FP variants (49).

3.4 Conclusion

We have developed a method for high-throughput screening of biosensor libraries with many hundreds of different linker combinations. This approach should be applicable to the optimisation of any genetically encoded biosensor for PTM, provided the gene encoding the enzyme activity of interested can be functionally expressed in *E. coli*. We have demonstrated this technology by undertaking the optimisation of a biosensor for detection of methylation of H3K27. We expect that expression of a chromatin-localised version of this optimised biosensor may facilitate efforts to spatially resolve H3K27-trimethylation patterns in living cells. Furthermore, when compared to existing *in vitro* assays for histone methyltransferase activity (169-171), this homogenous biosensor-based assay is notable for requiring the least number of reagents and liquid handling steps. Accordingly, H3K27-MetBio3 may be particularly useful in automated high-throughput screening efforts aimed at the identification of H3K27 methyltransferase inhibitors that could serve as chemical probes (172) or as leads for the development of new cancer therapies (173).

Chapter 4:

Conclusions and future work

In this dissertation, we have demonstrated a new methodology for rapid and convenient improvement of the response dynamic range of FRET-based biosensors intended for the detection of PTM enzymatic activities. Our methodology is based on *in colony* screening of libraries of variants of a given FRET-based biosensor under conditions of controlled expression of the biosensor protein and the PTM enzyme of interest achieved by placing the coding sequences of the biosensor construct and that of the PTM enzyme downstream of the two differentially induced promoters P_{tac} and P_{BAD}, respectively. As a model system, we chose to optimize the response dynamic range of a methylation biosensor designed to detect histone H3 Lys27 methyltransferase enzymatic activity. For the enzyme, we used *Paramecium bursaria chlorella virus* 1 vSET enzyme (174) because of its known catalytic activity of trimethylating Lys27 of histone H3 protein and the ability to be recombinantly produced it in its active form in *E. coli*.

Optimized plate composition was determined empirically by growing colonies of *E. coli* ElectroMax DH10B strain transformed with pUADE on LB-agar plates with different concentrations of L-arabinose and/or glucose, and IPTG followed by assessing the expression level of the two recombinant proteins by SDS-PAGE, Western blot, or fluorescence intensity.

As a proof-of-concept, we screened a small library (lib1) the members of which were constructed using different MREs composed of one of seven different binding domains and a histone H3 Lys27 substrate motif. Sequencing results of the constructs that exhibited the highest FRET ΔR/R% suggested that H3K27-MetBio1 based on Cbx7 as the binding domain was the best candidate biosensor

for the detection of histone H3 Lys27 trimethylation. These results were in accord with the *in vitro* methylation assay in which H3K27-MetBio1 exhibited a FRET ratio change of 29% after prolonged treatment with the enzyme. This biosensor was then used as a template for the construction of two subsequent libraries (lib2 and lib3) in which the linker regions were targeted for length modifications. Screening of these libraries led to the identification of H3K27-MetBio3 that exhibited an emission ratio change (66%) that was 2.3x improved relative to that of the initially constructed H3K27-MetBio1 (29%). However, the experimental set-up used up this point (protocol-1) limits the screening throughput of the assay to libraries of few hundred variants. In addition, this protocol suffered from a number of disadvantages that were described in Chapter 3. In an attempt to improve the throughput of our screening methodology, we formulated an alternative protocol (protocol-2) that involved FRET ΔR/R% analysis of fluorescence emission intensities acquired from the same colony before and after L-arabinose induction – as opposed to intensities acquired from identical clones grown on two plates with different compositions in protocol-1. This alternative protocol allows for screening hundreds of variants per plate, and thus greatly enhances the throughput of the screening methodology. Unfortunately, protocol-2 suffered from heterogeneity in L-arabinose application. It would be desirable to optimize an analogous protocol that allows for temporal control over vSET expression or its on-set of reaction.

Once an optimized protocol for screening has been established, we reckon that this system can be efficiently used to optimize the dynamic range of other FRET-

based biosensors designed to detect various PTM enzymatic activities, provided that these enzymes can be produced in their catalytically active forms in *E. coli*. In the following sections, we will provide experimental suggestions that may improve the throughput of our screening methodology. Also a number of other potential applications of this system will be presented.

4.1 Reoptimization of the screening conditions

The primary requirement for this screening methodology is to achieve a very tight control over the PTM enzyme expression or, alternatively, the onset of its catalytic activity. Although glucose provided satisfactory inhibition of vSET expression, it negatively impacted the expression level of the biosensor protein in colonies grown under repressive conditions resulting in disparate levels of biosensor accumulation between colonies grown on glucose plates and those grown on glucose/ L-arabinose plates. The same problem was also observed in the alternative protocol where colonies accumulated more of the biosensor protein during the incubation period following L-arabinose application (as evinced by the intensity of the acceptor mCitrine upon direct excitation, data not shown). The disparity of the level of the biosensor protein accumulation between colonies grown on the two different plate compositions in protocol-1 or before and after L-arabinose application in protocol-2 was reduced by O.N. incubation of the plates at 4 °C prior to data acquisition, without any compromise to the inhibition of vSET expression under repressive conditions. However, both protocols suffered from a number of disadvantages as was described in Chapter 3.

Ideally, the requirement of the tight control over the enzyme expression or activity can be satisfied without any impact on the level of the biosensor accumulation under repressive and inductive conditions. Moreover, a faster catalytic activity of the enzyme would greatly reduce the waiting period required to ensure maximal modification of the biosensor protein, allowing for data acquisition in a narrower time window before colonies accumulate more of the biosensor protein, and therefore circumvent any complications that may arise from comparing the FRET ratio changes from two different biosensor concentrations. In this section, we will provide some suggestions that may lead to satisfying these two conditions: Temporal control over the PTM reaction, and faster catalytic activity.

4.1.1 Temporal control over the PTM enzymatic activity

In colony PTM enzymatic activity can be temporally controlled by either: Growing the bacteria under conditions that maintain tight repression of the enzyme expression followed by alleviating these conditions; or allowing the bacteria to accumulate a 'modifiable', and otherwise inactive, form of the enzyme (or its substrate) that could be rendered active (or accessible to the enzyme, in case of the substrate) upon applying a chemical reagent or changing an environmental condition.

4.1.1.1 Controlling the onset of vSET expression

Tight transcription control can be achieved in *E. coli* by a number of promoter/activity modulator protein systems (175). One of the systems that allows for full transcription repression is the bacteriophage lambda promoter

whose activity is modulated by the thermolabile repressor protein cI857 (176, 177). Briefly, the repressor protein cI857 binds the operator regions upstream of the lambda promoter in the temperature range of 28-32 °C inhibiting mRNA transcription of the gene downstream of the promoter. Once the temperature of the culture is raised above 37 °C, the thermolabile cI857 is released from the operator sites allowing for the initiation of transcription and subsequent production of the recombinant protein (176, 177).

To ensure maximal control over the PTM enzyme expression, a two-tier transcription control can be used. For example, the previous lambdaP/cI857 can be used in conjunction with the transcription control mechanism that involves *attB* and *attP* site-specific DNA recombination by the bacteriophage lambda integrase (178, 179), which was described by Podhajska *et al.* in 1985 (180). By placing the lambda integrase transcription under the control of lambdaP/cI857 system, its production can be achieved by subjecting the bacterial culture to an elevated temperature. Upon its production, lambda integrase then catalyzes the inversion of the DNA sequence flanked by the *attB* and *attP* sites. By cloning either the promoter of choice (178) or the coding sequence of the protein (179) in an inverted position relative to the other, restoration to transcription-inductive configuration can only be obtained after the production of the integrase enzyme.

4.1.1.2 Controlling the onset of the enzymatic reaction

Another way to gain temporal control over the PTM reaction *in colony* can be achieved by controlling the enzyme catalytic activity or the accessibility of the

substrate to the enzyme. This could be performed by applying the technology of non-natural amino acid – specifically, amino acids with photocaged side chains – incorporation into the recombinant protein sequence in response to an amber codon substituting that of the targeted residue (181).

For example, Deiters *et al.* have demonstrated the utility of this method in controlling the β-galactosidase enzyme activity in *E. coli* (182). In their work, the incorporation of *o*-nitrobenzyl photocaged tyrosine into the enzyme sequence in lieu of the Tyr503 greatly diminished the activity of the enzyme which was restored upon illuminating the culture with 364 nm light (182).

Mutational analysis of vSET revealed that the three tyrosine residues Tyr16, Tyr106, and Tyr109 are important for the HKMT activity of vSET (183). Substituting one of these residues for the photocaged *o*-nitrobenzyl tyrosine is likely to inhibit the vSET catalytic activity. Alternatively, Lys27 residue in the substrate motif can be targeted for such substitution by a photocaged lysine rendering the substrate temporarily inaccessible for PTM activity of vSET. Practical examples of using photocaged lysine residues was demonstrated by Gautier *et al.* in 2010 (184) and Gautier *et al.* in 2011 (185).

By using this technology, the colonies are allowed to accumulate both of the recombinant proteins however PTM reaction does not proceed until the culture is subjected to a light source of suitable wavelength that allows for the 'decaging' of the side chain of the protected amino acid. This method does not require any specific promoter. Moreover, complications arising from manual application of

chemical reagents or subjecting the culture to an environmental stressor such as elevated temperatures can be completely avoided.

4.1.2 Obtaining a vSET variant with faster kinetics

After optimizing the experimental conditions that allows for controlling the onset of the PTM reaction *in colony*, it may be advantageous to obtain a vSET variant with faster kinetics. This could be achieved by constructing a library of vSET mutants followed by *in colony* screening for variants with enhanced catalytic speed. A vSET mutant with faster catalytic activity can reduce the waiting period required to achieve maximal substrate modification and as a consequence decrease the disparity of the biosensor concentrations at the times of image acquisition (before and after the enzymatic modification).

BiFC-based H3K27 biosensor may be more convenient than FRET in this screening experiment. A suggested experimental setup is as follows: a split HKMT biosensor construct can be encoded by a DNA sequence that can be transcribed to a polycistronic mRNA translatable to the two fragments of the biosensor, where the substrate motif is fused to an FP fragment the complementary part of which is fused to the Cbx7 binding domain. After initiating PTM reaction *in colony*, either by vSET transcription induction or light-induced uncaging of the enzyme or substrate, images of fluorescence intensity of colonies expressing the vSET variants acquired at a given time interval can be used to compare the rates of fluorescence build-up induced by the enzyme mutants with that induced by the wt vSET expressed by control colonies.

4.2 Screening the domain-shuffled libraries

Once optimized screening conditions are established, the system may be efficiently used to screen very large FRET-based biosensor libraries for improved response dynamic range. The previous chapter describes the construction of sixteen different libraries composed of thousands of variants arising from different linker combinations between the domains of the biosensor, in addition to the unique arrangement of the biosensor parts in each of these libraries. Moreover, these libraries were constructed with each domain flanked by two unique restriction sites, a setup that allows for convenient excision and replacement of the inter-domain linkers, shall the amino acid composition of these linkers be targeted for optimization. It might also be worthwhile to examine the effect circularly permuted variants of FPs might have on the response dynamic range when used in these libraries.

Bibliography

1. Demchenko AP (2010) The concept of λ-ratiometry in fluorescence sensing and imaging. *J Fluoresc* 20: 1099-1128.

2. Kricka LJ & Fortina P (2009) Analytical ancestry: "Firsts" in fluorescent labeling of nucleosides, nucleotides, and nucleic acids. *Clin Chem* 55: 670-683.

3. Prober JM, *et al* (1987) A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 238: 336-341.

4. Piatek AS, *et al* (1998) Molecular beacon sequence analysis for detecting drug resistance in *Mycobacterium tuberculosis*. *Nat Biotechnol* 16: 359-363.

5. Liang Y, *et al* (2011) Aptamer beacons for visualization of endogenous protein HIV-1 reverse transcriptase in living cells. *Biosens Bioelectron* 28: 270-276.

6. Yamamoto R & Kumar PKR (2000) Molecular beacon aptamer fluoresces in the presence of Tat protein of HIV-1. *Genes Cells* 5: 389-396.

7. Lakowicz JR (2006) in *Principles of fluorescence spectroscopy,* (Springer, Singapore), pp 63-95.

8. Chudakov DM, Matz MV, Lukyanov S & Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol Rev* 90: 1103-1163.

9. Newman RH, Fosbrink MD & Zhang J (2011) Genetically encodable fluorescent biosensors for tracking signaling dynamics in living cells. *Chem Rev* 111: 3614-3666.

10. Shimomura O, Johnson FH & Saiga Y (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. *J Cell Comp Physiol* 59: 223-239.

11. Johnson FH, *et al* (1962) Quantum efficiency of *Cypridina* luminescence, with a note on that of *Aequorea*. *J Cell Comp Physiol* 60: 85-103.

12. Wampler JE, Hori K, Lee JW & Cormier MJ (1971) Structured bioluminescence. two emitters during both the *in vitro* and the *in vivo* bioluminescence of the sea pansy,*Renilla*. *Biochemistry* 10: 2903-2909.

13. Tsien RY (1998) The Green Fluorescent Protein. *Annu Rev Biochem* 67: 509-544.

14. Ward WW, Cody CW, Hart RC & Cormier MJ (1980) Spectrophotometric identity of the energy transfer chromophores in *Renilla* and *Aequorea* green-fluorescent proteins. *Photochem Photobiol* 31: 611-615.

15. Morise H, Shimomura O, Johnson FH & Winant J (1974) Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* 13: 2656-2662.

16. Shimomura O (1979) Structure of the chromophore of *Aequorea* green fluorescent protein. *FEBS Lett* 104: 220-222.

17. Cody CW, Prasher DC, Westler WM, Prendergast FG & Ward WW (1993) Chemical structure of the hexapeptide chromophore of the *Aequorea* green fluorescent protein. *Biochemistry* 32: 1212-1218.

18. Chalfie M (1995) Green fluorescent protein. *Photochem Photobiol* 62: 651-656.

19. Prasher DC, Eckenrode VK, Ward WW, Prendergast FG & Cormier MJ (1992) Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* 111: 229-233.

20. Delagrave S, Hawtin RE, Silva CM, Yang MM & Youvan DC (1995) Red-shifted excitation mutants of the green fluorescent protein. *Nat Biotechnol* 13: 151-154.

21. Heim R, Cubitt AB & Tsien RY (1995) Improved green fluorescence . *Nature* 373: 663-664.

22. Ehrig T, O'Kane DJ & Prendergast FG (1995) Green-fluorescent protein mutants with altered fluorescence excitation spectra. *FEBS Lett* 367: 163-166.

23. Heim R, Prasher DC & Tsien RY (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci U S A* 91: 12501-12504.

24. Cubitt AB, *et al* (1995) Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci* 20: 448-455.

25. Rosenow MA, Huffman HA, Phail ME & Wachter RM (2004) The crystal structure of the Y66L variant of green fluorescent protein supports a cyclization-oxidation-dehydration mechanism for chromophore maturation. *Biochemistry* 43: 4464-4472.

26. Zhang L, Patel HN, Lappe JW & Wachter RM (2006) Reaction progress of chromophore biogenesis in green fluorescent protein. *J Am Chem Soc* 128: 4766-4772.

27. Pouwels LJ, Zhang L, Chan NH, Dorrestein PC & Wachter RM (2008) Kinetic isotope effect studies on the *de novo* rate of chromophore formation in fast- and slow-maturing GFP variants. *Biochemistry* 47: 10111-10122.

28. Strack RL, Strongin DE, Mets L, Glick BS & Keenan RJ (2010) Chromophore formation in DsRed occurs by a branched pathway. *J Am Chem Soc* 132: 8496-8505.

29. Wachter RM, Watkins JL & Kim H (2010) Mechanistic diversity of red fluorescence acquisition by GFP-like proteins. *Biochemistry* 49: 7417-7427.

30. Matz MV, *et al* (1999) Fluorescent proteins from nonbioluminescent *Anthozoa* species. *Nat Biotechnol* 17: 969-973.

31. Baird GS, Zacharias DA & Tsien RY (2000) Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc Natl Acad Sci U S A* 97: 11984-11989.

32. Gross LA, Baird GS, Hoffman RC, Baldridge KK & Tsien RY (2000) The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proc Natl Acad Sci U S A* 97: 11990-11995.

33. Berg J, Hung YP & Yellen G (2009) A genetically encoded fluorescent reporter of ATP:ADP ratio. *Nat Methods* 6: 161-166.

34. Pakhomov AA & Martynov VI (2008) GFP family: Structural insights into spectral tuning. *Chem Biol* 15: 755-764.

35. Mena MA, Treynor TP, Mayo SL & Daugherty PS (2006) Blue fluorescent proteins with enhanced brightness and photostability from a structurally targeted library. *Nat Biotechnol* 24: 1569-1571.

36. Ormö M, *et al* (1996) Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* 273: 1392-1395.

37. Yang F, Moss LG & Phillips Jr. GN (1996) The molecular structure of green fluorescent protein. *Nat Biotechnol* 14: 1246-1251.

38. Chalfie M, Tu Y, Euskirchen G, Ward WW & Prasher DC (1994) Green fluorescent protein as a marker for gene expression. *Science* 263: 802-805.

39. Morris MC (2010) Fluorescent biosensors of intracellular targets from genetically encoded reporters to modular polypeptide probes. *Cell Biochem Biophys* 56: 19-37.

40. Frommer WB, Davidson MW & Campbell RE (2009) Genetically encoded biosensors based on engineered fluorescent proteins. *Chem Soc Rev* 38: 2833-2841.

41. Förster T (1946) Energiewanderung und fluoreszenz. *Naturwissenschaften* 33: 166-175.

42. Didenko VV (2001) DNA probes using fluorescence resonance energy transfer (FRET): Designs and applications. *BioTechniques* 31: 1106-1121.

43. Kaláb P & Soderholm J (2010) The design of Förster (fluorescence) resonance energy transfer (FRET)-based molecular sensors for ran GTPase. *Methods* 51: 220-232.

44. Sahoo H, Roccatano D, Zacharias M & Nau WM (2006) Distance distributions of short polypeptides recovered by fluorescence resonance energy transfer in the 10 Å domain. *J Am Chem Soc* 128: 8118-8119.

45. Bánóczi Z, Alexa A, Farkas A, Friedrich P & Hudecz F (2008) Novel cell-penetrating calpain substrate. *Bioconjug Chem* 19: 1375-1381.

46. Charbonnière LJ, Hildebrandt N, Ziessel RF & Löhmannsröben H (2006) Lanthanides to quantum dots resonance energy transfer in time-resolved fluoro-immunoassays and luminescence microscopy. *J Am Chem Soc* 128: 12800-12809.

47. Gonzalez JE & Tsien RY (1995) Voltage sensing by fluorescence resonance energy transfer in single cells. *Biophys J* 69: 1272-1280.

48. González JE & Tsien RY (1997) Improved indicators of cell membrane potential that use fluorescence resonance energy transfer. *Chem Biol* 4: 269-277.

49. Campbell RE (2009) Fluorescent-protein-based biosensors: Modulation of energy transfer as a design principle. *Anal Chem* 81: 5972-5979.

50. Deuschle K, *et al* (2005) Construction and optimization of a family of genetically encoded metabolite sensors by semirational protein engineering. *Protein Sci* 14: 2304-2314.

51. Biondi RM, Baehler PJ, Reymond CD & Véron M (1998) Random insertion of GFP into the cAMP-dependent protein kinase regulatory subunit from *Dictyostelium discoideum*. *Nucleic Acids Res* 26: 4946-4952.

52. Siegel MS & Isacoff EY (1997) A genetically encoded optical probe of membrane voltage. *Neuron* 19: 735-741.

53. Baird GS, Zacharias DA & Tsien RY (1999) Circular permutation and receptor insertion within green fluorescent proteins. *Proc Natl Acad Sci U S A* 96: 11241-11246.

54. Topell S, Hennecke J & Glockshuber R (1999) Circularly permuted variants of the green fluorescent protein. *FEBS Lett* 457: 283-289.

55. Li X, *et al* (1997) Deletions of the *Aequorea victoria* green fluorescent protein define the minimal domain required for fluorescence. *J Biol Chem* 272: 28545-28549.

56. Deuschle K, *et al* (2006) Rapid metabolism of glucose detected with FRET glucose nanosensors in epidermal cells and intact roots of *Arabidopsis* RNA-silencing mutants. *Plant Cell* 18: 2314-2325.

57. Lager I, Looger LL, Hilpert M, Lalonde S & Frommer WB (2006) Conversion of a putative *Agrobacterium* sugar-binding protein into a FRET sensor with high selectivity for sucrose. *J Biol Chem* 281: 30875-30883.

58. Takanaga H, Chaudhuri B & Frommer WB (2008) GLUT1 and GLUT9 as major contributors to glucose influx in HepG2 cells identified by a high sensitivity intramolecular FRET glucose sensor. *Biochim Biophys Acta* 1778: 1091-1099.

59. Tsai M, *et al* (2009) Real-time monitoring of human enterovirus (HEV)-infected cells and anti-HEV 3C protease potency by fluorescence resonance energy transfer. *Antimicrob Agents Chemother* 53: 748-755.

60. Sabariegos R, *et al* (2009) Fluorescence resonance energy transfer-based assay for characterization of hepatitis C virus NS3-4A protease activity in live cells. *Antimicrob Agents Chemother* 53: 728-734.

61. Aye-Han N, Ni Q & Zhang J (2009) Fluorescent biosensors for real-time tracking of post-translational modification dynamics. *Curr Opin Chem Biol* 13: 392-397.

62. Harvey CD, *et al* (2008) A genetically encoded fluorescent sensor of ERK activity. *Proc Natl Acad Sci U S A* 105: 19264-19269.

63. Lin C & Ting AY (2004) A genetically encoded fluorescent reporter of histone phosphorylation in living cells. *Angew Chem Int Edit* 43: 2940-2943.

64. Lin C, Jao CY & Ting AY (2004) Genetically encoded fluorescent reporters of histone methylation in living cells. *J Am Chem Soc* 126: 5982-5983.

65. Sasaki K, Ito T, Nishino N, Khochbin S & Yoshida M (2009) Real-time imaging of histone H4 hyperacetylation in living cells. *Proc Natl Acad Sci U S A* 106: 16257-16262.

66. Carrillo LD, Krishnamoorthy L & Mahal LK (2006) A cellular FRET-based sensor for β-O-GIcNAc, a dynamic carbohydrate modification involved in signaling. *J Am Chem Soc* 128: 14768-14769.

67. Bermejo C, Ewald JC, Lanquar V, Jones AM & Frommer WB (2011) *In vivo* biochemistry: Quantifying ion and metabolite levels in individual cells or cultures of yeast. *Biochem J* 438: 1-10.

68. Okumoto S (2010) Imaging approach for monitoring cellular metabolites and ions using genetically encoded biosensors. *Curr Opin Biotechnol* 21: 45-54.

69. Deuschle K, *et al* (2005) Genetically encoded sensors for metabolites. *Cytometry A* 64: 3-9.

70. Dwyer MA & Hellinga HW (2004) Periplasmic binding proteins: A versatile superfamily for protein engineering. *Curr Opin Struct Biol* 14: 495-504.

71. Lager I, Fehr M, Frommer WB & Lalonde S (2003) Development of a fluorescent nanosensor for ribose. *FEBS Lett* 553: 85-89.

72. Hires SA, Zhu Y & Tsien RY (2008) Optical measurement of synaptic glutamate spillover and reuptake by linker optimized glutamate-sensitive fluorescent reporters. *Proc Natl Acad Sci U S A* 105: 4411-4416.

73. Kaper T, *et al* (2007) Nanosensor detection of an immunoregulatory tryptophan influx/kynurenine efflux cycle. *PLoS Biol* 5: 2201-2210.

74. Gu H, *et al* (2006) A novel analytical method for *in vivo* phosphate tracking. *FEBS Lett* 580: 5885-5893.

75. McCombs JE & Palmer AE (2008) Measuring calcium dynamics in living cells with genetically encodable calcium indicators. *Methods* 46: 152-159.

76. Mank M & Griesbeck O (2008) Genetically encoded calcium indicators. *Chem Rev* 108: 1550-1564.

77. Łukasiewicz S, Faron-Górecka A, Dobrucki J, Polit A & Dziedzicka-Wasylewska M (2009) Studies on the role of the receptor protein motifs possibly

involved in electrostatic interactions on the dopamine D_1 and D_2 receptor oligomerization. *FEBS J* 276: 760-775.

78. Rives M, *et al* (2009) Crosstalk between $GABA_B$ and mGlu1a receptors reveals new insight into GPCR signal integration. *EMBO J* 28: 2195-2208.

79. Marcaggi P, Mutoh H, Dimitrov D, Beato M & Knöpfel T (2009) Optical measurement of mGluR1 conformational changes reveals fast activation, slow deactivation, and sensitization. *Proc Natl Acad Sci U S A* 106: 11388-11393.

80. Chun W & Johnson GVW (2007) Activation of glycogen synthase kinase 3β promotes the intermolecular association of tau. the use of fluorescence resonance energy transfer microscopy. *J Biol Chem* 282: 23410-23417.

81. Levitt JA, Matthews DR, Ameer-Beg SM & Suhling K (2009) Fluorescence lifetime and polarization-resolved imaging in cell biology. *Curr Opin Biotechnol* 20: 28-36.

82. Orthaus S, *et al* (2008) Assembly of the inner kinetochore proteins CENP-A and CENP-B in living human cells. *Chembiochem* 9: 77-92.

83. Kerppola TK (2008) Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu Rev Biophys* 37: 465-487.

84. Kerppola TK (2008) Bimolecular fluorescence complementation: Visualization of molecular interactions in living cells. *Method Cell Biol* 85: 431-470.

85. Hemerka JN, *et al* (2009) Detection and characterization of influenza A virus PA-PB2 interaction through a bimolecular fluorescence complementation assay. *J Virol* 83: 3944-3955.

86. Vincenz C & Kerppola TK (2008) Different polycomb group CBX family proteins associate with distinct regions of chromatin using nonhomologous protein sequences. *Proc Natl Acad Sci U S A* 105: 16572-16577.

87. Kodama Y & Wada M (2009) Simultaneous visualization of two protein complexes in a single plant cell using multicolor fluorescence complementation analysis. *Plant Mol Biol* 70: 211-217.

88. Vidi P, Chemel BR, Hu C & Watts VJ (2008) Ligand-dependent oligomerization of dopamine D_2 and adenosine A_{2A} receptors in living neuronal cells. *Mol Pharmacol* 74: 544-551.

89. Souslova EA, *et al* (2007) Single fluorescent protein-based Ca^{2+} sensors with increased dynamic range. *BMC Biotechnol* 7:37.

90. Mizuno T, Murao K, Tanabe Y, Oda M & Tanaka T (2007) Metal-ion-dependent GFP emission *in vivo* by combining a circularly permutated green fluorescent protein with an engineered metal-ion-binding coiled-coil. *J Am Chem Soc* 129: 11378-11383.

91. Nausch LWM, Ledoux J, Bonev AD, Nelson MT & Dostmann WR (2008) Differential patterning of cGMP in vascular smooth muscle cells revealed by single GFP-linked biosensors. *Proc Natl Acad Sci U S A* 105: 365-370.

92. Patterson GH, Knobel SM, Sharif WD, Kain SR & Piston DW (1997) Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys J* 73: 2782-2790.

93. Llopis J, McCaffery JM, Miyawaki A, Farquhar MG & Tsien RY (1998) Measurement of cytosolic, mitochondrial, and golgi pH in single living cells with green fluorescent proteins. *Proc Natl Acad Sci U S A* 95: 6803-6808.

94. Johnson DE, *et al* (2009) Red fluorescent protein pH biosensor to detect concentrative nucleoside transport. *J Biol Chem* 284: 20499-20511.

95. Nakabayashi T, Wang H, Kinjo M & Ohta N (2008) Application of fluorescence lifetime imaging of enhanced green fluorescent protein to intracellular pH measurements. *Photochem Photobiol Sci* 7: 668-670.

96. Miesenböck G, De Angelis DA & Rothman JE (1998) Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* 394: 192-195.

97. Serresi M, Bizzarri R, Cardarelli F & Beltram F (2009) Real-time measurement of endosomal acidification by a novel genetically encoded biosensor. *Anal Bioanal Chem* 393: 1123-1133.

98. Bizzarri R, Serresi M, Luin S & Beltram F (2009) Green fluorescent protein based pH indicators for *in vivo* use: A review. *Anal Bioanal Chem* 393: 1107-1122.

99. Hanson GT, *et al* (2004) Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. *J Biol Chem* 279: 13044-13053.

100. Gutscher M, *et al* (2008) Real-time imaging of the intracellular glutathione redox potential. *Nat Methods* 5: 553-559.

101. Piston DW & Kremers G (2007) Fluorescent protein FRET: The good, the bad and the ugly. *Trends Biochem Sci* 32: 407-414.

102. Kotera I, Iwasaki T, Imamura H, Noji H & Nagai T (2010) Reversible dimerization of *Aequorea victoria* fluorescent proteins increases the dynamic range of FRET-based indicators. *ACS Chem Biol* 5: 215-222.

103. Nguyen AW & Daugherty PS (2005) Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat Biotechnol* 23: 355-360.

104. Nagai T, Yamada S, Tominaga T, Ichikawa M & Miyawaki A (2004) Expanded dynamic range of fluorescent indicators for Ca^{2+} by circularly permuted yellow fluorescent proteins. *Proc Natl Acad Sci U S A* 101: 10554-10559.

105. Lissandron V, *et al* (2005) Improvement of a FRET-based indicator for cAMP by linker design and stabilization of donor-acceptor interaction. *J Mol Biol* 354: 546-555.

106. Shimozono S & Miyawaki A (2008) Engineering FRET Constructs using CFP and YFP. *Method Cell Biol* 85: 381-393.

107. Piljić A, De Diego I, Wilmanns M & Schultz C (2011) Rapid development of genetically encoded FRET reporters. *ACS Chem Biol* 6: 685-691.

108. Seo J & Lee K (2004) Post-translational modifications and their biological functions: Proteomic analysis and systematic approaches. *J Biochem Mol Biol* 37: 35-44.

109. Pearson G, *et al* (2001) Mitogen-activated protein (MAP) kinase pathways: Regulation and physiological functions. *Endocr Rev* 22: 153-183.

110. Zarubin T & Han J (2005) Activation and signaling of the p38 MAP kinase pathway. *Cell Res* 15: 11-18.

111. Kerppola TK (2009) Visualization of molecular interactions using bimolecular fluorescence complementation analysis: Characteristics of protein fragment complementation. *Chem Soc Rev* 38: 2876-2886.

112. Kanno T, *et al* (2004) Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. *Mol Cell* 13: 33-43.

113. Zhang J, Ma Y, Taylor SS & Tsien RY (2001) Genetically encoded reporters of protein kinase A activity reveal impact of substrate tethering. *Proc Natl Acad Sci U S A* 98: 14997-15002.

114. Zhang J, Hupfeld CJ, Taylor SS, Olefsky JM & Tsien RY (2005) Insulin disrupts β-adrenergic signalling to protein kinase A in adipocytes. *Nature* 437: 569-573.

115. Allen MD & Zhang J (2006) Subcellular dynamics of protein kinase A activity visualized by FRET-based reporters. *Biochem Biophys Res Commun* 348: 716-721.

116. Ai H, Henderson JN, Remington SJ & Campbell RE (2006) Directed evolution of a monomeric, bright and photostable version of *Clavularia* cyan fluorescent protein: Structural characterization and applications in fluorescence imaging. *Biochem J* 400: 531-540.

117. Schägger H (2006) Tricine-SDS-PAGE. *Nat Protoc* 1: 16-22.

118. Ibraheem A, Yap H, Ding Y & Campbell RE (2011) A bacteria colony-based screen for optimal linker combinations in genetically encoded biosensors. *BMC Biotechnol* 11: 105.

119. Whitlow M, *et al* (1993) An improved linker for single-chain Fv with reduced aggregation and enhanced proteolytic stability. *Protein Eng* 6: 989-995.

120. Li Y, Sierra AM, Ai H & Campbell RE (2008) Identification of sites within a monomeric red fluorescent protein that tolerate peptide insertion and testing of corresponding circular permutations. *Photochem Photobiol* 84: 111-119.

121. Cheng Z & Campbell RE (2006) Assessing the structural stability of designed β-hairpin peptides in the cytoplasm of live cells. *Chembiochem* 7: 1147-1150.

122. Kornberg RD (1977) Structure of chromatin. *Annu Rev Biochem* 46: 931-954.

123. Luger K, Mäder AW, Richmond RK, Sargent DF & Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389: 251-260.

124. Kim JK, Samaranayake M & Pradhan S (2009) Epigenetic mechanisms in mammals. *Cell Mol Life Sci* 66: 596-612.

125. Kouzarides T (2007) Chromatin modifications and their function. *Cell* 128: 693-705.

126. Martin C & Zhang Y (2005) The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* 6: 838-849.

127. Peter CJ & Akbarian S (2011) Balancing histone methylation activities in psychiatric disorders. *Trends Mol Med* 17: 372-379.

128. Albert M & Helin K (2010) Histone methyltransferases in cancer. *Semin Cell Dev Biol* 21: 209-220.

129. Nielsen PR, *et al* (2002) Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* 416: 103-107.

130. Jacobs SA & Khorasanizadeh S (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* 295: 2080-2083.

131. Min J, Zhang Y & Xu R (2003) Structural basis for specific binding of polycomb chromodomain to histone H3 methylated at Lys27. *Genes Dev* 17: 1823-1828.

132. Fischle W, *et al* (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by polycomb and HP1 chromodomains. *Genes Dev* 17: 1870-1881.

133. Kim J, *et al* (2006) Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep* 7: 397-403.

134. Bernstein E, *et al* (2006) Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol Cell Biol* 26: 2560-2569.

135. Kaustov L, *et al* (2011) Recognition and specificity determinants of the human Cbx chromodomains. *J Biol Chem* 286: 521-529.

136. Yap KL, *et al* (2010) Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* 38: 662-674.

137. Griesbeck O, Baird GS, Campbell RE, Zacharias DA & Tsien RY (2001) Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J Biol Chem* 276: 29188-29194.

138. Mujtaba S, *et al* (2008) Epigenetic transcriptional repression of cellular genes by a viral SET protein. *Nat Cell Biol* 10: 1114-1122.

139. Qian C, *et al* (2006) Structural insights of the specificity and catalysis of a viral histone H3 lysine 27 methyltransferase. *J Mol Biol* 359: 86-96.

140. Wei H & Zhou M (2010) Dimerization of a viral SET protein endows its function. *Proc Natl Acad Sci U S A* 107: 18433-18438.

141. Fischle W, Franz H, Jacobs SA, Allis CD & Khorasanizadeh S (2008) Specificity of the chromodomain Y chromosome family of chromodomains for lysine-methylated ARK(S/T) motifs. *J Biol Chem* 283: 19626-19635.

142. Franz H, Mosch K, Soeroes S, Urlaub H & Fischle W (2009) Multimerization and H3K9me3 binding are required for CDYL1b heterochromatin association. *J Biol Chem* 284: 35049-35059.

143. Kimura RH, Steenblock ER & Camarero JA (2007) Development of a cell-based fluorescence resonance energy transfer reporter for *Bacillus anthracis* lethal factor protease. *Anal Biochem* 369: 60-70.

144. De Boer HA, Comstock LJ & Vasser M (1983) The tac promoter: A functional hybrid derived from the trp and lac promoters. *Proc Natl Acad Sci U S A* 80: 21-25.

145. Guzman LM, Belin D, Carson MJ & Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J Bacteriol* 177: 4121-4130.

146. Miyada CG, Stoltzfus L & Wilcox G (1984) Regulation of the araC of *Escherichia coli*: Catabolite repression, autoregulation, and effect on *araBAD* expression. *Proc Natl Acad Sci U S A* 81: 4120-4124.

147. Dong H, Nilsson L & Kurland CG (1995) Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *J Bacteriol* 177: 1497-1504.

148. Zacharias DA, Violin JD, Newton AC & Tsien RY (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* 296: 913-916.

149. Wang Y, *et al* (2005) Visualizing the mechanical activation of Src. *Nature* 434: 1040-1045.

150. Brumbaugh J, Schleifenbaum A, Gasch A, Sattler M & Schultz C (2006) A dual parameter FRET probe for measuring PKC and PKA activity in living cells. *J Am Chem Soc* 128: 24-25.

151. Schleifenbaum A, Stier G, Gasch A, Sattler M & Schultz G (2004) Genetically encoded FRET probe for PKC activity based on pleckstrin. *J Am Chem Soc* 126: 11786-11787.

152. Jost CA, Reither G, Hoffmann C & Schultz C (2008) Contribution of fluorophores to protein kinase C FRET probe performance. *Chembiochem* 9: 1379-1384.

153. Miyawaki A & Tsien RY (2000) Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein. *Method Enzymol* 327: 472-500.

154. Van der Krogt GNM, Ogink J, Ponsioen B & Jalink K (2008) A comparison of donor-acceptor pairs for genetically encoded FRET sensors: Application to the epac cAMP sensor as an example. *PLoS One* 3:

155. Fehr M, Lalonde S, Lager I, Wolff MW & Frommer WB (2003) *In vivo* imaging of the dynamics of glucose uptake in the cytosol of COS-7 cells by fluorescent nanosensors. *J Biol Chem* 278: 19127-19133.

156. Okumoto S, *et al* (2005) Detection of glutamate release from neurons by genetically encoded surface-displayed FRET nanosensors. *Proc Natl Acad Sci U S A* 102: 8740-8745.

157. Okada S, Ota K & Ito T (2009) Circular permutation of ligand-binding module improves dynamic range of genetically encoded FRET-based nanosensor. *Protein Sci* 18: 2518-2527.

158. Mank M, *et al* (2006) A FRET-based calcium biosensor with fast signal kinetics and high fluorescence change. *Biophys J* 90: 1790-1796.

159. Vinkenborg JL, Evers TH, Reulen SWA, Meijer EW & Merkx M (2007) Enhanced sensitivity of FRET-based protease sensors by redesign of the GFP dimerization interface. *Chembiochem* 8: 1119-1121.

160. Ouyang M, Sun J, Chien S & Wang Y (2008) Determination of hierarchical relationship of Src and Rac at subcellular locations with FRET biosensors. *Proc Natl Acad Sci U S A* 105: 14353-14358.

161. Evers TH, Appelhof MAM, de Graaf-Heuvelmans PTHM, Meijer EW & Merkx M (2007) Ratiometric detection of Zn(II) using chelating fluorescent protein chimeras. *J Mol Biol* 374: 411-425.

162. Van Dongen EMWM, *et al* (2007) Variation of linker length in ratiometric fluorescent sensor proteins allows rational tuning of Zn(II) affinity in the picomolar to femtomolar range. *J Am Chem Soc* 129: 3494-3495.

163. Palmer AE & Tsien RY (2006) Measuring calcium signaling using genetically targetable fluorescent indicators. *Nat Protoc* 1: 1057-1065.

164. Pham E, Chiang J, Li I, Shum W & Truong K (2007) A computational tool for designing FRET protein biosensors by rigid-body sampling of their conformational space. *Structure* 15: 515-523.

165. Chiang JJ & Truong K (2006) Computational modeling of a new fluorescent biosensor for caspase proteolytic activity improves dynamic range. *IEEE Trans Nanobioscience* 5: 41-45.

166. Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166: 557-580.

167. Ha J, *et al* (2007) Design and application of highly responsive fluorescence resonance energy transfer biosensors for detection of sugar in living *Saccharomyces cerevisiae* cells. *Appl Environ Microbiol* 73: 7408-7414.

168. Baneyx F & Mujacic M (2003) Cold-inducible promoters for heterologous protein expression. *Methods Mol Biol* 205: 1-18.

169. Collazo E, Couture J, Bulfer S & Trievel RC (2005) A coupled fluorescent assay for histone methyltransferases. *Anal Biochem* 342: 86-92.

170. Gowher H, Zhang X, Cheng X & Jeltsch A (2005) Avidin plate assay system for enzymatic characterization of a histone lysine methyltransferase. *Anal Biochem* 342: 287-291.

171. Rathert P, Cheng X & Jeltsch A (2007) Continuous enzymatic assay for histone lysine methyltransferases. *BioTechniques* 43: 602-608.

172. Cole PA (2008) Chemical probes for histone-modifying enzymes. *Nat Chem Biol* 4: 590-597.

173. Simon JA & Lange CA (2008) Roles of the EZH2 histone methyltransferase in cancer epigenetics. *Mutat Res* 647: 21-29.

174. Yamada T, Onimatsu H & Van Etten JL (2006) *Chlorella* Viruses. *Adv Virus Res* 65: 293-336.

175. Makrides SC (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev* 60: 512-538.

176. Remaut E, Stanssens P & Fiers W (1981) Plasmid vectors for high-efficiency expression controlled by the P_L promoter of coliphage lambda. *Gene* 15: 81-93.

177. Valdez-Cruz NA, Caspeta L, Pérez NO, Ramírez OT & Trujillo-Roldán MA (2010) Production of recombinant proteins in *E. coli* by the heat inducible expression system based on the phage lambda pL and/or pR promoters. *Microb Cell Fact* 9:

178. Wülfmg C & Plückthun A (1993) A versatile and highly repressible *Escherichia coli* expression system based on invertible promoters: Expression of a gene encoding a toxic product. *Gene* 136: 199-203.

179. Sektas M, Hasan N & Szybalski W (2001) Expression plasmid with a very tight two-step control: *Int/att*-mediated gene inversion with respect to the stationary promoter. *Gene* 267: 213-220.

180. Podhajska AJ, Hasan N & Szybalski W (1985) Control of cloned gene expression by promoter inversion *in vivo*: Construction of the heat-pulse-activated *att-nutL-p-att*-N module. *Gene* 40: 163-168.

181. Liu CC & Schultz PG (2010) Adding new chemistries to the genetic code. *Annu Rev Biochem* 79: 413-444.

182. Deiters A, Groff D, Ryu Y, Xie J & Schultz PG (2006) A genetically encoded photocaged tyrosine. *Angew Chem Int Edit* 45: 2728-2731.

183. Manzur KL, *et al* (2003) A dimeric viral SET domain methyltransferase specific to Lys27 of histone H3. *Nat Struct Biol* 10: 187-196.

184. Gautier A, *et al* (2010) Genetically encoded photocontrol of protein localization in mammalian cells. *J Am Chem Soc* 132: 4086-4088.

185. Gautier A, Deiters A & Chin JW (2011) Light-activated kinases enable temporal dissection of signaling networks in living cells. *J Am Chem Soc* 133: 2124-2127.

186. Ibraheem A & Campbell RE (2010) Designs and applications of fluorescent protein-based biosensors. *Curr Opin Chem Biol* 14: 30-36.

Appendix

Synthetic oligonucleotide sequences

Primer	Sequence (5' to 3')
Trugon1 Forward1 (Seq)	CTGCAGACTGATCATATGGTGAGCAAGGGCGAGGAG
Trugon1 Backward1 (Compl)	TTTTTGTTCCAGATCTAGCTTGTACAGCTCGTCCATGCCGTC
Trugon2 Forward2 (Seq)	AGAAATTACCCTAGGCTGCAGACTGATCATATGGTGAG
Trugon2 Backward2 (Compl)	ATCCTCTTCTGAGATGAGTTTTTGTTCCAGATCTAGCTTGTAC AG
Trugon3 Forward3 (Seq)	AATATCTCGAGCTCTAGAAATTACCCTAGGCTGCAGA
Trugon3 Backward3 (Compl)	TACATGGAATTCTTACAGATCCTCTTCTGAGATGAGTTTTTGT TC
TFP-XhoI-FD-SEQ	CTATATGACCTCGAGCATGGTGAGCAAGGGCGAGGAGCTGTT
TFP-XbaI-BK-CPL	ACGCCAAAACAGCCATCTAGACTTGTACAGCTCGTCCATGCC GAGAGTGAT
1-XbaI-H3K9-YFP-SEQ	GCTAGAAAGTCTACAGGAGGAAAGGCTATGGTGAGCAAGGG CGAGGAGCTGTT
2-XbaI-H3K9-YFP-SEQ	CGTGACCATTATGCCTCTAGAATGGCCCGTACTAAGCAGACT GCTAGAAAGTCTACAGGAGGAAAG
1-XbaI-H3K27-SEQ	AGAAAGTCAGCACCTGCAACGGGAATGGTGAGCAAGGGCGA GGAGCTGTT
2-XbaI-H3K27-SEQ	CATAATCATTATGTATCTAGACTGGCCACCAAGGCGGCCAGA AAGTCAGCACCTGCAACG
YF-BK-CPL	ACGCCAAAACAGCCAAGCTTCTTACTTGTACAGCTCGTCCATG CCGAGAGTGAT
2-YF-BK-CPL	CCCGCTCTCCACGGAACGCCAAAACAGCCAAGCTT
pGEX-For-TAC (Seq)	CTATAAAGGATCCTTATCGACTGCACGGTGCACCAATGCTTC
pGEX-Back-TAC (Comp)	GATTTTATCCATGGTCATGAATACTGTTTCCTGTGTGAAATTG TTATCCGCTCAC

Forward-BAD-transfer	CTATAAAACATGTTACATCATCATCATCATCACGGTATGGCTAGCATGACT
Backward-BAD-transfer	CATTAAACATAGATCTCACCGACAAACAACAGATAAAACGAAAGGCCCAG
POLY-Prmr1(pet)-Frwrd	GTATTTCTAAAACACGTGATCCGGATATAGTTCCTCCTTTCAGCAAAAAACCCCTCAAGACCCGTTTAGAGG
POLY-Prmr1(pet)-Bckwrd	CGTCATTATACACGTGATAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACT
MCS-SEQUENCE	TCGACCGAATTCTCCCCATTGGGTGGATATGGTGTATTTGCGAGAAAATCTTTCGAGAAGGGAGAACTTGTTGAAGAATGTTTGTGTATAGTGCGCCATAATGATGATTGGGGGACCGCCCTTGAAGATTATTTGTTTTCGAGAAAGAATATGTCTGCAATGGCTCTTGGTTTTGGTGCAATTTTTAACCATAGATCTGC
MCS-COMPLEMENTARY	AGCTGCAGATCTATGGTTAAAAATTGCACCAAAACCAAGAGCCATTGCAGACATATTCTTTCTCGAAAACAAATAATCTTCAAGGGCGGTCCCCCAATCATCATTATGGCGCACTATACACAAACATTCTTCAACAAGTTCTCCCTTCTCGAAAGATTTTCTCGCAAATACACCATATCCACCCAATGGGGAGAATTCGG
vSET-EcoRI-Frwd (seq)	GATATATTGAATTCATGTTTAATGACAGAGTCATCGTGAAAAAATCCCCATTG
vSET-BglII-Bckd (Compl)	CATTTATTAGATCTTTAATTTTGTGTTAATCTAGGTCTCGACAACCAGTAGTCATC
TFP-Primer 1	ACGCTAAGCTCGAGCATGGTGAGCAAGGGCGAGGAGACCACAATGGGC
(-4)-BkWd-KpnI(RF1)-TFP	TATCATGGTACCGTTGCGGGCCACGGCGCTCTCGTA
(-3)-BkWd-KpnI(RF1)-TFP	TATCATGGTACCGGAGTTGCGGGCCACGGCGCTCTC
(-2)-BkWd-KpnI(RF1)-TFP	TATCATGGTACCGGTGGAGTTGCGGGCCACGGCG
(-1)-BkWd-KpnI(RF1)-TFP	TATCATGGTACCGTCGGTGGAGTTGCGGGCCACGGC
0-BkWd-KpnI(RF1)-TFP	GCTCATGGTACCGCCGTCGGTGGAGTTGCGGGCCAC
1-BkWd-KpnI(RF1)-TFP	GATATAGGTACCGCTGCCGTCGGTGGAGTTGCGGGCCAC
2-BkWd-KpnI(RF1)-TFP	GATGCAGGTACCTGATCCGCCGTCGGTGGAGTTGCGGGCCAC
3-BkWd-KpnI(RF1)-TFP	GATTAAGGTACCTGTTGATCCGCCGTCGGTGGAGTTGCGGGCCAC

4-BkWd-KpnI(RF1)-TFP	GTATGCGGTACCTGCTGTTGATCCGCCGTCGGTGGAGTTGCGGGCCAC
5-BkWd-KpnI(RF1)-TFP	GTATATGGTACCTGATGCTGTTGATCCGCCGTCGGTGGAGTTGCGGGCCAC
6-BkWd-KpnI(RF1)-TFP	GATCAGGGTACCTGATCCTGCTGTTGATCCGCCGTCGGTGGAGTTGCGGGCCAC
8-BkWd-KpnI(RF1)-TFP	GTAATCGGTACCTGTTCCTGATCCTGCTGTTGATCCGCCGTCGGTGGAGTTGCGGGCCAC
10-BkWd-KpnI(RF1)-TFP	CATATAGGTACCTGCTCCTGTTCCTGATCCTGCTGTTGATCCGCCGTCGGTGGAGTTGCGGGCCAC
14-BkWd-KpnI(RF1)-TFP	TTTAATGGTACCTGCTGTACCTGATGCTCCTGTTCCTGATCCTGCTGTTGATCCGCCGTCGGTGGAGTTGCGGGCCAC
(-3)-FrWd-EagI(RF3)-YFP	CTATAATCGGCCGGCACCGGGGTGGTGCCCATCCTGGTCGAG
(-2)-FrWd-EagI(RF3)-YFP	CTATAATCGGCCGGCTTCACCGGGGTGGTGCCCATCCTGGTC
(-1)-FrWd-EagI(RF3)-YFP	CTATAATCGGCCGGCCTGTTCACCGGGGTGGTGCCCATCCTGGTC
0-FrWd-EagI(RF3)-YFP	GTAGCATCGGCCGGCGAGCTGTTCACCGGGGTGGTGCCCATCCTG
1-FrWd-EagI(RF3)-YFP	CTAATATCGGCCGGCAGTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
2-FrWd-EagI(RF3)-YFP	CTGATCTCGGCCGGCGGTAGTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
3-FrWd-EagI(RF3)-YFP	CTAATCTCGGCCGGCTCAGGTAGTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
4-FrWd-EagI(RF3)-YFP	GATACATCGGCCGGCGCAACGGGTAGTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
6-FrWd-EagI(RF3)-YFP	CTATCATCGGCCGGCGGAAGCGCAACGGGTAGTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
8-FrWd-EagI(RF3)-YFP	CAATAATCGGCCGGCACGGCTGGAAGCGCAACGGGTAGTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
10-FrWd-EagI(RF3)-YFP	CATATATCGGCCGGCACGGGAACGGCTGGAAGCGCAACGGGTAGTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
14-FrWd-EagI(RF3)-YFP	TATTTATCGGCCGGCGCTGGCAGTGGAACGGGAACGGCTGGAAGCGCAACGGGTAGTGAGCTGTTCACCGGGGTGGTGCCCATCCTG

YFP-Primer 4	GTCTGCAAGCTTCTTACTTGTACAGCTCGTCCATGCCGAGAGT GATCCCGGCG
LigCbx7-H3K27-forward	GATAGCGGTACCATGGAGCTGTCA
LigCbx7-H3K27-Backward	GATGATAGCCGGCCGTTCCTGTAG
SalI(RF2)-H3K27-Sense	CATGTAGGGTCGACACTTGCTACAAAGGCAGCACGCAAGAGC GCACCTGCTACAGGAACGGCCGGCTATCATC
EagI(RF3)-H3K27-Antisense	GATGATAGCCGGCCGTTCCTGTAGCAGGTGCGCTCTTGCGTGC TGCCTTTGTAGCAAGTGTCGACCCTACATG
FrWd-KpnI(RF1)-Cbx7	GATAGCGGTACCATGGAGCTGTCAGCCATAGGCGAGCAGGTG
0-BkWd-SalI(RF2)-Cbx7	GGAACCTGTCGACCCCTCCTCCTTCTCCTCGTAGGCCATGACA AGGCG
2-BkWd-SalI(RF2)-Cbx7	AGTCGCTGTCGACCCTGATCCCTCCTCCTTCTCCTCGTAGGCC ATGACAAGGCG
4-BkWd-SalI(RF2)-Cbx7	GTGGCGTGTCGACCCTGCTGTTGATCCCTCCTCCTTCTCCTCGT AGGCCATGACAAGGCG
6-BkWd-SalI(RF2)-Cbx7	GCTTCGTGTCGACCCTGATCCTGCTGTTGATCCCTCCTCCTTCT CCTCGTAGGCCATGACAAGGCG
8-BkWd-SalI(RF2)-Cbx7	GCTCGATGTCGACCCTGTTCCTGATCCTGCTGTTGATCCCTCC TCCTTCTCCTCGTAGGCCATGACAAGGCG
10-BkWd-SalI(RF2)-Cbx7	GAGCATTGTCGACCCTGCTCCTGTTCCTGATCCTGCTGTTGAT CCCTCCTCCTTCTCCTCGTAGGCCATGACAAGGCG
14-BkWd-SalI(RF2)-Cbx7	GTGTACTGTCGACCCTGCTGTACCTGATGCTCCTGTTCCTGAT CCTGCTGTTGATCCCTCCTCCTTCTCCTCGTAGGCCATGACAA GGCG
15-BkWd-SalI(RF2)-Cbx7	AGTTACTGTCGACCCTGTTGCTGTACCTGATGCTCCTGTTCCT GATCCTGCTGTTGATCCCTCCTCCTTCTC
20-BkWd-SalI(RF2)-Cbx7	GTTTTATGTCGACCCTCCTGTTCCTCTTCCTGCTGCTGTACCTG ATGCTCCTGTTCCTGATCCTGCTGTTGATCCCTCCTCCTTCTCC TCGTAGGCCATGACAAGGCG
25-BkWd-SalI(RF2)-Cbx7	AGTTACTGTCGACCCAGTACCAGCAGATCCTCCTGTTCCTCTT CCTGCTGCTGTACCTGATGCTCCTGTTCCTGA
30-BkWd-SalI(RF2)-Cbx7	AGTTATTGTCGACCCTGCTGTTCCTGATGCAGTACCAGCAGAA GTTCCTGTTCCTCTTCCTGCTGCTGTACCTGATGCTCCTGTTCC TGA

ACDB-Sense	TCGATCCTCGAGATGGTGGGTACCCCCGGGATAAACAGGGCC GGCCTAACGGCCGGCGATACAAAGCTTCCTAGGATCAAGGGG TCGACATAAT
BDCA-Sense	TCGATCCCTAGGATCAAGGGGTCGACAACGGCCGGCGATACA AAGCTTCCCGGGATAAACAGGGCCGGCCTACTCGAGATGGTG GGTACCTAAT
ACBD-Sense	TCGATCCTCGAGATGGTGGGTACCCCCGGGATAAACAGGGCC GGCCTACCTAGGATCAAGGGGTCGACAACGGCCGGCGATACA AAGCTTTAAT
BDAC-Sense	TCGATCCCTAGGATCAAGGGGTCGACAACGGCCGGCGATACA AAGCTTCTCGAGATGGTGGGTACCCCCGGGATAAACAGGGCC GGCCTATAAT
ABCD–Sense	TCGATCCTCGAGATGGTGGGTACCCCTAGGATCAAGGGGTCG ACACCCGGGATAAACAGGGCCGGCCTAACGGCCGGCGATAC AAAGCTTTAAT
ABDC–Sense	TCGATCCTCGAGATGGTGGGTACCCCTAGGATCAAGGGGTCG ACAACGGCCGGCGATACAAAGCTTCCCGGGATAAACAGGGCC GGCCTATAAT
BACD–Sense	TCGATCCCTAGGATCAAGGGGTCGACACTCGAGATGGTGGGT ACCCCCGGGATAAACAGGGCCGGCCTAACGGCCGGCGATACA AAGCTTTAAT
BADC–Sense	TCGATCCCTAGGATCAAGGGGTCGACACTCGAGATGGTGGGT ACCACGGCCGGCGATACAAAGCTTCCCGGGATAAACAGGGCC GGCCTATAAT
CABD–Sense	TCGATCCCCGGGATAAACAGGGCCGGCCTACTCGAGATGGTG GGTACCCCTAGGATCAAGGGGTCGACAACGGCCGGCGATACA AAGCTTTAAT
DBAC–Sense	TCGATCACGGCCGGCGATACAAAGCTTCCTAGGATCAAGGGG TCGACACTCGAGATGGTGGGTACCCCCGGGATAAACAGGGCC GGCCTATAAT
CDAB-Sense	TCGATCCCCGGGATAAACAGGGCCGGCCTAACGGCCGGCGAT ACAAAGCTTCTCGAGATGGTGGGTACCCCTAGGATCAAGGGG TCGACATAAT

DCAB–Sense	TCGATCACGGCCGGCGATACAAAGCTTCCCGGGATAAACAGG GCCGGCCTACTCGAGATGGTGGGTACCCCTAGGATCAAGGGG TCGACATAAT
CDBA–Sense	TCGATCCCCGGGATAAACAGGGCCGGCCTAACGGCCGGCGAT ACAAAGCTTCCTAGGATCAAGGGGTCGACACTCGAGATGGTG GGTACCTAAT
DCBA–Sense	TCGATCACGGCCGGCGATACAAAGCTTCCCGGGATAAACAGG GCCGGCCTACCTAGGATCAAGGGGTCGACACTCGAGATGGTG GGTACCTAAT
CADB–Sense	TCGATCCCCGGGATAAACAGGGCCGGCCTACTCGAGATGGTG GGTACCACGGCCGGCGATACAAAGCTTCCTAGGATCAAGGGG TCGACATAAT
DBCA–Sense	TCGATCACGGCCGGCGATACAAAGCTTCCTAGGATCAAGGGG TCGACACCCGGGATAAACAGGGCCGGCCTACTCGAGATGGTG GGTACCTAAT
ACDB–Antisense	AGCTATTATGTCGACCCCTTGATCCTAGGAAGCTTTGTATCGC CGGCCGTTAGGCCGGCCCTGTTTATCCCGGGGGTACCCACCAT CTCGAGGA
BDCA–Antisense	AGCTATTAGGTACCCACCATCTCGAGTAGGCCGGCCCTGTTTA TCCCGGGAAGCTTTGTATCGCCGGCCGTTGTCGACCCCTTGAT CCTAGGGA
ACBD–Antisense	AGCTATTAAAGCTTTGTATCGCCGGCCGTTGTCGACCCCTTGA TCCTAGGTAGGCCGGCCCTGTTTATCCCGGGGGTACCCACCAT CTCGAGGA
BDAC–Antisense	AGCTATTATAGGCCGGCCCTGTTTATCCCGGGGGTACCCACCA TCTCGAGAAGCTTTGTATCGCCGGCCGTTGTCGACCCCTTGAT CCTAGGGA
ABCD–Antisense	AGCTATTAAAGCTTTGTATCGCCGGCCGTTAGGCCGGCCCTGT TTATCCCGGGTGTCGACCCCTTGATCCTAGGGGTACCCACCAT CTCGAGGA
ABDC–Antisense	AGCTATTATAGGCCGGCCCTGTTTATCCCGGGAAGCTTTGTAT CGCCGGCCGTTGTCGACCCCTTGATCCTAGGGGTACCCACCAT CTCGAGGA

BACD–Antisense	AGCTATTAAAGCTTTGTATCGCCGGCCGTTAGGCCGGCCCTGT TTATCCCGGGGGTACCCACCATCTCGAGTGTCGACCCCTTGAT CCTAGGGA
BADC–Antisense	AGCTATTATAGGCCGGCCCTGTTTATCCCGGGAAGCTTTGTAT CGCCGGCCGTGGTACCCACCATCTCGAGTGTCGACCCCTTGAT CCTAGGGA
CABD–Antisense	AGCTATTAAAGCTTTGTATCGCCGGCCGTTGTCGACCCCTTGA TCCTAGGGGTACCCACCATCTCGAGTAGGCCGGCCCTGTTTAT CCCGGGGA
DBAC–Antisense	AGCTATTATAGGCCGGCCCTGTTTATCCCGGGGGTACCCACCA TCTCGAGTGTCGACCCCTTGATCCTAGGAAGCTTTGTATCGCC GGCCGTGA
CDAB–Antisense	AGCTATTATGTCGACCCCTTGATCCTAGGGGTACCCACCATCT CGAGAAGCTTTGTATCGCCGGCCGTTAGGCCGGCCCTGTTTAT CCCGGGGA
DCAB–Antisense	AGCTATTATGTCGACCCCTTGATCCTAGGGGTACCCACCATCT CGAGTAGGCCGGCCCTGTTTATCCCGGGAAGCTTTGTATCGCC GGCCGTGA
CDBA–Antisense	AGCTATTAGGTACCCACCATCTCGAGTGTCGACCCCTTGATCC TAGGAAGCTTTGTATCGCCGGCCGTTAGGCCGGCCCTGTTTAT CCCGGGGA
DCBA–Antisense	AGCTATTAGGTACCCACCATCTCGAGTGTCGACCCCTTGATCC TAGGTAGGCCGGCCCTGTTTATCCCGGGAAGCTTTGTATCGCC GGCCGTGA
CADB–Antisense	AGCTATTATGTCGACCCCTTGATCCTAGGAAGCTTTGTATCGC CGGCCGTGGTACCCACCATCTCGAGTAGGCCGGCCCTGTTTAT CCCGGGGA
DBCA–Antisense	AGCTATTAGGTACCCACCATCTCGAGTAGGCCGGCCCTGTTTA TCCCGGGTGTCGACCCCTTGATCCTAGGAAGCTTTGTATCGCC GGCCGTGA
L-S-Sub-Sense	CCGGGCTTGCTACAAAGGCAGCACGCAAGAGCGCACCTGCTA CAGGAACGGGAACGGCTGGAAGCGCAACGGGTAGTAGGGCC GG

S-L-Sub-Sense	CCGGGACGGGAACGGCTGGAAGCGCAACGGGTAGTCTTGCTACAAAGGCAGCACGCAAGAGCGCACCTGCTACAGGAAGGGCCGG
S-S-Sub-Sense	CCGGGCTTGCTACAAAGGCAGCACGCAAGAGCGCACCTGCTACAGGAAGGGCCGG
L-L-Sub-Sense	CCGGGACGGGAACGGCTGGAAGCGCAACGGGTAGTCTTGCTACAAAGGCAGCACGCAAGAGCGCACCTGCTACAGGAGGATCAACAGCAGGATCAGGAACAGGAGCAAGGGCCGG
L-S-Sub-AntiSense	CCCTACTACCCGTTGCGCTTCCAGCCGTTCCCGTTCCTGTAGCAGGTGCGCTCTTGCGTGCTGCCTTTGTAGCAAGC
S-L-Sub-AntiSense	CCCTTCCTGTAGCAGGTGCGCTCTTGCGTGCTGCCTTTGTAGCAAGACTACCCGTTGCGCTTCCAGCCGTTCCCGTC
S-S-Sub-AntiSense	CCCTTCCTGTAGCAGGTGCGCTCTTGCGTGCTGCCTTTGTAGCAAGC
L-L-Sub-AntiSense	CCCTTGCTCCTGTTCCTGATCCTGCTGTTGATCCTCCTGTAGCAGGTGCGCTCTTGCGTGCTGCCTTTGTAGCAAGACTACCCGTTGCGCTTCCAGCCGTTCCCGTC
0-FrWd-XhoI(RF1)-TFP	CTCGTACTCGAGATGGTGAGCAAGGGCGAGGAGACCACAATGGGC
-2-FrWd-XhoI(RF1)-TFP	GTCGCTCTCGAGAGCAAGGGCGAGGAGACCACAATGGGCGTAATCAAGCC
-4-FrWd-XhoI(RF1)-TFP	GTCATGCTCGAGGGCGAGGAGACCACAATGGGCGTAATCAAGCCCGAC

-6-FrWd-XhoI(RF1)-TFP	GACTCGCTCGAGGAGACCACAATGGGCGTAATCAAGCCCGACATGAAGATCAAGCTG
-7-FrWd-XhoI(RF1)-TFP	GCGGCGCTCGAGACCACAATGGGCGTAATCAAGCCCGACATGAAGATCAAGCTGAAGATGGAG
-8-FrWd-XhoI(RF1)-TFP	GCTCCTCTCGAGACAATGGGCGTAATCAAGCCCGACATGAAGATCAAGCTGAAGATGGAGGGC
0-BkWd-HindIII(RF1)-YFP	GGCGATAAGCTTCTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGC
-2-BkWd-HindIII(RF1)-YFP	CCTAGCAAGCTTCAGCTCGTCCATGCCGAGAGTGATCCCGGC
-4-BkWd-HindIII(RF1)-YFP	TACAGCAAGCTTGTCCATGCCGAGAGTGATCCCGGCGGC
-6-BkWd-HindIII(RF1)-YFP	ATCCAGAAGCTTGCCGAGAGTGATCCCGGCGGCGGTCAC
-7-BkWd-HindIII(RF1)-YFP	CCATGCAAGCTTGAGAGTGATCCCGGCGGCGGTCACGAACTC
-8-BkWd-HindIII(RF1)-YFP	GGAAGCAAGCTTAGTGATCCCGGCGGCGGTCACGAACTCCAG
0-FrWd-AvrII-(RF1)-Cbx7	GATAGCCCTAGGATGGAGCTGTCAGCCATAGGCGAGCAGGTG
5-FrWd-AvrII-(RF1)-Cbx7	GATAGCCCTAGGGGATCAACAGCAGGAATGGAGCTGTCAGCCATAGGCGAGCAGGTG

10-FrWd-AvrII-(RF1)-Cbx7	GATATCCCTAGGAGTGGAACGGGAACGGCTGAAAGCGCAAC GGGTAGTATGGAGCTGTCAGCCATAGGCGAGCAGGTG