Novel genetically encoded biosensors for cGMP

Dissertation an der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

Arne Fabritius
München, 2016

Diese Dissertation wurde angefertigt
unter der Leitung von Dr. Oliver Griesbeck
am Max Planck Institut für Neurobiologie

Erstgutachter:	Dr. Oliver Grisbeck
Zweitgutachter:	Prof. Elisabeth Weiß
Tag der Abgabe:	7. April 2016
Tag der mündlichen Prüfung:	2. Dezember 2016

Erklärung

Hiermit versichere ich an Eides statt, dass meine Dissertation selbständig und ohne unerlaubte Hilfsmittel angefertigt worden ist. Die vorliegende Dissertation wurde weder ganz, noch teilweise bei einer anderen Prüfungskommission vorgelegt. Ich habe noch zu keinem früheren Zeitpunkt versucht, eine Dissertation einzureichen oder an einer Doktorprüfung teilzunehmen.

München, den 7. April 2016

in memory of my mom,
who passed away the day I was given the opportunity of this project
I finished it
you are missed by all of us

Abstract

Cyclic guanosine monophosphate (cGMP) is an important second messenger and plays a key role in the maintenance of cardio-vascular homeostasis. In the central nervous system cGMP signalling has been linked to long term potentiation (LTP) in the hippocampus and the amygdala, hinting at its involvement in the processes of learning and memory formation. Despite the complexity of the systems cGMP is involved in, the standard methods for monitoring cGMP oscillations still rely on destructive, *ex vivo* concentration determination in competitive binding assays. First attempts have been made by several groups to develop more sophisticated methods based on genetically encoded probes. However, these first generation tools show severe limitations, impeding their widespread use *in vivo*.

Here we systematically addressed these limitations in a methodical approach and developed a bright, genetically encoded FRET sensor for cGMP based on a cyan/yellow pair. Our sensor exhibits a large dynamic range, high selectivity and simple binding kinetics, outperforming established cGMP biosensors.

Furthermore we developed single emission biosensors based on FRET of a red fluorescent protein to a dark acceptor, demonstrating the validity of this approach for cGMP sensors. This class of red fluorescent biosensors can harness more favourable imaging conditions in the red part of the visible spectrum, such as reduced scattering and auto fluorescence, while allowing the combination with other optogenetic tools.

In the process of these efforts we established a generic pipeline for fast biosensor prototyping and assessment, relying on modern assembly cloning methods in a harmonized vector system. Furthermore, we describe the construction of a low-cost automated screening set-up for directed evolution of genetically encoded biosensors and fluorescent proteins in bacteria. This screening set-up and the streamlined prototyping approach lower the entry hurdles in the field of biosensor development and are applicable for a wide variety of potential target molecules. Our pipeline allows other researchers to develop essential tools to elucidate the inner workings of cellular signalling circuits.

Table of Contents

Table of Figures

Table of Tables

Abbreviations

Table 0-1 General abbreviations

Abbreviation	Meaning
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ANP	atrial natriuretic peptide
API	application programming interface
ATP	adenosine triphosphate
avGFP	aequoria victoria GFP
BNP	B-type natriuretic peptide
Ca^{2+}	calcium ion
cAMP	cyclic adenosine monophosphate
CCD	charge-coupled device
CFP	cyan fluorescent protein
cGK	cGMP dependent protein kinase
cGMP	cyclic guanosine monophosphate
CNB	cyclic nucleotide binding site
CNC	computerized numeric control
CNG	cGMP gated ion channels
CNP	C-type natriuretic peptide
CP	chromoprotein
crp	cAMP receptor protein
CVD	cardio-vascular diseases
dH_2O	distilled dihydrogen monoxide
DNA	deoxyribonucleic acid
dNTP	deoxy nucleotide triphosphate
DoE	design of experiments
dsDNA	double strand DNA
DTT	dithiothreitol
e.g.	exempli gratia, for example
EC_{50}	half maximal effective concentration
ECFP	enhanced CFP
EGFP	enhanced GFP
ELISA	enzyme-linked immunosorbent assay
eNOS	endothelial NO synthase
EROS	erectile optogenetic stimulator
et al.	et alii, and others
EtBr	ethidium bromide
EYFP	enhanced YFP
FDM	fused deposition modeling
FDR	false discovery rate

Table 0-1 General abbreviations (continued)

Abbreviation	Meaning
FLIM	fluorescence life time imaging
FP	fluorescent protein
FRET	Förster resonance energy transfer
GECI	genetically encoded calcium indicator
GFP	green fluorescent protein
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
GuHCl	guanidinium chloride
H2O	dihydrogen monoxide
HCN	hyperpolarization-activated cyclic nucleotide gated ion channel
i.e.	id est, that is
K_d	dissociation constant
LB	lysogenic broth
LC	liquid chromatography
LED	light emitting diode
LTD	long term depression
LTP	long term potentiation
m/z	mass divided by charge
MCS	multiple cloning site
MDF	medium density fiber plate
MS	mass spectrometry
Ni^{2+}	nickel ion
NMDA	N-methyl-D-aspartate
nNOS	neuronal NO synthase
NO	nitric oxide
NP	natriuretic peptide
OD	optical density
OE-PCR	overlap extension polymerase chain reaction
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDE	phosphodiesterase
pGC	particulate guanylate cyclase
pH	negative decadic logarythm of the hydrogen ion concentration
PP_i	pyrophosphate
RAA	renin/angiotensin/aldosterone system
RBS	ribosome binding site
RIA	radioimmunoassay
RT	room temperature, ~21 °C
SD	sensor domain
sGC	soluable guanylate cyclase
SLiCE	seamless ligation cloning extract

Table 0-1 General abbreviations (continued)

Abbreviation	Meaning
ssDNA	Single strand DNA
UV	ultra violet
v/v	volume per volume (concentration ratio)
w/v	weight per volume (concentration ratio)
YFP	yellow fluorescent protein

Table 0-2 Units and prefixes

Unit	Meaning	
L	liter	
g	gram	
M	molar	
m	meter	
bp	base pairs	
s	second	
min	minutes	
h	hour	
U	unit [arbitrary definition of enzyme activity]	
a.u.	arbitrary units	
x *g*	multiple of standard acceleration due to gravity	
rpm	rounds per minute	
W	watt	
Prefix	Meaning	[factor]
P	peta	[10^{15}]
T	tera	[10^{12}]
G	giga	[10^{9}]
M	mega	[10^{6}]
K	kilo	[10^{3}]
c	centi	[10^{-2}]
m	milli	[10^{-3}]
µ	micro	[10^{-6}]
n	nano	[10^{-9}]
p	pico	[10^{-12}]
f	femto	[10^{-15}]

Table 0-3 Abbreviations in nucleotide sequences

Abbreviation	Nucleotides		
A	**A**denine		
C	**C**ytosine		
G	**G**uanosine		
T	**T**hymidine		
U	**U**racil	[only RNA]	
K	Guanosine or Thymidine	[**K**etones]	G, T
M	Adenine or Cytosine	[A**m**ines]	A, C
R	Adenine or Guanosine	[Pu**r**ines]	A, G
Y	Cytosine or Thymidine	[P**y**rimidines]	C, T
S	Cytosine or Guanosine	[**S**trong bonds]	C, G
W	Adenine or Thymidine	[**W**eak bonds]	A, T
B	All but **A**denine		C, G, T
D	All but **C**ytosine		A, G, T
H	All but **G**uanosine		A, C, T
V	All but **T**hymidine		A, C, G
N	A**n**y nucleotide		A, C, G, T

1 Introduction

1.1 Fluorescent Proteins

1.1.1 Physical Principles of Fluorescence

The phenomena of fluorescence has been amply discussed by the literature (for review see Valeur and Berberan-Santos (2012)) and its relevant processes are best described by the Perrin–Jablonski diagram (see Figure 1.1). Energy of an electromagnetic wave is first absorbed by a chromophore. This energy is transferred to an electron, which is then excited from the ground state S_0 to one of the vibrational levels of S_1 or S_2. These electrons then undergo a fast, internal conversion and fall to the lowest level of the excited state S_1 in a non-radiative fashion. De-excitation can now occur through several processes, which stand in direct competition, and highly depend on the structure of the chromophore, its environment, and the relative position of its energy levels, namely: (1) Non-radiative relaxation; energy of the excited state is lost from the system through vibration and contact with the solvent without the emission of a photon. (2) Inter system crossing; energy is transferred to the excited triplet state T_1 and subsequently emitted through phosphorescence. (3) Fluorescence, the energy of the excited state is emitted as a photon. Eventually the excited electron returns to the ground state S_0.

The emitted light of fluorescence highly depends on the light which can be absorbed by the chromophore. Only photons, which match the energy difference between the ground state S_0 and the excited states S_1 and S_2 and their vibrational levels can be absorbed. The fluorescence emission spectrum is red-shifted in relation to the absorption spectrum, because part of the initially absorbed energy is removed from the system during internal conversion. Initially

excited electrons are lifted to unstable vibrational levels of S_1 and S_2 and subsequently drop to the lowest level of S_1, which is the only level from which fluorescence emission occurs. Emitted photons carry only the energy difference between S_1 and S_0, but not of the initial drop during internal conversion.

An important characteristic of a given chromophore is its quantum yield Φ. It is defined as the fraction of emitted photons to the number of absorbed photons. From an electron standpoint, it can be seen as the number of excited electrons returning to the ground state S_0 by emission of a photon divided by the total of all de-excited electrons. Fluorescence competes with non-radiative de-excitation, usually achieved by molecular vibration or collision with solvent molecules. Chromophores, less able to vibrate due to being rigid or stabilized, therefore have higher quantum yields, because more energy has to be emitted through photons.

In fluorescent proteins, typically π-electrons composing the conjugated system of the chromophore, are elevated to their excited state π^*. In general, the extent of the conjugated π-system determines the required energy for the transition into the excited state; larger π-systems lower the difference between the energy levels and can therefore get excited by photons with longer wavelengths.

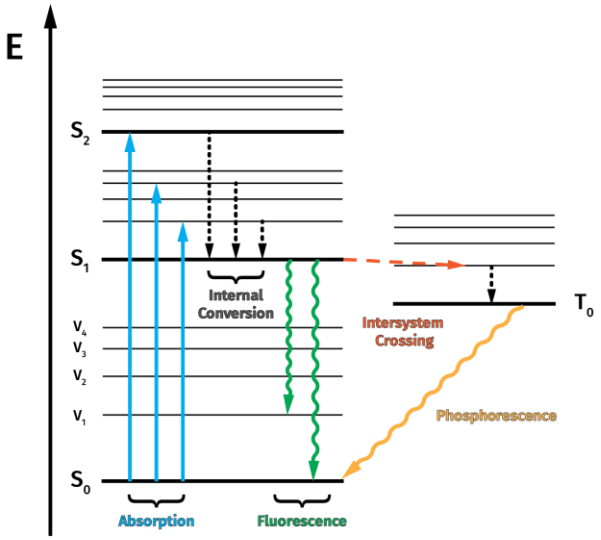

Figure 1.1 **Perrin-Jablonski diagram**

1.1.2 Physical Principles of Förster Resonance Energy Transfer

The theory behind Förster resonance energy transfer (FRET) was first laid down by the physical chemist Theodor Förster in 1946 and further developed in later works (Förster 1946, Förster 1948, Förster 1951). (for in depth review see Valeur and Berberan-Santos (2012)).

FRET is a phenomenon in which energy from an excited donor chromophore is transferred non-radiatively to an acceptor chromophore in close vicinity, exciting it in the process. FRET is a long-range dipole–dipole coupling; the excitation of the donor leads to a charge transfer within the donor chromophore, which subsequently induces a charge transfer in the acceptor chromophore, mimicking the excitation through a photon. This energy transfer is a non-radiative relaxation from the perspective of the excited donor, because the excited electrons are transferred to the ground state S_0, without the emission of a photon. In turn, the acceptor molecule gets excited through this process and can relax this excitation through fluorescence if its structure and environmental conditions are permissive.

The efficiency of this energy transfer (Φ_T) highly depends on the photo-physical properties of the involved chromophores, their distance, orientation and environment. Φ_T is given as:

$$\Phi_T = \frac{1}{1 + (r/R_0)^6} \tag{1.1}$$

r is the distance between donor and acceptor
R_0 is the Förster radius for a specific pair of donor and acceptor

Formula 1.1 reveals that the efficiency of transfer is highly dependent on the distance between chromophores, as it falls off with the inverse to the power of six of r. Typically this distance is in the range of 1–10 nm.

The Förster radius R_0 is defined as the distance at which the efficiency of energy transfer Φ_T is 50 %, for a specific pair of donor and acceptor chromophore. It can be calculated as:

$$R_0 = \left(\frac{9(ln10)\kappa^2\Phi_D^0}{128\pi^5 N_A n^4} J\right)^{1/6} \quad (1.2)$$

κ^2 is the orientation factor between donor and acceptor chromophore
Φ_D^0 is the quantum yield of the donor chromophore in the absence of the acceptor
J is the spectral overlap between the emission spectrum of the donor and the excitation spectrum of the acceptor
N_a is Avogadro's constant
n is the refractive index of the medium

From formula 1.2 several important dependencies become apparent: (1) the orientation of the chromophores is quantified by the scalar factor κ^2. When the transitional dipole moments of the chromophores are orthogonal, κ^2 = 0. If they are parallel κ^2 becomes 1 and if they are collinear κ^2 = 4. In freely rotating chromophores, κ^2 averages out to $2/3$. This shows the dramatic effect the orientation of the chromophore can have; orthogonal transitional dipoles reduce the Förster radius to zero, consequently making FRET impossible. (2) The quantum yield of the donor chromophore Φ_D^0 is an important characteristic of the FRET pair. As with κ^2, low values for Φ_D^0 reduce the Förster radius and thus FRET efficiency for a given distance. Φ_D^0 is a measure of the amount of excited state energy that can be lost to vibrational relaxation and therefore does not take part in the dipole coupling. This means in turn, that care needs to be taken when picking a donor chromophore for a specific FRET pair, since low quantum yields result in reduced potential for FRET. (3) A further factor to consider in the choice of FRET partners is the spectral overlap J. It is the integral overlap of the donor emission spectrum and the acceptor excitation spectrum. Counterintuitively this suggests that a photon is emitted by the donor chromophore and is then subsequently absorbed by the acceptor, but this is in fact not the case, since FRET is a non-radiative interaction. The spectral overlap is rather a measure of the compatibility of the energy levels of the excited state of the donor and the acceptor. Energy transition can only occur, if energy levels are matched, much like photon energy needs to match the difference between ground state and excited state.

The spectral overlap J is calculated by formula 1.3.

$$J = \int_0^\infty F_\lambda(\lambda)\varepsilon(\lambda)\lambda^4 d\lambda \tag{1.3}$$

F_λ is the normalized emission spectrum of the donor
ε is the absorption coefficient of the acceptor

The spectral overlap integral of a given FRET pair depends not only on the shape of the excitation spectrum of the acceptor chromophore, but also on its extinction coefficient. Therefore, ε is an important factor to consider, when picking potential acceptors. Two fluorescent proteins may, for example, have very similar excitation spectra, but very different extinction coefficients, making the fluorescent protein with the higher extinction coefficient more suitable as an acceptor. Taken together, to achieve large Förster radii, donors should display very high quantum yields and acceptors high extinction coefficients, while emission spectra of the donors should match the excitation spectrum of the acceptors.

FRET efficiency can be indirectly measured by exciting the donor chromophore and monitoring the emission of the donor and the emission of the acceptor. Since FRET decreases the emission of the donor and results in fluorescence of the acceptor, the ratio of these two values approximates FRET efficiency. As described above, the distance of the chromophores and their orientation has a large influence on FRET. Therefore, the observed FRET ratio can be used to estimate both of these variables. Especially in intramolecular FRET, conformational changes in the molecule that harbors both donor and acceptor can be easily monitored.

1.1.3 The Green Fluorescent Protein

The green fluorescent protein (GFP) was first described by Shimomura et al. (1962) and further characterized (Johnson et al. 1962) as a by-product in the extraction procedure of the luminescent protein aequorin from the jellyfish *Aequorea Victoria.* It was mentioned in a mere side note in the original publication. The gene, encoding aequoria GFP (*av*GFP), was cloned by Prasher et al. (1992) and subsequently demonstrated to be sufficient, to produce recombinant fluorescent protein without co-factors or additional proteins by Chalfie et al. (1994), Inouye and Tsuji (1994). Since then, GFP and its engineered variants have become some of the most important and versatile tools in cell biology and spawned a plethora of applications. For their contributions to this development, Osamu Shimomura, Martin Chalfie and Roger Y. Tsien were awarded the Nobel Prize for Chemistry in 2008 for the discovery, recombinant expression and engineering of the green fluorescent protein, respectively.

The 238 amino acids of GFP form a 11 stranded β-barrel tertiary structure with a central α-helix harboring the chromophore (Ormö et al. 1996, Yang et al. 1996) (see Figure 1.2).

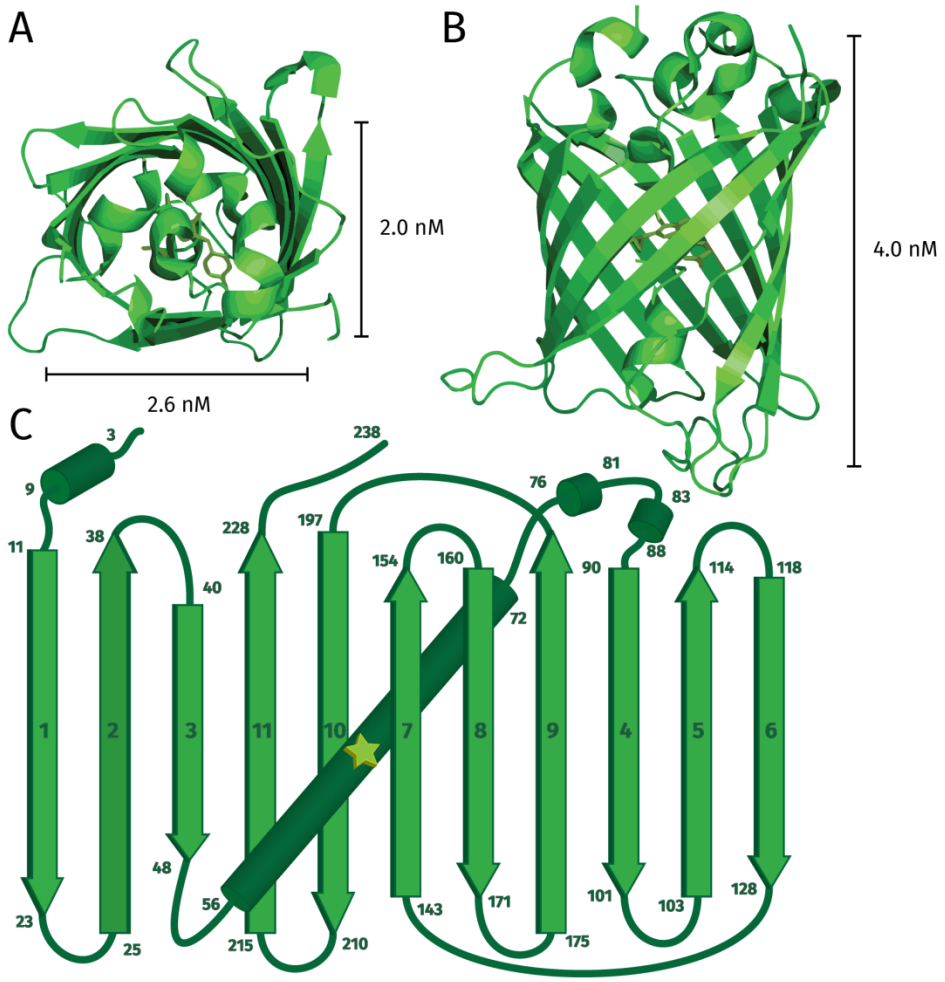

Figure 1.2 Structure of EGFP

Crystal structure of EGFP (Arpino et al. 2012) A) top view B) side view C) topology map of EGFP

The chromophore itself is formed by an autocatalytic conversion of three amino acid side chains, namely Serine 65, Tyrosine 66 and Glycine 67. The formation of the chromophore is depicted in Figure 1.3. First the amid-group of Gly67 performs a nucleophilic attack towards the carbonyl group of Ser65, cycling the backbone. An imidazolinone is then formed by eliminating a water molecule. Finally, the α–β bond of Tyr66 is dehydrogenated by molecular oxygen, creating a molecule of hydrogen peroxide in the process (Tsien 1998). The requirement of oxygen for this dehydrogenation is a limiting factor in the use of GFP in anaerobic organisms (Inouye and Tsuji 1994). This step introduces a conjugated double bond between the phenol group of tyrosine and the newly formed imidazolinone, joining their conjugated π-electron systems. This larger π-system gives rise to the photo-physical properties of GFP. The isolated chromophore (a p-hydroxybenzyl ideneimidazolinone) is barely fluorescent in solution, but is stabilized within the protein structure by the surrounding scaffold, which prevents non-radiative relaxation of the excited state (Niwa et al. 1996).

GFP shows a strong absorption band at 488 nm and a fluorescence emission maximum of 508 nm. The wild-type also exhibits a very strong absorption band at 397 nm, representing the protonated form of the tyrosine residue inside the chromophore.

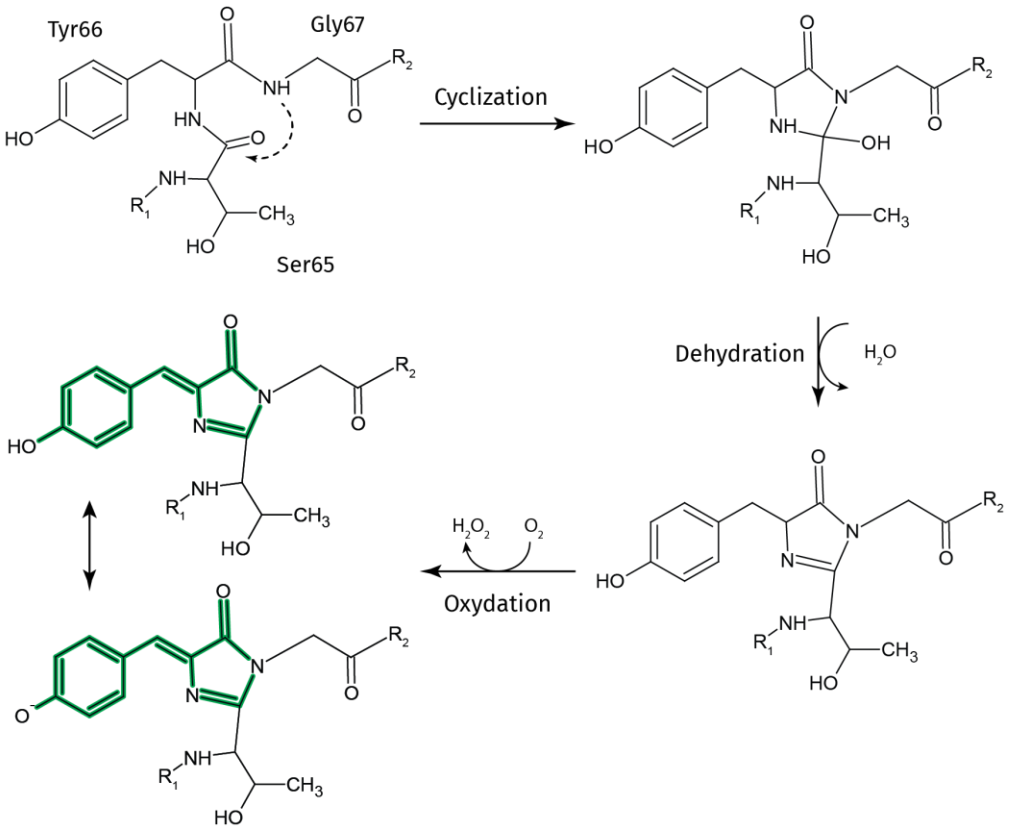

Figure 1.3 Chromophore formation in GFP

1.1.4 Engineered Fluorescent Proteins

Since the first successful cloning of *av*GFP in 1992 (Prasher et al.) and subsequent expression in heterologous organisms, much effort was put into improving the photo-physical properties of the wild-type protein through mutagenesis. Substitutions of aromatic amino acids at the Tyr66 residue, for example, lead to blue shifted variants, while the introduction of aromatic residues at position 203 promoted π-stacking with the chromophore, which in turn shifted the emission spectrum towards yellow emission (Tsien 1998). Besides color variants many other improvements were made, which addressed different problems of the wild-type protein. These improvements include: (1) increase in quantum efficiency of fluorescence emission, (2) increase in absorption coefficient, (3) photo-stability, (4) folding efficiency, (5) pH-stability, (6) monomerization, and (7) decrease in cytotoxicity.

The discovery of other fluorescent proteins in corals, homologue to GFP, which emit in the red part of the visible spectrum, by Matz et al. (1999) lead to an expansion of the color pallet. Since then, many GFP-like fluorescent proteins have been found in different anthozoas and even vertebrates (Shaner et al. 2013). These new fluorescent proteins and GFP variants are the foundation of an ongoing engineering process taking places for the last 20 years. To date, many improvements have been made by various groups, leading to hundreds of variants of fluorescent proteins. Going into detail about all of them, or even just a fraction, is beyond the scope of this work. Instead, Table 1-1, provides a non-inclusive list of some notable engineered fluorescent proteins and their photo-physical properties.

Table 1-1 Fluorescent proteins and their photo-physical properties

Protein	λ_{ex} [nm]	λ_{em} [nm]	ε [$mM^{-1}cm^{-1}$]	Φ	ε x Φ	pK_a	Bleaching [s]	Maturation [min]	Reference
Blue Fluorescent Proteins									
Azurite	383	450	26.2	0.55	14.4	5.0	33		(Mena et al. 2006)
EBFP2	383	448	32.0	0.56	18.0	4.5	55		(Ai et al. 2007)
mKalama1	385	456	36.0	0.45	16.0	5.5			(Ai et al. 2007)
mTagBFP2	399	454	50.6	0.64	32.4	2.7	53		(Subach et al. 2011)
mTagBFP	402	457	52.0	0.63	32.8	2.7	34		(Subach et al. 2008)
Cyan Fluorescent Proteins									
ECFP	433	475	32.5	0.4	13.0	4.7			(Heim et al. 1994)
Cerulean	433	475	43.0	0.62	26.7	4.7	36		(Rizzo et al. 2004)
mCerulean3	433	475	40.0	0.8	32.0	4.7			(Markwardt et al. 2011)
SCFP3A	433	474	30.0	0.56	16.8	4.5			(Kremers et al. 2006)
CyPet	435	477	35.0	0.51	17.8	5.02			(Nguyen and Daugherty 2005)
mTurquoise	434	474	30.0	0.84	25.2				(Goedhart et al. 2010)
mTurquoise2	434	474	30.0	0.93	27.9	3.1	90		(Goedhart et al. 2012)
mTFP1	462	492	64.0	0.85	54.0	4.3			(Ai et al. 2006)
monomeric Midoriishi-Cyan	470	496	22.2	0.7	15.5	7.0			(Karasawa et al. 2004)
Aquamarine	430	474	26.0	0.89	23.1	3.3			(Erard et al. 2013)
Green Fluorescent Proteins									
TurboGFP	482	502	70.0	0.53	37.1	5.2			(Evdokimov et al. 2006)
TagGFP2	483	506	56.5	0.6	33.9	4.7			(Subach et al. 2008)
mUKG	483	499	60.0	0.72	43.2	5.2			(Tsutsui et al. 2008)
Superfolder GFP	485	510	83.3	0.65	54.1				(Pédelacq et al. 2006)
Emerald	487	509	57.5	0.68	37.3	6.0	101		(Cubitt et al. 1999)
EGFP	488	507	56.0	0.6	33.6	6.0	174	25	(Yang et al. 1996)
Monomeric Azami Green	492	505	55.0	0.74	40.7	5.8			(Karasawa et al. 2003)
mWasabi	493	509	70.0	0.8	56.0	6.0	93		(Ai et al. 2008)
Clover	505	515	111.0	0.76	84.4	6.1	50	30	(Lam et al. 2012)
mClover3	506	518	109.0	0.78	85.0	6.5	80		(Bajar et al. 2016)
mNeonGreen	506	517	116.0	0.8	92.8	5.7	158	10	(Shaner et al. 2013)
Yellow Fluorescent Proteins									
EYFP	513	527	83.4	0.61	50.9	6.9	60		(Ormö et al. 1996)
Topaz	514	527	94.5	0.6	56.7				(Cubitt et al. 1999)
Venus	515	528	92.2	0.57	52.5	6.0	15		(Nagai et al. 2002)
SYFP2	515	527	101.0	0.68	68.7	6.0			(Kremers et al. 2006)
Citrine	516	529	77.0	0.76	58.5	5.7	49		(Griesbeck et al. 2001)
Ypet	517	530	104.0	0.77	80.1	5.6			(Nguyen and Daugherty 2005)
mPapaya1	530	541	43.0	0.83	35.7	6.8			(Hoi et al. 2013)

Table 1-2 Fluorescent proteins and their photo-physical properties (continued)

Protein	λ_{ex} [nm]	λ_{em} [nm]	ε [$mM^{-1}cm^{-1}$]	Φ	ε x Φ	pK_a	Bleaching [s]	Maturation [min]	Reference
Orange Fluorescent Proteins									
Monomeric Kusabira-Orange	548	559	51.6	0.6	31.0	5.0			(Karasawa et al. 2004)
tdTomato	554	581	138.0	0.69	95.2	4.7	70	60	(Shaner et al. 2004)
mOrange	548	562	71.0	0.69	49.0	6.5	9	150	(Shaner et al. 2004)
mOrange2	549	565	58.0	0.6	34.8	6.5			(Shaner et al. 2008)
mKOκ	551	563	105.0	0.61	64.0	4.2			(Tsutsui et al. 2008)
mKO_2	551	565	63.8	0.62	39.6	5.5			(Sakaue-Sawano et al. 2008)
Red Fluorescent Proteins									
TagRFP	555	584	100.0	0.48	49.0	3.8	48	100	(Merzlyak et al. 2007)
TagRFP-T	555	584	81.0	0.41	33.2	4.6	337	100	(Shaner et al. 2008)
mRuby	558	605	112.0	0.35	39.2	4.4			(Kredel et al. 2009)
mRuby2	559	600	113.0	0.38	43.0	5.3	123	150	(Lam et al. 2012)
mRuby3	558	592	128.0	0.45	58.0	4.8	349		(Bajar et al. 2016)
mTangerine	568	585	38.0	0.3	11.4	5.7			(Shaner et al. 2004)
mApple	568	592	75.0	0.49	36.7	6.5			(Shaner et al. 2008)
mStrawberry	574	596	90.0	0.29	26.1	4.5	15	50	(Shaner et al. 2004)
FusionRed	580	608	95.0	0.19	18.1	4.6	150	130	(Shemiakina et al. 2012)
mCherry	587	610	72.0	0.22	15.8	4.5	96	40	(Shaner et al. 2004)
mNectarine	558	578	58.0	0.45	26.1	6.9	11	30	(Johnson et al. 2009)
Far Red Fluorescent Proteins									
mKate	588	635	45.0	0.33	14.9	6.0			(Shcherbo et al. 2007)
mKate2	588	633	62.5	0.4	25.0	5.4	84	20	(Shcherbo et al. 2009)
mLumin	587	621	70.0	0.46	32.2	4.7	166		(Chu et al. 2009)
mPlum	590	649	41.0	0.1	4.1	4.5	53	100	(Wang et al. 2004)
mRaspberry	598	625	86.0	0.15	12.9				(Wang et al. 2004)
mNeptune	600	650	67.0	0.2	13.4	5.4	255	35	(Lin et al. 2009)
mNeptune2.5	599	643	95.0	0.28	26.6	5.8	506	26	(Chu et al. 2014)
eqFP650	592	650	65.0	0.24	15.6	5.7	190		(Shcherbo et al. 2010)
eqFP670	605	670	70.0	0.06	4.2	4.5	1289		(Shcherbo et al. 2010)
TagRFP657	611	657	34.0	0.1	3.4	5.0			(Morozova et al. 2010)
TagRFP675	598	675	46.0	0.08	3.7	5.7		25	(Piatkevich et al. 2013)
mCardinal	604	659	87.0	0.19	16.5	5.3	730	27	(Chu et al. 2014)

Table 1-2 Fluorescent proteins and their photo-physical properties (continued)

Protein	λ_{ex} [nm]	λ_{em} [nm]	ε [$mM^{-1}cm^{-1}$]	Φ	ε x Φ	pK_a	Bleaching [s]	Maturation [min]	Reference
Sapphire Fluorescent Proteins									
Sapphire	399	511	29.0	0.64	18.6				(Cubitt et al. 1999)
T-Sapphire	399	511	44.0	0.6	26.4	4.9	25		(Zapata-Hommer and Griesbeck 2003)
mAmetrine	406	526	45.0	0.58	26.1		2.8		(Ai et al. 2008)
Long Stokes Shift Fluorescent Proteins									
mKeima Red	440	620	14.4	0.24	3.5	6.5			(Kogure et al. 2006)
mBeRFP	446	611	65.0	0.27	17.6	5.6			(Yang et al. 2013)
LSS-mKate2	460	605	26.0	0.17	4.4	2.7			(Piatkevich et al. 2010)
LSS-mKate1	463	624	31.2	0.08	2.5	3.2			(Piatkevich et al. 2010)
LSSmOrange	437	572	52.0	0.45	23.4	5.7		138	(Shcherbakova et al. 2012)

1.1.5 Chromoproteins

Chromoproteins are defined as compounds consisting of a protein and a pigment, which is usually, but not always, bound as a co-factor (Fearon 1940). For the purpose of this work, a narrower definition of chromoproteins is used: Chromoproteins are a sub-class of the GFP-like protein family and possess a pigment, which is formed through auto-catalysis from its poly-peptide chain. These proteins absorb strongly in the visible spectrum and are therefore intensely colored. On the other hand, they exhibit very small quantum yields and consequently little to no fluorescence. Synonymously, the term pocilloporin has been used for these proteins (Dove et al. 1995), referring to the family of corals this class of proteins had been first isolated from.

Many chromoproteins were isolated from anthozoa species, specifically corals and sea anemones (Gurskaya et al. 2001, Labas et al. 2002, Shagin et al. 2004, Chan et al. 2006, Alieva et al. 2008). In these species chromoproteins function as a pigment and are the source of their coloration (Dove et al. 2001). Typically, they exist as dimers, tetramers or other higher order oligomers (Alieva et al. 2008); which is detrimental to their use in biosensors, as these oligomers tend to aggregate and can lead to cell toxicity when recombinantly expressed (Shaner et al. 2005). Recently, a monomeric chromoprotein, Ultramarine, was engineered by Pettikiriarachchi et al. (2012) from the oligomeric precursor Rtms5, originating from the reef building coral *Montipora efflorescens*.

In contrast, efforts were made to engineer fluorescent proteins with very low quantum yields (Ganesan et al. 2006, Murakoshi et al. 2008); which were used as dark acceptors for FRET sensors in fluorescent lifetime imaging (FLIM). FLIM utilizes the fact, that the fluorescent lifetime of a donor fluorophore is altered by the presence of a suitable FRET acceptor. Since only the lifetime of the donor has to be monitored, the spectrum can be freed up by using a dark acceptor. However, these dark acceptors still retain some fluorescence, which can potentially interfere with other probes (Ganesan et al. 2006, Murakoshi et al. 2008).

Pettikiriarachchi et al. (2012) demonstrated, that the engineered chromoprotein Ultramarine can be successfully used as a dark acceptor in a FRET pair, overcoming both the potential problems from oligomerization and retained fluorescence. This strategy seemed very promising and was therefore applied in the development of a FRET based cGMP sensor. Besides FLIM, such a FRET sensor can also be used in standard imaging as a single fluorophore biosensor, in which the brightness of a fluorescent protein is modulated by the presence of a ligand molecule.

1.2 Optogenetics

"I suggest that neuroscientists should tell molecular biologists what their difficulties are, in the hope that this will stimulate the production of useful new biological tools"

Francis Crick (1999)

Optogenetics is the combination of genetic targeting of specific cells with optical methods, which either monitor or control these cells with light (Miesenböck 2009). The field of optogenetics originated from neuroscience to address the problem of analyzing complex neuronal circuits in the brain in a non-invasive manner. Due to their broad applicability, optogenetical methods are now spreading to the wider field of cell biology (Toettcher et al. 2011).

In principle, the optogenetic approach incorporates the recombinant expression of engineered light sensitive proteins under the control of cell specific promoters, to address a cell population of interest. These cell populations are then selectively probed and manipulated with light, which can be monitored and applied in a very spatially and temporally controlled manner. Broadly speaking, there are two types of optogenetic tools: actuators and reporters (Miesenböck 2009).

Optogenetic actuators are light sensitive proteins, which change the internal state of the cell expressing them upon illumination. These proteins were initially often light sensitive ion channels, which let ions pass after being activated by light, stimulating or inhibiting the host cell in the process (Nagel et al. 2005, Deisseroth 2015). To date, a wide variety of actuators is available, which stimulate processes as diverse as gene expression (Motta-Mena et al. 2014), second messenger synthesis (Stierl et al. 2011) and light activated degradation (Gasser et al. 2014).

1.2.1 Genetically Encoded Biosensors

The main focus of this work is on reporters, the second type of optogenetic tools. Optogenetic reporters are engineered proteins, which are sensitive to an internal state of a cell and modulate a light signal according to changes in that particular state. Most often these reporters are modified fluorescent proteins, which are read out via fluorescent microscopy (Dugué et al. 2012). In general, genetically encoded biosensors consist of two types of domains: the sensor domain, which interacts with the molecule to detect and the reporter domain, a single or pair of fluorescent proteins, which translate this change into a detectable light signal (Tantama et al. 2012). The main advantage of these types of biosensors over other detection methods, like chemical dyes, is that they are self-contained within a single poly-peptide chain and can be genetically targeted to specific cell populations and even compartments within cells (Mank and Griesbeck 2008).

The development of these molecular probes is exemplified by genetically encoded calcium indicators (GECIs). GECIs were the first genetically encoded fluorescent biosensors (Miyawaki et al. 1997) and have seen extensive engineering efforts over the last two decades to bring them to perfection (Mank and Griesbeck 2008, Rose et al. 2014, Miyawaki and Niino 2015). GECIs are biosensors, which bind intracellular calcium and modulate their fluorescence intensity according to the calcium concentration. Because changes of the internal calcium concentration can be directly linked to neuronal activity (Mank and Griesbeck 2008), sensors which can detect these changes were very important for neuroscience, likely encouraging their early development and intensive improvements. Besides calcium, sensors haven been developed for a wide range of inorganic and organic compounds: zinc (Hessels and Merkx 2014), chloride (Arosio and Ratto 2014), cyclic nucleotides (Nikolaev and Lohse 2009, Paramonov et al. 2015), glutamate (Liang et al. 2015), ATP (Tantama et al. 2013) and many more (Tantama et al. 2012, Miyawaki and Niino 2015). Also, biophysical and biochemical states can be probed by genetically encoded sensors; examples are: pH (Benčina 2013), temperature (Sakaguchi et al. 2015), redox potential (Chiu et al. 2014), membrane potential (Storace et al. 2015) and tension (Guo et al. 2014).

Focusing on single chain fluorescent biosensors, i.e. sensors which are encoded by a single gene, they can be broadly categorized into two groups: single and dual emission fluorescent sensors (ratio metric sensors). Mechanisms, advantages and disadvantages of these types of sensors are discussed in the following sections.

1.2.1.1 Single Emission Fluorescent Biosensors

Single fluorophore sensors modulate the emission of a fluorescent protein according to a change of an internal state of a cell (e.g. concentration of an analyte). This modulation can be achieved by different means: (1) utilizing the innate environment sensitivity of certain fluorescent proteins, (2) inserting sensor domains into the β-barrel of a fluorescent protein, thus exposing and destabilizing its chromophore and (3) using dark acceptors in FRET sensors.

The fluorescence emission of Tyr66 based chromophores, such as GFP and YFP, is inherently sensitive to changes in their environment, because the hydroxyl group of tyrosine can be protonated, which changes the π-system of the chromophore, blue-shifting its absorption spectrum and quenching its fluorescence (Bizzarri et al. 2009). This property has been used in the past as the basis for fluorescent biosensors, which can detect environmental changes, such as changes in pH (Sankaranarayanan et al. 2000) or chloride concentration (Grimley et al. 2013). However, this strategy only works for very specific molecules and states and is therefore often not applicable.

This environment sensitivity can be enhanced by directly exposing the chromophore to the solvent. Usually, the chromophore is efficiently shielded by the strands of the β-barrel. By opening up the β-barrel directly at the chromophore position and inserting a sensor domain, exposure to the solvent and thus fluorescence intensity can be modulated. Baird et al. (1999) discovered that fluorescent proteins can tolerate the insertion of sensor domains into the β-barrel. The authors inserted the calcium binding protein calmodulin in between residue 144 and 145 of EYFP, resulting in the calcium sensor camgaroo1.

This strategy was further refined with the use of circular permutated fluorescent proteins. In circular permutated fluorescent proteins, the DNA encoding the fluorescent protein is split inside the sequence encoding a β-strand or a loop and the two resulting DNA strands are then swapped in their order, fusing the native N-term with a short linker to the native C-term (Baird et al. 1999). This results in a fluorescent protein, which has a similar overall structure to the wild-type FP, with the exception of having its new N- and C-term very close together and in a different position. If these new termini are located close to the chromophore, even small changes in their relative location can lead to exposure of the chromophore to the solvent. This strategy was successfully implemented by Nakai et al. (2001) to create a calcium sensor, by fusing the calcium binding protein calmodulin and the M13 peptide to an at position 144–149 circular permutated version of GFP, transforming calcium binding of the calmodulin domain into modulation of fluorescence emission. This calcium sensor, named G-CaMP, was the basis for a whole family of single fluorophore GECIs (GCaMP2 (Tallini et al. 2006), GCaMP3 (Tian et al. 2009), GCaMP5 (Akerboom et al. 2012), G-CaMP6/7/8 (Ohkura et al. 2012), GCaMP6s/m/f (Chen et al. 2013)). Similar

approaches were employed to produce single fluorophore biosensors for cGMP (Nausch et al. 2008).

Another strategy for single emission biosensors was demonstrated by Pettikiriarachchi et al. (2012). Here the authors showed that the fluorescence intensity of a fluorescent protein can be modulated by the presence of a chromoprotein, which acts as a dark acceptor in FRET (see 3.3.4). This strategy will be further explored in this work.

The advantage of single emission fluorescent biosensors is the simplicity of the imaging procedure, since only a single fluorescent protein has to be monitored (Miyawaki and Niino 2015), which also allows to collect more light from the whole emission spectrum. Furthermore, they are more readily combined with either optogenetic actuators or other fluorescent biosensors, because of their smaller spectral bandwidth. On the other side, there are several disadvantages of this type of sensor (Rose et al. 2014): the intensity of emission is not only modulated by the presence of an analyte, but also depends on concentration of the fluorescent probe. This complicates the interpretation of results, when biosensor expression levels vary between cells, or when cells move in and out of the focal plane. Furthermore, sensors which are dark in the basal state are difficult to locate in the tissue, which makes long term experiments with several revisits of the same region more challenging.

1.2.1.2 Dual Emission Fluorescent Biosensors

In general, there are two types of dual emission fluorescent biosensors: (1) single emission biosensors, which are fused to a reference fluorescent protein, (2) FRET based biosensors with two fluorescent proteins. The focus of this work is on unimolecular FRET based biosensors, i.e. sensors which encode two fluorescent proteins within the same polypeptide chain.

In FRET based biosensors the emission of two fluorescent proteins is modulated through the binding of a ligand to a sensor domain. This binding event causes a conformational change within the molecule, rearranging the orientation and distance of the fluorophores (Miyawaki et al. 1997). This in turn has an impact on the FRET efficiency, as the distance and orientation of the transition moments within the chromophores is a major factor in FRET efficiency (see 1.1.2). FRET efficiency can be estimated by exciting the donor fluorophore and monitoring donor and acceptor emission. The ratio between these two intensities then allows inference of ligand concentration.

The main advantage of these types of sensors is that measurements are ratiometric and therefore, in principle, independent of sensor concentration (Tantama et al. 2012). This means that variations in sensor concentration which are caused by differences in expression levels over time or in between cells can be accounted and adjusted for. Furthermore, artefacts introduced by movement

of specimen or cells, shifting in and out of the focal plane and therefore changing the brightness of fluorophores, can also be corrected, as only the ratio of emission is related to ligand concentration, but not the intensity itself (Mues et al. 2013). Furthermore, FRET sensors are always visible when recombinantly expressed, because at least one fluorescent protein will emit, making long term experiments with multiple revisits of the same cells easier, as cells are always sufficiently labeled (Rose et al. 2014).

This comes at the cost of more complex imaging set-ups, as two fluorescent proteins have to be monitored simultaneously (Tantama et al. 2012). Furthermore, because of their larger spectral footprint, as they are consisting of two fluorescent proteins, they typically occupy a larger part of the visible spectrum than single emission sensors, making it more difficult to accommodate other optogenetic tools.

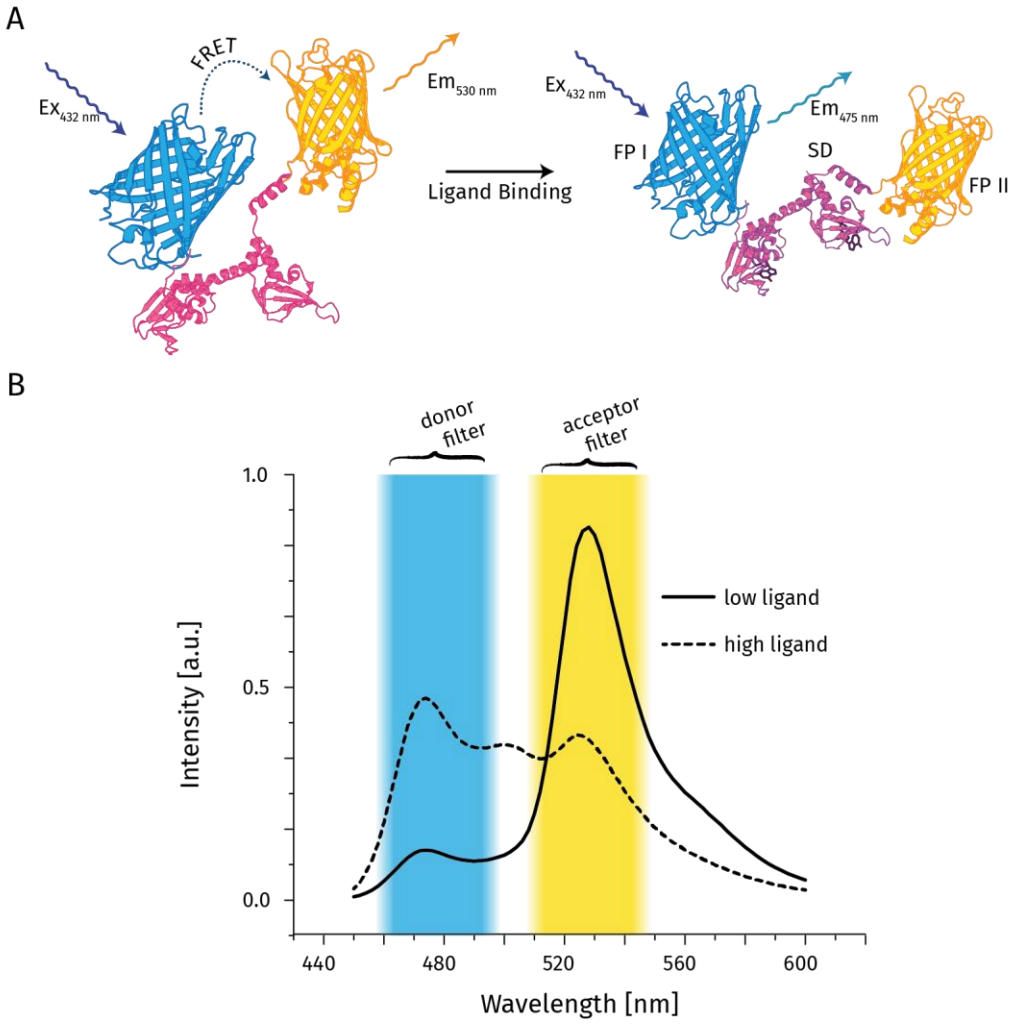

Figure 1.4 Model of FRET-based biosensor

A) Two fluorescent proteins (FP I and FP II) are fused via a sensor domain (SD). In the unbound state the donor FP transfers energy to the acceptor FP. Upon ligand binding a conformational change within the molecule occurs, rearranging the chromophores and reducing FRET efficiency. B) Emission spectra of FRET based biosensors in ligand-bound and ligand-free state with indicated imaging windows for fluorescence microscopy

1.3 cGMP

1.3.1 cGMP, a Cellular Secondary Messenger

Cyclic guanosine–3',5'–monophosphate (cGMP) is a secondary messenger in cellular signaling. It is formed through the activity of guanylate cyclases, catalyzing the conversion of the nucleotide GTP into a cyclic phosphodiester and releasing inorganic pyrophosphate (PP_i) (Figure 1.5 A). PP_i in turn is degraded into two phosphates through the activity of pyrophosphatase, driving the cyclization of cGMP by removing one of the products, rendering this reaction irreversible (Yang and Wensel 1992).

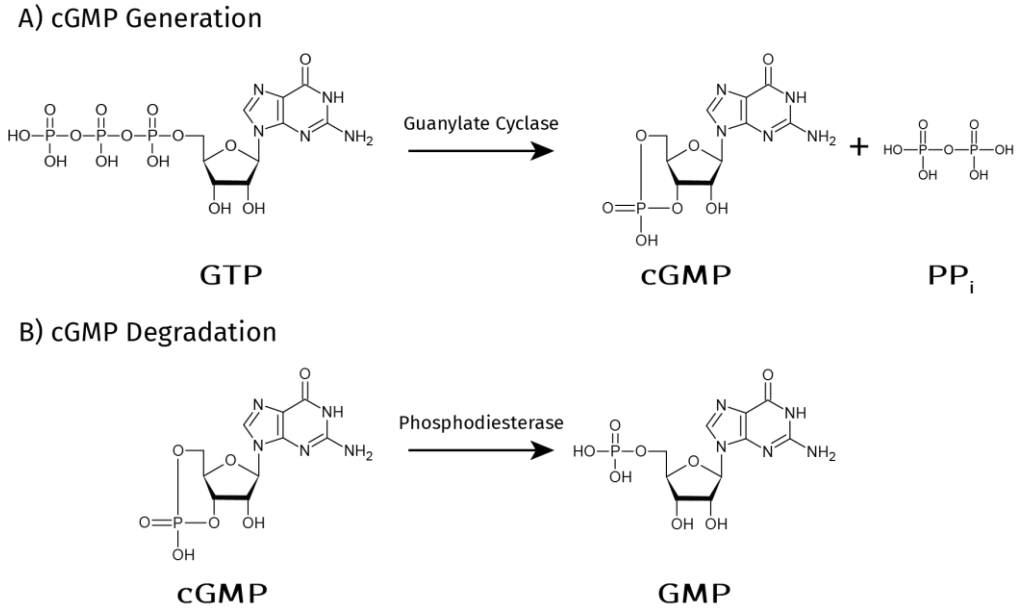

Figure 1.5 Enzyme reactions involved in cGMP generation and degradation

There are two major classes of guanylate cyclases: membrane-bound particulate guanylate cyclases (pCG) and soluble guanylate cyclases (sGC). They represent two distinct signaling pathways, through which cGMP is generated.

pCGs are homo-dimeric transmembrane receptors, which are typically activated by extracellular ligands (Kuhn 2009). These ligands are peptide hormones called natriuretic peptides. They bind to the extracellular receptor domain and cause an intramolecular rotation, which is transduced into the cytoplasm, leading to the synthesis of cGMP in the cyclase domain (Misono et al. 2005). There are seven pCG receptors in mammals (CG-A to CG-F) and five natriuretic peptides have been identified as their ligands (ANP, BNP, CNP, guanylin, uroguanylin) so far (Kuhn 2009).

sGC are hetero-dimeric receptors in the cytosol, consisting of an α and β subunit. They are activated by nitric oxide (NO), which is generated by nitric oxide synthases from L-arginine and can freely diffuse through membranes, enabling

cell-signaling (Derbyshire and Marletta 2009). Nitric oxide synthases and soluble guanylate cyclases form the NO/cGMP signaling pathway.

The downstream targets of cGMP signaling are cGMP dependent protein kinases (cGK), cGMP gated ion channels (CNG) and, to an extent, phosphodiesterases (PDE). In mammals, there are three isoforms of cGKs (cGKIα, cGKIβ and cGKII).These are serine/threonine kinases, which phosphorylate their respective targets, typically activating them in the process (Hofmann et al. 2009, Schlossmann and Desch 2009). CNGs, on the other hand, are ion channels, which are modulated by the binding of cGMP, subsequently changing the membrane potential or increasing the intracellular Ca^{2+} concentration (Biel and Michalakis 2009).

Phosphodiesterases degrade cyclic nucleotides into their monophosphate counterparts and therefore regulate the cGMP signal (Figure 1.5 B). There are 21 PDE genes in the human genome, grouped into 11 families (Conti and Beavo 2007). Besides degrading cyclic nucleotides, some PDEs are also allosterically regulated by them. These allosterically regulated PDEs are often bi-specific for cAMP and cGMP, allowing cross-talk between the two signaling pathways (Kleppisch 2009).

1.3.2 Physiological Functions of cGMP Signaling

cGMP signaling plays an important role in maintaining cardiovascular homeostasis. In particular, the action of the neurotic peptide (NP)/NO/cGMP pathway counterbalances the renin/angiotensin/aldosterone system (RAA) in the regulation of blood pressure. Under high cardiac load arterial natriuretic peptide (ANP) and B-type natriuretic peptide (BNP) are secreted into the bloodstream by the heart (Boerrigter et al. 2009). ANP and BNP then bind to the receptor GC-A and stimulate cGMP synthesis (Kuhn 2003). There are several hypotensive effects of cGMP stimulation: smooth muscle relaxation leads to vascular dilation; blood volume is reduced through increased renal blood flow and increased excretion of sodium and water (Kuhn 2009). By virtue of these effects the cGMP signaling pathway is pharmacologically addressed to treat cardiac diseases and hypertension. This was traditionally done through NO donors (stimulating sGC) like nitroglycerin (Brunton 1867) and more recently through artificial natriuretic peptides and phosphodiesterase inhibitors, all of which elevate intracellular cGMP concentration (Boerrigter et al. 2009). The introduction of the PDE5 inhibitor sildenafil (Viagra®) furthermore revealed the role of the cGMP signaling pathway in erectile dysfunction (Sandner et al. 2009). Besides this, cGMP signaling is involved in processes as diverse as bone growth and remodeling (Tsuji and Kunieda 2005), platelet aggregation (Walter and Gambaryan 2009) and sperm motility (Miraglia et al. 2011), demonstrating its ubiquitous importance.

Furthermore, cGMP plays a key role as a second messenger in the central nervous system and is therefore of great interest as a target molecule for optogenetic investigation. Nitric oxide (NO)/cGMP signaling has been implicated in synaptic long term potentiation (LTP) in the hippocampus and the amygdala and in synaptic long term depression (LTD) in the cerebellum (for review see Kleppisch and Feil (2009)).

In neurons in the hippocampus and amygdala N-methyl-D-aspartate (NMDA) receptor mediated calcium influx activates the Ca^{2+}/calmodulin-regulated neuronal NO synthase (nNOS), which in turn generates NO (see Figure 1.6 A). The gaseous NO can freely diffuse through membranes and permeates retrograde into the presynapse. There it stimulates soluble guanylate cyclases (sGC), which synthesize cGMP. The downstream targets for this cGMP signal are cGMP dependent protein kinases (cGK), cyclic nucleotide gated ion channels (CNG) and hyperpolarization-activated cyclic nucleotide gated ion channels (HCN). In the hippocampus, it is hypothesized that cGKI localizes to the presynaptic membrane and then phosphorylates various target proteins. This modulates vesicle clustering and eventually leads to an increased glutamate release. The exact mechanism for this process and the resulting LTP is still not clear and subject for further studies. CGKI knock outs in mice showed reduced LTP in the amygdala, which lead to impaired long-term fear memory in behavioral experiments,

linking this form of synaptic plasticity to learning and memory formation (Paul et al. 2007).

Conversely, in synapses between parallel fibers and purkinje neurons in the cerebellum synaptic plasticity is evoked through long term depression (LTD)(see Figure 1.6 B). This is process is mediated, in contrast to the hippocampus and the amygdala, through an anterograde NO signal. Nitric oxide is generated in the presynapse, crosses the synaptic cleft where it stimulates cGMP synthesis by sGCs. LTD is believed to be caused by cGKI which may trigger increased AMPA receptor internalization. LTD in the cerebellum was suggested to play a role in motor learning (Ito 2001).

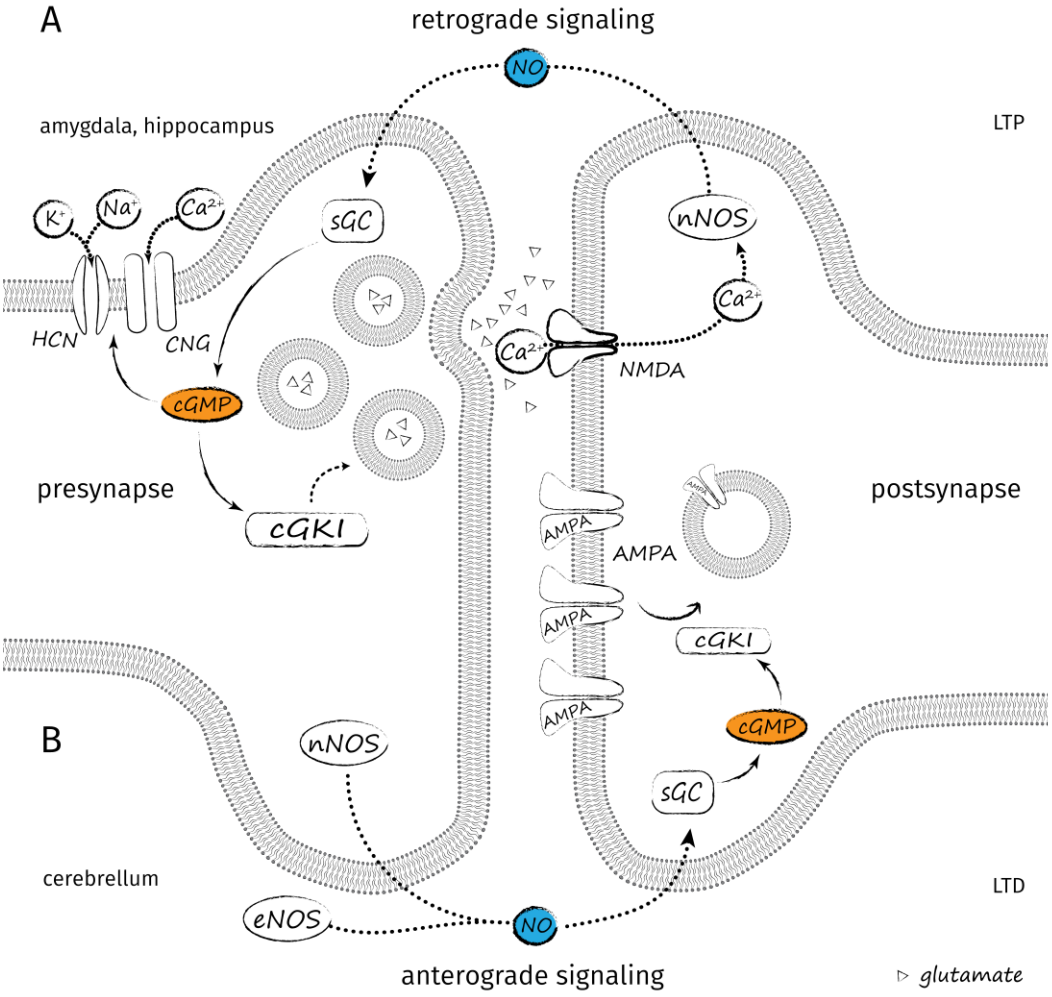

Figure 1.6 Current models of NO/cGMP signaling in synaptic plasticity

A) LTP through retrograde NO signaling in the amygdala and hippocampus B) LTD through anterograde NO signaling in the cerebellum

1.3.3 Optogenetic Tools for cGMP

The standard methods used for studying cGMP currently, rely on biochemical measurements of cGMP concentrations in either crude or purified cell extract. These methods utilize competitive binding of cGMP and a labeled cGMP analogue to a cGMP binding site (RIA, ELISA). Schmidt (2009) provides an excellent overview of all established assay types. While these assays are very sensitive to even femtomolar concentrations of cGMP, they all suffer from the same drawback: they all require cell extract as starting material, which leads to poor special and temporal resolution of the acquired measurements, since the tissue that is being observed is destroyed in the process (Nikolaev and Lohse 2009). cGMP signaling on the other hand is, as described above, very complex and dynamic and has been shown to be compartmentalized inside the cell (Fischmeister et al. 2006). Therefore, more sophisticated methods, that allow *in vivo* monitoring of cGMP oscillations in living tissue or even animals, have to be developed.

1.3.3.1 Genetically encoded Biosensors for cGMP

Several groups have developed genetically encoded biosensors (reviewed by Gorshkov and Zhang (2014)), to monitor cGMP concentrations *in vivo*. Table 1-2 contains a list of the best cGMP sensors to date and their properties. Most of the current sensors are FRET-based and utilize cGMP binding domains from either cGKI or PDE5A.

The first three in that collection (CGY-Del1, Cygnet 2.1 and cGES-DE5) can be seen as the first generation of cGMP sensors; which were developed independent of each other (Sato et al. 2000, Honda et al. 2001) or with a different binding domain (Nikolaev et al. 2006). This first generation, while an important step in the development, also exemplifies several limitations for their use for live cell imaging. CGY-Del1 for example displays poor selectivity between cGMP and cAMP. Cygnet 2.1 on the other hand was reported to exhibit very slow kinetics (Nikolaev et al. 2006). Both CGY-Del1 and Cygnet 2.1 use a large part of cGKIα, with only minor truncations of the N-terminus, resulting in a very large fusion protein. Since the catalytic domain of cGKIα is also retained, there might be potential for unwanted cross-reactivity of the sensor itself, causing artefacts in the observed specimen through phosphorylation. Furthermore, all three of these sensors have a very limited dynamic range, which impedes their usefulness in *in vivo* applications. Adding to that, the fluorescent proteins used in these sensors are very dim in comparison to more modern variants. Especially ECFP, with low quantum yield (0.4) and extinction coefficient (32.5 $mM^{-1}cm^{-1}$), is not only very dim but also limits the Förster radius and therefore potentially the dynamic range.

Table 1-2 cGMP sensors and their properties

Name	Type	Binding Domain	ΔR/R	EC_{50} cGMP	EC_{50} cAMP	selectivity	FPs	Reference
ratio metric sensors								
CGY-Del1	FRET	cGKIαΔ1-47	~10 %	20.0 nM	152.3 nM	7.6	ECFP / YFP 10C Q69K	(Sato et al. 2000)
Cygnet 2.1	FRET	cGKIαΔ1-77	~30 %	1.7 µM	>1 mM	>600	ECFP / Citrine	(Honda et al. 2001)
cGES-DE5	FRET	PDE5A GAF-A	~16 %	1.5 µM	630 µM	420	ECFP / EYFP	(Nikolaev et al. 2006)
Red cGES-DE5	FRET	PDE5A GAF-A	~20 %	40 nM	-	-	T-sapphireΔC11 dimer2 (DsRed)	(Niino et al. 2009)
cGi500	FRET	cGKI CNB-A CNB-B	77 %	470 nM	-	-	CFP W1B / YFP 10C	(Russwurm et al. 2007)
GFP^2-GAFa-Rluc	BRET	PDE5A GAF-A	~30 %	30 nM	-	-	Rluc / GFP^2	(Biswas et al. 2008)
Single fluorophore sensors			ΔF/F					
Cygnus	FRET	PDE5A GAF-A	~9 %	1 µM	400 µM	400	mTagBFP sREACh	(Niino et al. 2010)
δ-FlincG	Single FP	cGKIαΔ1-77	175 %	170 nM	48 µM	280	$cpEGFP^{145}$	(Nausch et al. 2008)

Red cGES-DE5, cGi500 and to an extent GFP^2-GAFa-Rluc and Cygnus are the second generation of cGMP sensors, which were engineered to overcome some of the described problems of the first generation. Specifically Russwurm et al. (2007) demonstrated a very methodical approach in sensor engineering, reducing the sensor domain to just CNB-A and CNB-B of cGKIα, while retaining good affinity and fair dynamic range. Red cGES-DE5 on the other hand was optimized to work in unison with a second FRET-indicator (Niino et al. 2009), shifting its spectral footprint dramatically to accommodate a CFP/YFP pair. This engineering effort also improved affinity and to a lesser extent the dynamic range. Still the problem of dim fluorescent proteins persists for both sensors. T-sapphire is, in terms of brightness (ε=44.0 $mM^{-1}cm^{-1}$, Φ=0.6, $\varepsilon * \Phi$=26.4), a slight improvement over ECFP (ε=32.5 $mM^{-1}cm^{-1}$, Φ=0.4, $\varepsilon * \Phi$=13), but it can only be effectively excited in the near UV, which limits its applicability. Dimer2 on the other hand is dimmer (ε=60 $mM^{-1}cm^{-1}$, Φ=0.69, $\varepsilon * \Phi$=41.4) than all of the YFP variants used in the first generation and is also, as the name suggests, a dimer (Campbell et al. 2002), which could potentially lead to aggregation and other artefacts.

GFP^2-GAFa-Rluc is an interesting twist on PDE5A based FRET sensors. Here the authors (Biswas et al. 2008) placed the GAF-A domain from PDE5A between an engineered version of GFP and the *Renilla* luciferase (Rluc), using a bioluminescent protein as the donor for resonance energy transfer (BRET). BRET has some advantages over FRET (Xu et al. 2003): since the sensor is the source of the light, no background florescence is induced through excitation light, reducing noise in the sample considerably. Furthermore bleaching of the chromophore and phototoxicity through intense excitation can be avoided. The downsides of BRET sensors on the other hand are, that bioluminescence is very dim and therefore requires very sensitive instruments and long exposure times. Besides that, luciferase requires a substrate (coelenterazine), which has to be externally supplied.(Xu et al. 2003) This puts conceptual limitations to BRET imaging, especially for *in vivo* applications.

Cygnus is an engineered version of cGES-DE5 (Niino et al. 2010), which replaces the CFP/YFP pair of the original with mTagBFP and the dark acceptor sREACh, creating a FRET based single fluorophore sensor, as demonstrated by Murakoshi et al. (2008) before. While this sensor demonstrates the feasibility of this approach for cGMP, the sensor itself has a very poor dynamic range and is thus only of limited use.

δ-FlincG is a single fluorophore cGMP sensor (Nausch et al. 2008), based on the first GCaMP (Nakai et al. 2001). While the maximal intensity change reported by Nausch et al. (2008) compares very well to other cGMP sensors, Bhargava et al. (2013) casted serious doubt on the accuracy of these findings. According to Bhargava et al. (2013) the sensors of the FlincG family were very pH-sensitive, a feature inherited from the first generation GCamP sensors, thus causing artefacts in their initial characterization. Furthermore, the group reported problems with the genetic stability of the published sensor construct and uncertainty about its identity. Besides that, it is not unreasonable to assume that δ-FlincG inherited other problems from GCaMP1, namely its dimness and fast bleaching rate (Mank and Griesbeck 2008).

Of the described sensors only cGi500 and Red cGES-DE5 found wider adaption so far, with transgenic mice being generated by Thunemann et al. (2013) and Götz et al. (2014).

Taken together cGMP imaging is still in its infancy and many problems with the pioneer sensors have to be overcome before they compare to their highly engineered counterparts used in Ca^{2+} imaging, like Twitch2B (Thestrup et al. 2014) and GCaMP6s (Chen et al. 2013). In this work, cGi500 was chosen as a base framework for further development, because of its relatively high dynamic range and large knowledgebase provided by Russwurm et al. (2007).

1.3.3.2 Genetically encoded Stimulators of cGMP Generation

More recently, the optogenetic toolbox for studying cGMP was enriched by the development of genetically encoded, light activated guanylate cyclases. Specifically, Ryu et al. (2010) mutated the light activated adenylate cyclase BlaC from the filamentous bacterium *Beggiatoa* sp., to use GTP as educt instead of ATP, yielding a light activated guanylate cyclase, denoted as BlgC. Kim et al. (2015) improved on this initial work, leading to the BlgC variant EROS, which was successfully used to treat erectile dysfunction in rats through optogenetic stimulation.

Furthermore a study by Gao et al. (2015) isolated a membrane bound light activated guanylate cyclase from the aquatic fungus *Blastocladiella emersonii* and established it as an optogenetic tool (BeCyclOp). The authors of the study demonstrated major improvements over EROS, with a 50 x higher production of cGMP and significantly less cAMP production after light stimulation.

With the availability of optogenetic activators and reporters for the cGMP pathway, the foundation has been laid for studying complex neurobiological questions through optical means, once these tools ripen to maturity.

1.4 Aims

The overarching aim of this study is to develop genetically encoded fluorescent biosensors for cGMP with improved characteristics, to make them suitable for a variety of *in vivo* imaging applications. The outlined key issues of first and second generation cGMP sensors will be addressed, with a focus on brighter fluorophores, larger dynamic range and higher cGMP selectivity. The base constructs for this process will be cGi500, the cGMP sensor with the best overall characteristics to date. Building on the insights gained from a host of recent structural studies, cGMP binding domains will be optimized for simpler binding kinetics and higher selectivity. Lessons learned from the development of genetically encoded calcium sensors will be generalized and transferred to cGMP sensors. Specifically, an important objective is the transfer of the development procedure into *E. coli*, to simplify the process in order to be able to implement changes more quickly and to eventually employ directed evolution approaches such as bacterial plate screening (Litzlbauer et al. 2015). In addition, cGMP sensors with red shifted spectral characteristics will be explored, in order to harness more favorable imaging conditions within the red part of the visible spectrum. In connection to that, it is the aim to assess the viability of a single emission cGMP biosensor based on a red fluorescent protein which is quenched by a chromoprotein through FRET.

In pursuit of the main goal, a secondary aim is to streamline the general biosensor development pipeline, to make this process quicker and more adaptable. A further objective in this spirit is the construction of an automated screening system for high throughput directed evolution approaches.

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals and Reagents

Table 2-1 Chemicals

Name	Vendor	Abbreviation	Catalogue #
2-Propanol (Isopropanol)	Sigma Aldrich		33539
Acetic acid	Sigma Aldrich		33209
Adenosine 3′,5′-cyclic monophosphate sodium salt monohydrate	Sigma Aldrich	cAMP	A6885
Agar-Agar	Carl Roth		2266.2
Agarose, high electroendosmosis	Biomol		1280.1
Ampicillin sodium salt	Carl Roth	Amp	K029.5
D(+) Saccharose	Carl Roth		4621.1
Dimethylsufloxid	Sigma Aldrich	DMSO	D8418
di-Sodium hydrogen phosphate	Merck	Na_2HPO_4	106580
Ethanol	Sigma Aldrich	EtOH	32205
Ethidium bromide solution 1 %	Carl Roth	EtBr	2218.2
Ethylenediaminetetraacetic acid	Merck	EDTA	108418
Glycerol 86 %	Carl Roth		4043.1
Guanosine 3′,5′-cyclic monophosphate sodium salt	Sigma Aldrich	cGMP	G6129
Imidazole	Merck		104716
Isopropyl β-D-1-thiogalactopyranoside	Carl Roth	IPTG	2316.4
Kanamycin sulphate	Carl Roth	Kan	T832.1
L-(+)-arabinose	Sigma Aldrich		A91906
Leupeptin	Sigma Aldrich	Leu	L0649
Magnesium chloride	Merck	$MgCl_2$	105833
Orange G (C.I. 16230)	Carl Roth		318.1
Pepstatin A	Sigma Aldrich	Pep A	P5318
Phenylmethanesulfonyl fluoride	Sigma Aldrich	PMSF	P7626
Polyethylene glycol 3000	Sigma Aldrich	PEG	81227
Polypropylene glycol P2000	Sigma Aldrich	PPG	81380
Potassium dihydrogen phosphate	Carl Roth	KH_2PO_4	P018.2
Sodium chloride	Merck	NaCl	106404
Streptomycin sulfate salt	Sigma Aldrich	Strep	S6501
Tetracycline hydrochloride	Carl Roth	Tet	237.1
Tris(hydroxymethyl)aminomethane	Carl Roth	Tris	2449.1
Tryptone	Carl Roth		8952.2
Yeast Extract	Carl Roth		2363.2

Table 2-2 Primer list

#	Name	Sequence 5’ → 3’	Mod.
1	T7	TAATACGACTCACTATAGG	
2	pRSETB_MCS_fwd	CGGGATCTGTACGACGATGACGATAAGGATCCG	
3	pRSETB_PPY_fwd	CGGGCTTTGTTAGCAGCCGGATCATTTAGGTGACACTATAGGAATTC	5’PO_4
4	pRSETB_PPY_rev	CTATTATCAGATATCCATGGTGGCCGGATCCTTATCGTCATCGTCGTAC	5’PO_4
5	pcDNA3_PPY_fwd	CGGATCCTTATCGTCATCGTCGTACAGATCCCCTATAGTGAGTCGTATTAATTTCGATAAG	5’PO_4
6	pcDNA3_PPY_rev	GCCACCATGGATATCTGATAATAGGAATTCCTATAGTGTCACCTAAATGCTAGAGCTCGC	5’PO_4
7	pSin_sacB_fwd	TAGTACATTTCATCTGACTAAGATAAGGATCCGGCCACCATGATTTAAATAAACTGCAGTTGACAACATAAAAAC	
8	pSin_sacB_rev	GTTTTGCTGGTCGGATCATTGCTATAGGAATTCCTATTATCAATTTAAATTTATTTGTTAACTGTTAATTG	
9	pSL_Short_fwd	GATAAGGATCCGGCCACCATG	
10	pSL_short_rev	CTATAGGAATTCCTATTATCA	
11	sacB_fwd_2	AGGATCCGGCCACCATGGATATCAAACTGCAGTTGACAACATAAAAACT	
12	sacB_rev_2	AGGAATTCCTATTATCAGATATCTTATTTGTTAACTGTTAATTGTCCTTG	
13	5CAT_fwd	ATATACATATGCGGGGTTCTCATCATCATCATCATCACCATCACCATCACCGCGGTATGGCTAGCATGACTGG	
14	preHis_rev	AGAACCCCGCATATGTATATC	
15	mKO2_SL_fwd	GACGATAAGGATCCGGCCACCATGGTGAGTGTGATTAAACCA	
16	mKO2_SL_rev	CTATAGGAATTCCTATTATCAGCTATGAGCTACTGCATCTTC	
17	GFP_SL_fwd	GACGATAAGGATCCGGCCACCATGGTGAGCAAGGGCGAGGAG	
18	GFP_trunk_rev	GGCGGCGGTCACGAACTCCAGCAGGAC	
19	GFP_fwd	ATGGTGAGCAAGGGCGAGGAGCTGTTC	
20	GFP_SL_rev	CTATAGGAATTCCTATTATCACTTGTACAGCTCGTCCATGCC	
21	GFPcGi500_rev	GGTAAATTTGCGAAATGCCTGCTTGTACAGCTCGTCCATGCC	
22	GFPTRUcGi500_rev	GGTAAATTTGCGAAATGCCTGGGCGGCGGTCACGAACTCCAGCAGGAC	
23	GFPcGi500_fwd	GCAGAAGCCAAAGCCAAATATATGGTGAGCAAGGGCGAGGAG	
24	Cloverd11_rev	GGCGGCGGTCACGAACTCCAG	
25	mCitrine_SL_rev	CTATAGGAATTCCTATTATCACTTGTATAGCTCGTCCATGCC	
26	YpetcGi500_fwd	GCAGAAGCCAAAGCCAAATATATGGTGAGCAAAGGCGAAGAG	
27	YPet_SL_rev	CTATAGGAATTCCTATTATCACTTATAGAGCTCGTTCATGCC	
28	cpVenuscGi500_f	GCAGAAGCCAAAGCCAAATATATGGGTGGGGTCCAGCTTGCAG	
29	cpCit174_SL_rev	CTATAGGAATTCCTATTATCAATCCTCAATGTTGTGACGGATC	
30	cpCitcGi500_fwd	GCAGAAGCCAAAGCCAAATATATGGGCGGCGTGCAGCTCGC	

Table 2-2 Primer list continued

#	Name	Sequence 5’ → 3’	Mod.
31	cpCit174_SL_rev	CTATAGGAATTCCTATTATCAGTCCTCGATGTTGTGGCGGATC	
32	mCitcGi500_rev	GGTAAATTTGCGAAATGCCTGCTTGTATAGCTCGTCCATGCC	
33	YPet_SL_fwd	GATAAGGATCCGGCCACCATGGTGAGCAAAGGCGAAGAGCTG	
34	YpetcGi500_rev	GGTAAATTTGCGAAATGCCTGCTTATAGAGCTCGTTCATGCC	
35	YpetTRUcGi500_r	GGTAAATTTGCGAAATGCCTGGGCGGCGGTCAGGAACTCCAGCAGCAC	
36	cpCit174_SL_fwd	GATAAGGATCCGGCCACCATGGGCGGCGTGCAGCTCGCCGAC	
37	cpCitcGi500_rev	GGTAAATTTGCGAAATGCCTGGTCCTCGATGTTGTGGCGGATC	
38	cpVenus174_SL_f	ATAAGGATCCGGCCACCATGGGTGGGGTCCAGCTTGCAGATC	
39	cpVenuscGi500_r	GGTAAATTTGCGAAATGCCTGATCCTCAATGTTGTGACGGATC	
40	cGi500_fwd	CAGGCATTTCGCAAATTTACCAAAAG	
41	cGi500_rev	ATATTTGGCTTTGGCTTCTGCATC	
42	N-CNB-B_fwd	ATGGAATTTCTGAAAAGCGTTCC	
43	N-CNB-A_rev	CAGGTTTTTCATAAAATCGTTATC	
44	doubleCNB-B_fwd	GATAACGATTTTATGAAAAACCTGCCGGAAGAAATTCTGAGCAAACTGG	
45	doubleCNB-B_rev	GGAACGCTTTTCAGAAATTCCATATATGCTTTATTGCTCACATC	
46	CNB-BII_Ypet_fwd	GAAAAAAGACATGCGAAGCGGATGGTGAGCAAAGGCGAAGAG	
47	mKO2_SL_fwd	GACGATAAGGATCCGGCCACCATGGTGAGTGTGATTAAACCA	
48	mKO3cGi500_rev	GGTAAATTTGCGAAATGCCTGGCTATGAGCTACTGCATCTTC	
49	v-cp18 f	GATAAGGATCCGGCCACCATGGGGCCAATGCCTTTCGCGTTT	
50	mKO3cp53cGi500_r	GGTAAATTTGCGAAATGCCTGCATTGGCCCGCCCTCGGCCATTG	
51	mKO2_SL_rev	CTATAGGAATTCCTATTATCAGCTATGAGCTACTGCATCTTC	
52	cGi500mKO3_fwd	GCAGAAGCCAAAGCCAAATATATGGTGAGTGTGATTAAACCA	
53	mKO3cp53cGi500_f	GCAGAAGCCAAAGCCAAATATATGGGGCCAATGCCTTTCGCG	
54	mKO r1A new e	CTATAGGAATTCCTATTATCACATTGGCCCGCCCTCGGCCAT	
55	mRuby2_SL_fwd	GATAAGGATCCGGCCACCATGGTGTCTAAGGGCGAAGAGCTG	
56	mRuby2cGi500_fwd	GCAGAAGCCAAAGCCAAATATATGGTGTCTAAGGGCGAAGAG	
57	ultramarine_SL_f	GATAAGGATCCGGCCACCATGTCTGTTATCGCCACCCAGATG	
58	UltracGi500_rev	GGTAAATTTGCGAAATGCCTGGGCCACCACAGGTTTGCGGGC	
59	CGi500ultra_fwd	GCAGAAGCCAAAGCCAAATATATGTCTGTTATCGCCACCCAG	
60	Ultra_SL_rev	TATAGGAATTCCTATTATCAGGCCACCACAGGTTTGCGGGCG	

2.1.2 Enzymes, Kits and Standards

Table 2-3 Enzymes, kits and standards

Name	Vendor	Catalogue #
1 kb DNA Ladder	New England Biolabs	N3232L
100 bp DNA Ladder	New England Biolabs	N3231L
Antarctic Phosphatase	New England Biolabs	M0289L
BamHI-HF®	New England Biolabs	R3136L
Deoxyribonuclease I (DNase I)	Sigma-Aldrich	D5025-150KU
DpnI	New England Biolabs	R0176L
EcoRI-HF®	New England Biolabs	R3101L
EcoRV-HF®	New England Biolabs	R3195L
Herculase II Fusion DNA Polymerase	Agilent	600679
Lysozyme from chicken egg white	Sigma-Aldrich	L6876
Nickel-IDA Agarose	Jena Bioscience	AC-310-500
NucleoSpin® Gel and PCR Clean-up	MACHEREY-NAGEL	740609.25
NucleoSpin® Plasmid EasyPure	MACHEREY-NAGEL	740727.25
PureYield™ Plasmid Midiprep System	Promega	A2495
Ribonuclease A (RNase A)	Sigma-Aldrich	R5503
SphI-HF®	New England Biolabs	R3182L
T4 DNA Ligase	New England Biolabs	M0202L
T4 DNA Ligase Reaction Buffer	New England Biolabs	B0202S

2.1.3 Devices and Appliances

Table 2-4 List of devices

Device type	Name	Vendor
CCD Camera	CoolSNAP-HQ	Visitron Systems GmbH
Centrifuge	RC-5B	Sorval
Centrifuge	3K30	SIGMA
Fluorescence spectrophotometer	Cary Eclipse	Varian
Gel documentation	GelDoc 200 Videosystem	BIO RAD
Micro-centrifuge	Centrifuge 5415 D	Eppendorf
Plate reader	Infinite® M200 PRO	Tecan
Power supply for electrophoresis	E143	Consort
Power supply for screening system	BT-305	BASETech
Stereo-fluorescence-microscope	Leica M205 FA	Leica Microsystems
Thermocycler	T3000	Analytic Jena AG
Thermocycler	FlexCycler2	Analytic Jena AG
Ultra-sonic water bath	ONOREX SUPER RK 510	Bandelin
UV-Vis spectrometer	NanoDrop 1000 spectrometer	Thermo Scientific

2.1.4 Strains and Plasmids

Table 2-5 Strains and plasmids

Name	Details	Source
E. coli XL1 blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacI*qZΔ*M15* Tn*10* (Tetr)]	Invitrogen
E. coli Bl21 (DE3) gold	*E. coli* B F– *ompT hsdS*(rB– mB–) *dcm*+ Tetr *gal* λ(DE3) *endA* Hte	Invitrogen
E. coli PPY	F⁻ *end*A1 *rec*A1 *gal*E15 *gal*K16 *nup*G *rps*LΔ*lac*X74 Φ*80lac*ZΔM15 *ara*D139Δ(*ara,leu*)7697 *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) *cyn*X::[*ara*C pBAD- *redα* EM7- *redβ* Tn5-*gam*]λ⁻	(Zhang et al. 2012)
Plasmids		
pcDNA3		Invitrogen
pRSET-B		Invitrogen
pSinRep5		Invitrogen

2.1.5 Buffers

Table 2-6 Buffer composition

Buffer	Component	Concentration	Notes
50 x TAE	Tris base	2 M	
	Acetic acid	1 M	
	EDTA	50 mM	
1 x TAE	Tris base	40 mM	
	Acetic acid	20 mM	
	EDTA	1 mM	
10 x Orange G DNA loading dye	EDTA	10 mM	
	Tris	100 mM	pH 7.5
	Glycerol	50 % (v/v)	
	Orange G	1 % (w/v)	
2xYT liquid medium	Tryptone	16 g/L	autoclave
	Yeast extract	10 g/L	
	NaCl	5 g/L (85.5 mM)	pH 7
LB (Lennox) liquid medium	Tryptone	10 g/L	autoclave
	Yeast extract	5 g/L	
	NaCl	5 g/L (85.5 mM)	pH 7
LB (Lennox) solid medium	Tryptone	10 g/L	autoclave
	Yeast extract	5 g/L	
	NaCl	5 g/L (85.5 mM)	
	Agar-Agar	10 g/L	pH 7
YTS solid medium	Yeast extract	5 g/L	autoclave
	Tryptone	10 g/L	
	D(+) Saccharose	100 g/L	
	Agar-Agar	10 g/L	pH 7
Autoinduction additive	D-(+)-glucose	1.25 % (w/v)	sterile filter
	Glycerol	15 % (v/v)	
	Lactose	5 % (w/v)	
Protein Resuspension Buffer I	Na_2HPO_4	20 mM	
	NaCl	300 mM	
	Imidazole	20 mM	pH 7.8 (HCl)
Protein Resuspension Buffer I	Na_2HPO_4	20 mM	
	NaCl	300 mM	
	Imidazole	55 mM	pH 7.8 (HCl)
Protein Elution Buffer	Na_2HPO_4	20 mM	
	NaCl	300 mM	
	Imidazole	250 mM	pH 7.8 (HCl)

Table 2-6 Buffer composition (continued)

Buffer	Component	Concentration	Notes
PBS	NaCl	137 mM	autoclave
	KCl	2.7 mM	
	Na_2HPO_4	10 mM	
	KH_2PO_4	2 mM	pH 7.4
TSS	Trypton	10 g/L	sterile filter
	Yeast Extract	5 g/L	
	NaCl	5 g/L (85.5 mM)	
	PEG	100 g/L	
	DMSO	5 % (v/v)	
	$MgCl_2$	50 mM	pH 6.5

2.2 Methods

2.2.1 Working with DNA

2.2.1.1 Polymerase Chain Reaction

"I was sagging as I walked out to my little silver Honda Civic, which never failed to start. Neither Fred, empty Becks bottles, nor the sweet smell of the dawn of the age of PCR could replace Jenny [girlfriend, who broke up with him]. I was lonesome."

Kary B. Mullis (1993)

Since the invention of the polymerase chain reaction (PCR) by Kary B. Mullis in 1983 it has become one of the most important and versatile tools in molecular biology and warranted him the Nobel Prize for Chemistry in 1993.

In brief, template DNA strands are cyclically heated and cooled in a thermocycler, melting and re-hybridizing them with complementary DNA. During this process two primers (short oligonucleotides) bind the start and the end of a desired DNA section and a thermostable polymerase elongates the primers, copying the template DNA. In each cycle both DNA strands are copied and can serve as templates in the next cycle, enabling a doubling of DNA strands with every cycle. This exponential reaction can relatively quickly produce large quantities of linear dsDNA starting from very few molecules of template DNA.

In this work PCR was used for several purposes: (1) to amplify DNA for construction of recombinant plasmids, (2) to introduce synthetic sequences into amplified DNA, (3) to fuse separate pieces of linear DNA, (4) to generate DNA libraries, and (5) to verify the length of resulting products from cloning.

To guarantee the integrity of DNA sequences during amplification a polymerase with proofreading capability was used. The enzyme of choice was Herculase II Fusion DNA Polymerase (Agilent), with an advertised error rate of 1 in 777000 bp and an elongation rate of 30 s/kbp. Reaction conditions were chosen according to the guidelines provided by the vendor. A typical PCR mix, with a reaction volume of 50 µL, is shown in Table 2-7. Components were pipetted in depicted order into a 0.2 mL reaction tube and carefully mixed. Primers were HPSF purified, custom synthesized oligonucleotides from Eurofins Genomics. Lyophilized oligonucleotides were resolubilized in autoclaved dH_2O to a stock concentration of 100 µM. Stocks were diluted 10 fold to obtain a 10 µM working solution. Both stock and working solutions were stored at – 20 °C. In case of problems with amplification of GC-rich targets, the PCR mix was augmented with up to 8 % (v/v) DMSO.

Table 2-7 Composition of standard PCR with Herculase II Fusion Polymerase

Compound	Concentration	Volume	Final Concentration
dH_2O, autoclaved		35 µL	
5 x Herculase II reaction buffer	5 x	10 µL	1 x
Primer 1	10 µM	1 µL	0.2 µM
Primer 2	10 µM	1 µL	0.2 µM
dNTP mix	25 mM (each dNTP)	1 µL	500 µM (each dNTP)
Template DNA	1–400 ng/µL	1 µL	0.02–8 ng/µL
Herculase II Fusion DNA Polymerase		1 µL	
	Finale Volume:	50 µL	

PCRs were performed in the thermocyclers T3000 or FlexCycler² from Analytic Jena AG. A standard PCR temperature program is depicted in Table 2-8. Two parameters of this basic program were adjusted to the specific requirements of each individual reaction: (1) elongation time, according to the length of the expected product, and (2) annealing temperature, according to the specific primer set in use.

Table 2-8 Temperature program for PCRs with Herculase II Fusion Polymerase

Step	Temperature [°C]	Time [s]	Cycles
Denaturation	98	120	Initial
Denaturation	98	20	30x
Annealing	65	20	
Elongation	72	30 per 1 kbp of product	
Elongation	72	180	Final

Annealing temperatures for the use of Herculase II Fusion polymerase can be calculated, according to the vendor, as the melting temperature of the primers –5 °C. Because most often complex primers with long non-complementary 5'-overhangs were used, the utility of these calculations is limited. An annealing temperature of 65°C was found to work for most primer sets. Annealing temperatures were optimized using a thermocycler with gradient option. These cyclers enable PCRs with different annealing temperatures, but otherwise identical parameters, to be performed in parallel within the same machine. In general annealing temperatures were either: (1) lowered, if little or no PCR product was observed, to facilitate better binding of the primers to the template, or (2) increased to reduce unspecific primer binding, if additional, unwanted PCR products were observed.

Introduction of short synthetic DNA Sequences

DNA polymerases are used in PCR to catalyze the extension of the template-bound primer at the 3'-end of the oligonucleotide. Depending on the sequence, 18–21 nucleotides are sufficient to specifically bind the template DNA within the operating temperature range of polymerases. Since only the 3'-end has to bind to the template DNA to achieve successful elongation, the 5'-end of a primer can be non-complementary to the template strand. This property can be used to introduce synthetic nucleotide sequences at the ends of PCR amplified linear DNA. As long as there are a sufficient number of complementary base pairs at the 3'-end of the primer, the 5'-"overhang" can be relatively long and is mainly limited by the synthesis length a supplier is able to deliver (~100 bp). In this work, this property was used to introduce a variety of different synthetic sequences to amplified DNA: (1) recognition sites for restriction endonucleases to facilitate classical cloning, (2) homology regions to facilitate SLiCE cloning, (3) functional genetic elements (e.g. promoters, RBS, Kozak-consensus sequence), (5) randomized sequences to generate DNA libraries and (6) short purification tags (e.g. His-tag) to enable purification of recombinant proteins.

Overlap extension PCR

Overlap extension PCR (OE-PCR) was first introduced by Higuchi et al. (1988), mainly as a method for site directed mutagenesis. In the context of this work it was used to fuse two smaller linear DNA fragments sequence specific (e.g. to generate fusion proteins).

The fusion was achieved by a two-step process: First the two fragments were amplified separately with standard PCR, as described above. During these PCRs homologue regions were introduced at the end of the first fragment and at the start of the second fragment through the use of primers with 5' overhangs. Typically, primers consisted of 21 base pairs for primer binding and 21 base pairs overhang. This generated two DNA fragments, which had a common overlap sequence of 42 base pairs. These two fragments were then purified as described in 2.2.1.3 and used as template for the second step. The fragments were equimolar combined into a reaction as shown in Table 2-9 and a PCR was performed using the program in Table 2-8. The two primers, which were used in this reaction, are complementary to the 5'- end of the first template and the 3'- end of the second template. These primers were functionalized through 5'- overhangs for subsequent cloning steps. In the initial cycles of the PCR two different hybridization products were formed. Of these two possible products, only one has the necessary 3'- OH groups, which are required for successful extension with DNA polymerases. These extended strands were amplified in subsequent PCR cycles, as they incorporated complementary regions for both primers. In practice it was found, that OE-PCRs often show 3 distinct bands: the high running fusion product and the lower running templates. Therefore OE-PCR products were always gel purified as described in 2.2.1.3, by cutting the desired band out of the agarose gel. Resulting purified products were then inserted into vector DNA by means of classical (see 2.2.1.6) or SliCE cloning (see 2.2.1.7).

Table 2-9 Composition of OE-PCR with Herculase II Fusion Polymerase

Compound	Concentration	Volume	Final Concentration
ddH_2O, autoclaved		34 µL	
5 x Herculase II reaction buffer	5 x	10 µL	1 x
Primer 1	10 µM	1 µL	0.2 µM
Primer 2	10 µM	1 µL	0.2 µM
dNTP mix	25 mM (each dNTP)	1 µL	500 µM (each dNTP)
Template 1	10 – 100 ng/µL	1 µL	0.2 – 2 ng/µL
Template 2	10 – 100 ng/µL	1 µL	0.2 – 2 ng/µL
Herculase II Fusion DNA Polymerase		1 µL	
	Finale Volume:	50 µL	

Colony PCR

Colony PCR is a fast and cost effective high throughput method, to assess whether recombinant DNA fragments were successfully integrated into target vectors. If target vectors are kept constant, as it was often the case in this work, parallelization of the cloning process is easily achieved. This allowed the verification of several dozen prototype biosensors at once.

In brief, potentially positive *E. coli* clones were tested for successful integration of the transgenic fragment by amplification of the integration site, using primers that are complementary to regions on the plasmid. Material from the colony, instead of purified plasmid DNA, was used as a template. This is what sets colony PCR apart from other means of verification (e.g. restriction digest of purified plasmid). As it can be performed with colonies harvested directly from agar selection plates after the cloning step. There is no requisite for another overnight growth cycle to produce plasmid DNA in a small liquid culture. Even though purified plasmid DNA is required eventually, clones can be tested right away and only positive candidates are subsequently inoculated for plasmid purification, thus cutting down costs for plasmid preparation kits.

The detail protocol was as follows: Primers for the colony PCR were designed, that two possible products can be observed, a short product, if there was no integration and a long product if integration was successful. Typically, priming sites were chosen ~150 bp up- and downstream from the targeted integration site, yielding a 300 bp product for empty vectors and an x + 300 bp product for positive clones (x = insert length).

After successful transformation of *E. coli* with a cloning mix, colonies were observed on selective agar plates after overnight incubation. Colonies were picked using sterile pipette tips and streaked on selective agar plates into a grid pattern that allowed later identification of individual clones. Right after streaking, the pipette tip was briefly transferred into 10 µL of PCR mix in 0.2 mL reaction tubes, carrying small amounts of bacteria over into the mix. The key factor in this process was, to minimize the amount of material used. Since bacterial colonies contain a huge variety of biological compounds, there are many potential PCR inhibiting factors (e.g. proteases) present. Because of this and the fact that PCR can amplify miniscule amounts DNA, counterintuitively, colony PCR worked best with very small colonies, since only small amounts of bacterial material were present to begin with. Selective plates with streaked-out colonies were incubated at 37° C.

Table 2-10 depicts a standard reaction mix for a colony PCR. The volume of the reaction is reduced to 10 µL in comparison to the standard PCR, because only a small amount of product is needed for subsequent analysis, further reducing the cost of this method. Several colonies were tested in parallel, warranting the production and subsequent distribution of a larger scale master mix. The temperature program used was the one shown in Table 2-8 with

appropriate annealing temperatures for the primer sets used and extension times for the products expected.

After the PCR program, products were analyzed by agarose gel electrophoresis as described in 2.2.1.2. Positive clones were then inoculated from the back-up selective agar plate into 2 mL selective liquid LB medium in 14 mL culture tubes, in the evening of the same day they were streaked out – even if growth on the plates was not yet visible. Back-up plates and culture tubes were incubated overnight at 37°C, the latter with vigorous agitation.

Lastly, colony PCR is only a rough verification method and was used as such, since it can only detect correct length (within the boundaries of gel electrophoresis resolution), but not sequence identity. Final verification was always achieved through sequencing.

Table 2-10 Composition of colony PCR mix with Herculase II Fusion Polymerase

Compound	Concentration	Volume	Final Concentration
ddH_2O, autoclaved		7.4 µL	
5 x Herculase II reaction buffer	5 x	2 µL	1 x
Primer 1	10 µM	0.2 µL	0.2 µM
Primer 2	10 µM	0.2 µL	0.2 µM
dNTP mix	25 mM (each dNTP)	0.1 µL	250 µM (each dNTP)
Herculase II Fusion DNA Polymerase		0.1 µL	
	Finale Volume:	10 µL	

2.2.1.2 Agarose Gel Electrophoresis

Agarose gel electrophoresis is the standard method to separate polynucleotides according to their length, which allows identification and isolation of specific polynucleotides after subsequent visualization of the separated specimen. Detailed description of this method can be found in Green and Sambrook (2012). In the context of this work agarose gel electrophoresis was used for the separation, analysis and isolation of DNA, linear and circular, between 100 bp and 15 kbp in length.

In the buffer conditions (1 x TAE, pH 8.0) used for gel electrophoresis DNA is negatively charged, due to the deprotonation of phosphate groups inside the DNA backbone. Placed into an electric field in a liquid environment, this negative charge accelerates the DNA strands and they move towards the anode. This movement is the basis of separation in gel electrophoresis, which is a type of chromatography. The mobile phase (DNA) interacts physically with a stationary phase (polymerized agarose gel), which leads to a separation of different types of molecules according to their physical properties. Specifically, polymerized agarose is a hydrocolloid gel, which forms pores, through which DNA strands move. Since larger DNA molecules present a larger surface, they interact stronger with the pores of the gel and therefor move slower than smaller DNA molecules. The electric mobility of a DNA molecule in such a gel depends linearly on the logarithm of its length. In time, different DNA molecules separate according to their size, with smaller DNA fragments being closer to the anode than larger ones.

Besides separation, DNA molecules also have to be visualized to allow any kind of analysis. To make DNA visible, ethidium bromide is incorporated into the gel. Ethidium bromide (EtBr) is an organic compound with a conjugated π- system, which acts a chromophore, absorbing in aqueous solutions strongly at 210 nm and 285 nm. EtBr intercalates with DNA, inserting itself between the DNA bases through van der Waals interactions. The planar molecules lie parallel to the planes of the DNA bases and are sandwiched between them. They are stabilized from above and below, which fixes the chromophore in place, allowing for less molecular movement. This causes an increase in the fluorescent quantum yield, because absorbed electromagnetic energy has fewer opportunities to leave the molecule through molecular movement and is therefore re-emitted as photons. This emission lies in the visible part of the spectrum with its peak at around 610 nm and can be either imaged with a camera or seen with the naked eye. The quantum yield of DNA bound EtBr is ~20-30 times higher than the free solution, making the DNA complex clearly visible over the background when excited with UV light (Green and Sambrook 2012).

Agarose gel electrophoresis was used for two purposes: (1) for analysis of DNA (PCR products, plasmids, restriction maps) and (2) the preparation of linear DNA

(linearized plasmids and PCR products for cloning). Custom made horizontal gel chambers with mobile trays and combs of different well-width were used (4mm for analytical gels, 12 mm for preparative gels). 1 % (w/v) agarose gels were made by suspending 0.5 g Agarose in 50 mL 1 x TAE buffer. The agarose suspension was melted by heating it in the microwave for 2 min at 800 W. The molten agarose was poured into a gel-tray and 2.5 µL of a 1 % (w/v) ethidium bromide solution were added (end concentration 0.5 µg/mL) and thoroughly mixed. The agarose gel was left to solidify for 30 min at room temperature, while a comb displaced part of the gel to create wells, in which the DNA could be pipetted in. The solidified gel was placed into the gel chamber and flooded with 1 x TAE buffer. Samples were mixed 1:6 with loading dye, which provides density to the sample, so it can sink into the pockets easily. Besides that, the loading dye gives a visual indication of: (1) the sample in the pocket and (2) the progress of the electrophoresis, since the dye molecules move through the gel at comparable rates to DNA. To estimate the molecular weight of gel-bands, at least one lane was loaded with either 1 kb or 100 bp DNA Ladder (New England Biolabs). Gels were run at 70 V–120 V and 50 mA for 30 min–1 h, depending on desired resolution and available time (higher voltage being faster, but with poorer resolution and potential for melting gels).

For analytical gel electrophoresis, gels were imaged, using a GelDoc 2000 (BIORAD) unit. Preparative gels were imaged using the UV table of the same unit, but in a slightly different manner. Gründemann and Schömig (1996) pointed out that DNA is severely damaged by the use of UV tables in preparative Gel extraction. The half-life of DNA, which is still functional in subsequent cloning steps, on a UV table is around 3–6 s, rendering 99 % of DNA useless after only a 45 s exposure. To prevent UV damage, preparative gels were cut “blind”. This was achieved by slicing a thin strip of the gel, containing the marker-lane and a small fraction of the lane with the DNA to purify. This thin strip was examined by eye on the UV-table and the band of interest was located and marked with a scalpel. Subsequently the strip was re-aligned with the non-exposed part of the gel and the band was cut out with the help of the marking; thus avoiding UV exposure altogether. The isolated gel-bands were then further processed as described in 2.2.1.3.

2.2.1.3 Purification of PCR Products and linear DNA

In this work purification of PCR products and linear DNA was mainly performed to isolate specific linear DNA fragments for subsequent use in the production of recombinant plasmids. To ensure good yields for subsequent enzyme reactions, high concentration of input DNA was prioritized. Empirically, it was found, that losses in gel purification were around 70–85 % of input material, while PCR clean-ups usually only retained 15–20 %. Therefore, PCRs were usually first analyzed via gel electrophoresis, as explained in the previous section. This analysis was used to determine if gel purification was necessary or if PCR clean-up would be sufficient. Gel purifications were only performed if major unspecific side-products were observed.

For the purification process of both methods the Nucleospin® Gel and PCR Clean-up kit was used according to the manufactures instructions. This kit is based on the following principle: DNA fragments are exposed to high salt buffers, containing chaotropic salts (such as guanidinium hydrochloride or guanidinium thiocyanate). These conditions destroy the hydrate shell around the DNA molecules and expose their negatively charged backbone phosphates. These are then able to bind negatively charged oxygen ions of the silica membrane, through the forming of cation bridges with the help of sodium ions. The membrane bound DNA is washed under high salt and ethanolic conditions, to remove enzymes, buffer components, nucleotides and primers (this is achieved because binding strength increases with the length of the DNA molecule, allowing the adjustment of conditions to a certain cut-off length). After that, the bound DNA is re-hydrated under low salt conditions, releasing the molecules from the membrane.

Specifically, PCR reactions were mixed 1:2 (v/v) with NTI buffer (high salt, chaotropic buffer). Cut-out gel bands, in turn, were weighted and mixed 1:2 (w/v) with NTI buffer, and agitated on a shaker (IKA-VIBRAX-VXR, 1,600 rpm), until the gel block was completely dissolved. Subsequent steps were performed similarly for both purification methods. DNA in high salt NTI buffer was transferred to the provided silica membrane columns and centrifuged (RT, 11,000 x *g*, 1 min) to facilitate binding to the membrane. The flow-through was discarded. Afterwards, columns were washed by application of 700 µL of Wash Buffer NT3 (high salt, 80 % ethanol) and centrifugation (RT, 11,000 x *g*, 1 min). Again, the flow-through was discarded. To remove residual ethanol, the columns were centrifuged again (RT, 11,000 x *g*, 2 min) and subsequently transferred to sterile Eppendorf reaction tubes. 50 µL of Elution Buffer NE (low salt, 5 mM Tris/HCl, pH 8.5) were applied to re-solubilize the DNA. Elution was achieved through an additional centrifugation step (RT, 11,000 x *g*, 1 min).

2.2.1.4 Determination of DNA Concentration and Purity

DNA concentration and purity was determined through UV absorption spectroscopy. DNA exhibits a strong absorption maximum at 260 nm, which enables the correlation of the $OD_{260\,nm}$ to the DNA concentration. Furthermore, the absorption at 230 nm and 280 nm was measured to determine the concentration of potential contaminants within the DNA preparation. Proteins have a strong absorption at 280 nm, from the π-systems of the aromatic amino acids (tryptophan, phenylalanine and tyrosine). Chaotropic salts, containing guanidinium ions, phenolates and thiocyanates, which are all commonly found in DNA preparation protocols on the other hand, exhibit strong absorption at around 230 nm. Therefore the ratios of $OD_{260\,nm}$:$OD_{280\,nm}$ and $OD_{260\,nm}$:$OD_{230\,nm}$ were used to assess whether these contaminants were present, with a $OD_{260\,nm}$:$OD_{280\,nm}$ ratio of ~1.8 being considered pure of proteins and a $OD_{260\,nm}$:$OD_{230\,nm}$ ratio of 1.8–2 being considered pure of substances absorbing at 230 nm. In practice $OD_{260\,nm}$:$OD_{230\,nm}$ ratios observed in DNA samples, purified with the Nucleospin® Gel and PCR Clean-up kit, were much lower than these values. This indicates the presence of residual guanidinium thiocyanate, which according to the manufacturer, does not influence subsequent enzyme reactions in the typically observed concentration range.

UV absorption measurements were performed using either a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific), or a NanoQuant Plate™ in combination with an Infinite® M200 PRO microplate reader (Tecan), according to the instructions provided by the respective manufacturer.

2.2.1.5 Restriction Digest of DNA

Restriction endonucleases (restriction enzymes) are enzymes, which recognize specific, often palindromic sequences in double stranded DNA. These motifs are typically between 4–8 bp long, determining the frequency, with which they are found in a random DNA sequence. After binding to their respective recognition site, these enzymes catalyze the cleavage of both DNA strands. Depending on the enzyme in question, different types of “ends” result from this reaction: (1) “blunt ends”, in which both strands are cut at the same position, and (2) “sticky ends”, in which DNA strands are cut in a staggered fashion, producing either 5’- or 3’-overhangs. This property of restriction enzymes has made them an invaluable tool in molecular biology; acting as a pair of sequence specific molecular scissors. Within the context of the work, restriction enzymes were used for several purposes: (1) to generate specific “sticky ends” for directional

“classic”-cloning, (2) to linearize circular plasmids for SLiCE cloning, (3) to inactivate plasmid DNA, which served as template in PCRs and (3) to verify the identity of plasmid DNA through digestion patterns.

For directional cloning two restriction enzymes with different recognition motifs and “sticky ends”, which were also present in the MCS of the target vector, were selected. Inserts were either cut-out from existing vectors with the same recognition sites, or amplified via PCR using primers with synthetic 5’-overhangs, as described in 2.2.1.1, and then cut. Typically, these overhangs consisted of the recognition site, a 5’-spacer of 3–5 bps, to facilitate efficient cleavage, and occasionally 1–2 bps between the recognition site and the priming site, to adjust for potential frameshifts. Plasmids and PCR products were first purified as described in 2.2.3.3 and 2.2.1.3 and then digested with restriction enzymes according to the instructions provided by the manufacturer. After digestion, linear DNA was again purified, using either gel extraction, if e.g. a fragment with a resolvable size-difference was cut-out of a vector, or PCR clean-up, if no visible difference between digested and undigested species can be achieved through electrophoresis.

Restriction enzymes were ordered from NEB, focusing on the use of HF® Restriction Endonucleases, which are engineered to operate within the same buffer conditions and exhibit less unspecific binding. A standard reaction mix for a double restriction digest is depicted in Table 2-11. Digests were incubated for 1–3 h at the appropriate temperature (typically 37 °C). If buffer conditions for two enzymes varied, digests were performed sequentially, starting with the enzyme requiring lower salt concentrations, topping-up the buffer after the first digest and continuing with the second enzyme.

Table 2-11 Reaction mix for double restriction digest with NEB HF®-enzymes

Compound	Concentration	Volume	Final Concentration
DNA	10–200 ng/µL	40 µL	0.2–4 ng/µL
H_2O		3 µL	
CutSmart® Buffer	10 x	5 µL	1 x
Enzyme A	20,000 U/mL	1 µL	0.4 U/µL
Enzyme B	20,000 U/mL	1 µL	0.4 U/µL
	Finale Volume:	50 µL	

As “classic” cloning through “sticky”-end hybridization was only rarely performed in favor of modern assembly cloning (see 2.2.1.7), restriction digest was mainly used to linearize plasmids for these techniques. To avoid re-ligation of linearized plasmids through complementary base-pairing of “sticky”-ends, a single blunt-cutting enzyme was selected for these tasks.

Some restriction endonucleases depend on specific methylation patterns within their recognition site. This property originates from their initial purpose in microorganisms as defense mechanisms against foreign DNA. Restriction endonucleases distinguish host DNA from foreign DNA, by specific methylation patterns, which can only be found in foreign DNA. Methylation only occurs if methyltransferases are present, as is the case for plasmid DNA produced in *E. coli* cultures, but not for regular PCRs. This can be used to selectively cut DNA of bacterial origin in mixtures were both PCR products and plasmids are present. Since the templates for PCRs in this work were often plasmids using the same backbone as the newly constructed plasmids, these templates could produce false positive colonies after being carried through from the PCR to the eventual transformation into chemical competent *E. coli*. To avoid false positive clones through this process, a restriction digest with DpnI was performed routinely after each PCR, whose products were used in subsequent constructions of new plasmids. This digest was performed by simply adding 1 µL of DpnI to a finished PCR reaction and incubating it at 37 °C for 2 h, without changing buffers. DpnI is a restriction enzyme with a four base pair recognition site (GATC), which can only cut in the presence of a methylated adenine. Its short recognition site assures that it is very frequently found within a random DNA sequence (statistically every 4^4bps = 256 bps). Therefor potentially contaminating template-plasmids are frequently cut into smaller linear pieces and thus rendered non-functional.

Furthermore, restriction digests were performed in order to verify the identity of plasmids. Typically, a frequently cutting enzyme (4 bp recognition site) was used to digest plasmid DNA. Afterwards an analytical gel electrophoresis was performed and the resulting pattern of observed gel bands was compared to simulated, digests of the expected sequence *in silico*.

2.2.1.6 Ligation and Dephosphorylation of DNA

Ligation

If restriction endonucleases are the scissors of molecular biology, then DNA ligases are the glue stick, which puts cut DNA strands back together. In the context of this work, T4 DNA ligase (NEB) was used, an enzyme which catalyzes the formation of phosphodiester bonds from 5' phosphate and 3' hydroxyl groups in the backbone of DNA strands. This reaction, which requires energy in form of ATP, is the reversal of a restriction digest. T4 DNA ligase was used to insert recombinant DNA into linearized plasmids. Vector and insert DNA were first cut with two different restriction enzymes, as described in 2.2.1.5, producing two different overhanging "sticky"-ends on each linear fragment. Overhanging ends of the vector and insert could then hybridize through complementary base pairing and therefore allow directional specificity. This hybridization was then fixed through the use of T4 DNA ligase, which forms covalent bonds between the vector- and insert-DNA, resulting in a functional circular plasmid. Ligations were performed according to the instructions provided by the vendor. Vector- and insert-DNA were supplied in a molecular ratio of 1:5 into ligation reactions. A standard ligation reaction-mix is depicted in Table 2-12. T4 Ligase Reaction Buffer was initially aliquoted into 4 µL lots, and then stored at –20 °C, to prevent ATP degradation through repeated freeze thaw cycles and contamination (aliquots were only defrosted once and then discarded). Ligations were incubated for 1 h at 16 °C and subsequently transformed into chemically competent *E. coli* XL1 Blue (see 2.2.3.4).

Table 2-12 Reaction mix for ligation with T4 DNA Ligase

Compound	Concentration	Volume	Final Concentration
Insert DNA	5 x µM	4 µL	2 x µM
Vector DNA	1 x µM	4 µL	0.4 x µM
T4 DNA Ligase Reaction Buffer	10 x	1 µL	1 x
T4 DNA Ligase	400,000 U/mL	1 µL	40 U/µL
	Finale Volume:	10 µL	

Dephosphorylation

The ligation of one linear vector and one linear insert molecule produces only one circular plasmid and therefore decreases entropy. The intramolecular circularization of linear vector DNA competes with the desired insertion of recombinant DNA into the linear vector and is thermodynamically favored, because no entropy is lost due to reduction of the number of resulting molecules. This circularization can occur through complementary base pairing if only one restriction enzyme was used, or even through non-complementary base pairing if two different restriction enzymes were used. The circularization of vector DNA leads to false positive colonies in the cloning process. To prevent circularization, linearized vector DNA was treated with Antarctic Phosphatase (NEB), an enzyme which catalyzes the dephosphorylation at the 5'-end of DNA strands. Since these 5' phosphate groups are required for ligation, vector DNA cannot circularize. In contrast, ligation of linear insert DNA with dephosphorylated linear vector DNA can still occur, because the insert DNA provides two phosphates at the 5'-ends. Instead of four covalent bonds between the four single strands, only two are formed, leading to a circular plasmid with two nicks. These nicked plasmids are still functional and will be repaired inside *E. coli* after transformation. Dephosphorylation was performed right after the restriction digest of vector DNA, without prior clean-up (reaction mix as shown in Table 2-13). The enzyme reaction was incubated for 30 min at 37 °C and subsequently purified as described in 2.2.1.3.

Table 2-13 Reaction mix for desphosphorylation with Antarctic Phosphatase

Compound	Concentration	Volume	Final Concentration
Vector restriction digest		44 µL	
Antarctic Phosphatase reaction buffer	10 x	5 µL	1 x
Antarctic Phosphatase	5,000 U/mL	1 µL	0.1 U/µL
	Finale Volume:	50 µL	

Introduction of synthetic DNA through Ligation

Besides regular cloning, ligations were also used to introduce short synthetic DNA sequences into plasmid DNA. Instead of ligating linear pieces of DNA, which were cut with restriction enzymes, long PCR products are circularized into a functional plasmid. Products of standard PCRs cannot be circularized, because the 5'-ends of primers usually lack the required phosphate groups. In order to facilitate this ligation, functionalized oligonucleotides, which are phosphorylated at the 5'-end, were used. A PCR with phosphorylated primers was performed, which amplified the whole plasmid. The 5'overhangs of these phosphorylated primers were used to incorporate new synthetic sequence information (see 2.2.1.1). PCR products were purified as described in 2.2.1.3, and circularized through intramolecular ligation using standard reaction mix and conditions and subsequently transformed into chemically competent *E. coli* XL1 Blue.

2.2.1.7 SLiCE Cloning

In recent years several new cloning technics have been developed, which use different approaches from traditional cloning with restriction enzymes (Valla and Lale (2014)). These new cloning techniques simplify cloning in many ways and overcome limitations traditional cloning poses. One of these new techniques, i.e. SLiCE cloning, was published by Zhang et al. (2012). SLiCE cloning (acronym for **s**eamless **li**gation **c**loning **e**xtract) is a DNA assembly method, which relies on *in vitro* recombination of short homologue sequences at the ends of linear DNA. This method was adapted and extensively used throughout this work.

Specifically, a crude cell-extract is prepared from an engineered *E. coli* strain. This cell-extract contains the enzymes which catalyze the *in vitro* recombination. The production of this crude cell-extract is very simple, cheap and results in large quantities of active enzyme mix, which is a big advantage, because there is no need for commercial enzymes. The cloning procedure itself is performed by the introduction of short homology regions through PCR, which are shared between the insert (or several inserts) and the vector DNA and subsequent recombination of these homology regions into a functional circular plasmid. Through this method up to seven inserts can be cloned into a target vector at the same time, in a directional and specific manner. This recombination is independent of the sequence of the homology region. Both of these properties are big advantages over traditional cloning. The ability to clone several inserts at the same time, with high efficiency, allows for quick cloning of fusion proteins, which is at the core of this work. The sequence independence is also important, as there is no longer a need for specific restriction sites that would

pose a problem specifically in large vectors, or the engineering of fusion proteins, as certain parts of the protein coding sequence are fixed by the requirement for restriction site sequences.

Seamless Ligation Cloning Extract

In detail, the protocol of Zhang et al. (2014), was adapted in the following way. An overnight culture of *E. coli* PPY was inoculated from a single colony into 5 mL of 2xYT liquid medium supplemented with 10 µg/mL streptomycin in sterile 14 mL round bottom culture tubes. Overnight cultures were agitated (220 rpm) at 37 °C in a round shaker incubator. On the next day a 200 mL main culture (2xYT medium, 10 µg/mL streptomycin) was inoculated with the overnight culture in a baffled Erlenmeyer-flask. The culture was incubated at 37 °C in a round shaker (220 rpm). Growth was monitored as described in 2.2.3.2 until an $OD_{600\ nm}$ of 5–5.5 was reached. Protein production was induced by addition of 1.1 mL of a 36 % (w/v) L-(+)-arabinose solution (final concentration 0.2 % (w/v)). After induction, the culture was incubated for 2 h at 37 °C and strong agitation (220 rpm), to facilitate protein production. After this, cells were harvested through centrifugation (5,000 x *g*, 20 min, 4 °C), the supernatant was discarded. To produce the cell extract from the cell-pellet, an adapted version of the large-scale protein purification was performed (see 2.2.2.2). First cells were re-suspended in 2.5 mL Protein Resuspension Buffer I and transferred to a 50 mL Falcon tube. To prevent proteases from digesting the active components of the enzyme extract, the cell suspension was supplemented with a cocktail of protease inhibitors (4 mM PMSF, 20 µg/mL Pepstatin A, 4 µg/mL Leupeptin; end concentration) and thoroughly vortexed. To break open *E. coli* cells, several physical and enzymatic steps were performed. First, the cell suspension was frozen solid at –80 °C and subsequently thawed again. This produces ice crystals within the cells, which rupture the cell-wall and membrane. After this step, the cell suspension was kept on ice, to reduce potential degradation. Next, 1 mg of lyophilized Lysozyme powder was added to the suspension, thoroughly vortexed and incubated for 30 min on ice. Lysozyme is an enzyme that catalyzes the hydrolysis of peptidoglycan in the bacterial cell wall. 50 µL of a 1 mg/mL RNAse solution was added (end concentration 20 µg/mL). Then, the suspension was incubated for 30 min in an ultrasound water bath. The ultrasonic agitation facilitates the physical rupture of already damaged cell-walls and membranes. After this step, the cell lysate was cleared through centrifugation (20,000 x g, 30 min, 4 °C). The supernatant was decanted, supplemented with glycerol to a final concentration of 50 % (v/v) and incubated at 4 °C overnight to facilitate protein folding. The supernatant was subsequently aliquoted into 100 µL lots, shock frozen in liquid nitrogen and stored at –80 °C. One batch contained between 4,000 and 6,000 reactions.

In vitro Recombination with SLiCE

The enzyme extract facilitates the recombination of short homology regions at the ends of linear DNA. These homology regions were introduced as described in 2.2.1.1 through 5'-overhangs on primers during PCR. Typically, 21 bp were added at the end of each insert, which were either homologue to the ends of the vector or to the next insert in line. This leads to homology regions of 21 bp between the vector and insert and 42 bp between multiple insert. In this work, up to 4 inserts were successfully cloned into a vector in this manner. PCR products were digested with DpnI, as described in 2.2.1.5, to inactivate potentially contaminating template plasmids. Vectors were either linearized through a single restriction digest (typically a "blunt" cutter), or through amplification with PCR. If PCR was used, it was also possible to introduce 21 bp regions, which are homologous to the insert(s). Linearized vector and insert DNA were purified (see 2.2.1.3) and DNA concentration was determined (see 2.2.1.4). A recombination reaction mix is shown in Table 2-14. Vector and insert DNA were supplied in a molar ratio of 1:5, respectively (for vector and two inserts the ratio was 1:5:5, etc.). Since the suggested reaction buffer of Zhang et al. (2014) was very similar to T4 Ligase reaction Buffer, this buffer was used, mainly for convenience. Recombination reactions were incubated for 15 min at 37 °C and subsequently transformed in chemically competent *E. coli* XL1 Blue.

Table 2-14 SLiCE reaction mix

Compound	Concentration	Volume	Final Concentration
Insert DNA	5 x µM	4 µL	2 x µM
Vector DNA	1 x µM	4 µL	0.4 x µM
T4 DNA Ligase Reaction Buffer	10 x	1 µL	1 x
SLiCE		1 µL	
	Finale Volume:	10 µL	

2.2.2 Working with Proteins

2.2.2.1 Overexpression of recombinant Proteins in E. coli

In order to analyze the performance and biophysical properties of a constructed biosensor, it was overexpressed in *E. coli* and then purified. Biosensors were cloned in two different derivatives of the expression vector pRSET B (Thermo Fisher Scientific), deemed pRSET SL II and pRSET SL III. Both vectors are based on the same backbone and are only different in the length of their His-tags (6 x His and 10 x His respectively). The overexpression of recombinant protein is achieved by making use of the inducible *lac*/T7 system. The transcription of the target gene is driven by the strong T7 promoter which is located downstream of the biosensor coding sequence on the vector. The T7 promoter is recognized by T7 RNA polymerase. In this work *E. coli* Bl21 (DE3) was used as an expression strain which harbors a copy of T7 RNA polymerase under the control of the *lacUV5* promoter in its genome. This promoter is part of the lac operon, which metabolically controls the expression of genes involved in lactose catabolism. It is activated by the presence of lactose (or compounds structurally similar to lactose) and therefore chemically inducible. Isopropyl β-D-1-thiogalactopyranoside (IPTG) is a lactose analogue, which can bind and release the lac repressor, but cannot be catabolized by *E. coli* and is therefore often used to induce protein expression. IPTG induction requires fairly precise timing and therefore constant, close monitoring of the culture to achieve optimal expression. Early induction leads to reduced cell density, because most energy is devoted to protein expression, and late induction to poor protein expression, because nutrients have already been spent on cell division. However, a study by Studier (2005) suggested the use of a more sophisticated method for metabolic control of induction of protein expression. The lac operon is actually antagonistically controlled by glucose and lactose. The presence of glucose inhibits protein expression even if lactose is present. Thus, a mixture of glucose and lactose can be used as a timer, which automatically induces protein expression. At first the presence of glucose inhibits the expression of recombinant protein under the control of the lac operon, allowing the culture to reach sufficient density. As glucose is depleted by the growing culture, the presence of lactose finally induces the production of T7 RNA polymerase, starting high rates of protein expression. In addition, glycerol is added to supply the culture with an additional energy source. This so-called auto induction medium was found to be very effective in the overexpression of biosensors.

In detail, the following procedure was performed, adapting the suggested protocol by Studier (2005). Chemically competent *E. coli* BL21 (DE3) were transformed with purified plasmid preparations of either pRSET SLII or pRSET SLIII, containing the biosensor of interest, as described in 2.2.3.4. Cells were immediately transferred to baffled 250 mL Erlenmeyer flasks containing 50 mL LB medium supplemented with 0.05 % (w/v) D-(+)-glucose, 0.2 % (w/v) lactose, 0.6 % glycerol and 100 µg/mL ampicillin. Furthermore, polypropylene glycol (PPG) was added to an end concentration of 0.1 % (v/v). PPG is used as an anti-foaming agent, as it does not readily mix with water, and exhibits phase separation, which prevents foam from occurring during continues shaking. Flasks were incubated for 72 h at room temperature under vigorous agitation (220 rpm) in a round shaker incubator. Owing to the fact that biosensors expressed in the context of this work always incorporated fluorescent proteins or chromoproteins, success of protein production was visually assessable by the color of either the suspension culture or the harvested cell pellet.

2.2.2.2 Purification of His-tagged Proteins

Proteins were expressed using the vectors pRSET SLII and pRSET SLIII (see 3.1.3), which add His-tags of different lengths (6 x His, 10 x His respectively) to the N-terminus of recombinant proteins. The basic protocol for extraction and purification of these differently tagged proteins was the same, using slightly adjusted buffer conditions for 10 x His-tagged proteins, denoted in square brackets. In principle, *E. coli* cells are lysed through various physical and enzymatic methods and the soluble fraction is collected through centrifugation. His-tagged recombinant proteins are then purified through metal-ion affinity chromatography. This chromatography is based on the coordination of multiple histidine side chains to Ni^{2+}-ions, which are immobilized through chelating agents on agarose beads. Elution of the stationary phase is achieved through competition with high concentrations of imidazole that similarly binds to Ni^{2+}-ions.

In detail, overexpression cultures were harvested by centrifugation in 50 mL Falcon tubes (5,000 x *g*, 10 min, 4 °C), discarding the supernatant. Cell pellets were then re-suspended in 10 mL Protein Resuspension Buffer I [Protein Resuspension Buffer II], supplemented with a protease inhibitor cocktail (4 mM PMSF, 20 µg/mL Pepstatin A, 4 µg/mL Leupeptin; end concentration) to prevent protein degradation by proteases released through cell disintegration. The cell suspension was frozen solid at –80 °C to promote cell-rupture through the formation of ice crystals. Suspensions were then defrosted and supplemented with 1 mg of lyophilized chicken lysozyme. Lysozyme catalyzes the hydrolysis of 1,4-beta-linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine residues of peptidoglycan within the bacterial cell wall, further weakening the structural integrity of the cells. Furthermore, 5 µg/mL DNAse and 10 µg/mL

RNAse were added to degrade nucleic acids, which are released from lysed cells and which otherwise create a sludge. The enzymatic degradation of the cell wall was promoted by incubating the suspension at 37 °C for 30 min. Afterwards Falcon tubes were transferred to an ultrasonic water bath (cooled by crushed ice to prevent heat induced protein denaturation) and sonicated for 30 min. To remove insoluble cell-debris, the suspension was centrifuged (20,000 x *g*, 30 min, 4 °C) and the supernatant was collected in 15 mL Falcon tubes. 120 µL of a 6 % (v/v) Nickel-IDA agarose bead suspension (Jena Bioscience) were added to the clarified lysate. This suspension was then incubated for 2 h under constant weak agitation on a rotator, to promote binding of His-tagged proteins to Ni^{2+} ions present in the agarose resin. Subsequently, the bead suspension was transferred to 1 mL polypropylene columns (Qiagen) that retained beads with bound protein. Beads were washed by passing 10 mL Protein Resuspension Buffer I [Protein Resuspension Buffer II] through the column. His-tagged proteins were eluted from the resin by addition of 600 µL of Protein Elution Buffer. The eluate was collected in 1.5 mL Eppendorf reaction tubes and incubated at 4 °C overnight, to promote protein folding and maturation. The purified protein solution was either used immediately in subsequent experiments or frozen at –80 °C for long term storage.

2.2.2.3 Dialysis of Protein Solutions

To free proteins from salts and other buffer components, which are incompatible with subsequent methods, they were dialyzed after elution. Dialysis is achieved by separating the high salt protein solution from a low salt dialysis solution by a semi-permeable cellulose membrane, which allows small molecules to pass but retains larger molecules. The salt ions then diffuse along their concentration gradient through the membrane into the low salt dialysis solution. This process reduces the salt concentration in the protein solution, while the protein concentration remains constant.

Specifically, a custom dialysis device was assembled: A 8 mm hole was drilled into the center of the lid of an 1.5 mL Eppendorf reaction tube, using an 8 mm wood drill bit and a cordless drill. These reaction tubes were then sterilized by autoclaving. As dialysis membrane ZelluTrans (Carl Roth) dialysis tubing with a molecular weight cut off of 8,000–10,000 Da was used. Dialysis tubing was cut into 2 cm single walled squares and then soaked in dialysis buffer (typically PBS) to revitalize the cellulose membranes. 400 µL eluted protein solution were transferred into reaction tubes with holes. The revitalized dialysis membranes were then placed over the opening of the reaction tube. Then the lid was closed in a manner that traps the cellulose membrane between the rim of the lid and the reaction tube. This covered the hole in the reaction tube with a single layer of membrane, allowing the protein solution to get in contact with dialysis solution through the lid. Dialysis devices assembled in such a way were then placed up-side down in a floating device and were then transferred into a beaker glass containing 500 mL PBS. Care was taken, that the protein solution was completely in contact with the cellulose membrane and that no air bubbles were trapped between the dialysis solution and the bottom of the membrane. The dialysis solution was stirred by a magnetic stirring device and incubated at 4 °C. Typically three dialysis steps were undertaken to increase difference in the salt gradient. For each step dialysis solution was replaced with fresh buffer and was incubated for 2 h, with one of the steps requiring overnight incubation.

This method is very cost effective and allows for parallel dialysis of several dozen samples.

2.2.2.4 Mass Spectrometry

Mass spectrometry methods were adapted from Hornburg et al. (2014) and performed in close collaboration with Daniel Hornburg from the Department of Proteomics and Signal Transduction of the Max Planck Institute for Biochemistry.

Sample Preparation

His-tagged proteins were purified via metal ion affinity chromatography as described in 2.2.2.2 and subsequently dialyzed against PBS (see 2.2.2.3). To 800 µL of purified protein solution 100 µL of 6 M guanidinium chloride (GuHCl) solution (final concentration 0.67 M) and 9 µL of 1 M dithiothreitol (DTT) (final concentration 10 mM) were added and thoroughly mixed. The sample was then incubated at 95 °C for 10 min in a thermos shaker at 190 rpm. The proteins within the sample were denatured by heat and the effect of GuHCl, a chaotropic salt, making their polypeptide chains more accessible to the later proteolytic digest. DTT on the other hand is a reducing agent, which was used to break disulfide bonds. Samples were then sonicated for 5 min, to re-solubilize any precipitate and then incubated for 25 min at room temperature. Reduced cysteine residues were then alkylated by adding 100 µL 0.5 M iodoacetamide (IAA) to a final concentration of 50 mM. The reaction was incubated for 45 min at room temperature in the dark and then stopped with the addition of 100 µL of a urea/thiourea (6 M/2 M) solution. Samples were centrifuged (16,000 x *g*, 5 min, RT) to pellet insoluble components and the supernatant was transferred into a new reaction tube. Denatured proteins were then digested by adding 4 µL of a Lys-C solution (0.5 µg/µL) and 4 µL of a Trypsin solution (0.5 µg/µL). Lys-C and Trypsin are proteases with very specific digestion patterns, always cleaving after either lysine or after arginine and lysine, respectively. This property allows the prediction of resulting peptides and therefore their expected masses when the protein sequence is known. The proteolytic digest was aided by the addition of 100 µL acetonitrile to make the proteins more ”fluffy” ([sic] Hornburg, 2015) and then incubated for 4 h at room temperature. Finally, samples were desalted on C18 Stage Tips (Rappsilber et al. 2007).

LC-MS/MS

Peptide mixtures ware separated using an EasynLC 1,000 HPLC system (Thermo Scientific) with columns (75 µm inner diameter, 20 cm length) packed in-house with 1.9 µm C18 particles. Samples were loaded in buffer A (0.5 % formic acid) and separated in an 85 min gradient from 2 % to 60 % buffer B (80 % acetonitrile, 0.5 %formic acid) with a flow rate of 250 nl/min and a column temperature of 40 °C. After separation, peptides were directly ionized by a nano-electrospray source, coupling the liquid chromatography (LC) directly to the mass spectrometer (MS). The mass spectrometer was a quadrupole Orbitrap (Q Exactive, Thermo Fisher Scientific,(Scheltema et al. 2014)) which was operated in data dependent mode. The survey scan range was between 300 and 1650 m/z, with a resolution of 60,000 at m/z 200. The 10 most abundant isotope patterns with a charge ≥2 were fragmented and further analyzed. Fragmentation was achieved by high-energy collision dissociation (normalized collision energy 25, isolation window 1.4 Th, resolution 15,000 at m/z 200). Repeated sequencing was limited by setting the dynamic exclusion of sequenced peptides to 20 s. Ion injection times and ion target value thresholds were 20 ms and 3 x 10^6 for survey scans and 50 ms and 10^5 for MS/MS scans, respectively. Data acquisition was performed using Xcalibur software (Thermo Scientific).

Data Analysis and Statistics

MS raw files were analyzed with MaxQuant software (v 1.5.2.22) (Cox and Mann 2008), using Andromeda (Cox et al. 2011) to search MS/MS spectra against a UniProt FASTA database of *E. coli* K12 proteins (UP000000625, 18.7.2015) and added sequences for the proteins of interest and common contaminants. Trypsin was used for enzyme specificity and N-terminal cleavage after proline and up to two miscleavages were allowed. The minimum peptide length was set to seven amino acids. N-terminal acetylation, methionine oxidation and deamidation of asparagine and glutamine were set as variable modifications, while carbamidomethylation was set as a fixed modification. The cutoff for the false discovery rate (FDR) was 1 % at the peptide and at the protein level. 4.5 ppm and 20 ppm mass deviation was allowed for the initial precursor mass and the fragment mass, respectively. The cutoff score for MS/MS spectra was 17. At least one unique peptide was required to identify a protein and homologous proteins which could not be discriminated were grouped.

2.2.2.5 Fluorescence Measurements

Fluorescence measurements were either performed in a Varian Cary Eclipse spectrophotometer or a Tecan Infinite® M200 PRO plate reader.

For spectrophotometer measurements, 20 µL of a purified protein solution were added to 970 µL PBS and thoroughly mixed in a 10 mm quartz cuvette. Emission- and excitation spectra were recorded using slow scanning speed and 1 nm steps. The photomultiplier voltage was adjusted to individual proteins and optimized so that spectral peaks fell into a window between 200 a.u. and 600 a.u. The excitation and emission slit width was usually kept at 5 nm, and only increased for very low concentrated or very dim fluorescent proteins. Ligands for biosensors were added in a volume of 10 µL (e.g. 100 mM cGMP for cGMP sensors) and thoroughly mixed. Mixing of protein solutions was found to be a very crucial step, which can introduce errors into subsequent measurements. On the one hand, concentrated protein solutions have to be thoroughly mixed to avoid uneven distribution; on the other hand, it was found that recombinant protein binds to plastic pipette tips and even glass Pasteur pipettes and is therefore removed with each mixing step, leading to lower fluorescence intensities. To mitigate this problem, protein solutions were initially thoroughly mixed through pipetting with a 1 mL Eppendorf pipette. After recording of spectra and addition of ligand solutions, samples were mixed through prolonged shaking of cuvettes containing them. Measurements were either performed at room temperature or in a special cuvette holder connected to a temperature controlled water bath, to measure at elevated temperatures (e.g. 30 °C or 37 °C).

Plate reader fluorescence measurements were conducted in black, flat bottom 96 well plates (Brandt). Typically, 4 µL of purified protein solution were diluted in 196 µL PBS. Gain and z-height were optimized for one well, using the software integrated optimization function, and applied to all wells which were measured.

Spectra of FRET based biosensors were normalized to the isosbestic point. The FRET ratio (R) was calculated according to equation 8.4 for peak emissions. The dynamic range of FRET based sensors was calculated as $\Delta R/R$, given by equation 8.5, with R_0 and R_1 being the FRET ratio before and after addition of the maximum concentration of ligand molecule, respectively.

$$R = \frac{Acceptor[Em_{max}]}{Donor[Em_{max}]} \tag{8.4}$$

$$\Delta R/_R = \frac{(R_0 - R_1)}{R_1} \times 100\ \% \tag{8.5}$$

The dynamic range of single fluorophore sensors, on the other hand, was characterized by $\Delta F/F$, as given by equation 8.6, F_0 and F_1 being the fluorescence

at the emission maximum before and after addition of the maximum concentration of ligand molecule, respectively.

$$\Delta F/_F = \frac{F_1 - F_0}{F_0} \times 100\ \% \tag{8.6}$$

It should be noted, that $\Delta R/R$ and $\Delta F/F$ as measures of the dynamic range of fluorescent biosensors, although of similar nature, are not directly comparable, with $\Delta R/R$ being a ratio of ratios, while $\Delta F/F$ only being a simple ratio.

2.2.2.6 Determination of Biosensor K_d

The apparent K_d of cGMP biosensors to cGMP and cAMP was determined through titration. Two solutions of equal biosensor concentration were prepared, which contained either no cyclic nucleotide ($cNMP_0$) or the maximum concentration of cyclic nucleotide ($cNMP_{max}$) used in the titration (1 mM cGMP, 10 mM cAMP). 200 µL of $cNMP_{max}$ solution were added to the first well in a row of a 96 well plate, all subsequent wells in that row were filled with 150 µL of $cNMP_0$ solution. A serial dilution was performed, by transferring 50 µL from each well to its consecutive one, thoroughly mixing after each step and discarding pipette tips after each transfer. The concentration series achieved through this process spans six orders of magnitude, and was found to be sufficient to cover all observed K_ds. Fluorescence spectra of each well were recorded using a fluorescence plate reader (Tecan). Spectra of FRET-based biosensors were normalized to the isosbestic point and FRET ratios were calculated for each well according to equation 8.4. FRET ratios were plotted against the decadic logarithm of the cyclic nucleotide concentration, using the graphing software Origin 8. Plotted values were fitted sigmoidal, using a dose-response curve with a variable Hill slope. The apparent K_d was equated to the EC_{50} of the fitted dose-response curve.

2.2.3 Working with Bacteria

2.2.3.1 Culture and Strain Maintenance

E. coli strains were cultivated in complex liquid and solid media, containing antibiotics; typically, LB with 100 µg/mL ampicillin. Single perfectly round colonies on solid media were assumed to be all ancestors of a single bacterium and therefore genetically identical. To purify genetically mixed bacteria cultures, aliquots of these cultures were streaked onto selective plates using sterile inoculation needles, in a manner which diluted them until only single colonies were visible.

Cryo-conservation of Bacteria

To preserve strains for longer periods of time, cryo-stocks were established. A single colony of the strain to preserve was used to inoculate a 2 mL overnight culture (LB, with strain appropriate antibiotics) and incubated at 37 °C. 350 µL of sterile 30 % (v/v) glycerol solution were added to 850 µL of freshly grown culture (end concentration 8.75 %) in a 1.5 mL Eppendorf reaction tube and thoroughly mixed. Glycerol forms strong hydrogen bonds with water molecules and therefore interferes with the formation of ice crystals, because the water crystal lattice is disturbed. The cell suspension was shock frozen in liquid nitrogen, to further prevent ice crystals from forming, and then stored at −80 °C.

To revitalize a cryo-stock, a small portion of the stock was scratched of its surface with a sterile inoculation needle, without defrosting the rest of the stock, and used to inoculate fresh cultures.

2.2.3.2 Monitoring Growth

Bacterial growth was monitored photometrically, by measuring the optical density at 600 nm. Light is scattered on the interfaces of bacterial membranes. This scattering can be used, to an extent, as a measure of cell density.

To eliminate absorption of media components, growth medium was used as a reference and subsequently subtracted from further measurements. 1 mL cell suspension was transferred into a 1 mL plastic cuvette and the $OD_{600\,nm}$ was determined. The relationship between OD and cell density was assumed to be linear, in a value range from 0.1–0.5. If observed values exceeded this range, suspensions were diluted with growth medium and OD values were multiplied with the dilution factor.

2.2.3.3 Plasmid Isolation and Purification

Plasmids were isolated using the NucleoSpin® Plasmid QuickPure kit from Macherey-Nagel for small scale isolation ("mini-prep") and the PureYield™ Plasmid Midiprep System for large scale isolation ("midi-prep") according to the instructions provided by the manufacturers. The basic principle is as follows: *E. coli* cells are first lysed through an alkaline/SDS containing buffer. Then chromosomal DNA and proteins are denatured and precipitated, leaving plasmid DNA in solution. Precipitated components are separated through centrifugation. Subsequently plasmid DNA is bound to silica membranes, as described in in 2.2.1.3, washed and eluted.

Mini-prep

2 mL LB cultures with appropriate antibiotic were grown at 37 °C overnight into a dense cell suspension. Cells were pelleted through centrifugation (11,000 x *g*, 1 min, RT) in 2 mL Eppendorf reaction tubes. The supernatant was discarded and cell-pellets were suspended in 250 µL buffer A1. Lysis was initiated by the addition of 250 µL lysis buffer A2, containing SDS and sodium hydroxide. Proteins and chromosomal DNA were precipitated by adding 350 µL A3 buffer. The lysate was clarified through centrifugation (11,000 x *g*, 5 min, RT) and the resulting supernatant was transferred to the provided silica columns. The solution was passed through the column with a further centrifugation step (11,000 x *g*, 1 min, RT), discarding the flow through. The silica bound plasmid DNA was washed, by adding 450 µL buffer AQ, containing 80 % ethanol, and centrifuging (11,000 x *g*, 3 min, RT). Plasmid DNA was then eluded from the membrane by applying 50 µL elution buffer AE, incubating for 1 min and centrifuging once more (11,000 x *g*, 1 min, RT). Purified plasmid DNA was frozen at −20 °C for long term storage.

Midi-prep

200 mL LB cultures with appropriate antibiotic were grown overnight at 37 °C in 500 mL baffled Erlenmeyer flasks in a round shaker-incubator (220 rpm). Cells were harvested by centrifugation (5,000 x *g*, 10 min, 4 °C) and the supernatant was decanted. Cell pellets were resuspended in 6 mL Cell Resuspension Solution (CRA). To initialize lysis, 6 mL Cell Lysis Solution (CLA) were added, containing 1 % SDS and 200 mM NaOH, carefully mixed by slowly shaking the container and incubated for 3 min at room temperature. The lysate was neutralized by adding 10 mL Neutralization Solution (NSB) and mixed by shaking. This step precipitated cellular proteins and chromosomal DNA. The lysate was then clarified by centrifugation (15,000 x *g*, 15 min, RT). A column stack of a PureYield™ Binding

Column on the bottom and a PureYield™ Clearing Column on the top was assembled on a vacuum manifold. The supernatant of the centrifugation was transferred to the Clearing Column and vacuum was applied to the manifold until the solution passed both the Clearing and the Binding Column. The Clearing Column filtered remaining precipitate from the supernatant and passed the clarified lysate onto the Binding Column, where plasmid DNA was retained. After this step the Clearing Column was discarded. *E. coli* endotoxins were removed by washing the bound plasmid DNA by passing 5 mL Endotoxin Removal Wash, containing 40 % isopropanol, through the Binding Column, while applying vacuum. 20 mL Column Wash, containing 60 % ethanol, were passed through the Binding Column by applying vacuum, to remove salts and other impurities. The binding membrane was dried, by applying vacuum to the Binding Column for 1 min. The Binding Column was then removed from the vacuum manifold. Excess Column Wash buffer was carefully dried off with a paper towel and the column was then placed on an assembled Eluator™ Vacuum Elution Device, containing a 1.5 mL Eppendorf reaction tube. 600 µL of 70 °C warm Nuclease-Free Water were transferred onto the binding membrane inside the Binding Column and incubated for 2 min, to elute the plasmid DNA from the membrane. Vacuum was then applied to pass the DNA solution into the 1.5 mL Eppendorf reaction tube. Purified plasmid DNA was frozen at –20 °C for long term storage.

2.2.3.4 Transformation of E. coli

Plasmid DNA was introduced into *E. coli* via chemical transformation. Chemical transformation is based on preparing cell suspensions, which contain a high concentration of Ca^{2+} ions, the so called “competent” cells. These competent cells are then incubated with DNA and later exposed to elevated temperatures (heat shock), to promote DNA uptake. Although it is hypothesized, that the presence Ca^{2+} ions (1) generates pores in the bacterial membranes and (2) masks the charge of the DNA backbone, the exact mechanism of transformation remains still at large (Green and Sambrook 2012).

Production of chemically competent *E. coli*

The protocol used in the context of this work is an adapted version of the one published by Chung et al. (1989).

Specifically, a pre-culture was established by inoculating 10 mL LB medium (containing 10 µg/mL tetracycline) with cells from a cryo-stock in a 100 mL Erlenmeyer flask and incubating it overnight at 37 °C under constant shaking (220 rpm). On the next day, a 300 mL main culture (pre-warmed LB medium without antibiotics, 37 °C, 1 L Erlenmeyer flask) was inoculated with the entire volume of the pre-culture. The culture was incubated under constant agitation (220 rpm) at 37 °C until it reached an $OD_{600\,nm}$ of 0.5. Next, the flask with the culture was incubated on wet ice for 10 min. Cells were harvested by centrifugation (1,000 x *g*, 15 min, 4 °C) and the supernatant was discarded. Pellets were carefully re-suspended in 30 mL ice-cold TSS buffer and 7.5 mL ice-cold 86 % (v/v) glycerol were added (end concentration ~15 %). The cell suspension was aliquoted into 50 µL portions in 1.5 mL Eppendorf reaction tubes and shock frozen in liquid nitrogen. Competent cells were stored at –80 °C.

Transformation of chemically competent *E. coli*

For transformation, an aliquot of chemically competent cells was defrosted on wet ice. 1 µL of DNA solution (e.g. isolated plasmids, SLiCE ligations, etc.) was added to the cell suspension and incubated between 5 min–45 min. The incubation period was chosen according to the needed transformation-efficiency, short incubation for propagation of established plasmids and longer incubation for complex cloning steps. Next, the cell suspension was “heat-shocked” for 42 s at 42 °C in a water bath. The cell suspension was then cooled for 2 min on wet ice. Most plasmids used in this work, contained the β-lactamase gene (*bla*) as a resistance marker and were used in combination with liquid or solid media containing ampicillin as selecting agent. Ampicillin inhibits the synthesis of peptidoglycan and therefore prevents cell division, as no additional cell-wall can be made, while other cell processes remain intact (translation, protein synthesis). Because of this, in contrast to other antibiotics, resistance against ampicillin can be acquired by *E. coli* while already exposed to the selection agent. Thus, cells transformed with plasmids containing the *bla* cassette were either plated on selective agar plates or transferred into selective liquid media right away. If other selection markers were used, cells were transferred into 1 mL non-selective SOC medium and incubated for 1 h at 37 °C, to express the resistance-mediating proteins, before they were exposed to selective media.

2.2.3.5 Counter-selection using *sacB*

Empirically, it was determined, that the main source of false positive colonies from SLiCE cloning (as described in 2.2.1.7) was un-cut target vector, which was co-purified after the linearization and carried through all subsequent steps and then successfully transformed *E. coli* cells. To minimize this background, a counter selection gene (*sacB*) was established in vectors, which were frequently used (see 3.1.2). *SacB* encodes the levansucrase from *B. subtilis* and catalyzes the hydrolysis of sucrose and subsequent synthesis of levans, which are high molecular weight fructose polymers (Pelicic et al. 1996). Expression of *sacB* in *E. coli* is lethal, if exposed to 10 % sucrose in media containing low NaCl concentrations (Gay et al. 1985, Blomfield et al. 1991).

E. coli cells, which were transformed with vectors containing *sacB* as counter-selection marker, were plated on YTS-agar (0.5 % (w/v) yeast extract, 1 % (w/v) tryptone, 10 % (w/v) D-(+)-sucrose, 1 % (w/v) agar-agar) with the appropriate antibiotic to select for the presence of the newly assembled plasmid.

3 Results

3.1 Vector System for Biosensor Development

The biosensor development approach applied in this work required a lot of complex cloning, especially in the prototyping stage, in which different sensor and reporter domains were assembled and tested (see 3.3). Furthermore, prototype sensors were evaluated according to their performance *in vitro*, in cell culture and in brain slices. Therefore, it was necessary to be able to quickly and efficiently subclone candidate sensors into vectors, which can facilitate expression in these conditions. Having to rely on the availability of restriction sites was often an obstacle in the past for several reasons: (1) Suitable restriction sites, which were used in the assembly of the FRET sensor, were not available for subcloning anymore. (2) Restriction sites which were used within the coding sequence of a sensor were fixed sequences, which in turn were translated into fixed amino acids. As FRET sensors are assembled from fluorescent proteins and sensor domains, these fixed sequences were often located in the critical linker region between domains, which then were not available for linker diversification. (3) Planning the assembly of prototype sensors from multiple parts from various sources, became very complex and therefore time consuming, which reduced the number of potential sensors that could be made and tested. (4) Subcloning, while usually trivial, also became challenging because critical restriction sites were often already in use or appeared by chance naturally in the sequence. This was especially problematic in large viral vectors, as they usually

only have very limited multiple cloning sites for related reasons. Besides that, “classic” cloning can typically only facilitate direction selective insertion of one DNA fragment at a time efficiently, requiring multiple steps to assemble a complete biosensor from multiple parts. It was therefore desirable to establish an efficient cloning strategy, which allows sequence independent fast assembly of prototype sensors and their subsequent subcloning into a variety of vectors for further characterization.

Zhang et al. (2012) demonstrated a very efficient novel cloning method, which relies on *in vitro* homologous recombination and can facilitate the direction selective assembly of up to seven DNA fragments into a target vector. This approach simplified sensor construction considerably and was therefore the method of choice to achieve the stated goals. Specifically, sensors could be assembled from multiple parts in a one-step procedure in a sequence independent manner.

3.1.1 Harmonization of Vectors for quick Transfer

SLiCE cloning relies on *in vitro* recombination of 15–50 bp homologous regions at the ends of linear DNA (Zhang et al. 2012). This process is independent of the sequence itself, as long as consecutive DNA fragments exhibit the same sequence in the overlaps. In order to quickly subclone potential sensor candidates it was useful to generate a series of vectors which share the same homologous regions. Sensor constructs then could simply be amplified by priming form the constant homologous region, using standard primers and PCR conditions, generating the homologous regions required for subcloning in the process (see Figure 3.1 C). This step could be parallelized easily, as PCRs were standardized independent of the actual architecture of the sensor. This approach also had the advantage, that primers, which introduced vector-facing homologous ends, could be reused, as fluorescent proteins often share N-and C-terminal ends. This made it possible to turn sensor assembly into a modular system with reusable parts, enabling combinatorial cloning and quick exchange of individual components.

Three different vectors for very specific tasks were designed: (1) a small bacterial expression vector for cloning and recombinant expression in *E. coli*, (2) an eukaryotic expression vector for the transfection of cultured cells and (3) a large viral vector for transduction of brain slices. These three vectors were constructed based on pRSET B (Invitrogen), pcDNA3 (Invitrogen) and pSinRep5 (Invitrogen) respectively. The resulting harmonized vectors were named pRSET SL, pcDNA3 SL and pSinRep5 SL (SL stands for SLiCE). Since homologous regions would be shared in all three of those vectors, these sequences had to be designed very carefully to fulfill multiple tasks in different backgrounds. pRSET B, a small pUC based expression vector for *E. coli*, encodes a 34 amino acid N-terminal fusion tag upstream of the multiple cloning site (MCS). This fusion tag contains the 6 x Histidine sequence for protein purification, a transcript stabilizing sequence from gene 10 of the T7 phage, the Xpress™ epitope and an enterokinase recognition site, to enable the removal of this peptide from expressed protein. Since translation starts at the N-terminal fusion tag, no start codon is required when transferring genes of interest into the MCS of pRSET B. In pcDNA3 and pSinRep5 on the other hand, no such leader peptides are present and translation starts from the start codon provided by the gene of interest. Therefore, the upstream homologous region for the vector system was designed in such a way, that it could encode the leader peptide of pRSET B and also provide a valid translation start. Figure 3.1 A depicts both the up- and downstream homologous region of pRSET SL and pcDNA3 SL. The upstream homologous region partially encoded the pRSET leader peptide, which was only expressed in pRSET SL. Furthermore, it contained the Kozak-consensus sequence and a start codon, which served as a translation start in pcDNA3 SL and pSinRep5 SL; in

pRSET SL this sequence was translated into Ala-Thr-Met and part of the N-terminal fusion peptide. The 33 bp downstream homologous region was adapted from the downstream region of pcDNA3 and contained all three possible stop codons, providing translation termination for various organisms. To be able to linearize these vectors efficiently, an EcoRV restriction site was placed between the homologous regions in pRSET SL and pcDNA3 SL. The 6 bp recognition site was omitted during the SLiCE reaction and not integrated into resulting constructs. EcoRV was chosen, because it leaves “blunt” ends, reducing the possibility of hybridization of overhangs after the restriction digest. Furthermore, the restriction sites BamHI and EcoRI were retained, to allow compatibility with restriction cloning.

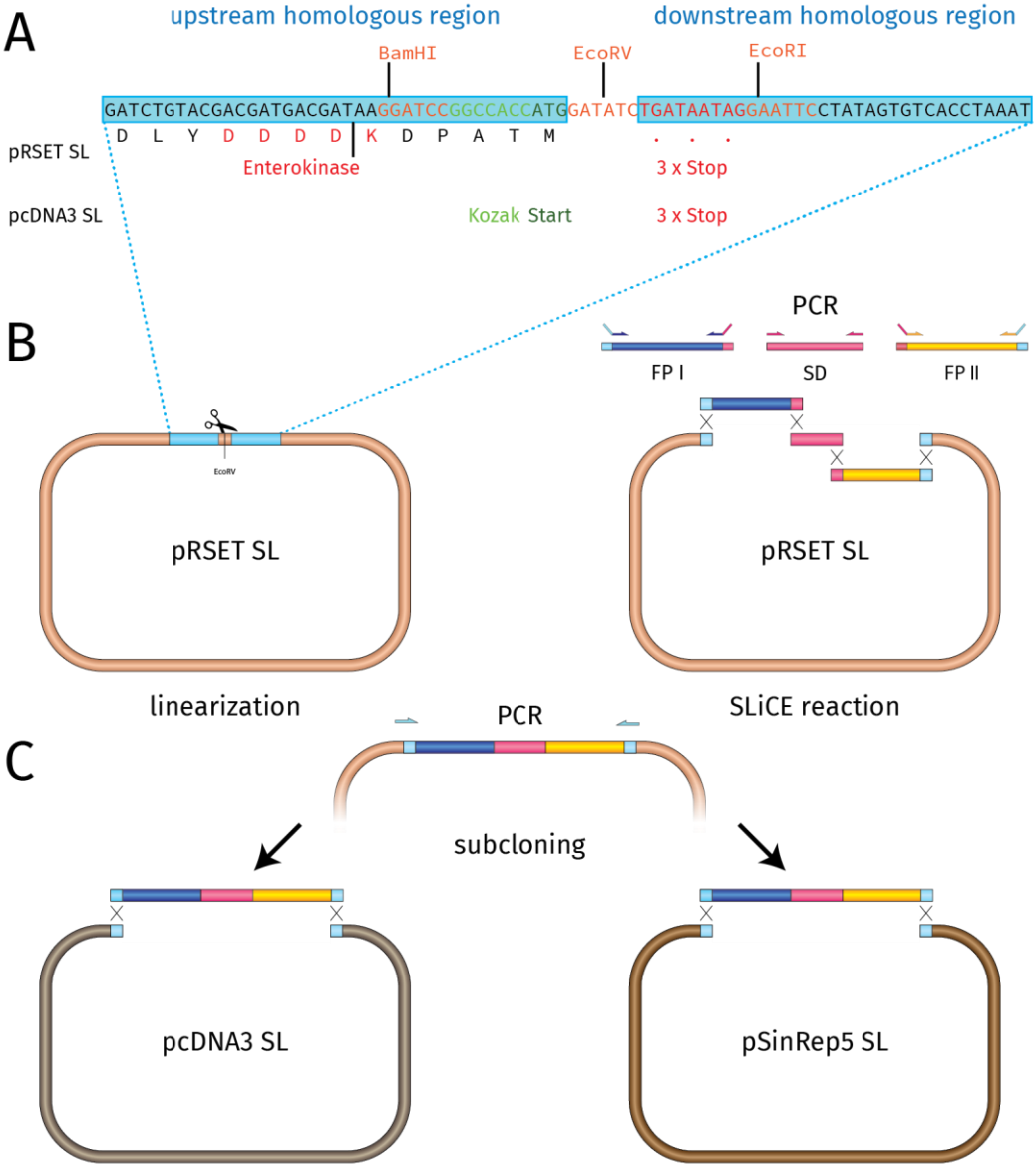

Figure 3.1 Harmonized vector system for biosensor development

A) Sequence of the homologous regions in pRSET SL and pcDNA3 SL B) SLiCE cloning with vector system: pRSET SL is linearized with EcoRV, PCR introduces overhangs in sensor fragments, one-step assembly in SLiCE reaction C) Subcloning of sensor prototypes into pcDNA3 SL and pSinRep5 SL with SLiCE reaction

The vectors pRSET SL and pcDNA3 SL were generated by amplifying pSRSET B and pcDNA3 with 5' phosphorylated primers containing the 5' overhangs, which introduced the synthetic sequence of the homologous regions (primer 3 and 4 for pRSET B and primer 5 and 6 pcDNA3, see Table 2-2). The resulting PCR products were circularized by ligation and sequence identity was confirmed by Sanger sequencing (sequence of pRSET SL under 6.1.1 and pcDNA3 SL under 6.1.4).

In the viral vector pSinRep5 SL, the homologous regions were shortened to 21 bp on either side, to minimize the impact of these foreign sequences on viral expression. Besides that, EcoRV was replaced with SwaI, as two EcoRV recognition sites were already present in this vector. To generate pSinRep5 SL, the counter selection cassette of pRSET SL II (see 3.1.2) was amplified with primers 7 and 8 (See Table 2-2) and inserted via SLiCE cloning into SphI linearized pSinRep5, omitting its MCS. The resulting plasmid was sequenced to confirm successful integration (for sequence of pSinRep5 SL see 6.1.6).

These three vectors and their iterations were very successfully applied throughout this work and proofed to be very reliable tools. In brief, a typical usage scenario was as follows (See Figure 3.1 B): pRSET SL was produced in large quantities in *E. coli*, isolated through "midi-prep", then linearized with EcoRV and purified. The linearized vector was produced in large batches and stored until needed. For sensor prototyping, two fluorescent proteins and a sensor domain were amplified from template, introducing homologous ends via 5' overhangs on primers. PCR products were purified and subsequently assembled with the linearized vector in a SLiCE reaction and transformed in *E. coli*. To confirm successful assembly, colony PCR was performed with primers 1 and 2 (see Table 2-2) and positive clones were inoculated for small scale plasmid purification. Isolated plasmids were then partially sequenced to demonstrate correct assembly. This protocol could be easily parallelized and performed for several dozen different sensor constructs at the same time. Typically, this procedure took only 2.5 days from amplification of sensor fragment to fully assembled plasmid, enabling many rounds of prototyping in relatively short time spans. Following successful cloning, sensor prototypes were expressed in *E. coli* and characterized *in vitro*.

Subcloning followed a similar strategy (see Figure 3.1 C). Characterized sensors were amplified from pRSET SL using primers 9 and 10, then assembled in a SLiCE reaction with linearized vector pcDNA3 SL or pSinRep5 SL and subsequently confirmed via colony PCR and sequencing.

3.1.2 Counter Selection

While the in 3.1.1 described vector system was very efficient in combination with SLiCE cloning, occasionally there was a considerable amount of false positive colonies. It was observed, that most false positive colonies harbored empty vector. Furthermore, the number of these “background” colonies increased when linearized vector was used, which was produced in large scale midi-preps. Therefore, it was suspected, that this background was the result of incompletely digested plasmid DNA, which was retained throughout the purification process. In support of this, there was no difference in length between the digested and undigested vectors and thus gel separation could be ineffective. To overcome this problem and reduce the number of false positive colonies, which had to be tested after cloning, a two layered strategy was employed to eliminate false positive colonies transformed by un-cut vector DNA.

Scholz et al. (2013) demonstrated a counter selection strategy for removing background from parental vector DNA, by constitutively expressing the toxic *ccdB* gene from the same plasmid. A similar strategy was utilized to eliminate background from un-cut vector. Instead of *ccdB*, the *sacB* gene of *bacillus subtilis* was chosen as a counter selection gene. *SacB* encodes a levansucrase which catalyzes the hydrolysis of sucrose and subsequent synthesis of levans, high molecular weight fructose polymers (Pelicic et al. 1996). Expression of *sacB* in *E. coli* is lethal if exposed to 10 % sucrose in media containing low NaCl concentrations (Gay et al. 1985, Blomfield et al. 1991). This had the advantage of a conditional counter selection, meaning plasmids encoding *sacB* could be propagated normally under standard conditions and only became lethal in sucrose containing media. *SacB* was constitutively expressed under the control of the lpp^{p}-5 promoter and placed between the up- and downstream homologous regions of the vectors described above (Figure 3.2). The lpp^{p}-5 promoter is an optimized version of the promoter of the *E. coli* outer membrane lipoprotein (*OmpA*), a very strong constitutive promoter (Inouye and Inouye 1985).

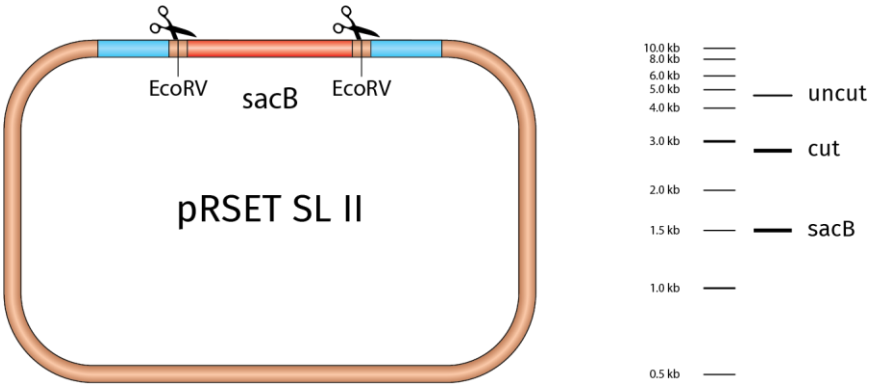

Figure 3.2 Counter selection with *sacB* in pRSET SL II

Incomplete linearization of pRSET SL II with EcoRV leads to 3 distinct gel bands, which could be selectively isolated through gel purification

The counter selection gene was flanked by two EcoRV (SwaI for pSinRep5) restriction sites. Digestion with EcoRV linearized the vectors much like their earlier iterations, removing the fragment encoding *sacB* in the process. The resulting vector fragments were identical to linearized vectors described in 3.1.1.

This enabled the reduction of background on two levels: During gel purification, completely digested vector DNA displayed a clear gel shift towards undigested or partially digested vector and could therefore be effectively isolated from these species (See Figure 3.2). Furthermore, the background was reduced by plating transformed *E. coli* on YTS agar, containing 10 % (w/v) sucrose, after SLiCE reactions. Bacteria which were transformed with un-cut plasmid DNA, which was carried through purification steps, were then unable to produce colonies.

This counter section strategy was developed in close collaboration with Tobias Kruse, who kindly provided the plasmid, encoding *sacB*. The counter selection cassette was amplified with primers 11 and 12 (See Table 2-2), purified and then assembled in a SLiCE reaction with linearized pRSET SL and pcDNA3 SL vectors, resulting in the plasmids pRSET SL II (see 6.1.2 for sequence) and pcDNA3 SL II (see 6.1.5 for sequence). To assess the effectivity of the counter section 270 ng of undigested pRSET SL II plasmid were transformed into *E. coli* XL1 Blue and plated either on selective LB agar plates or selective YTS agar plates. After overnight incubation at 37 °C LB agar plates displayed dense bacterial lawns, while YTS plates were empty. While these were relative extreme conditions, with very high concentrations of undigested vector, it demonstrated that this counter selection strategy is very effective even under adverse conditions.

The strategy of using SLiCE with harmonized vectors and *sacB* counter selection proofed to be very reliable and was used extensively throughout this work.

3.1.3 Improvement of His-tag Purification

During biosensor development prototype sensors were purified via metal ion affinity chromatography, in order to assess their biophysical properties. Sensor candidates were assembled in pRSET SL II, which introduced a 6 x His-tag at the N-terminus, mediating affinity towards immobilized nickel ions on agarose beads. The purification protocol had to fulfill two very important criteria: the procedure had to be as simple as possible in order to allow high throughput processing; and the purity of protein preparations had to be sufficient, to allow biophysical characterization assays. The purification was performed using Nickel-IDA Agarose bead suspensions (Jena Bioscience) in combination with gravity flow columns (Qiagen), closely following the guidelines provided by the manufacturers.

In order to assess the performance of this purification procedure, mKO_K, an orange fluorescent protein (Tsutsui et al. 2008), was cloned into pRSET SL II in a SLiCE reaction, using primers 15 and 16 to amplify the insert from template, and purified as described in 2.2.2.2. Then, the identity and quantity of proteins within the preparation was determined via a mass spectrometry based quantitative shotgun proteomics approach. Mass spectrometry and analysis of proteomics data were performed in close collaboration with Daniel Hornburg from the Department of Proteomics and Signal Transduction at the Max Planck Institute for Biochemistry. Figure 3.3 displays the composition of a standard purification of mKO_K, as determined by mass spectrometry.

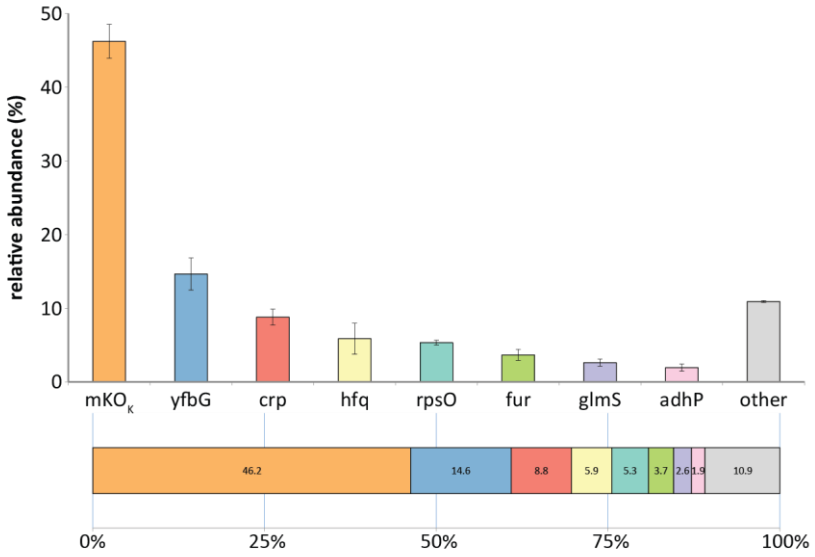

Figure 3.3 Composition of His-tag purified protein preparations of mKO_K

yfbG – *E. coli* formyl transferase
crp – *E. coli* cAMP receptor protein
hfq – *E. coli* host factor-I, RNA-binding protein, phage Qβ replication
rpsO – *E. coli* regulatory ribosomal protein
fur – *E. coli* ferric uptake regulator
glmS – *E. coli* Glucosamine-6-phosphat synthase
adhP – *E. coli* alcohol dehydrogenase, propanol-preferring

In total, 429 individual proteins were identified in the sample. Due to the logarithmic abundance distribution, only the eight most abundant proteins accounted for more than 1 % of the total protein amount. Critically, the target protein mKO_K constituted only 46 % of the sample.
The seven main co-purified proteins (yfbG, crp, hfq, rspO, fur, glmS and adhP) together accounted for 43 % and the remaining 421 identified proteins for 11 %. All seven main co-purified proteins were identified as *E. coli* proteins, which are common byproducts in metal ion chromatography of recombinantly expressed His-tagged proteins (Bolanos-Garcia and Davies 2006).

These proteins were enriched through the purification procedure, because they display surface clusters of histidine residues, competing with His-tagged proteins for nickel binding sites (Bolanos-Garcia and Davies 2006). While purity of fluorescent biosensors in purified protein preparation is important for the determination of the extinction coefficient (Tsien 1998), it was in general not considered to be a major concern for the determination other biophysical and photo-physical characteristics, as these sensors will be used in a cellular environment in the presence of thousands of other proteins, and therefore have to function regardless. However, the presence of the cAMP receptor protein of *E. coli* (crp) was identified as potential problem specifically for the development of biosensor for cGMP. Crp is a transcription regulator protein within *E. coli*, which is activated by cAMP binding (Busby and Ebright 1999). Because of its natural affinity to cAMP and its relative high abundance (~9 %) as a contaminant in purified protein samples, we concluded, that the presence of crp might interfere with determination of cAMP and cGMP binding constants and therefore should be ideally eliminated from samples. The structure of crp (Passner and Steitz 1997) revealed a N-terminal motive containing four histidine residues in close proximity, mimicking a His-tag, which is most likely the reason for its binding to metal ions. While these contaminants could be removed through other additional chromatographic separations like gel filtration, keeping sample preparation as streamlined as possible was highly prioritized. Therefore, the goal was set to adjust the parameters of the standard protocol for metal ion chromatography in order to remove crp as a contaminant and increase the overall purity of target protein.

In essence, in metal ion chromatography, histidine residues of His-tagged proteins bind non-covalently to immobilized nickel ions (Green and Sambrook 2012). Contaminant proteins, which naturally contain exposed histidine clusters, compete with recombinant protein for these binding sites (Bolanos-Garcia and Davies 2006). Imidazole, essentially the isolated side chain of histidine, is used to displace bound protein from these binding sites, to elute it from the column material. It was hypothesized, that there is a delicate equilibrium, between (1) the available binding sites in the nickel resin, (2) the affinity of His-tagged recombinant proteins, (3) the affinity of histidine rich contaminant proteins and (4) the concentration of imidazole in binding buffers, which could be shifted in

favor of the binding of recombinant protein. Three strategies were considered to shift this equilibrium. First, the binding affinity of His-tagged proteins could be increased by lengthening the His-tag, providing more potential binding sites for nickel. Second, increasing the concentration of imidazole in the binding buffer could displace contaminants with lower metal ion affinity, therefore reducing their binding to the resin. Lastly, if more binding sites than recombinant proteins are available, remaining binding sites will be filled by contaminant proteins, therefore column material has to be loaded in excess of His-tagged protein. Since recombinant protein was finite from a given culture volume, protein excess could be achieved by reducing the amount of column material instead of increasing the amount of protein.

In order to increase the affinity to nickel resin of recombinant protein, the 6 x His-tag of pRSET SL II was lengthened to ten histidine residues. The vector pRSET SL II was amplified with the primers 13 and 14, introducing a 10 x His-tag and a 20 bp region at one end of the linear product, which was homologous to the other end. The PCR product was circularized in a SLiCE reaction and sequence identity was confirmed by Sanger sequencing with primers 1 and 2. The resulting vector was named pRSET SL III (annotated sequence under 6.1.3). The orange fluorescent protein mKO_K was inserted into this new vector via SLiCE cloning, using primers 15 and 16 to amplify the insert from template. To compare binding affinity between 6 x His-tagged and 10 x His-tagged protein, mKO_K with both tags was expressed in *E. coli* and isolated as described in 2.2.2.2, using resuspension buffer with 20 mM imidazole. Clarified lysate was incubated with 300 µL Nickel-IDA Agarose slurry bead resin and then loaded onto gravity flow columns. Proteins were eluted from the column material using a step gradient with increasing concentrations of imidazole (1 mL per step) and eluates were fractionated. The concentration of fluorescent protein of each fraction was determined as measured by total fluorescence of the sample, calibrated against a standard curve. Figure 3.4 A depicts the amount of eluted protein after each step as a percentage of the total eluted protein. These results showed clearly, that the affinity of 10 x His-tagged mKO_K towards the nickel was greatly increased compared with 6 x His-tagged mKO_K. While 50 % of the protein with the 6 x His-tag had eluted already at 66 mM imidazole, the 10 x His-tagged protein reached 50 % elution at around 130 mM imidazole. Therefore, not only the affinity towards the raisin was increased, but also relatively high concentrations of imidazole could be used in the binding and wash buffer, as only a small fraction of 10 x His-tagged protein eluted at elevated concentrations, making more stringent washing conditions possible.

Next, the volume of column material was evaluated, to estimate at which point binding sites were saturated with recombinant protein. To that end, 10 x His-tagged mKO_K was expressed in a 600 mL culture, yielding 60 mL clarified bacterial lysate. The Imidazole concentrations in the binding and wash buffer were elevated to 40 mM.

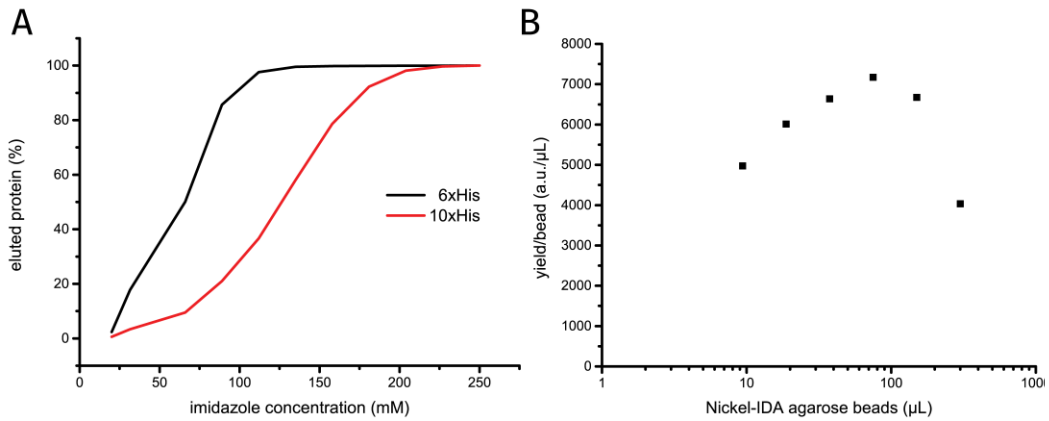

Figure 3.4 Optimization of binding conditions for His-tagged proteins

A) Relative concentration of eluted 6 x His and 10 x His-tagged mKO_K in imidazole step gradient fractions B) Relative yield of 10 x His-tagged mKO_K as a function of bead suspension volume

Six aliquots of 10 mL clarified lysate were incubated with decreasing volume of Nickel-IDA agarose bead suspension (300 µL, 150 µL, 75 µL, 37.5 µL, 18.75 µL, 9.375 µL), loaded onto gravity flow columns and subsequently washed with 10 mL wash buffer. Proteins were eluted with 1 mL Protein Elution Buffer and the concentration of fluorescent protein was determined, as measured by total fluorescence. Protein yields were set in relation to the bead volume, as seen in Figure 3.4 B. This experiment clearly demonstrated that binding sites in the column material were not saturated using 300 µL material, leaving room for contaminant proteins to bind. In fact, the total protein yield using 300 µL and 150 µL bead volume, showed no difference. Maximum yield per volume of beads was only reached at 75 µL under these conditions. Reducing the column material volume below 75 µL, reduced the yield again, possibly because with such small bead volumes losses due to retention on container walls are getting more significant.

These preliminary results outlined the parameter space, in which the protocol for purification of 10 x His-tagged proteins could be optimized. Specifically, the concentration of imidazole was considered in a range between 20 mM and 90 mM and the volume of Nickel-IDA agarose beads in a range between 20 µL and 150 µL.

Classical “one factor at a time” optimizations have the disadvantage that non-linear dependencies of variables cannot be discovered and usually many individual experiments have to be performed in order to find optimal conditions (Hibbert 2012). Because biological experiments are often very complex and require multi parameter optimization, the Design of Experiments (DoE) approach recently gained traction in the life science community (Hibbert 2012). DoE is a statistical method in which multiple parameters are varied at a time and then empirically modelled by linear, higher order functions and multiplicities in order to discover dependencies between variables. To optimize the protocol for 10 x His purification, DoE was employed, using the software tool MODDE 10 (Umetrics). Two parameters were optimized: imidazole concentration in Protein Resuspension Buffer (used for binding and washing) and volume of Nickel-IDA agarose column material, using protein purity, as measured by mass spectrometry, and yield, as estimated by total fluoresce, as performance read outs. For optimization, a central composite design for response surface modelling (RSM) was chosen. The range for each of the two factors were: 20 mM–90 mM, with a center point of 55 mM, for the imidazole concentration and 20 µL–150 µL, with a center point of 95 µL, for the volume of Nickle-IDA agarose bead suspension. In total 11 experiments were performed, combining minimum and maximum levels of each factor and a triplicate of the center point for estimation of robustness. Figure 3.5 A depicts the response surface of the DoE model with respect to purity. The optimum in this model was between 55 mM and 80 mM imidazole and between 80 µL and 150 µL bead suspension, displaying a purity > 99 %. When the yield was taken into account, the optimum range shrank considerably, as can be seen in Figure 3.5 B.

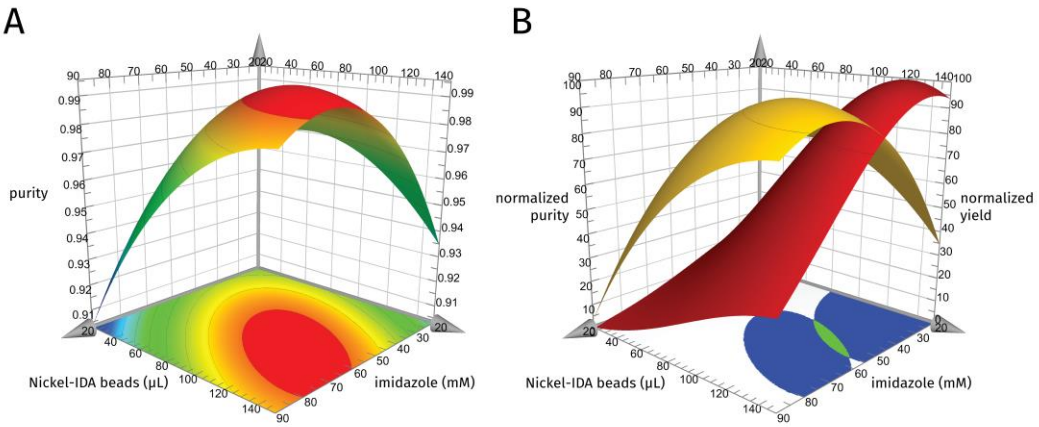

Figure 3.5 DoE results for 10 x His-tag purification protocol

A) Surface plot imidazole vs. Nickle-IDA bead volume with respect to protein purity. B) Sweet spot analysis: surface plots imidazole vs. Nickle-IDA bead volume for protein purity (yellow) and protein yield (red), 2 dimensional projections onto the base plane include conditions with purity > 0.99 and yield > 400 a.u. (green intersection)

From this analysis the optimal conditions were determined to be 55 mM imidazole concentration and 120 µL Nickel-IDA agarose bead suspension. Additionally, when looking more closely into the data of the individual experiments it became apparent that, even though depleted by a factor of ~300 compared to the standard 6 x His-tag purification, the cAMP receptor protein was still present within samples with low imidazole concentrations (20 mM). In samples with imidazole > 20 mM crp was not detected, underlining the validity of the assumptions initially made.

In conclusion, it can be said, that the purification protocol was improved very successfully, with adjustments of critical parameters and the switch to an elongated His-tag. Biosensors now could be purified with purities >99 %, while keeping the purification procedure simple and therefore more useful for high throughput applications. Furthermore, the most critical contamination for the development of biosensors detecting cyclic nucleotides was completely removed, accomplishing another important goal of this project. While these results were very encouraging, it needs pointing out, that the equilibrium between the individual compounds (His-tagged recombinant proteins, histidine rich contaminant proteins, imidazole and binding sites in Nickel-IDA agarose beads) is fragile and high expression yields of recombinant proteins are needed to push the it into the right direction.

3.2 Automated Set-up for Biosensor Screening

To refine biosensors it is often necessary to employ directed evolution and large scale screening techniques. These methods introduce variation into a given prototype sensor construct through mutagenesis, to create large libraries of different sensor variants. These libraries are then screened for improvements in performance and promising candidate variants are selected and further characterized. We recently demonstrated a technique for screening large libraries of biosensors in *E. coli*(Litzlbauer et al. 2015) colonies (Litzlbauer et al. 2015). In brief, biosensors were diversified through various PCR based methods *in vitro* (error prone PCR, introduction of degenerated linkers), cloned into expression vectors and transformed into competent *E. coli*, which are streaked onto selective agar plates. Each colony represented one variant, with each plate containing between 800–1,000 colonies. These colonies were then transferred to blotting paper and imaged with a custom-build wide field fluorescence imaging set-up, to assess performance of individual sensors. This assessment was carried out by monitoring the read-out variable of the biosensor (FRET ratio, fluorescence intensity) over time, while first applying permeation agents and then the analyte of interest with the help of a spray gun. Changes in the read-out variable were recorded and analyzed by custom software tools. The best performing variants were highlighted on a picture of the particular plate and later hand-picked for further characterization. This procedure enables high throughput screening of thousands of sensor variants. However, the main bottleneck of this method is the requirement for hand-picking colonies. This labor-intensive step limits the number of variants that can be analyzed and also potentially introduces errors, as extended periods of colony picking represent a challenge to body and mind.

While systems for automated colony picking are available for industrial applications, the selection criterion is the color of a colony (e.g. blue/white screening). These systems are not suited for fluorescence applications and would have to be retrofitted for this purpose. Adding to that, as these appliances are scaled for industrial purposes, they are relatively expensive. Therefore, the aim was to eliminate the bottleneck in the screening procedure, by building a low cost, integrated automated screening system, which provides online imaging, analysis and colony picking. This would increase throughput for directed evolution applications, while reducing human error. Furthermore, a focus was put on improving data quality, by enhancing wide field fluorescence imaging capability of this set-up.

This project was undertaken in a collaborative effort with David NG, who wrote the software controlling the imaging set-up.

3.2.1 Mechanical components

3.2.1.1 3D Manipulator

The colony picking unit of the integrated screening system had to fulfill the following criteria: (1) it had to be fast, to allow high throughput applications, (2) it had to be very precise and accurate in order to be able to successfully pick individual colonies, and (3) it had to be flexible to allow for different configurations (e.g. inoculating liquid media or re-plating on solid media). With these criteria in mind, the choice for the 3D manipulator fell on a delta robot design. The delta robot is a type of parallel robot and was first envisioned by Clavel (1988) for pick and place applications. It has three degrees of freedom in translation and allows for fast movement (Merlet 2006). Recently, delta robot designs have become very popular in the 3D printing community for the use in fused deposition modeling printers (FDM). Because FDM printers have similar requirements as the envisioned picking system (3D manipulator with high resolution, precision and speed), its design was adapted from an existing open source 3D printer design. Specifically, a 3D printer developed by Johann Rocholl in 2012, the “Rostock”, was used as the base for a colony picking system.

The colony picking system consisted of a triangular platform, harboring the pick head, which was held in place by three pairs of parallel arms made of carbon fiber tubes. The parallel arms in turn were mounted to three carriages, which were seated on vertical linear motion systems. The arms were connected with the platform and the carriages via custom build, magnetic, universal ball-joints, to ensure smooth motion in all directions with minimum tolerances. The linear motion systems, consisted of three pairs of 600 mm hardened steel rods fitted with linear bearings. They were positioned at 120° intervals on the edge of a circle with a radius of 210 mm, encompassing a cylindrical volume. The carriages seated with linear bearings on these rails were driven vertically by geared belts through the action of three NEMA 17 stepper motors mounted at the base of each individual column. Because of the fixed length of the arms and their parallel configuration, vertical movement of the three carriages could be converted into 3 degrees of translation movement of the platform harboring the pick head. This manipulator set-up allowed three-dimensional movement of the pick head within a cylindrical volume with a radius of 210 mm and a height of 200 mm, with a resolution of 0.05 mm in X/Y/Z.

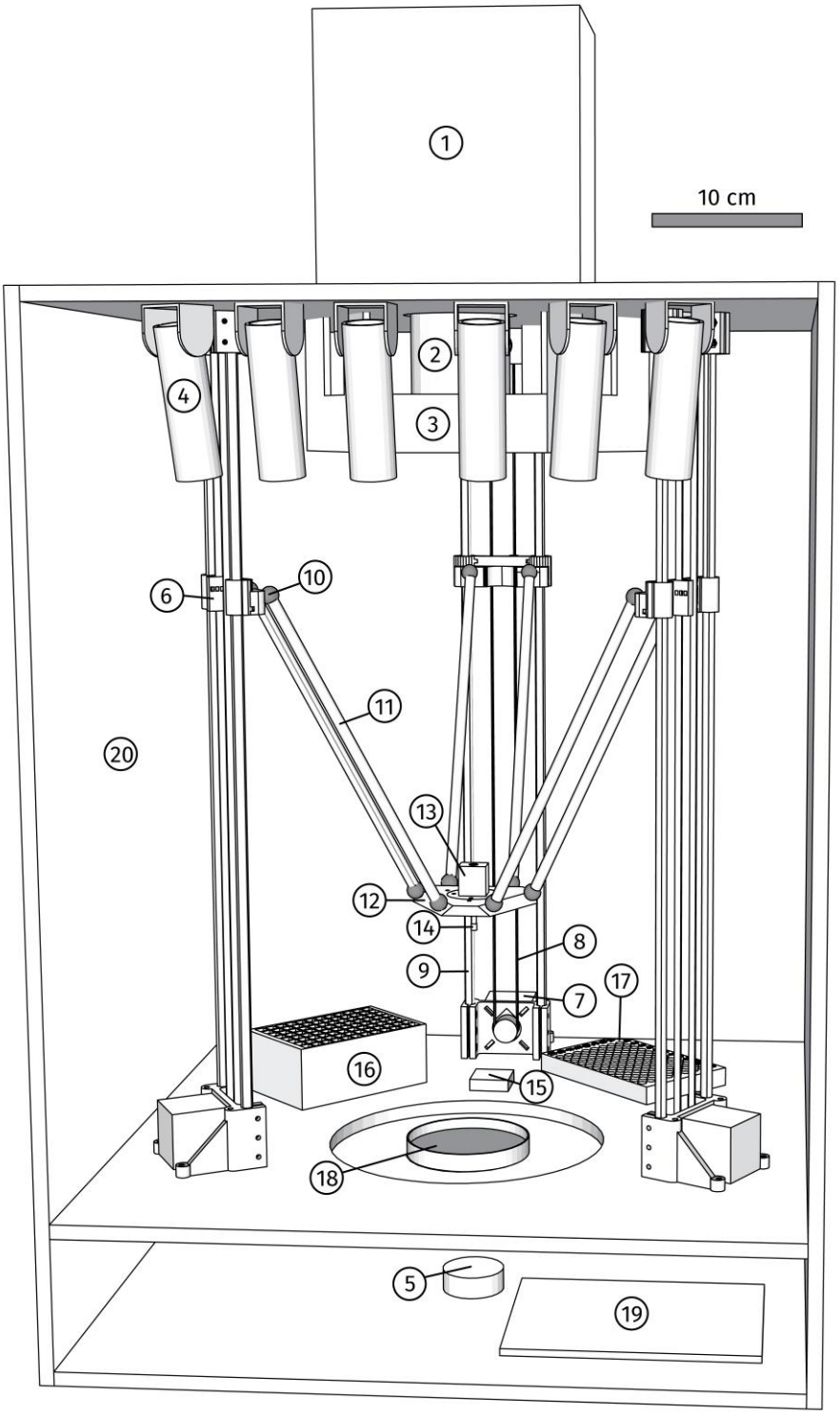

Figure 3.6 Automated screening set-up

1) camera, 2) zoom lens 3) filter wheel 4) illumination LED in tube 5) bottom illumination LED array 6) carriage 7) NEMA 17 stepper motor 8) gear belt 9) linear rail system 10) magnetic ball joints 11) parallel arms from carbon fiber rods 12) pick head 13) electro magnet 14) steel rod 15) sphere testing electrode 16) deep multi-well plate for inoculation 17) 96 well plate harboring steel spheres 18) agar plate 19) controller board 20) housing of MDF

3.2.1.2 Pick Head

For the pick head several important aspects had to be addressed. In order to avoid cross-contamination, tips had to be sterilized in between colonies in a fast manner. Colonies had to be picked from agar plates with varying heights and from blotting paper, rendering hard coding of height coordinates useless. Furthermore, it was desirable to be able to inoculate small liquid cultures in deep multi-well plates and to re-streak clones onto selective agar plates.

In order to address these issues, we designed a special pick head. Usually picking tips are sterilized either by heat or by dipping them into sterilization solutions. This procedure is speed limiting and would cause build-up of residues over time. Therefore, we opted for a replaceable tip, which is brought into contact with the bacterial colony, used for inoculation and then discarded. As a replaceable tips 2 mm steel spheres were used, which were placed individually in 96 well plates with V-bottom. The pick head itself consisted of a steel rod resting inside a copper coil. When put under load the coil induced a magnetic field inside the steel rod, lifting it vertically by around 5 mm (see Figure 3.7). The magnetized rod then was used to lift the steel spheres and transport them inside the volume of the 3D manipulator. When the electro magnet was turned off, the rod would drop into its initial position, ejecting the steel sphere downward in the process. This simple system allowed to quickly picking-up sterilized steel spheres from a 96 well plate, dipping them into a colony and then transferring the sphere into a deep well plate, inoculating liquid medium. For each colony, a fresh sterilized steel sphere was used, avoiding cross contamination.

Because of the varying heights of surfaces which were picked from, a responsive system was built to detect when the steel sphere touched the surface harboring colonies. For detection, an electrode was connected to either the agar plate or to a movable steel stage on which the blotting paper was placed. The other side of the detection circuit was connected to the steel rod. For picking, a steel sphere was picked up and then continuously lowered towards the agar plate. As soon as the sphere touched the agar plate, the circuit was closed, triggering the halt of the pick head. This system was also used to detect, if a steel sphere had been successfully picked up. This was achieved by first picking up a steel sphere and then lowering it over a test-electrode with a fixed height. The detection circuit could only be triggered if a sphere was present, if no sphere was detected, the pick head would return and try to pick up another sphere. To prevent the steel rod from triggering the detection circuit, its end was isolated using a silicone sleeve.

The standard picking routine was as follows: A fresh sphere was picked up by the pick head and then transferred to the testing electrode, to test whether a sphere was present. If no sphere was detected the head was sent to pick up another sphere Successful sphere detection would cause the pick head to move

above the selected colony. Then the head was lowered until the detection circuit was triggered, bringing the sphere in contact with bacterial material. The head was then lifted, and moved above well of a 24-deep well plate filled with selective liquid medium. For inoculation, the electromagnet was turned off, depositing the sphere into the well. If colonies were re-plated, the head was moved above a selective agar plate and lowered until the detection circuit was triggered. Then the head moved in a square pattern to streak out the bacteria. These squares were placed into a two-dimensional grid, to allow easy identification. After plating, the steel sphere was moved over a beaker glass filled with sterilization solution and discarded.

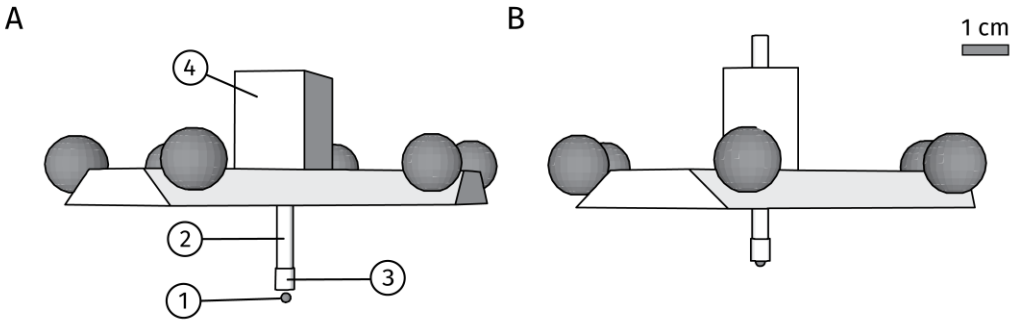

Figure 3.7 Pick head function

A) pick head with disengaged electromagnet; 1) steel sphere 2) steel rod 3) isolation 4) electromagnet B) pick head with engaged electromagnet

3.2.2 Imaging Set-up

3.2.2.1 Illumination

An important factor for biosensor assessment was the frequency of pictures that could be taken within a given time span. To reduce exposure time very bright single-color LEDs (Luxeon® Rebel, Phillips) were used as a light source. Another advantage of using LEDs was that they could be switched rapidly, further reducing time between pictures. LEDs were fitted with culminating lens assemblies and placed at the end of a 100 mm steel tubes. To reduce the emission bandwidth, 25 mm bandpass filters were fitted into the other end of the steel tubes and held in place by set-screws. The tubes were then fixed to the roof of the screening set-up and pointed towards the plate area. This assembly focused a homogenous light beam towards the plate area, while reducing scattering into other directions. In total, six light sources using different LEDS (royal blue, blue, cyan, green, amber and red) were installed in this manner, spanning the visible spectrum. Their emission spectrum could be adjusted by using different bandpass filters.

To increase the flexibility of the automated screening set-up it would be of use to be able to detect non-fluorescent colonies on agar plates. To achieve that, the plate area was fitted with a translucent, opaque bottom with a white LED array underneath. With this bottom illumination, colonies could be identified as dark silhouettes against a white background.

The complete screening set-up was housed in a 50 cm x 50 cm x 60 cm box made from medium density fiber plates, to exclude ambient light.

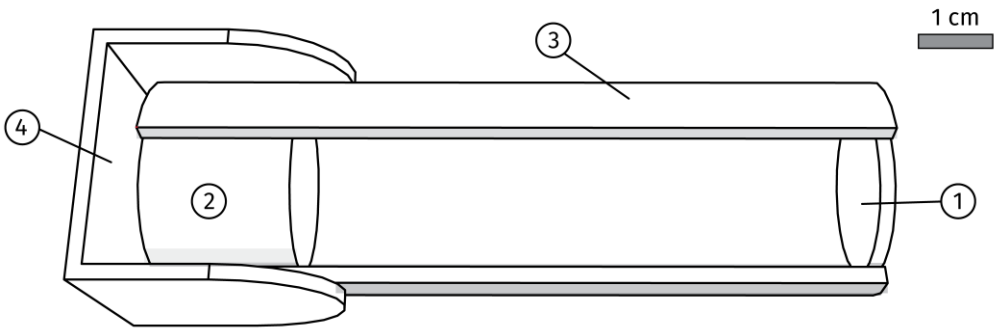

Figure 3.8 LED illumination system

1) bandpass filter 2) LED with culminating lens assembly 3) steel tube 4) holder

3.2.2.2 Filter and Camera

For imaging a CoolSNAP–HQ (Visitron Systems) CCD camera was used. In order to capture as much light as possible and to maximize the plate area on the camera sensor, the camera was fitted with a 50 mm manual zoom lens. This non-standard arrangement required a custom built filter wheel, which was placed in front of the zoom lens. The filter wheel contained seven 50 mm bandpass filters and was housed in a hexagonal box made from black acrylic. The wheel was turned by a NEMA 17 stepper motor and fitted with a 5 mm cube neodymium magnet. This magnet could be detected by a Hall-effect sensor in the housing and served as fixed homing point for the filter wheel.

3.2.3 Controls and Software

All mechanical components of the screening set-up (3D manipulator, pick head, filter wheel) and the LEDS were controlled by a Sanguinololu 1.3b microcontroller board. This controller board is Arduino based and was the developed by 3D printing community for the use in 3D printers and CNC machines. The firmware used, was Marlin 1.0.0.0. The board was connected via USB to the screening computer and operated through G-Code, a low-level programming language for machine tools. The Sanguinololu was fitted with 4 Polulu stepper driver boards, to control the four NEMA 17 stepper motors (three for the 3D manipulator, one for the filter wheel). The LEDs and the plate detection circuit were controlled using digital pins, provided by the board. The board was powered using a 15 V power supply.

On the software side, the screening set-up was controlled by a python program. This program provided high level functions (e.g. pick next colony, take picture etc.) and translated these functions into G-Code to communicate with the microcontroller board. The camera was controlled through the python API of µManager and also integrated into high level functions.

3.3 Novel genetically encoded Sensors for cGMP

From the existing genetically encoded biosensors for cGMP (see 1.3.3.1), cGi500 (Russwurm et al. 2007) was chosen as a starting point for the development of improved cGMP sensors. cGi500 was seen as the most promising candidate for several reasons: (1) it displayed a comparatively high dynamic range, (2) the systematic approach the authors took in establishing the sensor presented a solid knowledge foundation, (3) recently detailed structural insights into the function of the cGKI based binding domains became available (Kim et al. 2011, Huang et al. 2014, Huang et al. 2014) thanks to the continued efforts of the group of Choel Kim, and (4) the mechanism of decreasing FRET with cGMP binding was seen as favorable for the development of FRET based single emission sensors.

CGi500 was assembled *de novo*, based on the detailed description provided by Russwurm et al. (2007) and used as a reference construct. To that end CFP W1B and YFP 10C (Tsien 1998) were amplified using primers 17/18 and 19/20, respectively. The cGKI binding domain was *de novo* synthesized as a linear GeneArt™ String (Thermo Scientific) based on the bovine cGKI protein sequence (GenBank® accession X16086.1) and optimized for *E. coli* expression, using the algorithm provided by the GenerArt™ web interface. Linear fragments were assembled in pRSET SL II (and later transferred to pRSET SL III) via SLiCE reaction and confirmed with Sanger sequencing.

Contrary to the results published by (Russwurm et al. 2007) the dynamic range of this sensor was measured as 29.6 % (77 % in the original publication), when expressed and purified from *E. coli*. This discrepancy could have been caused by the N-terminal fusion purification tag in pRSET SL II. However, results obtained from transferring cGi500 into the mammalian expression vector pcDNA3 SL II, omitting the purification tag, and expressing the sensor in HEK 293 were in good agreement with the results from purified proteins from *E. coli*. Although not explicitly stated in the original publication, it was assumed that this dynamic range was based on measurements corrected for bleed through and cross excitation, which increases the nominal values. For the purpose of this work, bleed through and cross excitation were not taken into account and only uncorrected values are given, as corrected spectra are calculated from several error-prone measurements, introducing additional noise into the system while not adding immediately useful information. Despite this discrepancy, cGi500 in pRSET SL II was used as a reference construct, to compare new sensor variants to.

3.3.1 Improvement of Brightness of cGKI based FRET Sensors

The brightness of a fluorescent protein, as given by the product of quantum yield and extinction coefficient ($\Phi * \varepsilon$), is an important intrinsic property. As described in 1.1.2 both the quantum yield of the donor fluorophore and extinction coefficient of the acceptor fluorophore impact the Förster radius R_0 and therefore directly influence the performance of a FRET sensor. Furthermore, in imaging applications brighter fluorophores are preferred, as they increase the signal to noise ratio in biological samples with background fluorescence and allow reduced exposure times, thus increasing temporal resolution and decreasing photo toxic effects.

While ECFP and EYFP were very popular in early FRET sensors, these fluorescent proteins have considerable downsides (Kremers et al. 2006). Especially ECFP is very critical, as it is very dim ($\Phi * \varepsilon = 13$), with low quantum yield (0.4) and low extinction coefficient (32.5 $mM^{-1}cm^{-1}$) (Tsien 1998). While EYFP is much brighter than ECFP ($\Phi * \varepsilon = 50.9$) it is very environmentally sensitive (Griesbeck et al. 2001). Over the last decade many variants of these fluorescent proteins became available, which improved upon these downsides.

The first objective of this project was to find suitable cyan and yellow fluorescent proteins to replace ECFP and EYFP in cGi500. As potential donors, the bright cyan fluorescent protein variants mCerulean3 (Markwardt et al. 2011) and mTurqouise2 (Goedhart et al. 2012) were chosen, because of their very high quantum yields (0.8 and 0.93, respectively). As acceptors, the EYFP variants mVenus (Kremers et al. 2006), mCitrine (Griesbeck et al. 2001) and Ypet (Nguyen and Daugherty 2005) were picked, because of their reduced environmental sensitivity and increased extinction coefficients.

Furthermore, as the FRET efficiency is highly dependent on the orientation and distance between the fluorophores, this step was also used to methodically explore possible configurations of these fluorescent proteins, to harness potential improvements based on variations in these variables.

Nagai et al. (2004) demonstrated the improvement of the dynamic range of FRET sensors through the usage of circular permutated proteins, as their spatial orientations are shifted due to the relocation of the termini into loops connecting β-sheets. To explore this strategy, the circular permutated fluorescent proteins cpVenus^{CD}174 and cpCitine^{CD}174 (Thestrup et al. 2014) were also considered as potential acceptors.

Shimozono et al. (2006) showed that C-terminal truncation of N-terminally positioned fluorescent proteins can increase FRET efficiency by reducing the distance between the FRET partners. Therefore, N-terminally positioned fluorescent proteins were used both in an 11 amino acid truncated (CΔ11) and a non-truncated version.

Surprisingly, many unimolecular FRET sensors are constructed with an N-terminal donor fluorescent protein and a C-terminal acceptor fluorescent protein. Although this arrangement seems intuitive, it should not be a requirement for FRET. All combinations of fluorescent proteins were therefore assembled both in an N-terminal donor and a C-terminal donor configuration, to thoroughly assess which of these would work the best.

Sensors were constructed in a permutated modular manner, amplifying all variants for each individual component, and then assembling all possible combinations in pRSET SL II using SLiCE. This was achieved by keeping homologous regions for all sensors the same: N-terminal fluorescent proteins (FP I) were amplified with 21 bp homologous overhangs to the vector on the 5'-end and 21 bp homologous overhangs to the sensor domain on the 3'-end (with or without truncation). The sensor (SD) domain itself was amplified without overhangs, as a universal connector between varying fluorescent proteins. C-terminal fluorescent proteins (FP II) finally, were amplified with 21 bp homologous overhangs to the sensor domains on the 5'-end and 21 bp homologous overhangs to the downstream homology region in the vector. After cloning, correct assembly was confirmed via colony PCR. Sensor prototypes were expressed in *E. coli*, purified and their dynamic range was assessed via fluorescence spectroscopy.

Table 3-1 depicts the results of this initial screen of improved cyan/yellow FRET pairs. The wide variety of observed dynamic ranges suggested that this approach was very successful. Some configurations did not result in functional sensors at all, while others outperformed cGi500. Notably, sensors with circular permutated fluorescent proteins performed on average worse than their non-permutated counterparts, with no variant displaying a dynamic range above 21 %. Most of the sensors, which showed no measurable ratio change, had mTurquoise2 as a donor fluorescent protein. The four best performing sensors all used Ypet as acceptor, clearly demonstrating its superiority. Interestingly, sensors with C-terminal donors demonstrated better performance than sensors with N-terminal donors. The four best performing sensor prototypes, named A1–A4, were further characterized (see Table 3-2) and used as base constructs for further improvement efforts.

All four sensors displayed increased dynamic range compared to cGi500 and higher affinity towards cGMP, but also relatively high affinity to cAMP. Besides that, the intrinsic brightness of the sensors was improved significantly in all constructs, with the donor brightness being more than doubled (ECFP 13, mCerulean3 32, mTurquoise2 27.9) and the acceptor brightness increased 1.6 fold (EYFP 50.9, Ypet 80.1).

Table 3-1 Sensor variants with improved fluorescent proteins

FP I	P1,P2	SD	P1,P2	FP II	P1,P2	ΔR/R
mCerulean3	17,21	CNB-A CNB-B	40,41	mVenus	23,24	12.0 %
mCerulean3 CΔ11	17,22	CNB-A CNB-B	40,41	mVenus	23,24	18.7 %
mTurquoise2	17,21	CNB-A CNB-B	40,41	mVenus	23,24	18.6 %
mTurquoise2 CΔ11	17,22	CNB-A CNB-B	40,41	mVenus	23,24	18.5 %
mCerulean3	17,21	CNB-A CNB-B	40,41	mCitrine SF	23,25	15.9 %
mCerulean3 CΔ11	17,22	CNB-A CNB-B	40,41	mCitrine SF	23,25	20.1 %
mTurquoise2	17,21	CNB-A CNB-B	40,41	mCitrine SF	23,25	24.6 %
mTurquoise2 CΔ11	17,22	CNB-A CNB-B	40,41	mCitrine SF	23,25	0.5 %
mCerulean3	17,21	CNB-A CNB-B	40,41	Ypet	26,27	24.0 %
mCerulean3 CΔ11	17,22	CNB-A CNB-B	40,41	Ypet	26,27	38.7 %
mTurquoise2	17,21	CNB-A CNB-B	40,41	Ypet	26,27	0 %
mTurquoise2 CΔ11	17,22	CNB-A CNB-B	40,41	Ypet	26,27	23.7 %
mCerulean3	17,21	CNB-A CNB-B	40,41	cpVenus^{CD}174	28,29	11.8 %
mCerulean3 CΔ11	17,22	CNB-A CNB-B	40,41	cpVenus^{CD}174	28,29	17.6 %
mTurquoise2	17,21	CNB-A CNB-B	40,41	cpVenus^{CD}174	28,29	16.2 %
mTurquoise2 CΔ11	17,22	CNB-A CNB-B	40,41	cpVenus^{CD}174	28,29	18.4 %
mCerulean3	17,21	CNB-A CNB-B	40,41	cpCitine^{CD}174	30,31	15.4 %
mCerulean3 CΔ11	17,22	CNB-A CNB-B	40,41	cpCitine^{CD}174	30,31	16.5 %
mTurquoise2	17,21	CNB-A CNB-B	40,41	cpCitine^{CD}174	30,31	17.3 %
mTurquoise2 CΔ11	17,22	CNB-A CNB-B	40,41	cpCitine^{CD}174	30,31	20.5 %
mVenus	17,21	CNB-A CNB-B	40,41	mCerulean3	23,20	21.3 %
mVenus CΔ11	17,22	CNB-A CNB-B	40,41	mCerulean3	23,20	6.4 %
mVenus	17,21	CNB-A CNB-B	40,41	mTurquoise2	23,20	0 %
mVenus CΔ11	17,22	CNB-A CNB-B	40,41	mTurquoise2	23,20	0.6 %
mCitrine SF	17,32	CNB-A CNB-B	40,41	mCerulean3	23,20	23.3 %
mCitrine SF CΔ11	17,22	CNB-A CNB-B	40,41	mCerulean3	23,20	25.3 %
mCitrine SF	17,32	CNB-A CNB-B	40,41	mTurquoise2	23,20	31.6 %
mCitrine SF CΔ11	17,22	CNB-A CNB-B	40,41	mTurquoise2	23,20	36.7 %
Ypet	33,34	CNB-A CNB-B	40,41	mCerulean3	23,20	1.9 %
Ypet CΔ11	33,35	CNB-A CNB-B	40,41	mCerulean3	23,20	40.5 %
Ypet	33,34	CNB-A CNB-B	40,41	mTurquoise2	23,20	41.3 %
Ypet CΔ11	33,35	CNB-A CNB-B	40,41	mTurquoise2	23,20	49.7 %
cpVenus^{CD}174	38,39	CNB-A CNB-B	40,41	mCerulean3	23,20	18.5 %
cpVenus^{CD}174	38,39	CNB-A CNB-B	40,41	mTurquoise2	23,20	0.4 %
cpCitine^{CD}174	36,37	CNB-A CNB-B	40,41	mCerulean3	23,20	18.7 %
cpCitine^{CD}174	36,37	CNB-A CNB-B	40,41	mTurquoise2	23,20	0.2 %

P1,P2 Primer pair used to amplify the domain

Table 3-2 Characteristics of cGi500 based sensors with improved fluorophores

Name	FP I	SD	FP II	ΔR/R	K_d[cGMP]	K_d[cAMP]
cGi500	ECFP (W1B)	CNB-A CNB-B	EYFP (10C)	29.6 %	470 nM[1]	>100 µM[1]
A1	mCerulean3 CΔ11	CNB-A CNB-B	Ypet	38.7 %	377.0 nM	71.4 µM
A2	Ypet CΔ11	CNB-A CNB-B	mCerulean3	40.5 %	364.5 nM	127.8 µM
A3	Ypet	CNB-A CNB-B	mTurquoise2	41.3 %	300.2 nM	29.7 µM
A4	Ypet CΔ11	CNB-A CNB-B	mTurquoise2	49.7 %	377.5 nM	57.5 µM

[1]as reported by Russwurm et al. (2007)

3.3.2 Improvement of Binding Sites

An important aspect for the function or a biosensor for cGMP is its selectivity towards cGMP over the structurally similar cAMP, which is typically more abundant in cells (Nikolaev and Lohse 2009). The cGMP sensor cGi500 is based on a conformational change in the cyclic nucleotide binding domain of the bovine cGKI, when binding to cGMP (Russwurm et al. 2007) (see Figure 3.9). The cyclic nucleotide binding domain of cGKI consists of two binding sites: a high affinity binding site (CNB-A) and a low affinity binding site (CNB-B) (Kim et al. 2011). Recent studies showed, that CNB-A, while having a very high affinity, only displays poor selectivity (~2 fold) between cGMP and cAMP (Kim et al. 2011). In fact Huang et al. (2014) could demonstrate that the selectivity of cGKI for cGMP is mediated by the second, low affinity binding site (CNB-B). Taken together, the assumption was made, that in a cellular environment the combination of poor selectivity and high affinity of CNB-A and the relatively high concentration of cAMP would lead to a saturation of this binding site, reducing its contribution to the biosensor functionality.

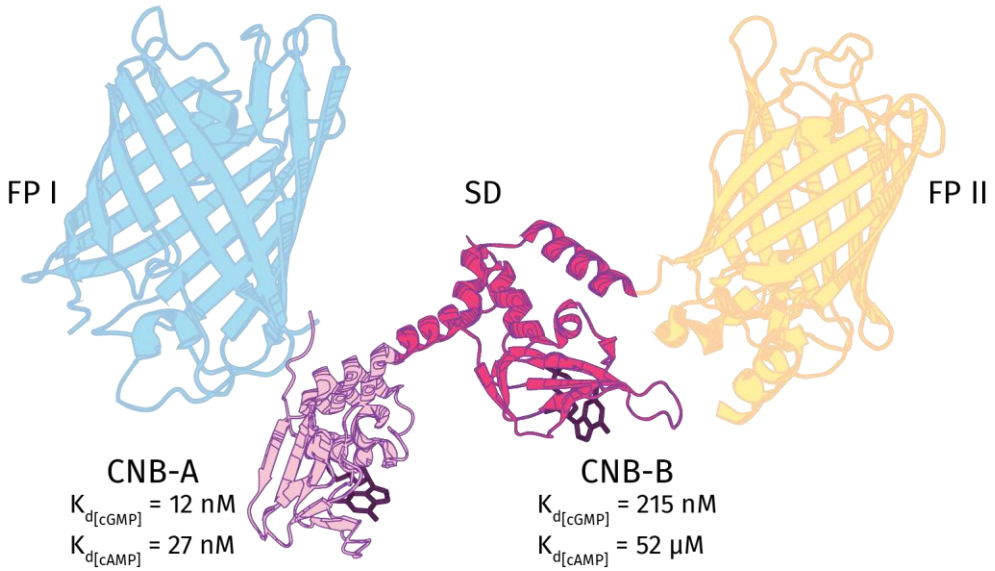

Figure 3.9 Structure scheme of cGi500-like fluorescent biosensors

Sensor domain (SD) is sandwiched between two fluorescent proteins (FPI, FPII) and consists of two binding sites (CNB-A, CNB-B). CNB-A: high affinity binding site with low selectivity (~2 fold) CNB-B: low affinity binding site with higher selectivity (~240 fold)(Huang et al. 2014)

3.3.2.1 Sensors with duplicated CNB-B Binding Site

To increase both the selectivity and the dynamic range of the initial prototype sensors, the sensor domain was modified to incorporate the knowledge gained from the structure of cGKI. While sharing only 37 % sequence homology CNB-A and CNB-B are structurally very similar, displaying two N-terminal α-helices, an 8 stranded anti-parallel β-barrel, which comprises the cyclic nucleotide binding pocket, and an α-Helix at the C-terminus, which is the effector of the conformational change upon ligand binding (Kim et al. 2011, Huang et al. 2014). It was reasoned, that exchanging the CNB-A binding site with a copy of CNB-B while keeping the overall sensor structure constant, would improve the functionality of the sensor, because: (1) both binding sites would be selective for cGMP and (2) both binding sites would contribute in the conformational change of the molecule.

In order to duplicate CNB-B, this binding site was amplified using primers 44/45. Then, the base constructs A1–A4 were all linearized with primers 42/43, omitting CNB-A. The linearized backbones were then assembled with the second binding site in a SLiCE reaction, yielding the constructs depicted in Table 3-3. This replaced CNB-A after its first α-helix with the homologous sequence of CNB-B, retaining a 29 amino acid chain from the N-terminus of the original cGi500 binding domain (see Figure 3.10).

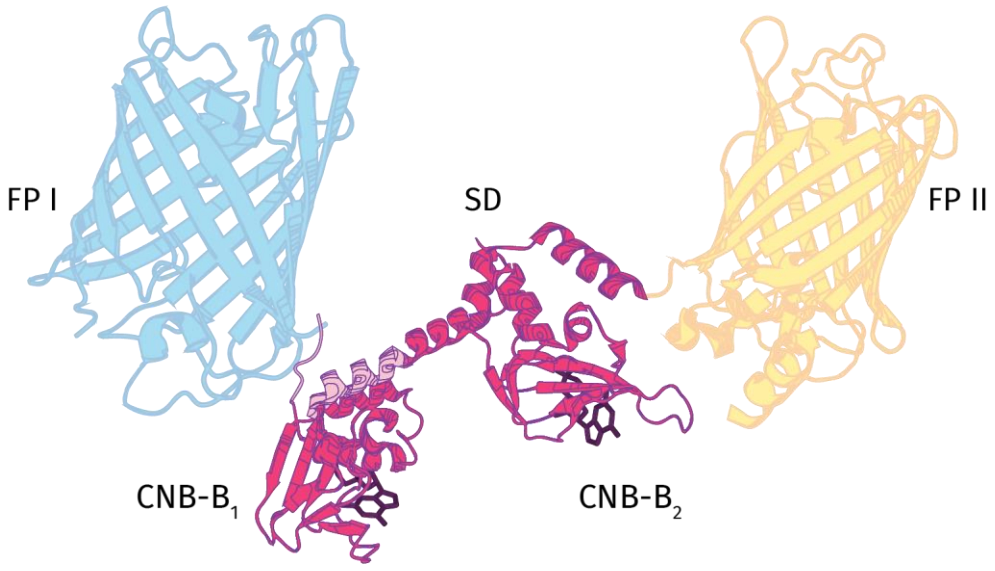

Figure 3.10 Structure scheme of biosensors with duplicated CNB-B binding site

CNB-A in cGi500 like sensors is replaced with a copy of the structurally homologous CNB-B binding site

The resulting prototype sensors were expressed in *E. coli*, purified and then characterized regarding their dynamic range and binding constants towards cGMP and cAMP (see Table 3-3). The results of this binding site graft supported the assumptions made, going into this project. All sensors with duplicated CNB-B binding site displayed improved selectivity for cGMP over cAMP. As expected the affinity for cGMP decreased slightly, because of CNB-Bs lower affinity towards cGMP.

Interestingly, the dynamic range of sensors incorporating mTurquoise2 as a donor increased only negligible, while sensors with mCerulean3 almost doubled in dynamic range. This underlined the observations from earlier experiments that the use of mCerulean3, although very similar to mTurquoise2, had an advantage in the context of the cGMP sensor. Furthermore, it was noted that the sensors B2 (see 6.2.1) and B3 (6.2.2) performed comparably, incorporating the same fluorophores but in opposite configuration with the N-terminal fluorophore being truncated. In this particular arrangement, the donor and acceptor positions were apparently interchangeable. Because of their improved dynamic range sensors B2 and B3 were transferred into pSinRep5 SL and further characterized in brain slices (see 3.3.5), to assess the comparability of *in vitro* data with *in vivo* conditions.

Table 3-3 Characteristics of cGMP sensors with duplicated CNB-B binding site

Name	FP I	SD	FP II	ΔR/R	K_d[cGMP]	K_d[cAMP]
cGi500	ECFP (W1B)	CNB-A CNB-B	EYFP (10C)	29.6 %	470 nM[1]	>100 µM[1]
B3	mCerulean3 CΔ11	2 x CNB-B	Ypet	72.6%	614.6 nM	148.7 µM
B2	Ypet CΔ11	2 x CNB-B	mCerulean3	71.8%	474.5 nM	203.6 µM
-	Ypet	2 x CNB-B	mTurquoise2	47.5%	n.d.	n.d.
B1	Ypet CΔ11	2 x CNB-B	mTurquoise2	53.5%	570.1 nM	251.4 µM

[1]as reported by Russwurm et al. (2007)

3.3.2.2 Sensors with single CNB-B Binding Site

The second option which was explored to improve on the cGi500 binding domain, was to delete the CNB-A binding site altogether. Besides potentially increasing specificity for cGMP, this approach had the additional advantage of simplifying the binding kinetics of resulting sensors. Biosensors with multiple binding sites often display cooperative binding, which poses considerable challenges, when cells with different base levels of analyte are compared (Rose et al. 2014). Furthermore, the biological impact of recombinantly expressed sensor can be reduced, by reducing the number of binding sites. Proteins which have affinity to an analyte will inadvertently buffer the intracellular concentration of this analyte, thus reducing its "free" concentration and therefore influencing the biological system (Rose et al. 2014).

In order to delete the CNB-A binding site, the base sensors A1–A4 were linearized by amplification with primers 43/44, omitting CNB-A. The linear PCR products were then circularized in a SLiCE reaction, yielding the constructs depicted in Table 3-4. Similarly to biosensors with duplicated CNB-B binding sites, the 29 amino acid N-terminal peptide of the original cGi500 binding domain was retained.

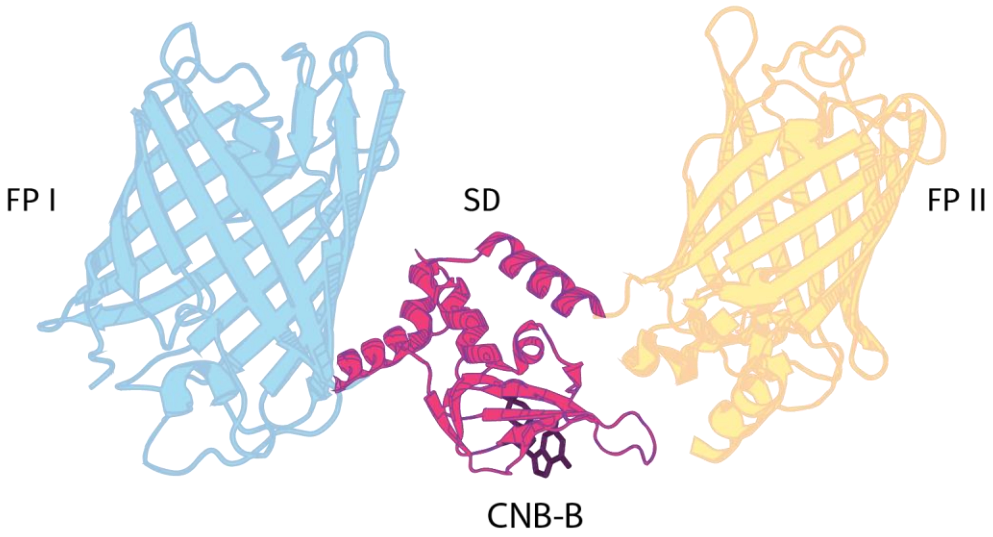

Figure 3.11 Structure scheme of biosensors with single CNB-B binding site

In contrast to the results obtained with duplicated CNB-B binding sites, single binding site sensors with C-terminal donors displayed clearly reduced dynamic ranges compared to the original prototypes. The construct C1 (see 6.2.3) with an N-terminal donor, on the other hand, showed a dramatically increased dynamic range. Donor and acceptor position apparently played an important role in the context of this sensor construct, as opposed to the sensors with duplicated CNB-B binding site. This underlined the necessity for in-depth exploration of possible sensor configurations during the biosensor development process, as results from different systems cannot necessarily be transferred.

Besides its superior dynamic range, the sensor variant C1 was also highly selective with a 1,900 fold discrimination between cGMP and cAMP (compared to 430 fold and 230 fold in B2 and B3 respectively). The loss in affinity was more pronounced than observed with duplicated CNB-B binding sites, indicating that the configuration of a cGMP biosensor can have considerable influence on its affinity, despite the usage of the same binding domain.

C1 was considered a very promising prototype, because of its large dynamic range and high selectivity and was therefore sub-cloned into pSinRep5 SL and further characterized in brain slices (see 3.3.5)

Table 3-4 Characteristics of cGMP sensors with single CNB-B binding site

Name	FP I	SD	FP II	ΔR/R	K_d[cGMP]	K_d[cAMP]
cGi500	ECFP (W1B)	CNB-A CNB-B	EYFP (10C)	29.6 %	470 nM[1]	>100 µM[1]
C1	mCerulean3 CΔ11	CNB-B	Ypet	81.4 %	1.7 µM	3.3. mM
-	Ypet	CNB-B	mTurquoise2	20.8 %	n.d.	n.d.
-	Ypet CΔ11	CNB-B	mTurquoise2	15.2 %	n.d.	n.d.
-	Ypet CΔ11	CNB-B	mCerulean3	25.4 %	n.d.	n.d.

[1]as reported by Russwurm et al. (2007)

3.3.2.3 cGKII based Sensors

Very recently Campbell et al. (2016) solved the crystal structure for the cyclic nucleotide binding domains of the human cGKII, determined the binding constants for its cyclic nucleotide binding sites and elucidated the mechanisms of cGMP binding. This study revealed, that similar to cGKI, the second cyclic nucleotide binding domain (CNB-BII) is responsible for the cGMP selectivity of this protein kinase. Strikingly, the selectivity for cGMP of CNB-BII is even higher (500 fold) than in CNB-B of cGKI (240 fold). The authors linked this selectivity to a different nucleotide binding mechanism, in which cGMP is completely encapsulated by the binding pocket. Adding to that, CNB-BII has a seven fold higher affinity towards cGMP than CNB-B (31.2 nM and 215 nM, respectively)(Campbell et al. 2016).

In order to marry high selectivity with high affinity, a sensor was constructed based on the prototype sensor C1, using the CNB-BII binding site of cGKII. To that end, the vector encoding C1 was linearized by amplification with primers 43/46. A linear DNA fragment encoding residues 291–419 of the human cGKII with 21 bp homologues regions towards the linearized vector on either end was *de novo* synthesized as a GeneArt™ String (Thermo Scientific). Linearized backbone and the fragment encoding CNB-BII were then assembled in a SLiCE reaction, yielding the sensor prototype D1 (see 6.2.4). The sensor was expressed in *E. coli*, purified and characterized *in vitro* (see Table 3-5).

The observed characteristics of this sensor were very promising. D1 exhibited outstanding selectivity for cGMP over cAMP. It was so high in fact, that the affinity for cAMP could only be estimated, because part of the binding curve lay outside of the solubility range for cAMP. Furthermore, D1 had a higher affinity for cGMP than C1. Although this improvement was not as striking as it would have been predicted from the affinity of CNB-BII, it was an improvement none the less. This furthermore underlined the susceptibility of the affinity of nucleotide binding domains to changes in their local environment that was observed before. The dynamic range of D1 was only 42.8 %, which is relatively low in comparison with the sensors B2, B3 and C1, but because of its other promising characteristics it was considered as a base construct for further improvement. D1 was transferred to pSinRep5 SL and evaluated in brain slices (see 3.3.5).

Table 3-5 Characteristics of cGMP sensors with single CNB-BII binding site

Name	FP I	SD	FP II	ΔR/R	K_d[cGMP]	K_d[cAMP]
cGi500	ECFP (W1B)	CNB-A CNB-B	EYFP (10C)	29.6 %	470 nM[1]	>100 µM[1]
C1	mCerulean3 CΔ11	CNB-B	Ypet	81.4 %	1.7 µM	3.3. mM
D1	mCerulean3 CΔ11	CNB-BII	Ypet	42.8 %	955 nM	~24 mM

3.3.3 Color Variants

FRET sensors based on the CFP/YFP pair have been very popular in the past and are still commonly used (compare FRET sensor list compiled by Miyawaki and Niino (2015)). However, this pair has a number of downsides, which make it desirable to establish FRET sensors with different fluorescent proteins. For example, compared to other fluorescent proteins, CFP variants are very dim, even the brightest are less than half as bright as other engineered modern fluorescent proteins (see Table 1-1 under 1.1.4). Furthermore, the 440 nm laser line, which is required to excite CFP/YFP pairs is not as commonly available as the 488 nm laser line, which is used for GFP excitation (Miyawaki and Niino 2015). For cGMP based FRET sensors specifically, a combination with optogenetic actuators like the light activated guanylate cyclase BlgC (Ryu et al. 2010), would be interesting. However, this cyclase is activated by blue light, which would selectively bleach CFP.

In order to provide potential alternatives for various applications, other potential FRET pairs for a cGMP sensor were investigated. The choice fell on green/orange pairs, because they are easily excitable with the 488 nm line, have less spectral overlap with BLUF domain based tools such as BlgC and recently, with Clover and mNeongreen, very bright green fluorescent proteins became available (Lam et al. 2012, Shaner et al. 2013).

Both Clover and mNeongreen were combined with mKO_3 and its circular permutated variant mKO_3 cp53. mKO_3 is a bright orange fluorescent protein without dead-end green intermediates common in orange FPs (Litzlbauer 2015). Prototype sensors were constructed in a modular fashion, amplifying individual components with homologous overhangs and subsequently assembling them in SLiCE reactions, as described before (see 3.3.1). Sensors were expressed in *E. coli*, purified and characterized *in vitro*.

The results of these experiments are listed in Table 3-6. Both of the earlier developed sensor domains (2 x CNB-B and CNB-B) were tested. The overall performance of all green/orange sensors was rather poor in comparison with CFP/YFP based sensors. This result was somewhat surprising, as the spectral compatibility between these pairs was comparable to CFP/YFP pairs used in earlier sensors (mCerulean3-Ypet R_0=10.1 nm, Clover-mKO_3 R_0=10.2 nm, mNeongreen-mKO_3 R_0=10.3 nm, with $\kappa^2 = 2/3$, calculated using a|e UV-Vis-IR Spectral Software 2.2, FloorTools). On the other hand, so far the sensor domains had been optimized using CFP/YFP FRET as a read-out. Since FRET efficiency is very susceptible to chances in orientation (see 1.1.2), it is plausible to assume that new fluorophores with different dipole transition moments are less effective for a given sensor configuration. To optimize the dynamic range of sensors using green/orange pairs, linker regions between sensor domains and fluorescent proteins would have to be investigated, to adjust the angles between functional elements to be more suitable.

Table 3-6 Green-orange cGMP sensor variants

FP I	P1,P2	SD	P1,P2	FP II	P1,P2	ΔR/R
Clover	17,21	CNB-B	40,41	mKO_3	51,52	25.1 %
Clover	17,21	CNB-B	40,41	mKO_3 cp53	53,54	12.6 %
mNeongreen	17,21	CNB-B	40,41	mKO_3	51,52	17.0 %
mNeongreen	17,21	CNB-B	40,41	mKO_3 cp53	53,54	–
mKO_3	47,48	CNB-B	40,41	Clover	23,20	10.0 %
mKO_3	47,48	CNB-B	40,41	mNeongreen	23,20	10.7 %
mKO_3 cp53	49,50	CNB-B	40,41	Clover	23,20	–
mKO_3 cp53	49,50	CNB-B	40,41	mNeongreen	23,20	12.1 %
Clover	17,21	2 x CNB-B	40,41	mKO_3	51,52	22.9 %
Clover	17,21	2 x CNB-B	40,41	mKO_3 cp53	53,54	13.3 %
mNeongreen	17,21	2 x CNB-B	40,41	mKO_3	51,52	21.0 %
mNeongreen	17,21	2 x CNB-B	40,41	mKO_3 cp53	53,54	11.6 %
mKO_3	47,48	2 x CNB-B	40,41	Clover	23,20	9.0 %
mKO_3	47,48	2 x CNB-B	40,41	mNeongreen	23,20	11.9 %
mKO_3 cp53	49,50	2 x CNB-B	40,41	Clover	23,20	6.8 %
mKO_3 cp53	49,50	2 x CNB-B	40,41	mNeongreen	23,20	14.8 %

P1,P2 Primer pair used to amplify the domain

3.3.4 Single Emission, red cGMP Sensors

Depending on the application, single emission biosensors are sometimes of advantage: (1) imaging is less complex than with dual emission sensors, as only a single fluorescent protein has to be monitored (Miyawaki and Niino 2015), (2) the use of a single fluorescent protein allows to image with very wide filters, thus gathering more light, enabling shorter exposure times (3) because of their smaller “spectral footprint” they can be combined with other ontogenetic tools. Furthermore, it would be of advantage to use red-shifted biosensors, as tissue penetration increases with longer wavelength due to reduced scattering (Helmchen and Denk 2005). This motivated the development of a red-shifted single emission biosensor for cGMP, as such a sensor could be combined with activity sensors, like Twich2B (Thestrup et al. 2014) or GCamP6s (Chen et al. 2013) or with cAMP sensors like Epac-S^{H134} (Polito et al. 2013).

Pettikiriarachchi et al. (2012) described the use of the monomeric chromoprotein Ultramarine as a dark acceptor in a FRET based biosensor. In such a sensor, a fluorescent protein (FP) is coupled to a chromoprotein (CP) via a sensor domain (SD) (see Figure 3.12). In the ligand-free state the fluorescence emission of the FP is quenched by the chromoprotein through FRET. Upon ligand binding, a conformational change in the sensor domain reduces the FRET efficiency through changes in orientation and distance, increasing the fluorescence emission in the process. Since all FRET sensors based on the cGK cGMP binding domains displayed a decrease in their FRET ratio after cGMP, it was seen as very suitable for this type of sensor.

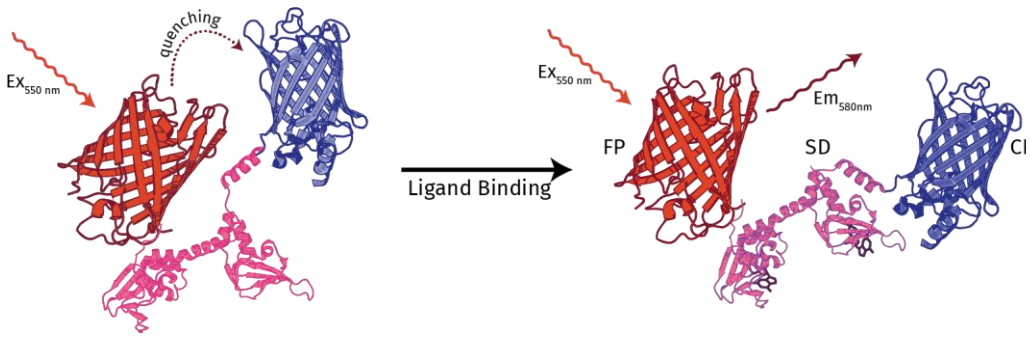

Figure 3.12 Model for FRET based single emission cGMP sensors

A sensor domain (SD) is sandwiched between a single donor fluorescent protein (FP) and an acceptor chromoprotein (CP). The brightness of the FP is modulated by ligand binding to the SD, reducing FRET efficiency through conformational change.

The DNA encoding Ultramarine was *de novo* synthesized as a linear GeneArt™ String (Thermo Scientific) based on the protein sequence in the crystal structure of Rtms5 (Beddoe et al. 2003) and the mutations reported by Pettikiriarachchi et al. (2012), which were introduced into Rtms5 to produce Ultramarine. The sequence was optimized for *E. coli* expression using the algorithm provided by the GeneArt™ web interface. The linear DNA fragment was cloned into pRSETSL II via SLiCE reaction and sequence identity was confirmed with Sanger sequencing.

Red single emission cGMP sensors were assembled in similar fashion as described before. Several donor fluorescent proteins were combined with ultramarine as dark acceptor, in N-terminal and C-terminal configuration using the sensor domains CNB-B and 2 x CNB-B. The following fluorescent proteins were used as donors: the red fluorescent protein mRuby2 (Lam et al. 2012), the afford mentioned orange fluorescent proteins mKO_3 and mKO_3 $cp5_3$ and one fragment of the tandem dimer tdTomato (Shaner et al. 2004) encoding a single fluorescent protein. Prototype sensors were assembled in SLiCE reactions, expressed in *E. coli* and characterized *in vitro*.

Table 3-7 depicts results obtained from the characterization of these biosensors. Overall sensor performance with respect to the change in fluorescence intensity was very promising. As a site note it should be pointed out, that ΔF/F, which is used as a measure for dynamic range in single emission sensors, is not directly comparable with ΔR/R, which is used to characterize FRET sensors. As a point of reference, the average change in fluorescence intensity per channel in the FRET sensor C1 was ~35 %. Interestingly, mRuby2 worked well in sensors with 2 x CNB-B in both N- and C-terminal configuration, similar to the sensors B2 and B3, but with reduced performance in sensors with single binding site.

Table 3-7 Red single emission cGMP sensors

FP I	P1,P2	SD	P1,P2	FP II	P1,P2	ΔF/F
mRuby2	55,21	CNB-B	40,41	Ultramarine	59,60	13.0 %
Tomato	17,21	CNB-B	40,41	Ultramarine	59,60	31.2 %
mKO_3	47,48	CNB-B	40,41	Ultramarine	59,60	21.6 %
mKO_3 cp53	49,50	CNB-B	40,41	Ultramarine	59,60	19.0 %
Ultramarine	57,58	CNB-B	40,41	mRuby2	56,20	15.8 %
Ultramarine	57,58	CNB-B	40,41	Tomato	23,20	12.9 %
Ultramarine	57,58	CNB-B	40,41	mKO_3	51,52	17.4 %
Ultramarine	57,58	CNB-B	40,41	mKO_3 cp53	53,54	6.7 %
mRuby2	55,21	2 x CNB-B	40,41	Ultramarine	59,60	20.7 %
Tomato	17,21	2 x CNB-B	40,41	Ultramarine	59,60	–
mKO_3	47,48	2 x CNB-B	40,41	Ultramarine	59,60	16.8 %
mKO_3 cp53	49,50	2 x CNB-B	40,41	Ultramarine	59,60	11.7 %
Ultramarine	57,58	2 x CNB-B	40,41	mRuby2	56,20	21.9 %
Ultramarine	57,58	2 x CNB-B	40,41	Tomato	23,20	17.4 %
Ultramarine	57,58	2 x CNB-B	40,41	mKO_3	51,52	14.2 %
Ultramarine	57,58	2 x CNB-B	40,41	mKO_3 cp53	53,54	11.6 %

P1,P2 Primer pair used to amplify the domain

It was furthermore noted, that mKO_3 and its circular permutated version mKO_3 cp53, performed similarly, in contrast to the results observed with circular permutated versions of YFP (see 3.3.1).

The best performing sensor, with respect to dynamic range, incorporated Tomato in N-terminal position with a single cGMP binding site (see 6.2.5). With a ΔF/F of 31.2 % it is comparable in performance to the sensors B2 and B3, which displayed an average intensity change of ~30 % per channel. The potential disadvantaged of this sensor could be aggregation, since Tomato is not a monomer.

Overall the results were very encouraging, demonstrating that the concept of such a sensor is applicable for cGMP sensors and yielding several prototype sensors as the base for further improvement.

3.3.5 Characterization of Sensor Variants in murine Brain Slices

To evaluate the performance of optimized sensors and the validity of *in vitro* data, prototype sensors were recombinantly expressed in brain slices and further characterized. Four promising sensor constructs and cGi500 were selected and transferred into the viral vector pSinRep5 SL: the single binding site sensors C1 (see 3.3.2.2) and D1 (see 3.3.2.3) and the 2 x CNB-B based sensors B2 and B3 (see 3.3.2.1).

The following experiments were performed by Dahdjim-Benoît Beltolngar and Pierre Vincent of the Université Pierre et Marie Curie. For in depth information on experimental procedures the reader is referred to Polito et al. (2013) and Polito et al. (2015) In brief, cranial murine brain slices were transduced using Sindbis virus particles to recombinantly express cGMP sensors in striatal neurons. CGMP oscillations were induced pharmacologically, applying the NO donor DEA/NO, to stimulate soluble guanylate cyclases (cGC), and the PDE inhibitor IBMX, to inhibit cGMP degradation.

Figure 3.13 depicts results of a typical cGMP imaging experiment in the murine striatum. The maximum dynamic range of cGMP sensors was estimated by first stimulating cGMP production through an NO donor and then inhibiting

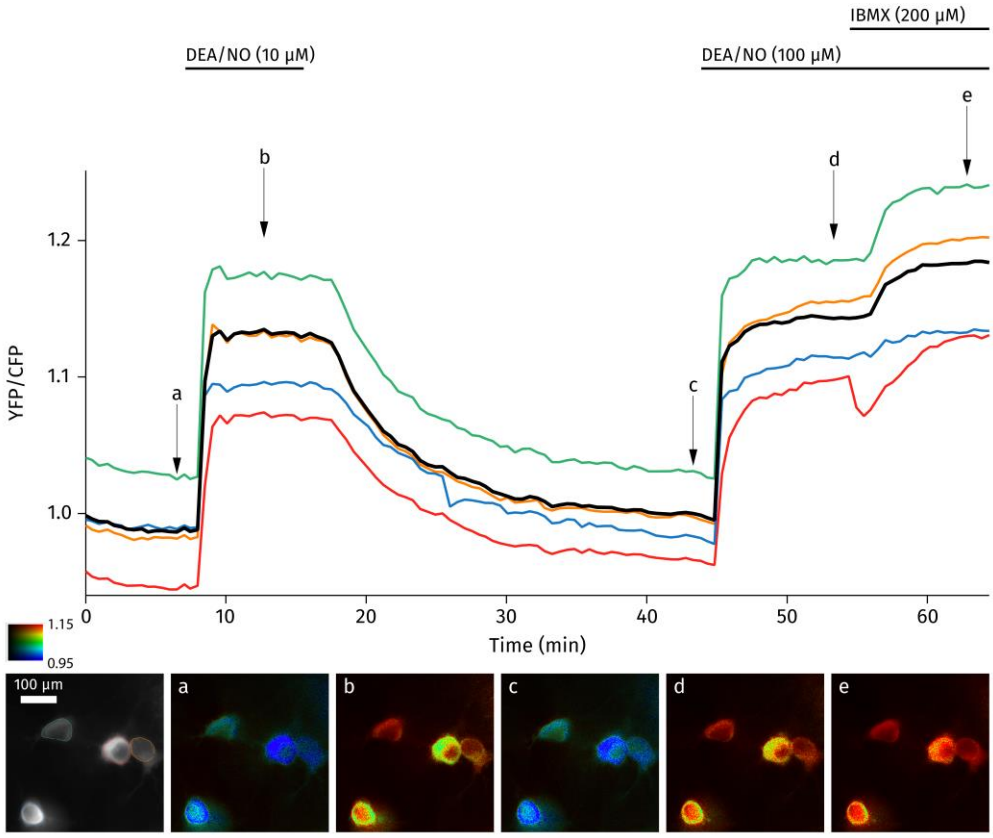

Figure 3.13 cGMP imaging with sensor B2 in the striatum of murine brain slices

sGC is stimulated with the NO donor DEA/NO, producing cGMP. PDEs are inhibited with IBMX, reducing cGMP degradation. Color of traces represents ROIs indicated in the left panel. a-e: images at time points indicated in the traces.

cGMP degradation through the application of an PDE inhibitor. The results of the sensor characterization are depicted in Figure 3.14. Overall, the performance observed in striatal neurons reflected data acquired *in vitro* very well, even if the exact values for dynamic range were not reached. This was attributed to the artificially optimal conditions in *in vitro* experiments and the unknown quantities regarding cGMP levels in the neuronal environment. Both C1 and B2 sensors displayed an around two fold larger dynamic range than cGi500, exemplifying the improvements made during sensor development. As predicted by the *in vitro* characterization the dynamic range of the cGKII based sensor was on par with cGi500. Surprisingly, the sensor B3 exhibited a lower dynamic range than B2, even though they performed equally well under *in vitro* conditions. This could be potentially attributed to the fact, that B3 displayed the highest affinity for cAMP of all depicted sensors and thus was partially saturated by basal cAMP levels. Nevertheless, this highlighted the importance of evaluations under typical experimental conditions and limitations of the *in vitro* model system.

Besides the dynamic range, the selectivity for cGMP over cAMP of individual sensors was evaluated. This was achieved, by first stimulating cAMP production through the application of forskolin, an adenylate cyclase agonist, recording the FRET ratio and then saturating the sensors through NO stimulation and PDE inhibition (see Figure 3.15 A). The selectivity for cGMP was estimated by the signal, which could be triggered by elevated cAMP levels, as a percentage of the maximum dynamic range.

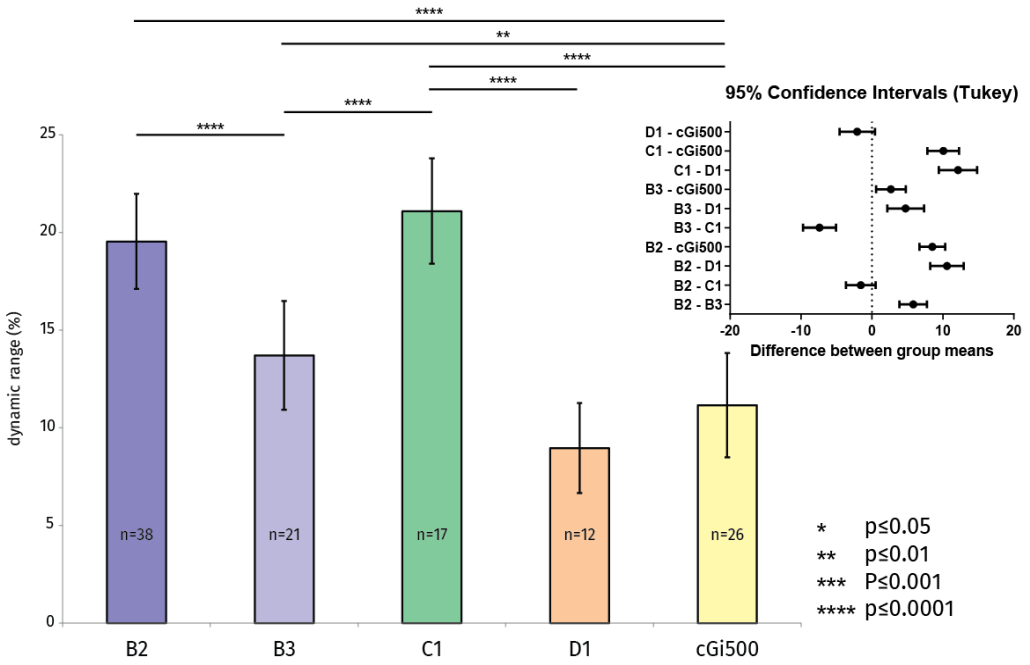

Figure 3.14 Maximum dynamic range of cGMP sensors in striatal neurons

Maximum dynamic range of engineered sensors B2, B3, C1 and D1 compared to cGi500. Murine striatal neurons were virally transfected with sensor constructs, imaged and stimulated by applying 100 µM DEA/No and 200 µM IBMX, to estimate maximum dynamic range of each sensor

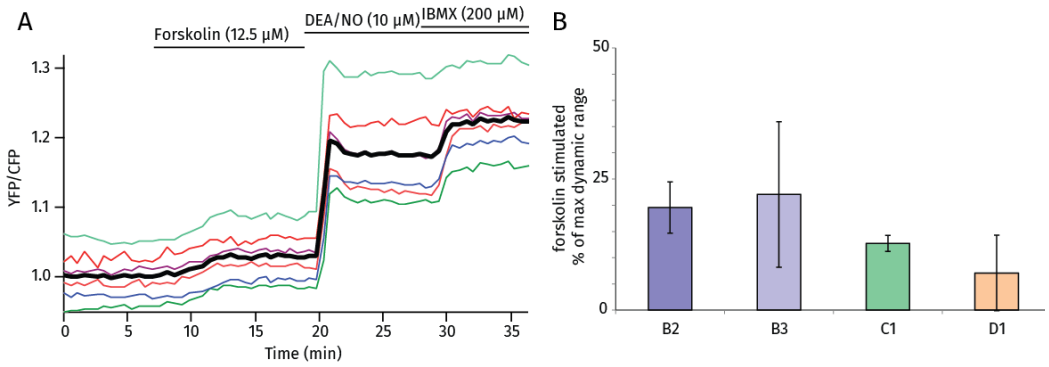

Figure 3.15 Evaluation of Biosensor Selectivity

A) Traces of a typical selectivity experiment using sensor C1 in striatal neurons. cAMP production was stimulated by addition of forskolin. Sensors were then saturated by NO donor stimulation and PDE inhibition to acquire the maximum FRET ratio. B) Maximum signal achieved by cAMP stimulation as a fraction of the maximum dynamic range.

The results obtained from these experiments (see Figure 3.15 B), were in good agreement with the *in vitro* characterization. Both B2 and B3 with relatively low *in vitro* discrimination factors (430 and 240, respectively), exhibited around 20 % sensor saturation by cAMP. C1 and D1 on the other hand showed only modest signals with cAMP, with D1 performing the best out of all of these sensors. This demonstrated that, while having a lower dynamic range than the other improved sensors, D1 and its CNB-BII sensor domain could be a valuable assets in future biosensors for cGMP.

Concluding from *in vitro* data and the results from brain slices, C1 was seen as the sensor with the best overall characteristics. It displayed a large dynamic range, two fold larger than cGi500, the best FRET sensor to date. Furthermore, it incorporated the brightest available versions of both CFP and YFP, improving imaging conditions. It displayed simple binding kinetics (Hill coefficient 1) and because of its single cGMP binding site only poses minimal buffering potential. Adding to that, it was one of the most selective cGMP sensors tested. Despite having a relatively low affinity for cGMP, it could still harness its full potential in a neuronal environment, demonstrating that cGMP levels in striatal neurons are within its operating range.

In broader consideration, these results established the validity of the abstraction of cGMP sensor development in a bacterial environment. *In vitro* data were good predictors of performance in neurons and could therefore be successfully used to improve sensor characteristics in a much simpler system. This is the basis for high throughput methods like bacterial plate screening (Litzlbauer et al. 2015), for further improvement of sensor performance through directed evolution.

4 Discussion and Outlook

In this work I developed a novel genetically encoded biosensor for cGMP, featuring exceptional brightness, large dynamic range, simple binding kinetics and high selectivity. In collaboration we functionally characterized this sensor in rodent brain slices. To date this probe is the best performing FRET based cGMP sensor. Furthermore, I provide the first proof of concept of a bright, red single emission probe for cGMP, based on the concept of a dark FRET acceptor.

In order to achieve these results I implemented several improvements in the biosensor development pipeline. These improvements include a harmonized vector system, which enables background free, multiple fragment cloning and rapid sequence independent transfer into mammalian expression systems and large viral vectors. Additionally, I established a streamlined, one-step protein purification protocol for *in vitro* biosensor characterization, which reaches purity of more than 99 % for the protein of interest. This enabled high parallelization and accurate measurements at the same time.

Furthermore, we designed and build a cost effective robotic screening system, which automates the labor intensive process of directed evolution through bacterial plate screening, increasing throughput and simplifying biosensor development.

4.1 Cyan/yellow cGMP FRET Sensor

One of the most intriguing questions in neurobiology is how the brain is able to learn, how information is stored in the brain, and how this information can be retrieved even after decades. This astounding phenomenon, which is so central to our self-conception, is still poorly understood.

The current model for learning and memory formation is based on activity dependent plasticity of synaptic transmission. Two forms of synaptic plasticity, long term potentiation (LTP) and long term depression (LTD) have been linked to learning processes. Nitric oxide(NO)/cGMP signaling has been demonstrated to be involved in LTP in the hippocampus and the amygdala, two very prominent brain regions for the formation and recall of memories.

The cGMP biosensors I developed in this work will help to address these fundamental questions. For instance, genetic targeting of these sensors to specific cell-types and brain regions is a powerful tool for deciphering intricate neuronal circuits. The ability to repeatedly revisit and interrogate the same neuron population is an important aspect for studying neuronal processes that last a long time. Unlike single emission biosensors, FRET sensors provide this ability, as cells are permanently labeled in at least one channel, revealing their morphology and identity to the observer. This aspect has been demonstrated for FRET based calcium sensors in the past (Aramuni and Griesbeck 2013) and will play an important role for investigating long term potentiation and depression using cGMP sensors.

In recent years, devices have been developed by several groups, which can be implanted into the central nervous system and allow the wireless stimulation and interrogation of optogenetic tools in freely behaving animals (Park et al. 2015, Wu et al. 2015). Once these technologies mature and are combined with optogenetic actuators and reporters for cGMP, we will be able to study memory formation in a natural environment, unobstructed by experimental set-ups and imaging equipment.

A more immediate application for genetically encoded cGMP sensors is their use in drug discovery. Cardio-vascular diseases (CVD), such as hypertension and coronary heart disease, are the leading cause of death worldwide (Mendis et al. 2011). cGMP is the central molecule in the regulation of cardio-vascular homeostasis. Therefore, proteins involved in cGMP signaling are potential drug targets. Genetically encoded cGMP sensors could be used as read-outs in cell-based high throughput assays for the screening of large candidate compound libraries. Further down the line these sensors can also be used to verify the effects of candidate drugs in *in vivo* models.

In order to achieve these long term goals many challenges have to be met. Compared to the performance of current genetically encoded calcium sensors,

the imaging of cGMP oscillations *in vivo* is still in its infancy. The engineering effort that went into the development of calcium sensors exemplifies the high bars for broad range *in vivo* applicability (Mank and Griesbeck 2008, Rose et al. 2014). The central challenge is the increase of the signal/noise ratio. Thus the dynamic range of cGMP sensors has to be improved significantly, to be able to resolve subtle cGMP transients. Furthermore, sensors have to be bright enough to elevate their signal above the level of background fluorescence. A biological source of noise is the interaction with the cAMP signaling cascade. Biosensors have to be highly specific for cGMP in order to avoid cross talk form cAMP, which is much higher concentrated in a cellular environment. In addition, the sensor affinity has to be fine-tuned for the cGMP concentration range found in specific cell types. Surprisingly little is known about this free cellular cGMP concentration, besides that it varies greatly between cell types.

In this work, I developed a cGMP biosensor which offers improvements on all of these fronts, overcoming previous limitations. Specifically, the construct C1 incorporates the brightest cyan and yellow fluorescent proteins available. Its dynamic range doubled in comparison to the established sensor cGi500 and it displayed a 1,900 fold selectivity for cGMP. Furthermore, results we obtained from experiments with the sensor construct D1, incorporating the cyclic nucleotide binding site of cGKII, showed that even higher specificities can be achieved. Noteworthy is also, that prototype sensors constructed during my optimization efforts, displayed a relatively large spectrum of affinities, even though the same binding sites were used. Taken together, this shows that it is possible to tailor these variables to specific applications and cell types.

The improvements of the biosensors presented here are the first step on the road towards the ideal biosensor for cGMP. Next, lessons learned from the development of genetically encoded calcium sensors will have to be applied to cGMP sensors, in order to meet the high standards required for *in vivo* imaging. The next logical step would be, to employ directed evolution and large scale screening to improve the performance of these sensors, as it has been done in the past for calcium sensors (Litzlbauer et al. 2015).

I demonstrated that characteristics of cGMP sensors engineered in a bacterial background can be directly correlated to their *in vivo* performance. This is a prerequisite for employing high throughput methods, like bacterial plate screening in the future. To this end we developed an automated screening system, allowing the quantitative evaluation of thousands of sensor variants per day.

The remaining challenge for implementing bacterial plate screening for cGMP sensors is the development of a screening protocol. Such a protocol would allow the controlled introduction of cGMP into *E. coli* in the context of bacterial colonies, in order to monitor the performance of individual sensor variants of large libraries.

Preliminary experiments not included in this work, showed that this is challenging, as cGMP does not readily permeate through membranes. One solution is the use of cGMP analogues, which are modified to increase penetration, such as 8-pCPT-cGMP. However, these compounds are prohibitively expensive when used in quantities necessary for bacterial plate screening. Adding to that, the use of analogues in a directed evolution approach could impose unwanted selection pressure, as sensors are adapted to a different molecule. Probably the best approach will be to employ small organic permeation agents such as isopropanol or toluene, which have been used for similar purposes before (Paoni and Koshland 1979, Litzlbauer et al. 2015). Here conditions have to be fine-tuned in order to guarantee, that cGMP can penetrate *E. coli* cells in colonies in an efficient manner without disturbing biosensor functionality or exerting toxicity in bacteria. These conditions can be rapidly optimized using the screening set-up developed in this work.

A more subtle challenge is that cAMP is a central signaling molecule within *E. coli* and regulates the response to diverse stimuli such as availability of specific carbon sources and osmotic conditions and might interact negatively with cGMP sensors, as it is present in high concentration of 20 µM -120 µM (Notley-McRobb et al. 1997, Balsalobre et al. 2006). It is important to evaluate cAMP signal contribution to avoid distorted results. This can be achieved by growing *E. coli* expressing candidate sensors on minimal media and then applying glucose while monitoring the FRET ratio.

The base construct for future directed evolution efforts should be the construct C1, developed in this work. Its major components, fluorescent proteins and sensor domain, have been optimized and proven functional in brain slices. Next, the sensor performance has to be carefully fine-tuned, with a focus on the dynamic range, as this is the key factor for resolving natural cGMP oscillations. This can be achieved by introducing randomized linkers in between functional domains, to adjust the angles and distance between fluorescent proteins and then screening for improved variants, as it has been done successfully in the past (Thestrup et al. 2014).

4.2 Red single Emission cGMP Sensor

In this work I furthermore developed single emission cGMP sensors, based on FRET from a red fluorescent protein to a chromoprotein. Even though, only a handful proof of concept sensors using this principle have been published so far, this approach has some advantages.

Most single emission biosensors are based on the modulation of the chromophore solvent accessibility. This strategy was successfully employed in the past to gain relatively large dynamic ranges. However, since this actively disturbs the chromophore environment, these achievements come at the cost of overall brightness and photo stability of the resulting sensor. These shortcomings then have to be addressed by additional mutation and screening steps on top of screening efforts for dynamic range and other variables. In contrast, FRET based single emission sensors have the advantage that the fluorophore stays intact. This allows the use of very bright, optimized fluorescent proteins, transferring superior photo characteristic onto resulting sensor constructs.

Adding to that, FRET based single emission biosensors cannot only be monitored through intensity based measurements, but also by fluorescence lifetime imaging (FLIM). Employing FLIM mitigates one of the major downsides of single fluorophore biosensors, namely the concentration dependency of the signal intensity. In contrast, the fluorescence life time solely depends on the local environment of the chromophore.

The main impendence for this approach is the availability of suitable dark acceptors. The first sensors using this concept were based on variants of YFP with low quantum yields (Ganesan et al. 2006, Murakoshi et al. 2008). However, these dark YFP variants still retained some fluorescence, thus diminishing the advantages of a single emission sensor. Only recently, the engineered chromoprotein Ultramarine has become available (Pettikiriarachchi et al. 2012), which behaved as a true dark acceptor in my experiments. To date, this chromoprotein remains the only viable dark acceptor, limiting this approach to orange and red fluorescent proteins.

A host of chromoproteins with different spectral characteristics from various coral species are known (Alieva et al. 2008). However, these chromoproteins are all either tetramers or higher order oligomers and thus not suitable for biosensor development. In order to expand this approach to the full spectral pallet of fluorescent proteins, wild type chromoproteins with suitable absorption spectra would first have to be converted into monomers. While this has been done before, the requirement for this additional engineering step probably deterred the wide spread implementation of this type of sensor.

Furthermore, red single emission cGMP sensors have a number of advantages for certain applications. Firstly, shifting into the red part of the visual spectrum

is beneficial for imagining in general. There is less auto-fluorescence in the red part of the spectrum. Besides that, one of the limitations of biosensors emitting in the blue-yellow part of the spectrum is the imaging depth, which only allows the investigation of surface features of the brain. This can be partially mitigated by using more red-shifted fluorescent proteins, as longer wavelengths experience reduced scattering, which increases penetration.

To harness these benefits completely it would be ideal to shift into the near infrared part of the electromagnetic spectrum, in which tissue is almost transparent. However, photo-physical characteristics of far-red fluorescent proteins decline drastically the longer their emission wavelengths is. Specifically quantum yields are extremely reduced, e.g. fluorescent proteins with emission maxima around 650 nm only display a quantum yield of 0.2, while all fluorescent proteins with emissions greater than 670 nm have quantum yields below 0.11.

This anti-correlation between emission wavelengths and quantum yield is probably caused by the fact that, as the wavelength increases, conjugated π-systems within the chromophore have to get extended in order absorb photons. This in turn has a negative impact on the quantum yield, as it becomes increasingly difficult to stabilize larger and larger chromophores within the confines of a protein structure.

Because of their low quantum yields, far-red and infrared fluorescent proteins are extremely dim and therefore not very suited for demanding imaging applications. FRET based biosensors in particular are negatively impacted by low quantum yields, as the Förster radius decreases significantly. Therefore, the use of orange and red fluorescent proteins, as demonstrated in this work, is a good compromise between tissue penetration on one side and sensor functionality on the other side.

Another aspect is related to possible application scenarios for cGMP sensors. When addressing questions related to long term potentiation of synaptic transmission, the need for combination of cGMP sensors with other optogenetic tools is conceivable. Specifically it would be interesting to combine cGMP imaging with calcium imaging, in order to correlate activity with cGMP oscillations, to elucidate LTP processes. Furthermore, the combination with optogenetic actuators could provide the ability to test hypothesis conceived during activity/cGMP imaging. However, the most optimized activity sensors (GCamP6s, Twitch2B) and actuators (ChR2) all operate in the blue-yellow part of the spectrum, potentially overlapping and interfering with cGMP sensors based on cyan/yellow fluorescent proteins. Therefore, the sensors based on single red fluorescent proteins, I developed in this work, presents an excellent alternative for these applications.

4.3 Improvements of Biosensor Development Pipeline

The engineering of biosensors for *in vivo* applications is a long and labor-intensive process. This is best exemplified by genetically encoded calcium sensors, which took the concerted efforts of several groups for more than 15 years from the first concepts to the refined tools we have today. Even after such a long development time, there is still room for improvement in current calcium sensors (Rose et al. 2014). Therefore, streamlining this development process is crucial, as there are many other potential targets for biosensors. The methods I developed and described in this work represent a generic pipeline for the engineering of genetically encoded biosensors. They focus on fast turnaround times, parallelization and elimination of bottlenecks.

The first stage of this development process is the prototyping phase, in which the major functional elements of a biosensor, such as fluorescent proteins and sensor domains, are selected and optimized. These prototypes are later refined through directed evolution and screening. As experience has shown it is crucial to optimize and fix the main functional domains first, as later changes of these components often render improvements made during screening obsolete. This is best exemplified by results presented in this work. Even exchanging a fluorescent protein for a slightly different variant of the same fluorescent protein had dramatic effects on sensor performance. The reason for this is that while biosensor engineering superficially looks like a modular building block system, it is still inherently unpredictable, because individual domains are not folding in isolation of each other.

One of the main bottlenecks in the prototyping phase is the cloning of individual components. Traditional cloning methods based on restriction enzymes usually require several steps for the assembly of a typical FRET sensor and are limited by the availability of restriction sites. The introduction of modern assembly cloning methods based on *in vitro* homologous recombination (e.g. SLiCE, Gibbson assembly), I present in this work, streamlined this process considerably. In addition, I established an effective counter-selection strategy in a harmonized vector system for biosensor development. In combination these two methods enable background free biosensor assembly from multiple fragments in one-step. This allows the parallel assembly of several dozen complex prototype sensors, greatly increasing throughput in this crucial phase. Further down the line, the harmonized vector system I established allows for rapid, efficient transfer of biosensor prototypes into large viral vectors, in order to obtain fast feedback from expression in mammalian systems.

These new DNA assembly methods are very powerful and could spawn innovations in the biosensor development pipeline. Specifically, they enable the introduction of recombination steps during directed evolution of randomized

linker libraries. Typically during the refinement of a FRET sensor the two linker regions between sensor domain and fluorescent proteins are randomized and then screened. The candidate with the best performance then becomes the new base construct and is further randomized. Usually, sensors which performed well but not as good as the best will be discarded. This is very unfortunate because these variants harbor valuable information. It is conceivable that they contain a good linker region which is paired with a suboptimal region.

To harness this information, well performing constructs would be first pooled and then their linker regions could be isolated via PCR, splitting paired linker regions. These two libraries of isolated linkers can then be reassembled using SLiCE, randomly recombining linker variants. This reassembly step is crucial and can only be performed with the needed efficiency employing these new cloning techniques. The resulting library of reassembled sensor variants has potentially a much higher chance of yielding variants that outperform the initially best sensor. This is because the fitness of the linker population is much higher than in random libraries, as they already underwent a selection procedure. In addition to that, these libraries are much less complex and could therefore conceivably be screened to completion. This introduces a recombination event into directed evolution which is analogue to sexual reproduction.

Another important bottleneck in the biosensor development pipeline is the screening of the sensor libraries. We previously described a method for the screening of large biosensor libraries in bacterial colonies (Litzlbauer et al. 2015). In brief, biosensor libraries are expressed in bacterial colonies. These colonies are then imaged in a wide field fluorescence imaging set-up and then exposed to the target molecule. Through this process up to 1,000 different variants can be assessed in parallel on one plate. Through software analysis the best performing variants are identified and later picked by hand. The requirement for manual colony picking represents a major bottleneck for this method and can potentially lead to errors, as colonies have to be visually identified.

Building on this approach we designed and build a low cost robotic screening system with enhanced imaging capabilities and automated colony picking. It is based on recent developments in the DIY 3D printing community and features an innovative picking mechanism, which allows fast colony picking without cross contamination. Imaging data is acquired and analyzed on the fly and promising variants can be picked in a continuous process.

Besides biosensor screening this robotic system can be used for other applications. Examples are the improvement of fluorescent protein characteristics, such as brightness, photo stability and pH stability or blue/white selection in high throughput cloning. Because of the low cost nature of this robotic system (all individual parts can be acquired for less than 1,000 €, excluding the camera), it can provide easy access to high throughput screening technology for less affluent laboratories, thus reducing entry hurdles into fields which require such instruments.

5 References

Ai, H.-w. W., K. L. Hazelwood, M. W. Davidson and R. E. Campbell (2008). "Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors." *Nature methods* 5(5): 401-403.

Ai, H.-w. W., J. N. Henderson, S. J. Remington and R. E. Campbell (2006). "Directed evolution of a monomeric, bright and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescence imaging." *The Biochemical journal* 400(3): 531-540.

Ai, H.-w. W., S. G. Olenych, P. Wong, M. W. Davidson and R. E. Campbell (2008). "Hue-shifted monomeric variants of Clavularia cyan fluorescent protein: identification of the molecular determinants of color and applications in fluorescence imaging." *BMC biology* 6: 13.

Ai, H.-w. W., N. C. Shaner, Z. Cheng, R. Y. Tsien and R. E. Campbell (2007). "Exploration of new chromophore structures leads to the identification of improved blue fluorescent proteins." *Biochemistry* 46(20): 5904-5910.

Akerboom, J., T.-W. W. Chen, T. J. Wardill, L. Tian, J. S. Marvin, S. Mutlu, N. C. Calderón, F. Esposti, B. G. Borghuis, X. R. Sun, A. Gordus, M. B. Orger, R. Portugues, F. Engert, J. J. Macklin, A. Filosa, A. Aggarwal, R. A. Kerr, R. Takagi, S. Kracun, E. Shigetomi, B. S. Khakh, H. Baier, L. Lagnado, S. S. Wang, C. I. Bargmann, B. E. Kimmel, V. Jayaraman, K. Svoboda, D. S. Kim, E. R. Schreiter and L. L. Looger (2012). "Optimization of a GCaMP calcium indicator for neural activity imaging." *The Journal of neuroscience* 32(40): 13819-13840.

Alieva, N. O., K. A. Konzen, S. F. Field, E. A. Meleshkevitch, M. E. Hunt, V. Beltran-Ramirez, D. J. Miller, J. Wiedenmann, A. Salih and M. V. Matz (2008). "Diversity and evolution of coral fluorescent proteins." *PloS one* 3(7).

References

Aramuni, G. and O. Griesbeck (2013). "Chronic calcium imaging in neuronal development and disease." *Experimental Neurology* 242: 50-56.

Arosio, D. and G. Ratto (2014). "Twenty years of fluorescence imaging of intracellular chloride." *Frontiers in cellular neuroscience* 8: 258.

Arpino, J. A., P. J. Rizkallah and D. D. Jones (2012). "Crystal structure of enhanced green fluorescent protein to 1.35 Å resolution reveals alternative conformations for Glu222." *PloS one* 7(10).

Baird, G. S., D. A. Zacharias and R. Y. Tsien (1999). "Circular permutation and receptor insertion within green fluorescent proteins." *Proceedings of the National Academy of Sciences* 96.

Bajar, B. T., E. S. Wang, A. J. Lam, B. B. Kim, C. L. Jacobs, E. S. Howe, M. W. Davidson, M. Z. Lin and J. Chu (2016). "Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting." *Scientific reports* 6: 20889.

Balsalobre, C., J. Johansson and B. E. Uhlin (2006). "Cyclic AMP-dependent osmoregulation of crp gene expression in Escherichia coli." *Journal of bacteriology* 188(16): 5935-5944.

Beddoe, T., M. Ling, S. Dove, O. Hoegh-Guldberg, R. J. Devenish, M. Prescott and J. Rossjohn (2003). "The production, purification and crystallization of a pocilloporin pigment from a reef-forming coral." *Acta crystallographica. Section D, Biological crystallography* 59(Pt 3): 597-599.

Benčina, M. (2013). "Illumination of the Spatial Order of Intracellular pH by Genetically Encoded pH-Sensitive Sensors." *Sensors* 13(12): 16736-16758.

Bhargava, Y., K. Hampden-Smith, K. Chachlaki, K. C. Wood, J. Vernon, C. K. Allerston, A. M. Batchelor and J. Garthwaite (2013). "Improved genetically-encoded, FlincG-type fluorescent biosensors for neural cGMP imaging." *Frontiers in molecular neuroscience* 6: 26.

Biel, M. and S. Michalakis (2009). Cyclic nucleotide-gated channels. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 111-136.

Biswas, K. H., S. Sopory and S. S. Visweswariah (2008). "The GAF domain of the cGMP-binding, cGMP-specific phosphodiesterase (PDE5) is a sensor and a sink for cGMP." *Biochemistry* 47(11): 3534-3543.

Bizzarri, R., M. Serresi, S. Luin and F. Beltram (2009). "Green fluorescent protein based pH indicators for in vivo use: a review." *Analytical and Bioanalytical Chemistry*.

Blomfield, I. C., V. Vaughn, R. F. Rest and B. I. Eisenstein (1991). "Allelic exchange in Escherichia coli using the Bacillus subtilis sacB gene and a temperature-sensitive pSC101 replicon." *Molecular Microbiology* 5(6): 1447-1457.

Boerrigter, G., H. Lapp and J. C. Burnett (2009). Modulation of cGMP in heart failure: a new therapeutic paradigm. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 485-506.

Bolanos-Garcia, V. and O. Davies (2006). "Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography." *Biochimica et Biophysica Acta* 1760(9): 1304-1313.

Brunton, L. T. (1867). "On the use of nitrite of amyl in angina pectoris." *The Lancet* 90(2291): 97-98.

Busby, S. and R. H. Ebright (1999). "Transcription activation by catabolite activator protein (CAP)." *Journal of Molecular Biology* 293(2): 199-213.

Campbell, J. C., J. J. Kim, K. Y. Li, G. Y. Huang, A. S. Reger, S. Matsuda, B. Sankaran, T. M. Link, K. Yuasa, J. E. Ladbury, D. E. Casteel and C. Kim (2016). "Structural Basis of Cyclic Nucleotide Selectivity in cGMP-Dependent Protein Kinase II." *The Journal of biological chemistry* 291(11): 5623-5633.

Campbell, R. E., O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias and R. Y. Tsien (2002). "A monomeric red fluorescent protein." *Proceedings of the National Academy of Sciences* 99(12): 7877-7882.

Chalfie, M., Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher (1994). "Green fluorescent protein as a marker for gene expression." *Science* 263(5148): 802-805.

Chan, M. C., S. Karasawa, H. Mizuno, I. Bosanac, D. Ho, G. G. Privé, A. Miyawaki and M. Ikura (2006). "Structural characterization of a blue chromoprotein and its yellow mutant from the sea anemone Cnidopus japonicus." *The Journal of biological chemistry* 281(49): 37813-37819.

Chen, T.-W. W., T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda and D. S. Kim (2013). "Ultrasensitive fluorescent proteins for imaging neuronal activity." *Nature* 499(7458): 295-300.

Chiu, W., A. Towheed and M. J. Palladino (2014). "Chapter Fourteen Genetically Encoded Redox Sensors." *Methods in enzymology* 542: 263-287.

Chu, J., R. D. Haynes, S. Y. Corbel, P. Li, E. González-González, J. S. Burg, N. J. Ataie, A. J. Lam, P. J. Cranfill, M. A. Baird, M. W. Davidson, H.-L. L. Ng, K. C. Garcia, C. H. Contag, K. Shen, H. M. Blau and M. Z. Lin (2014). "Non-invasive intravital imaging of cellular differentiation with a bright red-excitable fluorescent protein." *Nature methods* 11(5): 572-578.

Chu, J., Z. Zhang, Y. Zheng, J. Yang, L. Qin, J. Lu, Z.-L. L. Huang, S. Zeng and Q. Luo (2009). "A novel far-red bimolecular fluorescence complementation system that allows for efficient visualization of protein interactions under physiological conditions." *Biosensors & bioelectronics* 25(1): 234-239.

Chung, C. T., S. L. Niemela and R. H. Miller (1989). "One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution." *Proceedings of the National Academy of Sciences* 86(7): 2172-2175.

Clavel, R. (1988). Delta, a fast robot with parallel geometry. *Proceedings of the 18th International Symposium on Industrial Robots, Lausanne, 1988*.

References

Conti, M. and J. Beavo (2007). "Biochemistry and physiology of cyclic nucleotide phosphodiesterases: essential components in cyclic nucleotide signaling." *Annual review of biochemistry* 76: 481-511.

Cox, J. and M. Mann (2008). "MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification." *Nature biotechnology* 26(12): 1367-1372.

Cox, J., N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen and M. Mann (2011). "Andromeda: a peptide search engine integrated into the MaxQuant environment." *Journal of proteome research* 10(4): 1794-1805.

Crick, F. (1999). "The impact of molecular biology on neuroscience." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 354(1392): 2021-2025.

Cubitt, A. B., L. A. Woollenweber and R. Heim (1999). "Understanding structure-function relationships in the Aequorea victoria green fluorescent protein." *Methods in cell biology* 58: 19-30.

Deisseroth, K. (2015). "Optogenetics: 10 years of microbial opsins in neuroscience." *Nature neuroscience* 18(9): 1213-1225.

Derbyshire, E. R. and M. A. Marletta (2009). Biochemistry of soluble guanylate cyclase. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 17-31.

Dove, S. G., O. Hoegh-Guldberg and S. Ranganathan (2001). "Major colour patterns of reef-building corals are due to a family of GFP-like proteins." *Coral reefs* 19(3): 197-204.

Dove, S. G., M. Takabayashi and O. Hoegh-Guldberg (1995). "Isolation and partial characterization of the pink and blue pigments of pocilloporid and acroporid corals." *The Biological Bulletin* 189(3).

Dugué, G. P., W. Akemann and T. Knöpfel (2012). "A comprehensive concept of optogenetics." *Progress in brain research* 196: 1-28.

Erard, M., A. Fredj, H. Pasquier, D.-B. B. Beltolngar, Y. Bousmah, V. Derrien, P. Vincent and F. Merola (2013). "Minimum set of mutations needed to optimize cyan fluorescent proteins for live cell imaging." *Molecular bioSystems* 9(2): 258-267.

Evdokimov, A. G., M. E. Pokross, N. S. Egorov, A. G. Zaraisky, I. V. Yampolsky, E. M. Merzlyak, A. N. Shkoporov, I. Sander, K. A. Lukyanov and D. M. Chudakov (2006). "Structural basis for the fast maturation of Arthropoda green fluorescent protein." *EMBO reports* 7(10): 1006-1012.

Fearon, W. R. (1940). An introduction to biochemistry. St. Louis,, The C.V. Mosby company.

Fischmeister, R., L. R. Castro, A. Abi-Gerges, F. Rochais, J. Jurevicius, J. Leroy and G. Vandecasteele (2006). "Compartmentation of cyclic nucleotide signaling in the heart: the role of cyclic nucleotide phosphodiesterases." *Circulation research* 99(8): 816-828.

Förster, T. (1946). "Energy transport and fluorescence." *Naturwissenschaften* 6: 166-175.

Förster, T. (1948). "Zwischenmolekulare Energiewanderung und Fluoreszenz." *Annalen der Physik* 437(1-2): 55-75.

Förster, T. (1951). Fluorescence of organic compounds. Göttingen, Vandenhoek & Ruprecht.

Ganesan, S., S. M. Ameer-Beg, T. T. Ng, B. Vojnovic and F. S. Wouters (2006). "A dark yellow fluorescent protein (YFP)-based Resonance Energy-Accepting Chromoprotein (REACh) for Förster resonance energy transfer with GFP." *Proceedings of the National Academy of Sciences* 103(11): 4089-4094.

Gao, S., J. Nagpal, M. W. Schneider, V. Kozjak-Pavlovic, G. Nagel and A. Gottschalk (2015). "Optogenetic manipulation of cGMP in cells and animals by the tightly light-regulated guanylyl-cyclase opsin CyclOp." *Nature communications* 6: 8046.

Gasser, C., S. Taiber, C.-M. M. Yeh, C. H. Wittig, P. Hegemann, S. Ryu, F. Wunder and A. Möglich (2014). "Engineering of a red-light-activated human cAMP/cGMP-specific phosphodiesterase." *Proceedings of the National Academy of Sciences* 111(24): 8803-8808.

Gay, P., D. Le Coq, M. Steinmetz, T. Berkelman and C. I. Kado (1985). "Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria." *Journal of bacteriology* 164(2): 918-921.

Goedhart, J., L. van Weeren, M. A. Hink, N. O. Vischer, K. Jalink and T. W. Gadella (2010). "Bright cyan fluorescent protein variants identified by fluorescence lifetime screening." *Nature methods* 7(2): 137-139.

Goedhart, J., D. von Stetten, M. Noirclerc-Savoye, M. Lelimousin, L. Joosen, M. A. Hink, L. van Weeren, T. W. Gadella and A. Royant (2012). "Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%." *Nature communications* 3: 751.

Gorshkov, K. and J. Zhang (2014). "Visualization of cyclic nucleotide dynamics in neurons." *Frontiers in cellular neuroscience* 8: 395.

Götz, K. R., J. U. Sprenger, R. K. Perera, J. H. Steinbrecher, S. E. Lehnart, M. Kuhn, J. Gorelik, J.-L. L. Balligand and V. O. Nikolaev (2014). "Transgenic mice for real-time visualization of cGMP in intact adult cardiomyocytes." *Circulation research* 114(8): 1235-1245.

Green, M. R. and J. Sambrook (2012). Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press New York.

Griesbeck, O., G. S. Baird, R. E. Campbell, D. A. Zacharias and R. Y. Tsien (2001). "Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications." *The Journal of biological chemistry* 276(31): 29188-29194.

Grimley, J. S., L. Li, W. Wang, L. Wen, L. S. Beese, H. W. Hellinga and G. J. Augustine (2013). "Visualization of synaptic inhibition with an optogenetic sensor

developed by cell-free protein engineering automation." *The Journal of neuroscience* 33(41): 16297-16309.

Gründemann, D. and E. Schömig (1996). "Protection of DNA during preparative agarose gel electrophoresis against damage induced by ultraviolet light." *Biotechniques* 21(5): 898-903.

Guo, J., F. Sachs and F. Meng (2014). "Fluorescence-based force/tension sensors: a novel tool to visualize mechanical forces in structural proteins in live cells." *Antioxidants & redox signaling* 20(6): 986-999.

Gurskaya, N. G., A. F. Fradkov, A. Terskikh, M. V. Matz, Y. A. Labas, V. I. Martynov, Y. G. Yanushevich, K. A. Lukyanov and S. A. Lukyanov (2001). "GFP-like chromoproteins as a source of far-red fluorescent proteins." *FEBS letters* 507(1): 16-20.

Heim, R., D. C. Prasher and R. Y. Tsien (1994). "Wavelength mutations and posttranslational autoxidation of green fluorescent protein." *Proceedings of the National Academy of Sciences* 91(26): 12501-12504.

Helmchen, F. and W. Denk (2005). "Deep tissue two-photon microscopy." *Nature methods* 2(12): 932-940.

Hessels, A. M. and M. Merkx (2014). "Genetically-encoded FRET-based sensors for monitoring Zn(2+) in living cells." *Metallomics* 7(2): 258-266.

Hibbert, D. B. (2012). "Experimental design in chromatography: a tutorial review." *Journal of chromatography* 910: 2-13.

Higuchi, R., B. Krummel and R. K. Saiki (1988). "A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions." *Nucleic acids research* 16(15): 7351-7367.

Hofmann, F., D. Bernhard, R. Lukowski and P. Weinmeister (2009). cGMP regulated protein kinases (cGK). Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 137-162.

Hoi, H., E. S. Howe, Y. Ding, W. Zhang, M. A. Baird, B. R. Sell, J. R. Allen, M. W. Davidson and R. E. Campbell (2013). "An engineered monomeric Zoanthus sp. yellow fluorescent protein." *Chemistry & biology* 20(10): 1296-1304.

Honda, A., S. R. Adams, C. L. Sawyer, V. Lev-Ram, R. Y. Tsien and W. R. Dostmann (2001). "Spatiotemporal dynamics of guanosine 3',5'-cyclic monophosphate revealed by a genetically encoded, fluorescent indicator." *Proceedings of the National Academy of Sciences* 98(5): 2437-2442.

Hornburg, D., C. Drepper, F. Butter, F. Meissner, M. Sendtner and M. Mann (2014). "Deep proteomic evaluation of primary and cell line motoneuron disease models delineates major differences in neuronal characteristics." *Molecular & cellular proteomics* 13(12): 3410-3420.

Huang, G. Y., O. O. Gerlits, M. P. Blakeley, B. Sankaran, A. Y. Kovalevsky and C. Kim (2014). "Neutron diffraction reveals hydrogen bonds critical for cGMP-selective activation: insights for cGMP-dependent protein kinase agonist design." *Biochemistry* 53(43): 6725-6727.

Huang, G. Y., J. J. Kim, A. S. Reger, R. Lorenz, E.-W. W. Moon, C. Zhao, D. E. Casteel, D. Bertinetti, B. Vanschouwen, R. Selvaratnam, J. W. Pflugrath, B. Sankaran, G. Melacini, F. W. Herberg and C. Kim (2014). "Structural basis for cyclic-nucleotide selectivity and cGMP-selective activation of PKG I." *Structure* 22(1): 116-124.

Inouye, S. and M. Inouye (1985). "Up-promoter mutations in the lpp gene of Escherichia coli." *Nucleic acids research* 13(9): 3101-3110.

Inouye, S. and F. I. Tsuji (1994). "Aequorea green fluorescent protein. Expression of the gene and fluorescence characteristics of the recombinant protein." *FEBS letters* 341(2-3): 277-280.

Ito, M. (2001). "Cerebellar long-term depression: characterization, signal transduction, and functional roles." *Physiological reviews* 81(3): 1143-1195.

Johnson, D. E., H.-W. W. Ai, P. Wong, J. D. Young, R. E. Campbell and J. R. Casey (2009). "Red fluorescent protein pH biosensor to detect concentrative nucleoside transport." *The Journal of biological chemistry* 284(31): 20499-20511.

Johnson, F. H., O. Shimomura, Y. Saiga, L. C. Gershman, G. T. Reynolds and J. R. Waters (1962). "Quantum efficiency of Cypridina luminescence, with a note on that of Aequorea." *Journal of cellular and comparative physiology* 60(1): 85-103.

Karasawa, S., T. Araki, T. Nagai, H. Mizuno and A. Miyawaki (2004). "Cyan-emitting and orange-emitting fluorescent proteins as a donor/acceptor pair for fluorescence resonance energy transfer." *The Biochemical journal* 381(Pt 1): 307-312.

Karasawa, S., T. Araki, M. Yamamoto-Hino and A. Miyawaki (2003). "A Green-emitting Fluorescent Protein from Galaxeidae Coral and Its Monomeric Version for Use in Fluorescent Labeling." *The Journal of biological chemistry* 278(36): 34167-34171.

Kim, J. J., D. E. Casteel, G. Huang, T. H. Kwon, R. K. Ren, P. Zwart, J. J. Headd, N. G. Brown, D.-C. C. Chow, T. Palzkill and C. Kim (2011). "Co-crystal structures of PKG Iβ (92-227) with cGMP and cAMP reveal the molecular details of cyclic-nucleotide binding." *PloS one* 6(4).

Kim, T., M. Folcher, M. D. Baba and M. Fussenegger (2015). "A Synthetic Erectile Optogenetic Stimulator Enabling Blue-Light-Inducible Penile Erection." *Angewandte Chemie* 54(20): 5933-5938.

Kleppisch, T. (2009). Phosphodiesterases in the central nervous system. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 71-92.

Kleppisch, T. and R. Feil (2009). cGMP signalling in the mammalian brain: role in synaptic plasticity and behaviour. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 549-579.

Kogure, T., S. Karasawa, T. Araki, K. Saito, M. Kinjo and A. Miyawaki (2006). "A fluorescent variant of a protein from the stony coral Montipora facilitates dual-color single-laser fluorescence cross-correlation spectroscopy." *Nature biotechnology* 24(5): 577-581.

Kredel, S., F. Oswald, K. Nienhaus, K. Deuschle, C. Röcker, M. Wolff, R. Heilker, G. U. Nienhaus and J. Wiedenmann (2009). "mRuby, a bright monomeric red fluorescent protein for labeling of subcellular structures." *PloS one* 4(2).

Kremers, G.-J. J., J. Goedhart, E. B. van Munster and T. W. Gadella (2006). "Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Förster radius." *Biochemistry* 45(21): 6570-6580.

Kuhn, M. (2003). "Structure, regulation, and function of mammalian membrane guanylyl cyclase receptors, with a focus on guanylyl cyclase-A." *Circulation research* 93(8): 700-709.

Kuhn, M. (2009). Function and dysfunction of mammalian membrane guanylyl cyclase receptors: lessons from genetic mouse models and implications for human diseases. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 47-69.

Labas, Y. A., N. G. Gurskaya, Y. G. Yanushevich, A. F. Fradkov, K. A. Lukyanov, S. A. Lukyanov and M. V. Matz (2002). "Diversity and evolution of the green fluorescent protein family." *Proceedings of the National Academy of Sciences* 99(7): 4256-4261.

Lam, A. J., F. St-Pierre, Y. Gong, J. D. Marshall, P. J. Cranfill, M. A. Baird, M. R. McKeown, J. Wiedenmann, M. W. Davidson, M. J. Schnitzer, R. Y. Tsien and M. Z. Lin (2012). "Improving FRET dynamic range with bright green and red fluorescent proteins." *Nature methods* 9(10): 1005-1012.

Liang, R., G. J. Broussard and L. Tian (2015). "Imaging chemical neurotransmission with genetically encoded fluorescent sensors." *ACS chemical neuroscience* 6(1): 84-93.

Lin, M. Z., M. R. McKeown, H.-L. L. Ng, T. A. Aguilera, N. C. Shaner, R. E. Campbell, S. R. Adams, L. A. Gross, W. Ma, T. Alber and R. Y. Tsien (2009). "Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals." *Chemistry & biology* 16(11): 1169-1179.

Litzlbauer, J. (2015). Engineering and screening of genetically encoded FRET Calcium indicators. Dr. rer. nat. Dissertation, LMU München, *Ludwig Maximilians Universität München*.

Litzlbauer, J., M. Schifferer, D. Ng, A. Fabritius, T. Thestrup and O. Griesbeck (2015). "Large Scale Bacterial Colony Screening of Diversified FRET Biosensors." *PloS one* 10(6).

Mank, M. and O. Griesbeck (2008). "Genetically encoded calcium indicators." *Chemical reviews* 108(5): 1550-1564.

Markwardt, M. L., G.-J. J. Kremers, C. A. Kraft, K. Ray, P. J. Cranfill, K. A. Wilson, R. N. Day, R. M. Wachter, M. W. Davidson and M. A. Rizzo (2011). "An improved cerulean fluorescent protein with enhanced brightness and reduced reversible photoswitching." *PloS one* 6(3).

Matz, M. V., A. F. Fradkov, Y. A. Labas, A. P. Savitsky, A. G. Zaraisky, M. L. Markelov and S. A. Lukyanov (1999). "Fluorescent proteins from nonbioluminescent Anthozoa species." *Nature biotechnology* 17(10): 969-973.

Mena, M. A., T. P. Treynor, S. L. Mayo and P. S. Daugherty (2006). "Blue fluorescent proteins with enhanced brightness and photostability from a structurally targeted library." *Nature biotechnology* 24(12): 1569-1571.

Mendis, S., P. Puska and B. Norrving (2011). Global atlas on cardiovascular disease prevention and control, World Health Organization.

Merlet, J.-P. (2006). Parallel robots, Springer Science & Business Media.

Merzlyak, E. M., J. Goedhart, D. Shcherbo, M. E. Bulina, A. S. Shcheglov, A. F. Fradkov, A. Gaintzeva, K. A. Lukyanov, S. Lukyanov, T. W. Gadella and D. M. Chudakov (2007). "Bright monomeric red fluorescent protein with an extended fluorescence lifetime." *Nature methods* 4(7): 555-557.

Miesenböck, G. (2009). "The optogenetic catechism." *Science* 326(5951): 395-399.

Miraglia, E., F. De Angelis, E. Gazzano, H. Hassanpour, A. Bertagna, E. Aldieri, A. Revelli and D. Ghigo (2011). "Nitric oxide stimulates human sperm motility via activation of the cyclic GMP/protein kinase G signaling pathway." *Reproduction* 141(1): 47-54.

Misono, K. S., H. Ogawa, Y. Qiu and C. M. Ogata (2005). "Structural studies of the natriuretic peptide receptor: a novel hormone-induced rotation mechanism for transmembrane signal transduction." *Peptides* 26(6): 957-968.

Miyawaki, A., J. Llopis, R. Heim, J. M. McCaffery, J. A. Adams, M. Ikura and R. Y. Tsien (1997). "Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin." *Nature* 388(6645): 882-887.

Miyawaki, A. and Y. Niino (2015). "Molecular Spies for Bioimaging-Fluorescent Protein-Based Probes." *Molecular cell* 58(4): 632-643.

Morozova, K. S., K. D. Piatkevich, T. J. Gould, J. Zhang, J. Bewersdorf and V. V. Verkhusha (2010). "Far-red fluorescent protein excitable with red lasers for flow cytometry and superresolution STED nanoscopy." *Biophysical journal* 99(2): 5.

Motta-Mena, L. B., A. Reade, M. J. Mallory, S. Glantz, O. D. Weiner, K. W. Lynch and K. H. Gardner (2014). "An optogenetic gene expression system with rapid activation and deactivation kinetics." *Nature chemical biology* 10(3): 196-202.

Mues, M., I. Bartholomäus, T. Thestrup, O. Griesbeck, H. Wekerle, N. Kawakami and G. Krishnamoorthy (2013). "Real-time in vivo analysis of T cell activation in the central nervous system using a genetically encoded calcium indicator." *Nature medicine* 19(6): 778-783.

Murakoshi, H., S.-J. J. Lee and R. Yasuda (2008). "Highly sensitive and quantitative FRET-FLIM imaging in single dendritic spines using improved non-radiative YFP." *Brain cell biology* 36(1-4): 31-42.

Nagai, T., K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba and A. Miyawaki (2002). "A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications." *Nature biotechnology* 20(1): 87-90.

Nagai, T., S. Yamada, T. Tominaga, M. Ichikawa and A. Miyawaki (2004). "Expanded dynamic range of fluorescent indicators for Ca(2+) by circularly permuted yellow

fluorescent proteins." *Proceedings of the National Academy of Sciences* 101(29): 10554-10559.

Nagel, G., T. Szellas, S. Kateriya, N. Adeishvili, P. Hegemann and E. Bamberg (2005). "Channelrhodopsins: directly light-gated cation channels." *Biochemical Society transactions* 33(Pt 4): 863-866.

Nakai, J., M. Ohkura and K. Imoto (2001). "A high signal-to-noise Ca(2+) probe composed of a single green fluorescent protein." *Nature biotechnology* 19(2): 137-141.

Nausch, L. W., J. Ledoux, A. D. Bonev, M. T. Nelson and W. R. Dostmann (2008). "Differential patterning of cGMP in vascular smooth muscle cells revealed by single GFP-linked biosensors." *Proceedings of the National Academy of Sciences* 105(1): 365-370.

Nguyen, A. W. and P. S. Daugherty (2005). "Evolutionary optimization of fluorescent proteins for intracellular FRET." *Nature biotechnology* 23(3): 355-360.

Niino, Y., K. Hotta and K. Oka (2009). "Simultaneous live cell imaging using dual FRET sensors with a single excitation light." *PloS one* 4(6).

Niino, Y., K. Hotta and K. Oka (2010). "Blue fluorescent cGMP sensor for multiparameter fluorescence imaging." *PloS one* 5(2).

Nikolaev, V. O., S. Gambaryan and M. J. Lohse (2006). "Fluorescent sensors for rapid monitoring of intracellular cGMP." *Nature methods* 3(1): 23-25.

Nikolaev, V. O. and M. J. Lohse (2009). Novel techniques for real-time monitoring of cGMP in living cells. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 229-243.

Niwa, H., S. Inouye, T. Hirano, T. Matsuno, S. Kojima, M. Kubota, M. Ohashi and F. I. Tsuji (1996). "Chemical nature of the light emitter of the Aequorea green fluorescent protein." *Proceedings of the National Academy of Sciences* 93(24): 13617-13622.

Notley-McRobb, L., A. Death and T. Ferenci (1997). "The relationship between external glucose concentration and cAMP levels inside Escherichia coli: implications for models of phosphotransferase-mediated regulation of adenylate cyclase." *Microbiology* 143 (Pt 6): 1909-1918.

Ohkura, M., T. Sasaki, J. Sadakari, K. Gengyo-Ando, Y. Kagawa-Nagamura, C. Kobayashi, Y. Ikegaya and J. Nakai (2012). "Genetically encoded green fluorescent Ca2+ indicators with improved detectability for neuronal Ca2+ signals." *PloS one* 7(12).

Ormö, M., A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien and S. J. Remington (1996). "Crystal structure of the Aequorea victoria green fluorescent protein." *Science* 273(5280): 1392-1395.

Paoni, N. F. and D. E. Koshland (1979). "Permeabilization of cells for studies on the biochemistry of bacterial chemotaxis." *Proceedings of the National Academy of Sciences* 76(8): 3693-3697.

Paramonov, V. M., V. Mamaeva, C. Sahlgren and A. Rivero-Müller (2015). "Genetically-encoded tools for cAMP probing and modulation in living systems." *Frontiers in Pharmacology* 6.

Park, S. I., D. S. Brenner, G. Shin, C. D. Morgan, B. A. Copits, H. U. Chung, M. Y. Pullen, K. N. Noh, S. Davidson, S. J. Oh, J. Yoon, K.-I. I. Jang, V. K. Samineni, M. Norman, J. G. Grajales-Reyes, S. K. Vogt, S. S. Sundaram, K. M. Wilson, J. S. Ha, R. Xu, T. Pan, T.-I. I. Kim, Y. Huang, M. C. Montana, J. P. Golden, M. R. Bruchas, R. W. Gereau and J. A. Rogers (2015). "Soft, stretchable, fully implantable miniaturized optoelectronic systems for wireless optogenetics." *Nature biotechnology* 33(12): 1280.

Passner, J. M. and T. A. Steitz (1997). "The structure of a CAP–DNA complex having two cAMP molecules bound to each monomer." *Proceedings of the National Academy of Sciences* 94(7): 2843-2847.

Paul, S., P. Olausson, D. V. Venkitaramani, I. Ruchkina, T. D. Moran, N. Tronson, E. Mills, S. Hakim, M. W. Salter, J. R. Taylor and P. J. Lombroso (2007). "The striatal-enriched protein tyrosine phosphatase gates long-term potentiation and fear memory in the lateral amygdala." *Biological psychiatry* 61(9): 1049-1061.

Pédelacq, J.-D. D., S. Cabantous, T. Tran, T. C. Terwilliger and G. S. Waldo (2006). "Engineering and characterization of a superfolder green fluorescent protein." *Nature biotechnology* 24(1): 79-88.

Pelicic, V., J. M. Reyrat and B. Gicquel (1996). "Expression of the Bacillus subtilis sacB gene confers sucrose sensitivity on mycobacteria." *Journal of bacteriology* 178(4): 1197-1199.

Pettikiriarachchi, A., L. Gong, M. A. Perugini, R. J. Devenish and M. Prescott (2012). "Ultramarine, a chromoprotein acceptor for Förster resonance energy transfer." *PloS one* 7(7).

Piatkevich, K. D., J. Hulit, O. M. Subach, B. Wu, A. Abdulla, J. E. Segall and V. V. Verkhusha (2010). "Monomeric red fluorescent proteins with a large Stokes shift." *Proceedings of the National Academy of Sciences* 107(12): 5369-5374.

Piatkevich, K. D., V. N. Malashkevich, K. S. Morozova, N. A. Nemkovich, S. C. Almo and V. V. Verkhusha (2013). "Extended Stokes shift in fluorescent proteins: chromophore-protein interactions in a near-infrared TagRFP675 variant." *Scientific reports* 3: 1847.

Polito, M., E. Guiot, G. Gangarossa, S. Longueville, M. Doulazmi, E. Valjent, D. Hervé, J.-A. A. Girault, D. Paupardin-Tritsch, L. R. Castro and P. Vincent (2015). "Selective Effects of PDE10A Inhibitors on Striatopallidal Neurons Require Phosphatase Inhibition by DARPP-32(1,2,3)." *eNeuro* 2(4).

Polito, M., J. Klarenbeek, K. Jalink, D. Paupardin-Tritsch, P. Vincent and L. R. V. Castro (2013). "The NO/cGMP pathway inhibits transient cAMP signals through the activation of PDE2 in striatal neurons." *Frontiers in cellular neuroscience* 7: 211.

Prasher, D. C., V. K. Eckenrode, W. W. Ward, F. G. Prendergast and M. J. Cormier (1992). "Primary structure of the Aequorea victoria green-fluorescent protein." *Gene* 111(2): 229-233.

Rappsilber, J., M. Mann and Y. Ishihama (2007). "Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips." *Nature Protocols* 2(8): 1896-1906.

Rizzo, M. A., G. H. Springer, B. Granada and D. W. Piston (2004). "An improved cyan fluorescent protein variant useful for FRET." *Nature biotechnology* 22(4): 445-449.

Rose, T., P. Goltstein, R. Portugues and O. Griesbeck (2014). "Putting a finishing touch on GEC's." *Frontiers in molecular neuroscience* 7.

Russwurm, M., F. Mullershausen, A. Friebe, R. Jäger, C. Russwurm and D. Koesling (2007). "Design of fluorescence resonance energy transfer (FRET)-based cGMP indicators: a systematic approach." *The Biochemical journal* 407(1): 69-77.

Ryu, M.-H. H., O. V. Moskvin, J. Siltberg-Liberles and M. Gomelsky (2010). "Natural and engineered photoactivated nucleotidyl cyclases for optogenetic applications." *The Journal of biological chemistry* 285(53): 41501-41508.

Sakaguchi, R., S. Kiyonaka and Y. Mori (2015). "Fluorescent sensors reveal subcellular thermal changes." *Current opinion in biotechnology* 31: 57-64.

Sakaue-Sawano, A., H. Kurokawa, T. Morimura, A. Hanyu, H. Hama, H. Osawa, S. Kashiwagi, K. Fukami, T. Miyata, H. Miyoshi, T. Imamura, M. Ogawa, H. Masai and A. Miyawaki (2008). "Visualizing spatiotemporal dynamics of multicellular cell-cycle progression." *Cell* 132(3): 487-498.

Sandner, P., D. Neuser and E. Bischoff (2009). Erectile dysfunction and lower urinary tract. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 507-531.

Sankaranarayanan, S., D. Angelis, J. E. Rothman and T. A. Ryan (2000). "The Use of pHluorins for Optical Measurements of Presynaptic Activity." *Biophysical Journal* 79(4): 2199-2208.

Sato, M., N. Hida, T. Ozawa and Y. Umezawa (2000). "Fluorescent indicators for cyclic GMP based on cyclic GMP-dependent protein kinase Ialpha and green fluorescent proteins." *Analytical chemistry* 72(24): 5918-5924.

Scheltema, R. A., J.-P. P. Hauschild, O. Lange, D. Hornburg, E. Denisov, E. Damoc, A. Kuehn, A. Makarov and M. Mann (2014). "The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer." *Molecular & cellular proteomics* 13(12): 3698-3708.

Schlossmann, J. and M. Desch (2009). cGK substrates. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 163-193.

Schmidt, P. M. (2009). Biochemical Detection of cGMP From Past to Present: An Overview. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 195-228.

Scholz, J., H. Besir, C. Strasser and S. Suppmann (2013). "A new method to customize protein expression vectors for fast, efficient and background free parallel cloning." *BMC biotechnology* 13: 12.

Shagin, D. A., E. V. Barsova, Y. G. Yanushevich, A. F. Fradkov, K. A. Lukyanov, Y. A. Labas, T. N. Semenova, J. A. Ugalde, A. Meyers, J. M. Nunez, E. A. Widder, S. A. Lukyanov and M. V. Matz (2004). "GFP-like proteins as ubiquitous metazoan superfamily: evolution of functional features and structural complexity." *Molecular biology and evolution* 21(5): 841-850.

Shaner, N. C., R. E. Campbell, P. A. Steinbach, B. N. Giepmans, A. E. Palmer and R. Y. Tsien (2004). "Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein." *Nature biotechnology* 22(12): 1567-1572.

Shaner, N. C., G. G. Lambert, A. Chammas, Y. Ni, P. J. Cranfill, M. A. Baird, B. R. Sell, J. R. Allen, R. N. Day, M. Israelsson, M. W. Davidson and J. Wang (2013). "A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum." *Nature methods* 10(5): 407-409.

Shaner, N. C., M. Z. Lin, M. R. McKeown, P. A. Steinbach, K. L. Hazelwood, M. W. Davidson and R. Y. Tsien (2008). "Improving the photostability of bright monomeric orange and red fluorescent proteins." *Nature methods* 5(6): 545-551.

Shaner, N. C., P. A. Steinbach and R. Y. Tsien (2005). "A guide to choosing fluorescent proteins." *Nature methods* 2(12): 905-909.

Shcherbakova, D. M., M. A. Hink, L. Joosen, T. W. Gadella and V. V. Verkhusha (2012). "An orange fluorescent protein with a large Stokes shift for single-excitation multicolor FCCS and FRET imaging." *Journal of the American Chemical Society* 134(18): 7913-7923.

Shcherbo, D., E. M. Merzlyak, T. V. Chepurnykh, A. F. Fradkov, G. V. Ermakova, E. A. Solovieva, K. A. Lukyanov, E. A. Bogdanova, A. G. Zaraisky, S. Lukyanov and D. M. Chudakov (2007). "Bright far-red fluorescent protein for whole-body imaging." *Nature methods* 4(9): 741-746.

Shcherbo, D., C. S. Murphy, G. V. Ermakova, E. A. Solovieva, T. V. Chepurnykh, A. S. Shcheglov, V. V. Verkhusha, V. Z. Pletnev, K. L. Hazelwood, P. M. Roche, S. Lukyanov, A. G. Zaraisky, M. W. Davidson and D. M. Chudakov (2009). "Far-red fluorescent tags for protein imaging in living tissues." *The Biochemical journal* 418(3): 567-574.

Shcherbo, D., I. I. Shemiakina, A. V. Ryabova, K. E. Luker, B. T. Schmidt, E. A. Souslova, T. V. Gorodnicheva, L. Strukova, K. M. Shidlovskiy, O. V. Britanova, A. G. Zaraisky, K. A. Lukyanov, V. B. Loschenov, G. D. Luker and D. M. Chudakov (2010). "Near-infrared fluorescent proteins." *Nature methods* 7(10): 827-829.

Shemiakina, II, G. V. Ermakova, P. J. Cranfill, M. A. Baird, R. A. Evans, E. A. Souslova, D. B. Staroverov, A. Y. Gorokhovatsky, E. V. Putintseva, T. V. Gorodnicheva, T. V. Chepurnykh, L. Strukova, S. Lukyanov, A. G. Zaraisky, M. W. Davidson, D. M. Chudakov and D. Shcherbo (2012). "A monomeric red fluorescent protein with low cytotoxicity." *Nature communications* 3: 1204.

Shimomura, O., F. H. Johnson and Y. Saiga (1962). "Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea." *Journal of cellular and comparative physiology* 59: 223-239.

Shimozono, S., H. Hosoi, H. Mizuno, T. Fukano, T. Tahara and A. Miyawaki (2006). "Concatenation of cyan and yellow fluorescent proteins for efficient resonance energy transfer." *Biochemistry* 45(20): 6267-6271.

Stierl, M., P. Stumpf, D. Udwari, R. Gueta, R. Hagedorn, A. Losi, W. Gärtner, L. Petereit, M. Efetova, M. Schwarzel, T. G. Oertner, G. Nagel and P. Hegemann (2011). "Light modulation of cellular cAMP by a small bacterial photoactivated adenylyl cyclase, bPAC, of the soil bacterium Beggiatoa." *The Journal of biological chemistry* 286(2): 1181-1188.

Storace, D., M. S. Rad, Z. Han, L. Jin, L. B. Cohen, T. Hughes, B. J. Baker and U. Sung (2015). "Genetically Encoded Protein Sensors of Membrane Potential." *Advances in experimental medicine and biology* 859: 493-509.

Studier, F. W. (2005). "Protein production by auto-induction in high density shaking cultures." *Protein Expression and Purification* 41(1): 207-234.

Subach, O. M., P. J. Cranfill, M. W. Davidson and V. V. Verkhusha (2011). "An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore." *PloS one* 6(12).

Subach, O. M., I. S. Gundorov, M. Yoshimura, F. V. Subach, J. Zhang, D. Grüenwald, E. A. Souslova, D. M. Chudakov and V. V. Verkhusha (2008). "Conversion of red fluorescent protein into a bright blue probe." *Chemistry & biology* 15(10): 1116-1124.

Tallini, Y. N., M. Ohkura, B.-R. Choi, G. Ji, K. Imoto, R. Doran, J. Lee, P. Plan, J. Wilson, H.-B. Xin, A. Sanbe, J. Gulick, J. Mathai, J. Robbins, G. Salama, J. Nakai and M. I. Kotlikoff (2006). "Imaging cellular signals in the heart in vivo: Cardiac expression of the high-signal Ca2+ indicator GCaMP2." *Proceedings of the National Academy of Sciences* 103(12): 4753-4758.

Tantama, M., Y. Hung and G. Yellen (2012). "Optogenetic reporters: Fluorescent protein-based genetically encoded indicators of signaling and metabolism in the brain." *Progress in brain research* 196: 235-263.

Tantama, M., J. R. Martínez-François, R. Mongeon and G. Yellen (2013). "Imaging energy status in live cells with a fluorescent biosensor of the intracellular ATP-to-ADP ratio." *Nature communications* 4: 2550.

Thestrup, T., J. Litzlbauer, I. Bartholomäus, M. Mues, L. Russo, H. Dana, Y. Kovalchuk, Y. Liang, G. Kalamakis, Y. Laukat, S. Becker, G. Witte, A. Geiger, T. Allen, L. C. Rome, T.-W. W. Chen, D. S. Kim, O. Garaschuk, C. Griesinger and O. Griesbeck (2014). "Optimized ratiometric calcium sensors for functional in vivo imaging of neurons and T lymphocytes." *Nature methods* 11(2): 175-182.

Thunemann, M., L. Wen, M. Hillenbrand, A. Vachaviolos, S. Feil, T. Ott, X. Han, D. Fukumura, R. K. Jain, M. Russwurm, C. de Wit and R. Feil (2013). "Transgenic mice for cGMP imaging." *Circulation research* 113(4): 365-371.

Tian, L., S. A. Hires, T. Mao, D. Huber, M. E. Chiappe, S. H. Chalasani, L. Petreanu, J. Akerboom, S. A. McKinney, E. R. Schreiter, C. I. Bargmann, V. Jayaraman, K. Svoboda and L. L. Looger (2009). "Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators." *Nature methods* 6(12): 875-881.

Toettcher, J. E., C. A. Voigt, O. D. Weiner and W. A. Lim (2011). "The promise of optogenetics in cell biology: interrogating molecular circuits in space and time." *Nature methods* 8(1): 35-38.

Tsien, R. Y. (1998). "The green fluorescent protein." *Annual review of biochemistry* 67: 509-544.

Tsuji, T. and T. Kunieda (2005). "A loss-of-function mutation in natriuretic peptide receptor 2 (Npr2) gene is responsible for disproportionate dwarfism in cn/cn mouse." *The Journal of biological chemistry* 280(14): 14288-14292.

Tsutsui, H., S. Karasawa, Y. Okamura and A. Miyawaki (2008). "Improving membrane voltage measurements using FRET with new fluorescent proteins." *Nature methods* 5(8): 683-685.

Valeur, B. and M. N. Berberan-Santos (2012). Molecular fluorescence : principles and applications. Weinheim, Germany; Chichester, England, Wiley-VCH ; Wiley-VCH Verlag GmbH & Co. KGaA.

Valla, S. and R. Lale (2014). DNA cloning and assembly methods, Springer.

Walter, U. and S. Gambaryan (2009). cGMP and cGMP-dependent protein kinase in platelets and blood cells. Handbook of experimental pharmacology; cGMP: Generators, Effectors and Therapeutic Implications 533-548.

Wang, L., W. C. Jackson, P. A. Steinbach and R. Y. Tsien (2004). "Evolution of new nonantibody proteins via iterative somatic hypermutation." *Proceedings of the National Academy of Sciences* 101(48): 16745-16749.

Wu, F., E. Stark, P.-C. C. Ku, K. D. Wise, G. Buzsáki and E. Yoon (2015). "Monolithically Integrated µLEDs on Silicon Neural Probes for High-Resolution Optogenetic Studies in Behaving Animals." *Neuron* 88(6): 1136-1148.

Xu, Y., A. Kanauchi, A. G. von Arnim, D. W. Piston and C. H. Johnson (2003). "Bioluminescence resonance energy transfer: monitoring protein-protein interactions in living cells." *Methods in enzymology* 360: 289-301.

Yang, F., L. G. Moss and G. N. Phillips (1996). "The molecular structure of green fluorescent protein." *Nature biotechnology* 14(10): 1246-1251.

Yang, J., L. Wang, F. Yang, H. Luo, L. Xu, J. Lu, S. Zeng and Z. Zhang (2013). "mBeRFP, an improved large stokes shift red fluorescent protein." *PloS one* 8(6).

Yang, T. T., L. Cheng and S. R. Kain (1996). "Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein." *Nucleic acids research* 24(22): 4592-4593.

Yang, Z. and T. G. Wensel (1992). "Inorganic pyrophosphatase from bovine retinal rod outer segments." *The Journal of biological chemistry* 267(34): 24634-24640.

Zapata-Hommer, O. and O. Griesbeck (2003). "Efficiently folding and circularly permuted variants of the Sapphire mutant of GFP." *BMC biotechnology* 3: 5.

Zhang, Y., U. Werling and W. Edelmann (2012). "SLiCE: a novel bacterial cell extract-based DNA cloning method." *Nucleic acids research* 40(8): e55.

References

Zhang, Y., U. Werling and W. Edelmann (2014). "Seamless Ligation Cloning Extract (SLiCE) cloning method." *Methods in Molecular Biology* 1116: 235-244.

6 Appendix

6.1 Plasmid Sequences

6.1.1 pRSET SL

GenBank Header

```
LOCUS       pRSET_SL2887 bp       DNA    SYN    06-APR-2016
DEFINITION  pRSET_SL
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     promoter        20..38
                     /label=T7_promoter
     misc_feature    75..91
                     /label=T7_transl_en_RBS
     misc_feature    112..129
                     /label=6xHis
     misc_feature    133..165
                     /label=T7_gene10_leader
     terminator      254..382
                     /label=T7_terminator
     rep_origin      471..777
                     /label=f1_origin
     promoter        970..998
                     /label=AmpR_promoter
     gene            1040..1900
                     /gene="beta lactamase"
     rep_origin      2055..2674
                     /label=pBR322_origin
     misc_feature    214..222
                     /note="3xSTOP"
      misc_feature    169..207
                     /note="upstream homologous region"
     misc_feature    214..246
                     /note="downstream homologous region"
     misc_feature    199..207
                     /note="Kozak_Start"
     source          1..2895
                     /dnas_title="pRSET_SL"
```

DNA

```
ORIGIN
        1 GATCTCGATC CCGCGAAATT AATACGACTC ACTATAGGGA GACCACAACG GTTTCCCTCT
       61 AGAAATAATT TTGTTTAACT TTAAGAAGGA GATATACATA TGCGGGGTTC TCATCATCAT
      121 CATCATCATG GTATGGCTAG CATGACTGGT GGACAGCAAA TGGGTCGGGA TCTGTACGAC
      181 GATGACGATA AGGATCCGGC CACCATGGAT ATCTGATAAT AGGAATTCCT ATAGTGTCAC
      241 CTAAATGATC CGGCTGCTAA CAAAGCCCGA AAGGAAGCTG AGTTGGCTGC TGCCACCGCT
      301 GAGCAATAAC TAGCATAACC CCTTGGGGCC TCTAAACGGG TCTTGAGGGG TTTTTTGCTG
      361 AAAGGAGGAA CTATATCCGG ATCTGGCGTA ATAGCGAAGA GGCCCGCACC GATCGCCCTT
      421 CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGGACGCGCC CTGTAGCGGC GCATTAAGCG
      481 CGGCGGGTGT GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG
      541 CTCCTTTCGC TTTCTTCCCT TCCTTTCTCG CCACGTTCGC CGGCTTTCCC CGTCAAGCTC
      601 TAAATCGGGG GCTCCCTTTA GGGTTCCGAT TTAGTGCTTT ACGGCACCTC GACCCCAAAA
      661 AACTTGATTA GGGTGATGGT TCACGTAGTG GGCCATCGCC CTGATAGACG GTTTTTCGCC
      721 CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT GTTCCAAACT GGAACAACAC
      781 TCAACCCTAT CTCGGTCTAT TCTTTTGATT TATAAGGGAT TTTGCCGATT TCGGCCTATT
      841 GGTTAAAAAA TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAACAAA ATATTAACGC
      901 TTACAATTTA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT GTTTATTTTT
      961 CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA CCCTGATAAA TGCTTCAATA
     1021 ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT GTCGCCCTTA TTCCCTTTTT
     1081 TGCGGCATTT TGCCTTCCTG TTTTTGCTCA CCCAGAAACG CTGGTGAAAG TAAAAGATGC
     1141 TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT
     1201 CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTTCTGCT
     1261 ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA
     1321 CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG
     1381 CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA
     1441 CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG
     1501 GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA
     1561 CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAAC TATTAACTGG
     1621 CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG CGGATAAAGT
     1681 TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG
     1741 AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC
     1801 CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
     1861 GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC AAGTTTACTC
     1921 ATATATACTT TAGATTGATT TAAAACTTCA TTTTTAATTT AAAAGGATCT AGGTGAAGAT
     1981 CCTTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC
     2041 AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT TTTTTTCTGC GCGTAATCTG
     2101 CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG ATCAAGAGCT
     2161 ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA ATACTGTTCT
     2221 TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC CTACATACCT
     2281 CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT GTCTTACCGG
     2341 GTTGGACTCA AGACGATAGT TACCGGATAA GGCGCAGCGG TCGGGCTGAA CGGGGGGTTC
     2401 GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA
     2461 GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC CGGTAAGCGG
     2521 CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAAACGCCT GGTATCTTTA
     2581 TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG
     2641 GGGGCGGAGC CTATGGAAAA ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG
     2701 CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCCT GATTCTGTGG ATAACCGTAT
     2761 TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC GCAGCGAGTC
     2821 AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC
     2881 GATTCATTAA TGCAG
//
```

6.1.2 pRSET SL II

GenBank header

```
LOCUS       pRSET_SL_II    2887 bp       DNA    SYN    06-APR-2016
DEFINITION  pRSET_SL_II
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     promoter        20..38
                     /label=T7_promoter
     misc_feature    75..91
                     /label=T7_transl_en_RBS
     misc_feature    112..129
                     /label=6xHis
     misc_feature    133..165
                     /label=T7_gene10_leader
     terminator      1737..1865
                     /label=T7_terminator
     rep_origin      1954..2260
                     /label=f1_origin
     promoter        2453..2481
                     /label=AmpR_promoter
     gene            2523..3383
                     /gene="beta lactamse"
     rep_origin      3538..4157
                     /label=pBR322_origin
     misc_feature    1697..1705
                     /note="3xSTOP"
     misc_feature    169..207
                     /note="upstream homologous region"
     misc_feature    1697..1729
                     /note="downstream homologous region"
     misc_feature    199..207
                     /note="Kozak_Start"
     Promoter        214..268
                     /label=LPP5 promoter
     gene            269..1690
                     /gene="sacB"
     source          1..4378
                     /dnas_title="pRSET_SL_II"
```

DNA

```
ORIGIN
        1 GATCTCGATC CCGCGAAATT AATACGACTC ACTATAGGGA GACCACAACG GTTTCCCTCT
       61 AGAAATAATT TTGTTTAACT TTAAGAAGGA GATATACATA TGCGGGGTTC TCATCATCAT
      121 CATCATCATG GTATGGCTAG CATGACTGGT GGACAGCAAA TGGGTCGGGA TCTGTACGAC
      181 GATGACGATA AGGATCCGGC CACCATGGAT ATCAAACTGC AGTTGACAAC ATAAAAACTT
      241 TGTGTTATAC TTGTAACGTA AGGAGGTAAT GAACATCAAA AAGTTTGCAA AACAAGCAAC
      301 AGTATTAACC TTTACTACCG CACTGCTGGC AGGAGGCGCA ACTCAAGCGT TTGCGAAAGA
      361 AACGAACCAA AAGCCATATA AGGAAACATA CGGCATTTCC CATATTACAC GCCATGATAT
      421 GCTGCAAATC CCTGAACAGC AAAAAAATGA AAAATATCAA GTTCCTGAAT TCGATTCGTC
      481 CACAATTAAA AATATCTCTT CTGCAAAAGG CCTGGACGTT TGGGACAGCT GGCCATTACA
      541 AAACGCTGAC GGCACTGTCG CAAACTATCA CGGCTACCAC ATCGTCTTTG CATTAGCCGG
      601 AGATCCTAAA AATGCGGATG ACACATCGAT TTACATGTTC TATCAAAAAG TCGGCGAAAC
      661 TTCTATTGAC AGCTGGAAAA ACGCTGGCCG CGTCTTTAAA GACAGCGACA AATTCGATGC
      721 AAATGATTCT ATCCTAAAAG ACCAAACACA AGAATGGTCA GGTTCAGCCA CATTTACATC
      781 TGACGGAAAA ATCCGTTTAT TCTACACTGA TTTCTCCGGT AAACATTACG GCAAACAAAC
      841 ACTGACAACT GCACAAGTTA ACGTATCAGC ATCAGACAGC TCTTTGAACA TCAACGGTGT
      901 AGAGGATTAT AAATCAATCT TTGACGGTGA CGGAAAAACG TATCAAAATG TACAGCAGTT
      961 CATCGATGAA GGCAACTACA GCTCAGGCGA CAACCATACG CTGAGAGATC CTCACTACGT
     1021 AGAAGATAAA GGCCACAAAT ACTTAGTATT TGAAGCAAAC ACTGGAACTG AAGATGGCTA
     1081 CCAAGGCGAA GAATCTTTAT TTAACAAAGC ATACTATGGC AAAAGCACAT CATTCTTCCG
     1141 TCAAGAAAGT CAAAAACTTC TGCAAAGCGA TAAAAAACGC ACGGCTGAGT TAGCAAACGG
     1201 CGCTCTCGGT ATGATTGAGC TAAACGATGA TTACACACTG AAAAAAGTGA TGAAACCGCT
     1261 GATTGCATCT AACACAGTAA CAGATGAAAT TGAACGCGCG AACGTCTTTA AAATGAACGG
     1321 CAAATGGTAC CTGTTCACTG ACTCCCGCGG ATCAAAAATG ACGATTGACG GCATTACGTC
     1381 TAACGATATT TACATGCTTG GTTATGTTTC TAATTCTTTA ACTGGCCCAT ACAAGCCGCT
     1441 GAACAAAACT GGCCTTGTGT TAAAAATGGA TCTTGATCCT AACGATGTAA CCTTTACTTA
     1501 CTCACACTTC GCTGTACCTC AAGCGAAAGG AAACAATGTC GTGATTACAA GCTATATGAC
     1561 AAACAGAGGA TTCTACGCAG ACAAACAATC AACGTTTGCG CCAAGCTTCC TGCTGAACAT
     1621 CAAAGGCAAG AAAACATCTG TTGTCAAAGA CAGCATCCTT GAACAAGGAC AATTAACAGT
     1681 TAACAAATAA GATATCTGAT AATAGGAATT CCTATAGTGT CACCTAAATG ATCCGGCTGC
     1741 TAACAAAGCC CGAAAGGAAG CTGAGTTGGC TGCTGCCACC GCTGAGCAAT AACTAGCATA
     1801 ACCCCTTGGG GCCTCTAAAC GGGTCTTGAG GGGTTTTTTG CTGAAAGGAG GAACTATATC
     1861 CGGATCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC CTTCCCAACA GTTGCGCAGC
     1921 CTGAATGGCG AATGGGACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG TGTGGTGGTT
     1981 ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTTCTTC
     2041 CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCGTCAAG CTCTAAATCG GGGGCTCCCT
     2101 TTAGGGTTCC GATTTAGTGC TTTACGGCAC CTCGACCCCA AAAAACTTGA TTAGGGTGAT
     2161 GGTTCACGTA GTGGGCCATC GCCCTGATAG ACGGTTTTTC GCCCTTTGAC GTTGGAGTCC
     2221 ACGTTCTTTA ATAGTGGACT CTTGTTCCAA ACTGGAACAA CACTCAACCC TATCTCGGTC
     2281 TATTCTTTTG ATTTATAAGG GATTTTGCCG ATTTCGGCCT ATTGGTTAAA AAATGAGCTG
     2341 ATTTAACAAA AATTTAACGC GAATTTTAAC AAAATATTAA CGCTTACAAT TTAGGTGGCA
     2401 CTTTTCGGGG AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA
     2461 TGTATCCGCT CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA
     2521 GTATGAGTAT TCAACATTTC CGTGTCGCCC TTATTCCCTT TTTTGCGGCA TTTTGCCTTC
     2581 CTGTTTTTGC TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG
     2641 CACGAGTGGG TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTCGCC
     2701 CCGAAGAACG TTTTCCAATG ATGAGCACTT TTAAAGTTCT GCTATGTGGC GCGGTATTAT
     2761 CCCGTATTGA CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACTATTCT CAGAATGACT
     2821 TGGTTGAGTA CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT
     2881 TATGCAGTGC TGCCATAACC ATGAGTGATA ACACTGCGGC CAACTTACTT CTGACAACGA
     2941 TCGGAGGACC GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT GTAACTCGCC
     3001 TTGATCGTTG GGAACCGGAG CTGAATGAAG CCATACCAAA CGACGAGCGT GACACCACGA
     3061 TGCCTGTAGC AATGGCAACA ACGTTGCGCA AACTATTAAC TGGCGAACTA CTTACTCTAG
     3121 CTTCCCGGCA ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACTTCTGC
     3181 GCTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT GAGCGTGGGT
     3241 CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT
     3301 ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG
     3361 CCTCACTGAT TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG
     3421 ATTTAAAACT TCATTTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA
     3481 TGACCAAAAT CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA
     3541 TCAAAGGATC TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA
     3601 AACCACCGCT ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA
     3661 AGGTAACTGG CTTCAGCAGA GCGCAGATAC CAAATACTGT TCTTCTAGTG TAGCCGTAGT
     3721 TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT
     3781 TACCAGTGGC TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT
     3841 AGTTACCGGA TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT
     3901 TGGAGCGAAC GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA
     3961 CGCTTCCCGA AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG
     4021 AGCGCACGAG GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC
     4081 GCCACCTCTG ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA
     4141 AAAACGCCAG CAACGCGGCC TTTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA
     4201 TGTTCTTTCC TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG
     4261 CTGATACCGC TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG
     4321 AAGAGCGCCC AATACGCAAA CCGCCTCTCC CCGCGCGTTG GCCGATTCAT TAATGCAG//
```

6.1.3 pRSET SL III

GenBank Header

```
LOCUS       pRSET_SL_III   2887 bp       DNA    SYN    06-APR-2016
DEFINITION  pRSET_SL_III
ACCESSION
KEYWORDS
SOURCE
  ORGANISM  other sequences; artificial sequences; vectors.
FEATURES             Location/Qualifiers
     promoter        20..38
                     /label=T7_promoter
     misc_feature    75..91
                     /label=T7_transl_en_RBS
     misc_feature    148..180
                     /label=T7_gene10_leader
     terminator      1752..1880
                     /label=T7_terminator
     rep_origin      1969..2275
                     /label=f1_origin
     promoter        2468..2496
                     /label=AmpR_promoter
     gene            2538..3398
                     /gene="beta lactamase"
     rep_origin      3553..4172
                     /label=pBR322_origin
     misc_feature    1712..1720
                     /note="3xSTOP"
     misc_feature    184..222
                     /note="upstream homologous region"
     misc_feature    1712..1744
                     /note="downstream homologous region"
     misc_feature    214..222
                     /note="Kozak_Start"
     Promoter        229..283
                     /label=LPP5 promoter
     gene            284..1705
                     /gene="sacB"
     misc_feature    112..141
                     /note="10 x His "
     source          1..4393
                     /dnas_title="pRSET_SL_III"
```

DNA

```
ORIGIN
        1 GATCTCGATC CCGCGAAATT AATACGACTC ACTATAGGGA GACCACAACG GTTTCCCTCT
       61 AGAAATAATT TTGTTTAACT TTAAGAAGGA GATATACATA TGCGGGGTTC TCATCATCAT
      121 CATCATCACC ATCACCATCA CCGCGGTATG GCTAGCATGA CTGGTGGACA GCAAATGGGT
      181 CGGGATCTGT ACGACGATGA CGATAAGGAT CCGGCCACCA TGGATATCAA ACTGCAGTTG
      241 ACAACATAAA AACTTTGTGT TATACTTGTA ACGTAAGGAG GTAATGAACA TCAAAAAGTT
      301 TGCAAAACAA GCAACAGTAT TAACCTTTAC TACCGCACTG CTGGCAGGAG GCGCAACTCA
      361 AGCGTTTGCG AAAGAAACGA ACCAAAAGCC ATATAAGGAA ACATACGGCA TTTCCCATAT
      421 TACACGCCAT GATATGCTGC AAATCCCTGA ACAGCAAAAA AATGAAAAAT ATCAAGTTCC
      481 TGAATTCGAT TCGTCCACAA TTAAAAATAT CTCTTCTGCA AAAGGCCTGG ACGTTTGGGA
      541 CAGCTGGCCA TTACAAAACG CTGACGGCAC TGTCGCAAAC TATCACGGCT ACCACATCGT
      601 CTTTGCATTA GCCGGAGATC CTAAAAATGC GGATGACACA TCGATTTACA TGTTCTATCA
      661 AAAAGTCGGC GAAACTTCTA TTGACAGCTG GAAAAACGCT GGCCGCGTCT TTAAAGACAG
      721 CGACAAATTC GATGCAAATG ATTCTATCCT AAAAGACCAA ACACAAGAAT GGTCAGGTTC
      781 AGCCACATTT ACATCTGACG GAAAAATCCG TTTATTCTAC ACTGATTTCT CCGGTAAACA
      841 TTACGGCAAA CAAACACTGA CAACTGCACA AGTTAACGTA TCAGCATCAG ACAGCTCTTT
      901 GAACATCAAC GGTGTAGAGG ATTATAAATC AATCTTTGAC GGTGACGGAA AAACGTATCA
      961 AAATGTACAG CAGTTCATCG ATGAAGGCAA CTACAGCTCA GGCGACAACC ATACGCTGAG
     1021 AGATCCTCAC TACGTAGAAG ATAAAGGCCA CAAATACTTA GTATTTGAAG CAAACACTGG
     1081 AACTGAAGAT GGCTACCAAG GCGAAGAATC TTTATTTAAC AAAGCATACT ATGGCAAAAG
     1141 CACATCATTC TTCCGTCAAG AAAGTCAAAA ACTTCTGCAA AGCGATAAAA AACGCACGGC
     1201 TGAGTTAGCA AACGGCGCTC TCGGTATGAT TGAGCTAAAC GATGATTACA CACTGAAAAA
     1261 AGTGATGAAA CCGCTGATTG CATCTAACAC AGTAACAGAT GAAATTGAAC GCGCGAACGT
     1321 CTTTAAAATG AACGGCAAAT GGTACCTGTT CACTGACTCC CGCGGATCAA AAATGACGAT
     1381 TGACGGCATT ACGTCTAACG ATATTTACAT GCTTGGTTAT GTTTCTAATT CTTTAACTGG
     1441 CCCATACAAG CCGCTGAACA AAACTGGCCT TGTGTTAAAA ATGGATCTTG ATCCTAACGA
     1501 TGTAACCTTT ACTTACTCAC ACTTCGCTGT ACCTCAAGCG AAAGGAAACA ATGTCGTGAT
     1561 TACAAGCTAT ATGACAAACA GAGGATTCTA CGCAGACAAA CAATCAACGT TTGCGCCAAG
     1621 CTTCCTGCTG AACATCAAAG GCAAGAAAAC ATCTGTTGTC AAAGACAGCA TCCTTGAACA
     1681 AGGACAATTA ACAGTTAACA AATAAGATAT CTGATAATAG GAATTCCTAT AGTGTCACCT
     1741 AAATGATCCG GCTGCTAACA AAGCCCGAAA GGAAGCTGAG TTGGCTGCTG CCACCGCTGA
     1801 GCAATAACTA GCATAACCCC TTGGGGCCTC TAAACGGGTC TTGAGGGGTT TTTTGCTGAA
     1861 AGGAGGAACT ATATCCGGAT CTGGCGTAAT AGCGAAGAGG CCCGCACCGA TCGCCCTTCC
     1921 CAACAGTTGC GCAGCCTGAA TGGCGAATGG GACGCGCCCT GTAGCGGCGC ATTAAGCGCG
     1981 GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG CCAGCGCCCT AGCGCCCGCT
     2041 CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG GCTTTCCCCG TCAAGCTCTA
     2101 AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA CCCCAAAAAA
     2161 CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT TTTTCGCCCT
     2221 TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT TCCAAACTGG AACAACACTC
     2281 AACCCTATCT CGGTCTATTC TTTTGATTTA TAAGGGATTT TGCCGATTTC GGCCTATTGG
     2341 TTAAAAAATG AGCTGATTTA ACAAAAATTT AACGCGAATT TTAACAAAAT ATTAACGCTT
     2401 ACAATTTAGG TGGCACTTTT CGGGGAAATG TGCGCGGAAC CCCTATTTGT TTATTTTTCT
     2461 AAATACATTC AAATATGTAT CCGCTCATGA GACAATAACC CTGATAAATG CTTCAATAAT
     2521 ATTGAAAAAG GAAGAGTATG AGTATTCAAC ATTTCCGTGT CGCCCTTATT CCCTTTTTTG
     2581 CGGCATTTTG CCTTCCTGTT TTTGCTCACC CAGAAACGCT GGTGAAAGTA AAAGATGCTG
     2641 AAGATCAGTT GGGTGCACGA GTGGGTTACA TCGAACTGGA TCTCAACAGC GGTAAGATCC
     2701 TTGAGAGTTT TCGCCCCGAA GAACGTTTTC CAATGATGAG CACTTTTAAA GTTCTGCTAT
     2761 GTGGCGCGGT ATTATCCCGT ATTGACGCCG GGCAAGAGCA ACTCGGTCGC CGCATACACT
     2821 ATTCTCAGAA TGACTTGGTT GAGTACTCAC CAGTCACAGA AAAGCATCTT ACGGATGGCA
     2881 TGACAGTAAG AGAATTATGC AGTGCTGCCA TAACCATGAG TGATAACACT GCGGCCAACT
     2941 TACTTCTGAC AACGATCGGA GGACCGAAGG AGCTAACCGC TTTTTTGCAC AACATGGGGG
     3001 ATCATGTAAC TCGCCTTGAT CGTTGGGAAC CGGAGCTGAA TGAAGCCATA CCAAACGACG
     3061 AGCGTGACAC CACGATGCCT GTAGCAATGG CAACAACGTT GCGCAAACTA TTAACTGGCG
     3121 AACTACTTAC TCTAGCTTCC CGGCAACAAT TAATAGACTG GATGGAGGCG GATAAAGTTG
     3181 CAGGACCACT TCTGCGCTCG GCCCTTCCGG CTGGCTGGTT TATTGCTGAT AAATCTGGAG
     3241 CCGGTGAGCG TGGGTCTCGC GGTATCATTG CAGCACTGGG GCCAGATGGT AAGCCCTCCC
     3301 GTATCGTAGT TATCTACACG ACGGGGAGTC AGGCAACTAT GGATGAACGA AATAGACAGA
     3361 TCGCTGAGAT AGGTGCCTCA CTGATTAAGC ATTGGTAACT GTCAGACCAA GTTTACTCAT
     3421 ATATACTTTA GATTGATTTA AAACTTCATT TTTAATTTAA AAGGATCTAG GTGAAGATCC
     3481 TTTTTGATAA TCTCATGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC TGAGCGTCAG
     3541 ACCCCGTAGA AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC GTAATCTGCT
     3601 GCTTGCAAAC AAAAAAACCA CCGCTACCAG CGGTGGTTTG TTTGCCGGAT CAAGAGCTAC
     3661 CAACTCTTTT TCCGAAGGTA ACTGGCTTCA GCAGAGCGCA GATACCAAAT ACTGTTCTTC
     3721 TAGTGTAGCC GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT ACATACCTCG
     3781 CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT CTTACCGGGT
     3841 TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG GGGGGTTCGT
     3901 GCACACAGCC CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA CAGCGTGAGC
     3961 TATGAGAAAG CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA
     4021 GGGTCGGAAC AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCTGG TATCTTTATA
     4081 GTCCTGTCGG GTTTCGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC TCGTCAGGGG
     4141 GGCGGAGCCT ATGGAAAAAC GCCAGCAACG CGGCCTTTTT ACGGTTCCTG GCCTTTTGCT
     4201 GGCCTTTTGC TCACATGTTC TTTCCTGCGT TATCCCCTGA TTCTGTGGAT AACCGTATTA
     4261 CCGCCTTTGA GTGAGCTGAT ACCGCTCGCC GCAGCCGAAC GACCGAGCGC AGCGAGTCAG
     4321 TGAGCGAGGA AGCGGAAGAG CGCCCAATAC GCAAACCGCC TCTCCCCGCG CGTTGGCCGA
     4381 TTCATTAATG CAG //
```

6.1.4 pcDNA3 SL

GenBank Header

```
LOCUS       pcDNA3_SL               5390 bp    DNA             06-APR-2016
FEATURES             Location/Qualifiers
     misc_feature    0..0
                     /note="1. pcDNAII 3013bp, M13 ori/amp gene -> pcDNA3
                     5446bp"
     gene            2095..2889
                     /gene="Neomycin resistance gene"
     promoter        864..882
                     /note="T7_promoter"
     promoter        complement(943..960)
                     /note="PRO bacteriophage Sp6"
     rep_origin      3576..4249
                     /note="ORI E. coli pMB1 (ColE1 and pBR322)"
     gene            complement(4494..5254)
                     /gene="beta lactamase"
     rep_origin      1734..2059
                     /note="f1 origin"
     polyA_signal    962..1193
                     /note="BGA polyA"
     promoter        209..863
                     /note="CMV promoter"
     polyA_signal    2944..3316
                     /note="SV40 polyA"
     rep_origin      1928..2013
                     /note="SV40 ori"
     misc_feature    928..936
                     /note="3xSTOP"
     misc_recomb     883..921
                     /note="upstream homologous region"
     misc_recomb     928..960
                     /note="downstream homologous region"
     misc_recomb     913..921
                     /note="Kozak_Start"
     source          1..5390
                     /dnas_title="pcDNA3_SL"
```

DNA

```
ORIGIN
        1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC TGCTCTGATG
       61 CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
      121 CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
      181 TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
      241 GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
      301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
      361 CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
      421 ATTGACGTCA ATGGGTGGAC TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
      481 ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
      541 ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
      601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
      661 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
      721 AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
      781 GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
      841 CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGGATCTGTA CGACGATGAC
      901 GATAAGGATC CGGCCACCAT GGATATCTGA TAATAGGAAT TCCTATAGTG TCACCTAAAT
      961 GCTAGAGCTC GCTGATCAGC CTCGACTGTG CCTTCTAGTT GCCAGCCATC TGTTGTTTGC
     1021 CCCTCCCCCG TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA
     1081 AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG GGGTGGGGTG
     1141 GGGCAGGACA GCAAGGGGGA GGATTGGGAA GACAATAGCA GGCATGCTGG GGATGCGGTG
     1201 GGCTCTATGG CTTCTGAGGC GGAAAGAACC AGCTGGGGCT CTAGGGGGTA TCCCCACGCG
     1261 CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA
     1321 CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCCTTTCT CGCCACGTTC
     1381 GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGCATCCCTT TAGGGTTCCG ATTTAGTGCT
     1441 TTACGGCACC TCGACCCCAA AAAACTTGAT TAGGGTGATG GTTCACGTAG TGGGCCATCG
     1501 CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC
     1561 TTGTTCCAAA CTGGAACAAC ACTCAACCCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG
     1621 ATTTTGGGGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG
     1681 AATTAATTCT GTGGAATGTG TGTCAGTTAG GGTGTGGAAA GTCCCCAGGC TCCCCAGGCA
     1741 GGCAGAAGTA TGCAAAGCAT GCATCTCAAT TAGTCAGCAA CCAGGTGTGG AAAGTCCCCA
     1801 GGCTCCCCAG CAGGCAGAAG TATGCAAAGC ATGCATCTCA ATTAGTCAGC AACCATAGTC
     1861 CCGCCCCTAA CTCCGCCCAT CCCGCCCCTA ACTCCGCCCA GTTCCGCCCA TTCTCCGCCC
     1921 CATGGCTGAC TAATTTTTTT TATTTATGCA GAGGCCGAGG CCGCCTCTGC CTCTGAGCTA
     1981 TTCCAGAAGT AGTGAGGAGG CTTTTTTGGA GGCCTAGGCT TTTGCAAAAA GCTCCCGGGA
     2041 GCTTGTATAT CCATTTTCGG ATCTGATCAA GAGACAGGAT GAGGATCGTT TCGCATGATT
     2101 GAACAAGATG GATTGCACGC AGGTTCTCCG GCCGCTTGGG TGGAGAGGCT ATTCGGCTAT
     2161 GACTGGGCAC AACAGACAAT CGGCTGCTCT GATGCCGCCG TGTTCCGGCT GTCAGCGCAG
     2221 GGGCGCCCGG TTCTTTTTGT CAAGACCGAC CTGTCCGGTG CCCTGAATGA ACTGCAGGAC
     2281 GAGGCAGCGC GGCTATCGTG GCTGGCCACG ACGGGCGTTC CTTGCGCAGC TGTGCTCGAC
     2341 GTTGTCACTG AAGCGGGAAG GGACTGGCTG CTATTGGGCG AAGTGCCGGG GCAGGATCTC
     2401 CTGTCATCTC ACCTTGCTCC TGCCGAGAAA GTATCCATCA TGGCTGATGC AATGCGGCGG
     2461 CTGCATACGC TTGATCCGGC TACCTGCCCA TTCGACCACC AAGCGAAACA TCGCATCGAG
     2521 CGAGCACGTA CTCGGATGGA AGCCGGTCTT GTCGATCAGG ATGATCTGGA CGAAGAGCAT
     2581 CAGGGGCTCG CGCCAGCCGA ACTGTTCGCC AGGCTCAAGG CGCGCATGCC CGACGGCGAG
     2641 GATCTCGTCG TGACCCATGG CGATGCCTGC TTGCCGAATA TCATGGTGGA AAATGGCCGC
     2701 TTTTCTGGAT TCATCGACTG TGGCCGGCTG GGTGTGGCGG ACCGCTATCA GGACATAGCG
     2761 TTGGCTACCC GTGATATTGC TGAAGAGCTT GGCGGCGAAT GGGCTGACCG CTTCCTCGTG
     2821 CTTTACGGTA TCGCCGCTCC CGATTCGCAG CGCATCGCCT TCTATCGCCT TCTTGACGAG
     2881 TTCTTCTGAG CGGGACTCTG GGGTTCGAAA TGACCGACCA AGCGACGCCC AACCTGCCAT
     2941 CACGAGATTT CGATTCCACC GCCGCCTTCT ATGAAAGGTT GGGCTTCGGA ATCGTTTTCC
     3001 GGGACGCCGG CTGGATGATC CTCCAGCGCG GGGATCTCAT GCTGGAGTTC TTCGCCCACC
     3061 CCAACTTGTT TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC ACAAATTTCA
     3121 CAAATAAAGC ATTTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC ATCAATGTAT
     3181 CTTATCATGT CTGTATACCG TCGACCTCTA GCTAGAGCTT GGCGTAATCA TGGTCATAGC
     3241 TGTTTCCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA CAACATACGA GCCGGAAGCA
     3301 TAAAGTGTAA AGCCTGGGGT GCCTAATGAG TGAGCTAACT CACATTAATT GCGTTGCGCT
     3361 CACTGCCCGC TTTCCAGTCG GGAAACCTGT CGTGCCAGCT GCATTAATGA ATCGGCCAAC
     3421 GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCCGC TTCCTCGCTC ACTGACTCGC
     3481 TGCGCTCGGT CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT
     3541 TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG
     3601 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCCTGACG
     3661 AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT
     3721 ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA
     3781 CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCACGCT
     3841 GTAGGTATCT CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG CACGAACCCC
     3901 CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA
     3961 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG
     4021 TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG
     4081 TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT
     4141 GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA
     4201 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC
     4261 AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA
     4321 CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA
     4381 CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT
```

```
4441 TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT
4501 TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT
4561 TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT
4621 CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA
4681 ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG
4741 GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT
4801 TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG
4861 CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC ATGCCATCCG
4921 TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC
4981 GGCGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA
5041 CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA AGGATCTTAC
5101 CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT
5161 TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAAGG
5221 GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA TATTATTGAA
5281 GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA
5341 AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC
//
```

6.1.5 pcDNA3 SL II

GenBank Header

```
LOCUS       pcDNA3_SL_II            6873 bp    DNA             06-APR-2016

FEATURES             Location/Qualifiers

     gene            3578..4372
                     /gene="Neomycin resistance gene"
     promoter        864..882
                     /note="PRO bacteriophage T7"
     promoter        complement(2426..2443)
                     /note="PRO bacteriophage Sp6"
     rep_origin      5059..5732
                     /note="ORI E. coli pMB1 (ColE1 and pBR322)"
     gene            complement(5977..6737)
                     /gene="beta lactamase"
     rep_origin      3217..3542
                     /note="f1 origin"
     polyA_signal    2445..2676
                     /note="BGA polyA"
     promoter        209..863
                     /note="CMV promoter"
     polyA_signal    4427..4799
                     /note="SV40 polyA"
     rep_origin      3411..3496
                     /note="SV40 ori"
     misc_feature    2411..2419
                     /note="3xSTOP"
     misc_recomb     883..921
                     /note="upstream homologous region"
     misc_recomb     2411..2443
                     /note="downstream homologous region"
     misc_recomb     913..921
                     /note="Kozak_Start"
     Promoter        928..982
                     /label=LPP5 promoter
     gene            983..2404
                     /gene="sacB"
     source          1..6873
                     /dnas_title="pcDNA3_SL_II"
```

DNA

```
ORIGIN
        1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC TGCTCTGATG
       61 CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
      121 CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
      181 TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
      241 GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
      301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
      361 CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
      421 ATTGACGTCA ATGGGTGGAC TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
      481 ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
      541 ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
      601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
      661 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
      721 AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
      781 GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
      841 CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGGATCTGTA CGACGATGAC
      901 GATAAGGATC CGGCCACCAT GGATATCAAA CTGCAGTTGA CAACATAAAA ACTTTGTGTT
      961 ATACTTGTAA CGTAAGGAGG TAATGAACAT CAAAAAGTTT GCAAAACAAG CAACAGTATT
     1021 AACCTTTACT ACCGCACTGC TGGCAGGAGG CGCAACTCAA GCGTTTGCGA AAGAAACGAA
     1081 CCAAAAGCCA TATAAGGAAA CATACGGCAT TTCCCATATT ACACGCCATG ATATGCTGCA
     1141 AATCCCTGAA CAGCAAAAAA ATGAAAAATA TCAAGTTCCT GAATTCGATT CGTCCACAAT
     1201 TAAAAATATC TCTTCTGCAA AAGGCCTGGA CGTTTGGGAC AGCTGGCCAT TACAAAACGC
     1261 TGACGGCACT GTCGCAAACT ATCACGGCTA CCACATCGTC TTTGCATTAG CCGGAGATCC
     1321 TAAAAATGCG GATGACACAT CGATTTACAT GTTCTATCAA AAAGTCGGCG AAACTTCTAT
     1381 TGACAGCTGG AAAAACGCTG GCCGCGTCTT TAAAGACAGC GACAAATTCG ATGCAAATGA
     1441 TTCTATCCTA AAAGACCAAA CACAAGAATG GTCAGGTTCA GCCACATTTA CATCTGACGG
     1501 AAAAATCCGT TTATTCTACA CTGATTTCTC CGGTAAACAT TACGGCAAAC AAACACTGAC
     1561 AACTGCACAA GTTAACGTAT CAGCATCAGA CAGCTCTTTG AACATCAACG GTGTAGAGGA
     1621 TTATAAATCA ATCTTTGACG GTGACGGAAA AACGTATCAA AATGTACAGC AGTTCATCGA
     1681 TGAAGGCAAC TACAGCTCAG GCGACAACCA TACGCTGAGA GATCCTCACT ACGTAGAAGA
     1741 TAAAGGCCAC AAATACTTAG TATTTGAAGC AAACACTGGA ACTGAAGATG GCTACCAAGG
     1801 CGAAGAATCT TTATTTAACA AAGCATACTA TGGCAAAAGC ACATCATTCT TCCGTCAAGA
     1861 AAGTCAAAAA CTTCTGCAAA GCGATAAAAA ACGCACGGCT GAGTTAGCAA ACGGCGCTCT
     1921 CGGTATGATT GAGCTAAACG ATGATTACAC ACTGAAAAAA GTGATGAAAC CGCTGATTGC
     1981 ATCTAACACA GTAACAGATG AAATTGAACG CGCGAACGTC TTTAAAATGA ACGGCAAATG
     2041 GTACCTGTTC ACTGACTCCC GCGGATCAAA AATGACGATT GACGGCATTA CGTCTAACGA
     2101 TATTTACATG CTTGGTTATG TTTCTAATTC TTTAACTGGC CCATACAAGC CGCTGAACAA
     2161 AACTGGCCTT GTGTTAAAAA TGGATCTTGA TCCTAACGAT GTAACCTTTA CTTACTCACA
     2221 CTTCGCTGTA CCTCAAGCGA AAGGAAACAA TGTCGTGATT ACAAGCTATA TGACAAACAG
     2281 AGGATTCTAC GCAGACAAAC AATCAACGTT TGCGCCAAGC TTCCTGCTGA ACATCAAAGG
     2341 CAAGAAAACA TCTGTTGTCA AAGACAGCAT CCTTGAACAA GGACAATTAA CAGTTAACAA
     2401 ATAAGATATC TGATAATAGG AATTCCTATA GTGTCACCTA AATGCTAGAG CTCGCTGATC
     2461 AGCCTCGACT GTGCCTTCTA GTTGCCAGCC ATCTGTTGTT TGCCCCTCCC CCGTGCCTTC
     2521 CTTGACCCTG GAAGGTGCCA CTCCCACTGT CCTTTCCTAA TAAAATGAGG AAATTGCATC
     2581 GCATTGTCTG AGTAGGTGTC ATTCTATTCT GGGGGGTGGG GTGGGGCAGG ACAGCAAGGG
     2641 GGAGGATTGG GAAGACAATA GCAGGCATGC TGGGGATGCG GTGGGCTCTA TGGCTTCTGA
     2701 GGCGGAAAGA ACCAGCTGGG GCTCTAGGGG GTATCCCCAC GCGCCCTGTA GCGGCGCATT
     2761 AAGCGCGGCG GGTGTGGTGG TTACGCGCAG CGTGACCGCT ACACTTGCCA GCGCCCTAGC
     2821 GCCCGCTCCT TTCGCTTTCT TCCCTTCCTT TCTCGCCACG TTCGCCGGCT TTCCCCGTCA
     2881 AGCTCTAAAT CGGGGCATCC CTTTAGGGTT CCGATTTAGT GCTTTACGGC ACCTCGACCC
     2941 CAAAAAACTT GATTAGGGTG ATGGTTCACG TAGTGGGCCA TCGCCCTGAT AGACGGTTTT
     3001 TCGCCCTTTG ACGTTGGAGT CCACGTTCTT TAATAGTGGA CTCTTGTTCC AAACTGGAAC
     3061 AACACTCAAC CCTATCTCGG TCTATTCTTT TGATTTATAA GGGATTTTGG GGATTTCGGC
     3121 CTATTGGTTA AAAAATGAGC TGATTTAACA AAAATTTAAC GCGAATTAAT TCTGTGGAAT
     3181 GTGTGTCAGT TAGGGTGTGG AAAGTCCCCA GGCTCCCCAG GCAGGCAGAA GTATGCAAAG
     3241 CATGCATCTC AATTAGTCAG CAACCAGGTG TGGAAAGTCC CCAGGCTCCC CAGCAGGCAG
     3301 AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACCATA GTCCCGCCCC TAACTCCGCC
     3361 CATCCCGCCC CTAACTCCGC CCAGTTCCGC CCATTCTCCG CCCCATGGCT GACTAATTTT
     3421 TTTTATTTAT GCAGAGGCCG AGGCCGCCTC TGCCTCTGAG CTATTCCAGA AGTAGTGAGG
     3481 AGGCTTTTTT GGAGGCCTAG GCTTTTGCAA AAAGCTCCCG GGAGCTTGTA TATCCATTTT
     3541 CGGATCTGAT CAAGAGACAG GATGAGGATC GTTTCGCATG ATTGAACAAG ATGGATTGCA
     3601 CGCAGGTTCT CCGGCCGCTT GGGTGGAGAG GCTATTCGGC TATGACTGGG CACAACAGAC
     3661 AATCGGCTGC TCTGATGCCG CCGTGTTCCG GCTGTCAGCG CAGGGGCGCC CGGTTCTTTT
     3721 TGTCAAGACC GACCTGTCCG GTGCCCTGAA TGAACTGCAG GACGAGGCAG CGCGGCTATC
     3781 GTGGCTGGCC ACGACGGGCG TTCCTTGCGC AGCTGTGCTC GACGTTGTCA CTGAAGCGGG
     3841 AAGGGACTGG CTGCTATTGG GCGAAGTGCC GGGGCAGGAT CTCCTGTCAT CTCACCTTGC
     3901 TCCTGCCGAG AAAGTATCCA TCATGGCTGA TGCAATGCGG CGGCTGCATA CGCTTGATCC
     3961 GGCTACCTGC CCATTCGACC ACCAAGCGAA ACATCGCATC GAGCGAGCAC GTACTCGGAT
     4021 GGAAGCCGGT CTTGTCGATC AGGATGATCT GGACGAAGAG CATCAGGGGC TCGCGCCAGC
     4081 CGAACTGTTC GCCAGGCTCA AGGCGCGCAT GCCCGACGGC GAGGATCTCG TCGTGACCCA
     4141 TGGCGATGCC TGCTTGCCGA ATATCATGGT GGAAAATGGC CGCTTTTCTG GATTCATCGA
     4201 CTGTGGCCGG CTGGGTGTGG CGGACCGCTA TCAGGACATA GCGTTGGCTA CCCGTGATAT
     4261 TGCTGAAGAG CTTGGCGGCG AATGGGCTGA CCGCTTCCTC GTGCTTTACG GTATCGCCGC
     4321 TCCCGATTCG CAGCGCATCG CCTTCTATCG CCTTCTTGAC GAGTTCTTCT GAGCGGGACT
     4381 CTGGGGTTCG AAATGACCGA CCAAGCGACG CCCAACCTGC CATCACGAGA TTTCGATTCC
```

```
4441 ACCGCCGCCT TCTATGAAAG GTTGGGCTTC GGAATCGTTT TCCGGGACGC CGGCTGGATG
4501 ATCCTCCAGC GCGGGGATCT CATGCTGGAG TTCTTCGCCC ACCCCAACTT GTTTATTGCA
4561 GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT
4621 TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGTATA
4681 CCGTCGACCT CTAGCTAGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT
4741 TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG
4801 GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG
4861 TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG GAGAGGCGGT
4921 TTGCGTATTG GGCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC GGTCGTTCGG
4981 CTGCGGCGAG CGGTATCAGC TCACTCAAAG GCGGTAATAC GGTTATCCAC AGAATCAGGG
5041 GATAACGCAG GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG
5101 GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCCTG ACGAGCATCA CAAAAATCGA
5161 CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC GTTTCCCCCT
5221 GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC
5281 TTTCTCCCTT CGGGAAGCGT GGCGCTTTCT CAATGCTCAC GCTGTAGGTA TCTCAGTTCG
5341 GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTTCA GCCCGACCGC
5401 TGCGCCTTAT CCGGTAACTA TCGTCTTGAG TCCAACCCGG TAAGACACGA CTTATCGCCA
5461 CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG TGCTACAGAG
5521 TTCTTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGGA CAGTATTTGG TATCTGCGCT
5581 CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC
5641 ACCGCTGGTA GCGGTGGTTT TTTTGTTTGC AAGCAGCAGA TTACGCGCAG AAAAAAAGGA
5701 TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA
5761 CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTTAAAT
5821 TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT AAACTTGGTC TGACAGTTAC
5881 CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCGTTC ATCCATAGTT
5941 GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT
6001 GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG
6061 CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT
6121 ATTAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT
6181 GTTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC
6241 TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTTGTGCAA AAAAGCGGTT
6301 AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG
6361 GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTTCTGTG
6421 ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT
6481 TGCCCGGCGT CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC
6541 ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT
6601 TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTTACTTT CACCAGCGTT
6661 TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAAA AGGGAATAAG GGCGACACGG
6721 AAATGTTGAA TACTCATACT CTTCCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT
6781 TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTTCCG
6841 CGCACATTTC CCCGAAAAGT GCCACCTGAC GTC
//
```

6.1.6 pSinRep5 SL

GenBank Header

```
LOCUS       pSinRep5_SL            11426 bp    DNA      r      06-APR-2016
FEATURES             Location/Qualifiers
     misc_feature    11408..11426
                     /note="SP6"
     gene            60..7598
                     /note="p270 nonstructural polyprotein"
     promoter        7580..7603
                     /note="PSG subgenomic promoter"
     polyA_site      9472..9508
                     /note="poly A"
     gene            9702..10560
                     /note="beta lactamase"
     promoter        9632..9660
                     /note="Amp promoter"
     misc_feature    7629..7649
                     /note="upstream homologous region"
     misc_feature    7641..7649
                     /note="Kozak_Start"
     Promoter        7658..7712
                     /label=LPP5 promoter
     gene            7713..9134
                     /note="sacB"
                     /lgene=sacB
     misc_feature    9143..9151
                     /note="3xSTOP"
     misc_feature    9143..9163
                     /note="downstream homologous region"
     source          1..11426
                     /dnas_title="pSinRep5_SL"
     rep_origin      10717..11336
                     /note="pBR322"
```

DNA

```
ORIGIN
        1 ATTGACGGCG TAGTACACAC TATTGAATCA AACAGCCGAC CAATTGCACT ACCATCACAA
       61 TGGAGAAGCC AGTAGTAAAC GTAGACGTAG ACCCCCAGAG TCCGTTTGTC GTGCAACTGC
      121 AAAAAAGCTT CCCGCAATTT GAGGTAGTAG CACAGCAGGT CACTCCAAAT GACCATGCTA
      181 ATGCCAGAGC ATTTTCGCAT CTGGCCAGTA AACTAATCGA GCTGGAGGTT CCTACCACAG
      241 CGACGATCTT GGACATAGGC AGCGCACCGG CTCGTAGAAT GTTTTCCGAG CACCAGTATC
      301 ATTGTGTCTG CCCCATGCGT AGTCCAGAAG ACCCGGACCG CATGATGAAA TACGCCAGTA
      361 AACTGGCGGA AAAAGCGTGC AAGATTACAA ACAAGAACTT GCATGAGAAG ATTAAGGATC
      421 TCCGGACCGT ACTTGATACG CCGGATGCTG AAACACCATC GCTCTGCTTT CACAACGATG
      481 TTACCTGCAA CATGCGTGCC GAATATTCCG TCATGCAGGA CGTGTATATC AACGCTCCCG
      541 GAACTATCTA TCATCAGGCT ATGAAAGGCG TGCGGACCCT GTACTGGATT GGCTTCGACA
      601 CCACCCAGTT CATGTTCTCG GCTATGGCAG GTTCGTACCC TGCGTACAAC ACCAACTGGG
      661 CCGACGAGAA AGTCCTTGAA GCGCGTAACA TCGGACTTTG CAGCACAAAG CTGAGTGAAG
      721 GTAGGACAGG AAAATTGTCG ATAATGAGGA AGAAGGAGTT GAAGCCCGGG TCGCGGGTTT
      781 ATTTCTCCGT AGGATCGACA CTTTATCCAG AACACAGAGC CAGCTTGCAG AGCTGGCATC
      841 TTCCATCGGT GTTCCACTTG AATGGAAAGC AGTCGTACAC TTGCCGCTGT GATACAGTGG
      901 TGAGTTGCGA AGGCTACGTA GTGAAGAAAA TCACCATCAG TCCCGGGATC ACGGGAGAAA
      961 CCGTGGGATA CGCGGTTACA CACAATAGCG AGGGCTTCTT GCTATGCAAA GTTACTGACA
     1021 CAGTAAAAGG AGAACGGGTA TCGTTCCCTG TGTGCACGTA CATCCCGGCC ACCATATGCG
     1081 ATCAGATGAC TGGTATAATG GCCACGGATA TATCACCTGA CGATGCACAA AAACTTCTGG
     1141 TTGGGCTCAA CCAGCGAATT GTCATTAACG GTAGGACTAA CAGGAACACC AACACCATGC
     1201 AAAATTACCT TCTGCCGATC ATAGCACAAG GGTTCAGCAA ATGGGCTAAG GAGCGCAAGG
     1261 ATGATCTTGA TAACGAGAAA ATGCTGGGTA CTAGAGAACG CAAGCTTACG TATGGCTGCT
     1321 TGTGGGCGTT TCGCACTAAG AAAGTACATT CGTTTTATCG CCCACCTGGA ACGCAGACCT
     1381 GCGTAAAAGT CCCAGCCTCT TTTAGCGCTT TTCCCATGTC GTCCGTATGG ACGACCTCTT
     1441 TGCCCATGTC GCTGAGGCAG AAATTGAAAC TGGCATTGCA ACCAAAGAAG GAGGAAAAAC
     1501 TGCTGCAGGT CTCGGAGGAA TTAGTCATGG AGGCCAAGGC TGCTTTTGAG GATGCTCAGG
     1561 AGGAAGCCAG AGCGGAGAAG CTCCGAGAAG CACTTCCACC ATTAGTGGCA GACAAAGGCA
     1621 TCGAGGCAGC CGCAGAAGTT GTCTGCGAAG TGGAGGGGCT CCAGGCGGAC ATCGGAGCAG
     1681 CATTAGTTGA AACCCCGCGC GGTCACGTAA GGATAATACC TCAAGCAAAT GACCGTATGA
     1741 TCGGACAGTA TATCGTTGTC TCGCCAAACT CTGTGCTGAA GAATGCCAAA CTCGCACCAG
     1801 CGCACCCGCT AGCAGATCAG GTTAAGATCA TAACACACTC CGGAAGATCA GGAAGGTACG
     1861 CGGTCGAACC ATACGACGCT AAAGTACTGA TGCCAGCAGG AGGTGCCGTA CCATGGCCAG
     1921 AATTCCTAGC ACTGAGTGAG AGCGCCACGT TAGTGTACAA CGAAAGAGAG TTTGTGAACC
     1981 GCAAACTATA CCACATTGCC ATGCATGGCC CCGCCAAGAA TACAGAAGAG GAGCAGTACA
     2041 AGGTTACAAA GGCAGAGCTT GCAGAAACAG AGTACGTGTT TGACGTGGAC AAGAAGCGTT
     2101 GCGTTAAGAA GGAAGAAGCC TCAGGTCTGG TCCTCTCGGG AGAACTGACC AACCCTCCCT
     2161 ATCATGAGCT AGCTCTGGAG GGACTGAAGA CCCGACCTGC GGTCCCGTAC AAGGTCGAAA
     2221 CAATAGGAGT GATAGGCACA CCGGGGTCGG GCAAGTCAGC TATTATCAAG TCAACTGTCA
     2281 CGGCACGAGA TCTTGTTACC AGCGGAAAGA AAGAAAATTG TCGCGAAATT GAGGCCGACG
     2341 TGCTAAGACT GAGGGGTATG CAGATTACGT CGAAGACAGT AGATTCGGTT ATGCTCAACG
     2401 GATGCCACAA AGCCGTAGAA GTGCTGTACG TTGACGAAGC GTTCGCGTGC CACGCAGGAG
     2461 CACTACTTGC CTTGATTGCT ATCGTCAGGC CCCGCAAGAA GGTAGTACTA TGCGGAGACC
     2521 CCATGCAATG CGGATTCTTC AACATGATGC AACTAAAGGT ACATTTCAAT CACCCTGAAA
     2581 AAGACATATG CACCAAGACA TTCTACAAGT ATATCTCCCG GCGTTGCACA CAGCCAGTTA
     2641 CAGCTATTGT ATCGACACTG CATTACGATG GAAAGATGAA AACCACGAAC CCGTGCAAGA
     2701 AGAACATTGA AATCGATATT ACAGGGGCCA CAAAGCCGAA GCCAGGGGAT ATCATCCTGA
     2761 CATGTTTCCG CGGGTGGGTT AAGCAATTGC AAATCGACTA TCCCGGACAT GAAGTAATGA
     2821 CAGCCGCGGC CTCACAAGGG CTAACCAGAA AAGGAGTGTA TGCCGTCCGG CAAAAAGTCA
     2881 ATGAAAACCC ACTGTACGCG ATCACATCAG AGCATGTGAA CGTGTTGCTC ACCCGCACTG
     2941 AGGACAGGCT AGTGTGGAAA ACCTTGCAGG GCGACCCATG GATTAAGCAG CCCACTAACA
     3001 TACCTAAAGG AAACTTTCAG GCTACTATAG AGGACTGGGA AGCTGAACAC AAGGGAATAA
     3061 TTGCTGCAAT AAACAGCCCC ACTCCCCGTG CCAATCCGTT CAGCTGCAAG ACCAACGTTT
     3121 GCTGGGCGAA AGCATTGGAA CCGATACTAG CCACGGCCGG TATCGTACTT ACCGGTTGCC
     3181 AGTGGAGCGA ACTGTTCCCA CAGTTTGCGG ATGACAAACC ACATTCGGCC ATTTACGCCT
     3241 TAGACGTAAT TTGCATTAAG TTTTTCGGCA TGGACTTGAC AAGCGGACTG TTTTCTAAAC
     3301 AGAGCATCCC ACTAACGTAC CATCCCGCCG ATTCAGCGAG GCCGGTAGCT CATTGGGACA
     3361 ACAGCCCAGG AACCCGCAAG TATGGGTACG ATCACGCCAT TGCCGCCGAA CTCTCCCGTA
     3421 GATTTCCGGT GTTCCAGCTA GCTGGGAAGG GCACACAACT TGATTTGCAG ACGGGGAGAA
     3481 CCAGAGTTAT CTCTGCACAG CATAACCTGG TCCCGGTGAA CCGCAATCTT CCTCACGCCT
     3541 TAGTCCCCGA GTACAAGGAG AAGCAACCCG GCCCGGTCAA AAAATTCTTG AACCAGTTCA
     3601 AACACCACTC AGTACTTGTG GTATCAGAGG AAAAAATTGA AGCTCCCCGT AAGAGAATCG
     3661 AATGGATCGC CCCGATTGGC ATAGCCGGTG CAGATAAGAA CTACAACCTG GCTTTCGGGT
     3721 TTCCGCCGCA GGCACGGTAC GACCTGGTGT TCATCAACAT TGGAACTAAA TACAGAAACC
     3781 ACCACTTTCA GCAGTGCGAA GACCATGCGG CGACCTTAAA AACCCTTTCG CGTTCGGCCC
     3841 TGAATTGCCT TAACCCAGGA GGCACCCTCG TGGTGAAGTC CTATGGCTAC GCCGACCGCA
     3901 ACAGTGAGGA CGTAGTCACC GCTCTTGCCA GAAAGTTTGT CAGGGTGTCT GCAGCGAGAC
     3961 CAGATTGTGT CTCAAGCAAT ACAGAAATGT ACCTGATTTT CCGACAACTA GACAACAGCC
     4021 GTACACGGCA ATTCACCCCG CACCATCTGA ATTGCGTGAT TTCGTCCGTG TATGAGGGTA
     4081 CAAGAGATGG AGTTGGAGCC GCGCCGTCAT ACCGCACCAA AAGGGAGAAT ATTGCTGACT
     4141 GTCAAGAGGA AGCAGTTGTC AACGCAGCCA ATCCGCTGGG TAGACCAGGC GAAGGAGTCT
     4201 GCCGTGCCAT CTATAAACGT TGGCCGACCA GTTTTACCGA TTCAGCCACG GAGACAGGCA
     4261 CCGCAAGAAT GACTGTGTGC CTAGGAAAGA AAGTGATCCA CGCGGTCGGC CCTGATTTCC
     4321 GGAAGCACCC AGAAGCAGAA GCCTTGAAAT TGCTACAAAA CGCCTACCAT GCAGTGGCAG
     4381 ACTTAGTAAA TGAACATAAC ATCAAGTCTG TCGCCATTCC ACTGCTATCT ACAGGCATTT
```

```
4441 ACGCAGCCGG AAAAGACCGC CTTGAAGTAT CACTTAACTG CTTGACAACC GCGCTAGACA
4501 GAACTGACGC GGACGTAACC ATCTATTGCC TGGATAAGAA GTGGAAGGAA AGAATCGACG
4561 CGGCACTCCA ACTTAAGGAG TCTGTAACAG AGCTGAAGGA TGAAGATATG GAGATCGACG
4621 ATGAGTTAGT ATGGATTCAT CCAGACAGTT GCTTGAAGGG AAGAAAGGGA TTCAGTACTA
4681 CAAAAGGAAA ATTGTATTCG TACTTCGAAG GCACCAAATT CCATCAAGCA GCAAAAGACA
4741 TGGCGGAGAT AAAGGTCCTG TTCCCTAATG ACCAGGAAAG TAATGAACAA CTGTGTGCCT
4801 ACATATTGGG TGAGACCATG GAAGCAATCC GCGAAAAGTG CCCGGTCGAC CATAACCCGT
4861 CGTCTAGCCC GCCCAAAACG TTGCCGTGCC TTTGCATGTA TGCCATGACG CCAGAAAGGG
4921 TCCACAGACT TAGAAGCAAT AACGTCAAAG AAGTTACAGT ATGCTCCTCC ACCCCCCTTC
4981 CTAAGCACAA AATTAAGAAT GTTCAGAAGG TTCAGTGCAC GAAAGTAGTC CTGTTTAATC
5041 CGCACACTCC CGCATTCGTT CCCGCCCGTA AGTACATAGA AGTGCCAGAA CAGCCTACCG
5101 CTCCTCCTGC ACAGGCCGAG GAGGCCCCCG AAGTTGTAGC GACACCGTCA CCATCTACAG
5161 CTGATAACAC CTCGCTTGAT GTCACAGACA TCTCACTGGA TATGGATGAC AGTAGCGAAG
5221 GCTCACTTTT TTCGAGCTTT AGCGGATCGG ACAACTCTAT TACTAGTATG GACAGTTGGT
5281 CGTCAGGACC TAGTTCACTA GAGATAGTAG ACCGAAGGCA GGTGGTGGTG GCTGACGTTC
5341 ATGCCGTCCA AGAGCCTGCC CCTATTCCAC CGCCAAGGCT AAAGAAGATG GCCCGCCTGG
5401 CAGCGGCAAG AAAAGAGCCC ACTCCACCGG CAAGCAATAG CTCTGAGTCC CTCCACCTCT
5461 CTTTTGGTGG GGTATCCATG TCCCTCGGAT CAATTTTCGA CGGAGAGACG GCCCGCCAGG
5521 CAGCGGTACA ACCCCTGGCA ACAGGCCCCA CGGATGTGCC TATGTCTTTC GGATCGTTTT
5581 CCGACGGAGA GATTGATGAG CTGAGCCGCA GAGTAACTGA GTCCGAACCC GTCCTGTTTG
5641 GATCATTTGA ACCGGGCGAA GTGAACTCAA TTATATCGTC CCGATCAGCC GTATCTTTTC
5701 CACTACGCAA GCAGAGACGT AGACGCAGGA GCAGGAGGAC TGAATACTGA CTAACCGGGG
5761 TAGGTGGGTA CATATTTTCG ACGGACACAG GCCCTGGGCA CTTGCAAAAG AAGTCCGTTC
5821 TGCAGAACCA GCTTACAGAA CCGACCTTGG AGCGCAATGT CCTGGAAAGA ATTCATGCCC
5881 CGGTGCTCGA CACGTCGAAA GAGGAACAAC TCAAACTCAG GTACCAGATG ATGCCCACCG
5941 AAGCCAACAA AAGTAGGTAC CAGTCTCGTA AAGTAGAAAA TCAGAAAGCC ATAACCACTG
6001 AGCGACTACT GTCAGGACTA CGACTGTATA ACTCTGCCAC AGATCAGCCA GAATGCTATA
6061 AGATCACCTA TCCGAAACCA TTGTACTCCA GTAGCGTACC GGCGAACTAC TCCGATCCAC
6121 AGTTCGCTGT AGCTGTCTGT AACAACTATC TGCATGAGAA CTATCCGACA GTAGCATCTT
6181 ATCAGATTAC TGACGAGTAC GATGCTTACT TGGATATGGT AGACGGGACA GTCGCCTGCC
6241 TGGATACTGC AACCTTCTGC CCCGCTAAGC TTAGAAGTTA CCCGAAAAAA CATGAGTATA
6301 GAGCCCCGAA TATCCGCAGT GCGGTTCCAT CAGCGATGCA GAACACGCTA CAAAATGTGC
6361 TCATTGCCGC AACTAAAAGA AATTGCAACG TCACGCAGAT GCGTGAACTG CCAACACTGG
6421 ACTCAGCGAC ATTCAATGTC GAATGCTTTC GAAAATATGC ATGTAATGAC GAGTATTGGG
6481 AGGAGTTCGC TCGGAAGCCA ATTAGGATTA CCACTGAGTT TGTCACCGCA TATGTAGCTA
6541 GACTGAAAGG CCCTAAGGCC GCCGCACTAT TTGCAAAGAC GTATAATTTG GTCCCATTGC
6601 AAGAAGTGCC TATGGATAGA TTCGTCATGG ACATGAAAAG AGACGTGAAA GTTACACCAG
6661 GCACGAAACA CACAGAAGAA AGACCGAAAG TACAAGTGAT ACAAGCCGCA GAACCCCTGG
6721 CGACTGCTTA CTTATGCGGG ATTCACCGGG AATTAGTGCG TAGGCTTACG GCCGTCTTGC
6781 TTCCAAACAT TCACACGCTT TTTGACATGT CGGCGGAGGA TTTTGATGCA ATCATAGCAG
6841 AACACTTCAA GCAAGGCGAC CCGGTACTGG AGACGGATAT CGCATCATTC GACAAAAGCC
6901 AAGACGACGC TATGGCGTTA ACCGGTCTGA TGATCTTGGA GGACCTGGGT GTGGATCAAC
6961 CACTACTCGA CTTGATCGAG TGCGCCTTTG GAGAAATATC ATCCACCCAT CTACCTACGG
7021 GTACTCGTTT TAAATTCGGG GCGATGATGA AATCCGGAAT GTTCCTCACA CTTTTTGTCA
7081 ACACAGTTTT GAATGTCGTT ATCGCCAGCA GAGTACTAGA AGAGCGGCTT AAAACGTCCA
7141 GATGTGCAGC GTTCATTGGC GACGACAACA TCATACATGG AGTAGTATCT GACAAAGAAA
7201 TGGCTGAGAG GTGCGCCACC TGGCTCAACA TGGAGGTTAA GATCATCGAC GCAGTCATCG
7261 GTGAGAGACC ACCTTACTTC TGCGGCGGAT TTATCTTGCA AGATTCGGTT ACTTCCACAG
7321 CGTGCCGCGT GGCGGATCCC CTGAAAAGGC TGTTTAAGTT GGGTAAACCG CTCCCAGCCG
7381 ACGACGAGCA AGACGAAGAC AGAAGACGCG CTCTGCTAGA TGAAACAAAG GCGTGGTTTA
7441 GAGTAGGTAT AACAGGCACT TTAGCAGTGG CCGTGACGAC CCGGTATGAG GTAGACAATA
7501 TTACACCTGT CCTACTGGCA TTGAGAACTT TTGCCCAGAG CAAAAGAGCA TTCCAAGCCA
7561 TCAGAGGGGA AATAAAGCAT CTCTACGGTG GTCCTAAATA GTCAGCATAG TACATTTCAT
7621 CTGACTAAGA TAAGGATCCG GCCACCATGA TTTAAATAAA CTGCAGTTGA CAACATAAAA
7681 ACTTTGTGTT ATACTTGTAA CGTAAGGAGG TAATGAACAT CAAAAAGTTT GCAAAACAAG
7741 CAACAGTATT AACCTTTACT ACCGCACTGC TGGCAGGAGG CGCAACTCAA GCGTTTGCGA
7801 AAGAAACGAA CCAAAAGCCA TATAAGGAAA CATACGGCAT TTCCCATATT ACACGCCATG
7861 ATATGCTGCA AATCCCTGAA CAGCAAAAAA ATGAAAAATA TCAAGTTCCT GAATTCGATT
7921 CGTCCACAAT TAAAAATATC TCTTCTGCAA AAGGCCTGGA CGTTTGGGAC AGCTGGCCAT
7981 TACAAAACGC TGACGGCACT GTCGCAAACT ATCACGGCTA CCACATCGTC TTTGCATTAG
8041 CCGGAGATCC TAAAAATGCG GATGACACAT CGATTTACAT GTTCTATCAA AAAGTCGGCG
8101 AAACTTCTAT TGACAGCTGG AAAAACGCTG GCCGCGTCTT TAAAGACAGC GACAAATTCG
8161 ATGCAAATGA TTCTATCCTA AAAGACCAAA CACAAGAATG GTCAGGTTCA GCCACATTTA
8221 CATCTGACGG AAAAATCCGT TTATTCTACA CTGATTTCTC CGGTAAACAT TACGGCAAAC
8281 AAACACTGAC AACTGCACAA GTTAACGTAT CAGCATCAGA CAGCTCTTTG AACATCAACG
8341 GTGTAGAGGA TTATAAATCA ATCTTTGACG GTGACGGAAA AACGTATCAA AATGTACAGC
8401 AGTTCATCGA TGAAGGCAAC TACAGCTCAG GCGACAACCA TACGCTGAGA GATCCTCACT
8461 ACGTAGAAGA TAAAGGCCAC AAATACTTAG TATTTGAAGC AAACACTGGA ACTGAAGATG
8521 GCTACCAAGG CGAAGAATCT TTATTTAACA AAGCATACTA TGGCAAAAGC ACATCATTCT
8581 TCCGTCAAGA AAGTCAAAAA CTTCTGCAAA GCGATAAAAA ACGCACGGCT GAGTTAGCAA
8641 ACGGCGCTCT CGGTATGATT GAGCTAAACG ATGATTACAC ACTGAAAAAA GTGATGAAAC
8701 CGCTGATTGC ATCTAACACA GTAACAGATG AAATTGAACG CGCGAACGTC TTTAAAATGA
8761 ACGGCAAATG GTACCTGTTC ACTGACTCCC GCGGATCAAA AATGACGATT GACGGCATTA
8821 CGTCTAACGA TATTTACATG CTTGGTTATG TTTCTAATTC TTTAACTGGC CCATACAAGC
8881 CGCTGAACAA AACTGGCCTT GTGTTAAAAA TGGATCTTGA TCCTAACGAT GTAACCTTTA
8941 CTTACTCACA CTTCGCTGTA CCTCAAGCGA AAGGAAACAA TGTCGTGATT ACAAGCTATA
9001 TGACAAACAG AGGATTCTAC GCAGACAAAC AATCAACGTT TGCGCCAAGC TTCCTGCTGA
```

```
     9061 ACATCAAAGG CAAGAAAACA TCTGTTGTCA AAGACAGCAT CCTTGAACAA GGACAATTAA
     9121 CAGTTAACAA ATAAATTTAA ATTGATAATA GGAATTCCTA TAGCAATGAT CCGACCAGCA
     9181 AAACTCGATG TACTTCCGAG GAACTGATGT GCATAATGCA TCAGGCTGGT ACATTAGATC
     9241 CCCGCTTACC GCGGGCAATA TAGCAACACT AAAAACTCGA TGTACTTCCG AGGAAGCGCA
     9301 GTGCATAATG CTGCGCAGTG TTGCCACATA ACCACTATAT TAACCATTTA TCTAGCGGAC
     9361 GCCAAAAACT CAATGTATTT CTGAGGAAGC GTGGTGCATA ATGCCACGCA GCGTCTGCAT
     9421 AACTTTTATT ATTTCTTTTA TTAATCAACA AAATTTTGTT TTTAACATTT CAAAAAAAAA
     9481 AAAAAAAAAA AAAAAAAAAA AAAAAAAAGG GAATTCCTCG ATTAATTAAG CGGCCGCTCG
     9541 AGGGGAATTA ATTCTTGAAG ACGAAAGGGC CAGGTGGCAC TTTTCGGGGA AATGTGCGCG
     9601 GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
     9661 AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC
     9721 GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA
     9781 CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC
     9841 TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA
     9901 TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGTTGAC GCCGGGCAAG
     9961 AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC TCACCAGTCA
    10021 CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA
    10081 TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG AAGGAGCTAA
    10141 CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC
    10201 TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA
    10261 CGTTGCGCAA ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCGGCAA CAATTAATAG
    10321 ACTGGATGGA GGCGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT
    10381 GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC ATTGCAGCAC
    10441 TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA
    10501 CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT AAGCATTGGT
    10561 AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT CATTTTTAAT
    10621 TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC CCTTAACGTG
    10681 AGTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC
    10741 CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG
    10801 TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTTCCGAA GGTAACTGGC TTCAGCAGAG
    10861 CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT
    10921 CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG
    10981 GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC
    11041 GGTCGGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG ACCTACACCG
    11101 AACTGAGATA CCTACAGCGT GAGCATTGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG
    11161 CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG
    11221 GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC
    11281 GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGAGCT
    11341 CGTATGGACA TATTGTCGTT AGAACGCGGC TACAATTAAT ACATAACCTT ATGTATCATA
    11401 CACATACGAT TTAGGGGACA CTATAG
//
```

6.2 Sensor Reference Sequences

6.2.1 Sensor B2

GenBank Header

```
LOCUS       B2_pRSET_SL_III         2373 bp    DNA             07-APR-2016
FEATURES             Location/Qualifiers
     CDS             124..804
                     /dnas_title="YPetd11"
                     /product="YPet"
     misc_feature    49..81
                     /label=T7_gene10_leader
     misc_feature    85..108
                     /label=Xpress_EK
     misc_feature    85..123
                     /note="upstream homologous region"
     misc_feature    13..42
                     /note="10 x His"
     gene            805..1623
                     /note="2 x CNB-B"
     gene            1624..2340
                     /note="mCerulean3"
     misc_feature    2341..2373
                     /note="downstream homologous region"
     source          1..2373
                     /dnas_title="B2_pRSET_SL_III"
```

DNA

```
ORIGIN
        1 ATGCGGGGTT CTCATCATCA TCATCATCAC CATCACCATC ACCGCGGTAT GGCTAGCATG
       61 ACTGGTGGAC AGCAAATGGG TCGGGATCTG TACGACGATG ACGATAAGGA TCCGGCCACC
      121 ATGGTGAGCA AAGGCGAAGA GCTGTTCACC GGCGTGGTGC CCATCCTGGT GGAGCTGGAC
      181 GGCGACGTGA ACGGCCACAA GTTCAGCGTG AGCGGCGAGG GCGAGGGCGA CGCCACCTAC
      241 GGCAAGCTGA CCCTGAAGCT GCTGTGCACC ACCGGCAAGC TGCCCGTGCC CTGGCCCACC
      301 CTGGTGACCA CCCTGGGCTA CGGCGTGCAG TGCTTCGCCC GGTACCCCGA CCACATGAAG
      361 CAGCACGACT TCTTCAAGAG CGCCATGCCC GAGGGCTACG TGCAGGAGCG GACCATCTTC
      421 TTCAAGGACG ACGGCAACTA CAAGACCCGG GCCGAGGTGA AGTTCGAGGG CGACACCCTG
      481 GTGAACCGGA TCGAGCTGAA GGGCATCGAC TTCAAGGAGG ACGGCAACAT CCTGGGCCAC
      541 AAGCTGGAGT ACAACTACAA CAGCCACAAC GTGTACATCA CCGCCGACAA GCAGAAGAAC
      601 GGCATCAAGG CCAACTTCAA GATCCGGCAC AACATCGAGG ACGGCGGCGT GCAGCTGGCC
      661 GACCACTACC AGCAGAACAC CCCCATCGGC GACGGCCCCG TGCTGCTGCC CGACAACCAC
      721 TACCTGAGCT ACCAGAGCGC CCTGTTCAAG GACCCCAACG AGAAGCGGGA CCACATGGTG
      781 CTGCTGGAGT TCCTGACCGC CGCCCAGGCA TTTCGCAAAT TTACCAAAAG CGAACGTAGC
      841 AAAGATCTGA TCAAAGAAGC AATCCTGGAT AACGATTTTA TGAAAAACCT GCCGGAAGAA
      901 ATTCTGAGCA AACTGGCAGA TGTTCTGGAA GAAACCCATT ATGAAAACGG CGAATATATC
      961 ATTCGTCAGG GTGCACGTGG TGATACCTTT TTTATCATTA GCAAAGGCAA AGTGAACGTG
     1021 ACCCGTGAAG ATAGCCCGAA TGAAGATCCG GTTTTTCTGC GTACCCTGGG TAAAGGTGAT
     1081 TGGTTTGGTG AAAAAGCACT GCAGGGTGAA GATGTTCGTA CCGCCAATGT TATTGCAGCA
     1141 GAAGCAGTTA CCTGTCTGGT TATTGATCGT GATAGCTTTA AACATCTGAT TGGTGGTCTG
     1201 GATGATGTGA GCAATAAAGC ATATATGGAA TTTCTGAAAA GCGTTCCGAC CTTTCAGAGC
     1261 CTGCCGGAAG AAATTCTGAG CAAACTGGCA GATGTTCTGG AAGAAACCCA TTATGAAAAC
     1321 GGCGAATATA TCATTCGTCA GGGTGCACGT GGTGATACCT TTTTTATCAT TAGCAAAGGC
     1381 AAAGTGAACG TGACCCGTGA AGATAGCCCG AATGAAGATC CGGTTTTTCT GCGTACCCTG
     1441 GGTAAAGGTG ATTGGTTTGG TGAAAAAGCA CTGCAGGGTG AAGATGTTCG TACCGCCAAT
     1501 GTTATTGCAG CAGAAGCAGT TACCTGTCTG GTTATTGATC GTGATAGCTT TAAACATCTG
     1561 ATTGGTGGTC TGGATGATGT GAGCAATAAA GCATATGAAG ATGCAGAAGC CAAAGCCAAA
     1621 TATATGGTGA GCAAGGGCGA GGAGCTGTTC ACCGGGGTGG TGCCCATCCT GGTCGAGCTG
     1681 GACGGCGACG TAAACGGCCA CAAGTTCAGC GTGTCCGGCG AGGGCGAGGG CGATGCCACC
     1741 TACGGCAAGC TGACCCTGAA GTTCATCTGC ACCACCGGCA AGCTGCCCGT GCCCTGGCCC
     1801 ACCCTCGTGA CCACCCTGAG CTGGGGCGTG CAGTGCTTCG CCCGCTACCC CGACCACATG
     1861 AAGCAGCACG ACTTCTTCAA GTCCGCCATG CCCGAAGGCT ACGTCCAGGA GCGCACCATC
     1921 TTCTTCAAGG ACGACGGCAA CTACAAGACC CGCGCCGAGG TGAAGTTCGA GGGCGACACC
     1981 CTGGTGAACC GCATCGAGCT GAAGGGCATC GACTTCAAGG AGGACGGCAA CATCCTGGGG
     2041 CACAAGCTGG AGTACAACGC CATCCACGGC AACGTCTATA TCACCGCCGA CAAGCAGAAG
     2101 AACGGCATCA AGGCCAACTT CGGCCTCAAC TGCAACATCG AGGACGGCAG CGTGCAGCTC
     2161 GCCGACCACT ACCAGCAGAA CACCCCCATC GGCGACGGCC CCGTGCTGCT GCCCGACAAC
     2221 CACTACCTGA GCACCCAGTC CAAGCTGAGC AAAGACCCCA ACGAGAAGCG CGATCACATG
     2281 GTCCTGCTGG AGTTCGTGAC CGCCGCCGGG ATCACTCTCG GCATGGACGA GCTGTACAAG
     2341 TGATAATAGG AATTCCTATA GTGTCACCTA AAT
//
```

Protein Sequence Sensor B2

Leader Peptide
Donor FP
Sensor Domain
Acceptor FP

```
        1 MRGSHHHHHH HHHHRGMASM TGGQQMGRDL YDDDDKDPAT MVSKGEELFT GVVPILVELD
       61 GDVNGHKFSV SGEGEGDATY GKLTLKLLCT TGKLPVPWPT LVTTLGYGVQ CFARYPDHMK
      121 QHDFFKSAMP EGYVQERTIF FKDDGNYKTR AEVKFEGDTL VNRIELKGID FKEDGNILGH
      181 KLEYNYNSHN VYITADKQKN GIKANFKIRH NIEDGGVQLA DHYQQNTPIG DGPVLLPDNH
      241 YLSYQSALFK DPNEKRDHMV LLEFLTAAQA FRKFTKSERS KDLIKEAILD NDFMKNLPEE
      301 ILSKLADVLE ETHYENGEYI IRQGARGDTF FIISKGKVNV TREDSPNEDP VFLRTLGKGD
      361 WFGEKALQGE DVRTANVIAA EAVTCLVIDR DSFKHLIGGL DDVSNKAYME FLKSVPTFQS
      421 LPEEILSKLA DVLEETHYEN GEYIIRQGAR GDTFFIISKG KVNVTREDSP NEDPVFLRTL
      481 GKGDWFGEKA LQGEDVRTAN VIAAEAVTCL VIDRDSFKHL IGGLDDVSNK AYEDAEAKAK
      541 YMVSKGEELF TGVVPILVEL DGDVNGHKFS VSGEGEGDAT YGKLTLKFIC TTGKLPVPWP
      601 TLVTTLSWGV QCFARYPDHM KQHDFFKSAM PEGYVQERTI FFKDDGNYKT RAEVKFEGDT
      661 LVNRIELKGI DFKEDGNILG HKLEYNAIHG NVYITADKQK NGIKANFGLN CNIEDGSVQL
      721 ADHYQQNTPI GDGPVLLPDN HYLSTQSKLS KDPNEKRDHM VLLEFVTAAG ITLGMDELYK
```

6.2.2 Sensor B3

GenBank Header

```
LOCUS       B3_pRSET_SL_III         2373 bp    DNA             07-APR-2016
FEATURES             Location/Qualifiers
     gene            124..804
                     /note="mCerulean3 d11"
     misc_feature    49..81
                     /label=T7_gene10_leader
     misc_feature    85..108
                     /label=Xpress_EK
     misc_feature    85..123
                     /note="upstream homologous region"
     misc_feature    13..42
                     /note="10 x His"
     gene            805..1623
                     /note="2 x CNB-B"
     CDS             1624..2340
                     /dnas_title="YPet"
                     /product="YPet"
     misc_feature    2341..2373
                     /note="downstream homologous region"
     source          1..2373
                     /dnas_title="B3_pRSET_SL_III"
```

DNA

```
ORIGIN
        1 ATGCGGGGTT CTCATCATCA TCATCATCAC CATCACCATC ACCGCGGTAT GGCTAGCATG
       61 ACTGGTGGAC AGCAAATGGG TCGGGATCTG TACGACGATG ACGATAAGGA TCCGGCCACC
      121 ATGGTGAGCA AGGGCGAGGA GCTGTTCACC GGGGTGGTGC CCATCCTGGT CGAGCTGGAC
      181 GGCGACGTAA ACGGCCACAA GTTCAGCGTG TCCGGCGAGG GCGAGGGCGA TGCCACCTAC
      241 GGCAAGCTGA CCCTGAAGTT CATCTGCACC ACCGGCAAGC TGCCCGTGCC CTGGCCCACC
      301 CTCGTGACCA CCCTGAGCTG GGGCGTGCAG TGCTTCGCCC GCTACCCCGA CCACATGAAG
      361 CAGCACGACT TCTTCAAGTC CGCCATGCCC GAAGGCTACG TCCAGGAGCG CACCATCTTC
      421 TTCAAGGACG ACGGCAACTA CAAGACCCGC GCCGAGGTGA AGTTCGAGGG CGACACCCTG
      481 GTGAACCGCA TCGAGCTGAA GGGCATCGAC TTCAAGGAGG ACGGCAACAT CCTGGGGCAC
      541 AAGCTGGAGT ACAACGCCAT CCACGGCAAC GTCTATATCA CCGCCGACAA GCAGAAGAAC
      601 GGCATCAAGG CCAACTTCGG CCTCAACTGC AACATCGAGG ACGGCAGCGT GCAGCTCGCC
      661 GACCACTACC AGCAGAACAC CCCCATCGGC GACGGCCCCG TGCTGCTGCC CGACAACCAC
      721 TACCTGAGCA CCCAGTCCAA GCTGAGCAAA GACCCCAACG AGAAGCGCGA TCACATGGTC
      781 CTGCTGGAGT TCGTGACCGC CGCCCAGGCA TTTCGCAAAT TTACCAAAAG CGAACGTAGC
      841 AAAGATCTGA TCAAAGAAGC AATCCTGGAT AACGATTTTA TGAAAAACCT GCCGGAAGAA
      901 ATTCTGAGCA AACTGGCAGA TGTTCTGGAA GAAACCCATT ATGAAAACGG CGAATATATC
      961 ATTCGTCAGG GTGCACGTGG TGATACCTTT TTTATCATTA GCAAAGGCAA AGTGAACGTG
     1021 ACCCGTGAAG ATAGCCCGAA TGAAGATCCG GTTTTTCTGC GTACCCTGGG TAAAGGTGAT
     1081 TGGTTTGGTG AAAAAGCACT GCAGGGTGAA GATGTTCGTA CCGCCAATGT TATTGCAGCA
     1141 GAAGCAGTTA CCTGTCTGGT TATTGATCGT GATAGCTTTA AACATCTGAT TGGTGGTCTG
     1201 GATGATGTGA GCAATAAAGC ATATATGGAA TTTCTGAAAA GCGTTCCGAC CTTTCAGAGC
     1261 CTGCCGGAAG AAATTCTGAG CAAACTGGCA GATGTTCTGG AAGAAACCCA TTATGAAAAC
     1321 GGCGAATATA TCATTCGTCA GGGTGCACGT GGTGATACCT TTTTTATCAT TAGCAAAGGC
     1381 AAAGTGAACG TGACCCGTGA AGATAGCCCG AATGAAGATC CGGTTTTTCT GCGTACCCTG
     1441 GGTAAAGGTG ATTGGTTTGG TGAAAAAGCA CTGCAGGGTG AAGATGTTCG TACCGCCAAT
     1501 GTTATTGCAG CAGAAGCAGT TACCTGTCTG GTTATTGATC GTGATAGCTT TAAACATCTG
     1561 ATTGGTGGTC TGGATGATGT GAGCAATAAA GCATATGAAG ATGCAGAAGC CAAAGCCAAA
     1621 TATATGGTGA GCAAAGGCGA AGAGCTGTTC ACCGGCGTGG TGCCCATCCT GGTGGAGCTG
     1681 GACGGCGACG TGAACGGCCA CAAGTTCAGC GTGAGCGGCG AGGGCGAGGG CGACGCCACC
     1741 TACGGCAAGC TGACCCTGAA GCTGCTGTGC ACCACCGGCA AGCTGCCCGT GCCCTGGCCC
     1801 ACCCTGGTGA CCACCCTGGG CTACGGCGTG CAGTGCTTCG CCCGGTACCC CGACCACATG
     1861 AAGCAGCACG ACTTCTTCAA GAGCGCCATG CCCGAGGGCT ACGTGCAGGA GCGGACCATC
     1921 TTCTTCAAGG ACGACGGCAA CTACAAGACC CGGGCCGAGG TGAAGTTCGA GGGCGACACC
     1981 CTGGTGAACC GGATCGAGCT GAAGGGCATC GACTTCAAGG AGGACGGCAA CATCCTGGGC
     2041 CACAAGCTGG AGTACAACTA CAACAGCCAC AACGTGTACA TCACCGCCGA CAAGCAGAAG
     2101 AACGGCATCA AGGCCAACTT CAAGATCCGG CACAACATCG AGGACGGCGG CGTGCAGCTG
     2161 GCCGACCACT ACCAGCAGAA CACCCCCATC GGCGACGGCC CCGTGCTGCT GCCCGACAAC
     2221 CACTACCTGA GCTACCAGAG CGCCCTGTTC AAGGACCCCA ACGAGAAGCG GGACCACATG
     2281 GTGCTGCTGG AGTTCCTGAC CGCCGCCGGC ATCACCGAGG GCATGAACGA GCTCTATAAG
     2341 TGATAATAGG AATTCCTATA GTGTCACCTA AAT
//
```

Protein Sequence Sensor B3

Leader Peptide
Donor FP
Sensor Domain
Acceptor FP

```
        1 MRGSHHHHHH HHHHRGMASM TGGQQMGRDL YDDDDKDPAT MVSKGEELFT GVVPILVELD
       61 GDVNGHKFSV SGEGEGDATY GKLTLKFICT TGKLPVPWPT LVTTLSWGVQ CFARYPDHMK
      121 QHDFFKSAMP EGYVQERTIF FKDDGNYKTR AEVKFEGDTL VNRIELKGID FKEDGNILGH
      181 KLEYNAIHGN VYITADKQKN GIKANFGLNC NIEDGSVQLA DHYQQNTPIG DGPVLLPDNH
      241 YLSTQSKLSK DPNEKRDHMV LLEFVTAAQA FRKFTKSERS KDLIKEAILD NDFMKNLPEE
      301 ILSKLADVLE ETHYENGEYI IRQGARGDTF FIISKGKVNV TREDSPNEDP VFLRTLGKGD
      361 WFGEKALQGE DVRTANVIAA EAVTCLVIDR DSFKHLIGGL DDVSNKAYME FLKSVPTFQS
      421 LPEEILSKLA DVLEETHYEN GEYIIRQGAR GDTFFIISKG KVNVTREDSP NEDPVFLRTL
      481 GKGDWFGEKA LQGEDVRTAN VIAAEAVTCL VIDRDSFKHL IGGLDDVSNK AYEDAEAKAK
      541 YMVSKGEELF TGVVPILVEL DGDVNGHKFS VSGEGEGDAT YGKLTLKLLC TTGKLPVPWP
      601 TLVTTLGYGV QCFARYPDHM KQHDFFKSAM PEGYVQERTI FFKDDGNYKT RAEVKFEGDT
      661 LVNRIELKGI DFKEDGNILG HKLEYNYNSH NVYITADKQK NGIKANFKIR HNIEDGGVQL
      721 ADHYQQNTPI GDGPVLLPDN HYLSYQSALF KDPNEKRDHM VLLEFLTAAG ITEGMNELYK
```

6.2.3 Sensor C1

GenBank Header

```
LOCUS       C1_pRSET_SL_III         2002 bp    DNA             06-APR-2016
FEATURES             Location/Qualifiers
     gene            124..804
                     /note="mCerulean3 d11"
     misc_feature    49..81
                     /label=T7_gene10_leader
     misc_feature    85..108
                     /label=Xpress_EK
     misc_feature    85..123
                     /note="upstream homologous region"
     misc_feature    13..42
                     /note="10 x His
     misc_feature    805..1251
                     /note="1 x CNB-B"
     CDS             1252..1968
                     /dnas_title="YPet"
                     /product="YPet"
     misc_feature    1969..2001
                     /note="downstream homologous region"
     source          1..2002
                     /dnas_title="C1_pRSET_SL_III"
```

DNA

```
ORIGIN
        1 ATGCGGGGTT CTCATCATCA TCATCATCAC CATCACCATC ACCGCGGTAT GGCTAGCATG
       61 ACTGGTGGAC AGCAAATGGG TCGGGATCTG TACGACGATG ACGATAAGGA TCCGGCCACC
      121 ATGGTGAGCA AGGGCGAGGA GCTGTTCACC GGGGTGGTGC CCATCCTGGT CGAGCTGGAC
      181 GGCGACGTAA ACGGCCACAA GTTCAGCGTG TCCGGCGAGG GCGAGGGCGA TGCCACCTAC
      241 GGCAAGCTGA CCCTGAAGTT CATCTGCACC ACCGGCAAGC TGCCCGTGCC CTGGCCCACC
      301 CTCGTGACCA CCCTGAGCTG GGGCGTGCAG TGCTTCGCCC GCTACCCCGA CCACATGAAG
      361 CAGCACGACT TCTTCAAGTC CGCCATGCCC GAAGGCTACG TCCAGGAGCG CACCATCTTC
      421 TTCAAGGACG ACGGCAACTA CAAGACCCGC GCCGAGGTGA AGTTCGAGGG CGACACCCTG
      481 GTGAACCGCA TCGAGCTGAA GGGCATCGAC TTCAAGGAGG ACGGCAACAT CCTGGGGCAC
      541 AAGCTGGAGT ACAACGCCAT CCACGGCAAC GTCTATATCA CCGCCGACAA GCAGAAGAAC
      601 GGCATCAAGG CCAACTTCGG CCTCAACTGC AACATCGAGG ACGGCAGCGT GCAGCTCGCC
      661 GACCACTACC AGCAGAACAC CCCCATCGGC GACGGCCCCG TGCTGCTGCC CGACAACCAC
      721 TACCTGAGCA CCCAGTCCAA GCTGAGCAAA GACCCCAACG AGAAGCGCGA TCACATGGTC
      781 CTGCTGGAGT TCGTGACCGC CGCCCAGGCA TTTCGCAAAT TTACCAAAAG CGAACGTAGC
      841 AAAGATCTGA TCAAAGAAGC AATCCTGGAT AACGATTTTA TGAAAAACCT GCCGGAAGAA
      901 ATTCTGAGCA AACTGGCAGA TGTTCTGGAA GAAACCCATT ATGAAAACGG CGAATATATC
      961 ATTCGTCAGG GTGCACGTGG TGATACCTTT TTTATCATTA GCAAAGGCAA AGTGAACGTG
     1021 ACCCGTGAAG ATAGCCCGAA TGAAGATCCG GTTTTTCTGC GTACCCTGGG TAAAGGTGAT
     1081 TGGTTTGGTG AAAAAGCACT GCAGGGTGAA GATGTTCGTA CCGCCAATGT TATTGCAGCA
     1141 GAAGCAGTTA CCTGTCTGGT TATTGATCGT GATAGCTTTA AACATCTGAT TGGTGGTCTG
     1201 GATGATGTGA GCAATAAAGC ATATGAAGAT GCAGAAGCCA AAGCCAAATA TATGGTGAGC
     1261 AAAGGCGAAG AGCTGTTCAC CGGCGTGGTG CCCATCCTGG TGGAGCTGGA CGGCGACGTG
     1321 AACGGCCACA AGTTCAGCGT GAGCGGCGAG GGCGAGGGCG ACGCCACCTA CGGCAAGCTG
     1381 ACCCTGAAGC TGCTGTGCAC CACCGGCAAG CTGCCCGTGC CCTGGCCCAC CCTGGTGACC
     1441 ACCCTGGGCT ACGGCGTGCA GTGCTTCGCC CGGTACCCCG ACCACATGAA GCAGCACGAC
     1501 TTCTTCAAGA GCGCCATGCC CGAGGGCTAC GTGCAGGAGC GGACCATCTT CTTCAAGGAC
     1561 GACGGCAACT ACAAGACCCG GGCCGAGGTG AAGTTCGAGG GCGACACCCT GGTGAACCGG
     1621 ATCGAGCTGA AGGGCATCGA CTTCAAGGAG GACGGCAACA TCCTGGGCCA CAAGCTGGAG
     1681 TACAACTACA ACAGCCACAA CGTGTACATC ACCGCCGACA AGCAGAAGAA CGGCATCAAG
     1741 GCCAACTTCA AGATCCGGCA CAACATCGAG GACGGCGGCG TGCAGCTGGC CGACCACTAC
     1801 CAGCAGAACA CCCCCATCGG CGACGGCCCC GTGCTGCTGC CCGACAACCA CTACCTGAGC
     1861 TACCAGAGCG CCCTGTTCAA GGACCCCAAC GAGAAGCGGG ACCACATGGT GCTGCTGGAG
     1921 TTCCTGACCG CCGCCGGCAT CACCGAGGGC ATGAACGAGC TCTATAAGTG ATAATAGGAA
     1981 TTCCTATAGT GTCACCTAAA TG
//
```

Protein Sequence Sensor C1

Leader Peptide
Donor FP
Sensor Domain
Acceptor FP

```
        1 MRGSHHHHHH HHHHRGMASM TGGQQMGRDL YDDDDKDPAT MVSKGEELFT GVVPILVELD
       61 GDVNGHKFSV SGEGEGDATY GKLTLKFICT TGKLPVPWPT LVTTLSWGVQ CFARYPDHMK
      121 QHDFFKSAMP EGYVQERTIF FKDDGNYKTR AEVKFEGDTL VNRIELKGID FKEDGNILGH
      181 KLEYNAIHGN VYITADKQKN GIKANFGLNC NIEDGSVQLA DHYQQNTPIG DGPVLLPDNH
      241 YLSTQSKLSK DPNEKRDHMV LLEFVTAAQA FRKFTKSERS KDLIKEAILD NDFMKNLPEE
      301 ILSKLADVLE ETHYENGEYI IRQGARGDTF FIISKGKVNV TREDSPNEDP VFLRTLGKGD
      361 WFGEKALQGE DVRTANVIAA EAVTCLVIDR DSFKHLIGGL DDVSNKAYED AEAKAKYMVS
      421 KGEELFTGVV PILVELDGDV NGHKFSVSGE GEGDATYGKL TLKLLCTTGK LPVPWPTLVT
      481 TLGYGVQCFA RYPDHMKQHD FFKSAMPEGY VQERTIFFKD DGNYKTRAEV KFEGDTLVNR
      541 IELKGIDFKE DGNILGHKLE YNYNSHNVYI TADKQKNGIK ANFKIRHNIE DGGVQLADHY
      601 QQNTPIGDGP VLLPDNHYLS YQSALFKDPN EKRDHMVLLE FLTAAGITEG MNELYK
```

6.2.4 Sensor D1

GenBank Header

```
LOCUS       D1_pRSET_SL_III         2028 bp    DNA             07-APR-2016
FEATURES             Location/Qualifiers
     gene            124..804
                     /note="mCerulean3"
     misc_feature    49..81
                     /label=T7_gene10_leader
     misc_feature    85..108
                     /label=Xpress_EK
     misc_feature    85..123
                     /note="upstream homologous region"
     misc_feature    13..42
                     /note="10 x His"
     CDS             1279..1995
                     /dnas_title="YPet"
                     /product="YPet"
     misc_feature    1996..2028
                     /note="downstream homologous region"
     misc_feature    805..1278
                     /note="1 x CNB-BII"
     source          1..2028
                     /dnas_title="D1_pRSET_SL_III"
```

DNA

```
ORIGIN
        1 ATGCGGGGTT CTCATCATCA TCATCATCAC CATCACCATC ACCGCGGTAT GGCTAGCATG
       61 ACTGGTGGAC AGCAAATGGG TCGGGATCTG TACGACGATG ACGATAAGGA TCCGGCCACC
      121 ATGGTGAGCA AGGGCGAGGA GCTGTTCACC GGGGTGGTGC CCATCCTGGT CGAGCTGGAC
      181 GGCGACGTAA ACGGCCACAA GTTCAGCGTG TCCGGCGAGG GCGAGGGCGA TGCCACCTAC
      241 GGCAAGCTGA CCCTGAAGTT CATCTGCACC ACCGGCAAGC TGCCCGTGCC CTGGCCCACC
      301 CTCGTGACCA CCCTGAGCTG GGGCGTGCAG TGCTTCGCCC GCTACCCCGA CCACATGAAG
      361 CAGCACGACT TCTTCAAGTC CGCCATGCCC GAAGGCTACG TCCAGGAGCG CACCATCTTC
      421 TTCAAGGACG ACGGCAACTA CAAGACCCGC GCCGAGGTGA AGTTCGAGGG CGACACCCTG
      481 GTGAACCGCA TCGAGCTGAA GGGCATCGAC TTCAAGGAGG ACGGCAACAT CCTGGGGCAC
      541 AAGCTGGAGT ACAACGCCAT CCACGGCAAC GTCTATATCA CCGCCGACAA GCAGAAGAAC
      601 GGCATCAAGG CCAACTTCGG CCTCAACTGC AACATCGAGG ACGGCAGCGT GCAGCTCGCC
      661 GACCACTACC AGCAGAACAC CCCCATCGGC GACGGCCCCG TGCTGCTGCC CGACAACCAC
      721 TACCTGAGCA CCCAGTCCAA GCTGAGCAAA GACCCCAACG AGAAGCGCGA TCACATGGTC
      781 CTGCTGGAGT TCGTGACCGC CGCCCAGGCA TTTCGCAAAT TTACCAAAAG CGAACGTAGC
      841 AAAGATCTGA TCAAAGAAGC AATCCTGGAT AACGATTTTA TGAAAAACCT GCCTGAAGAT
      901 AAATTAACCA AGATCATTGA CTGCTTGGAA GTGGAATACT ATGACAAAGG AGATTACATC
      961 ATTAGAGAGG GCGAGGAAGG AAGTACCTTT TTCATTTTGG CAAAAGGAAA GGTAAAAGTA
     1021 ACACAGAGCA CAGAAGGCCA TGATCAACCA CAGCTGATAA AAACACTGCA GAAAGGAGAA
     1081 TACTTTGGAG AAAAAGCTCT TATCAGTGAT GATGTCAGGT CAGCTAACAT TATTGCTGAA
     1141 GAAAATGATG TTGCATGCCT GGTTATAGAT CGAGAAACAT TCAACCAAAC TGTCGGTACA
     1201 TTTGAAGAGC TGCAAAAATA CCTTGAAGGA TATGTGGCAA ACCTGAACCG TGATGATGAA
     1261 AAAAGACATG CGAAGCGGAT GGTGAGCAAA GGCGAAGAGC TGTTCACCGG CGTGGTGCCC
     1321 ATCCTGGTGG AGCTGGACGG CGACGTGAAC GGCCACAAGT TCAGCGTGAG CGGCGAGGGC
     1381 GAGGGCGACG CCACCTACGG CAAGCTGACC CTGAAGCTGC TGTGCACCAC CGGCAAGCTG
     1441 CCCGTGCCCT GGCCCACCCT GGTGACCACC CTGGGCTACG GCGTGCAGTG CTTCGCCCGG
     1501 TACCCCGACC ACATGAAGCA GCACGACTTC TTCAAGAGCG CCATGCCCGA GGGCTACGTG
     1561 CAGGAGCGGA CCATCTTCTT CAAGGACGAC GGCAACTACA AGACCCGGGC CGAGGTGAAG
     1621 TTCGAGGGCG ACACCCTGGT GAACCGGATC GAGCTGAAGG GCATCGACTT CAAGGAGGAC
     1681 GGCAACATCC TGGGCCACAA GCTGGAGTAC AACTACAACA GCCACAACGT GTACATCACC
     1741 GCCGACAAGC AGAAGAACGG CATCAAGGCC AACTTCAAGA TCCGGCACAA CATCGAGGAC
     1801 GGCGGCGTGC AGCTGGCCGA CCACTACCAG CAGAACACCC CCATCGGCGA CGGCCCCGTG
     1861 CTGCTGCCCG ACAACCACTA CCTGAGCTAC CAGAGCGCCC TGTTCAAGGA CCCCAACGAG
     1921 AAGCGGGACC ACATGGTGCT GCTGGAGTTC CTGACCGCCG CCGGCATCAC CGAGGGCATG
     1981 AACGAGCTCT ATAAGTGATA ATAGGAATTC CTATAGTGTC ACCTAAAT
//
```

Protein Sequence Sensor D1

Leader Peptide
Donor FP
Sensor Domain
Acceptor FP

```
        1 MRGSHHHHHH HHHHRGMASM TGGQQMGRDL YDDDDKDPAT MVSKGEELFT GVVPILVELD
       61 GDVNGHKFSV SGEGEGDATY GKLTLKFICT TGKLPVPWPT LVTTLSWGVQ CFARYPDHMK
      121 QHDFFKSAMP EGYVQERTIF FKDDGNYKTR AEVKFEGDTL VNRIELKGID FKEDGNILGH
      181 KLEYNAIHGN VYITADKQKN GIKANFGLNC NIEDGSVQLA DHYQQNTPIG DGPVLLPDNH
      241 YLSTQSKLSK DPNEKRDHMV LLEFVTAAQA FRKFTKSERS KDLIKEAILD NDFMKNLPED
      301 KLTKIIDCLE VEYYDKGDYI IREGEEGSTF FILAKGKVKV TQSTEGHDQP QLIKTLQKGE
      361 YFGEKALISD DVRSANIIAE ENDVACLVID RETFNQTVGT FEELQKYLEG YVANLNRDDE
      421 KRHAKRMVSK GEELFTGVVP ILVELDGDVN GHKFSVSGEG EGDATYGKLT LKLLCTTGKL
      481 PVPWPTLVTT LGYGVQCFAR YPDHMKQHDF FKSAMPEGYV QERTIFFKDD GNYKTRAEVK
      541 FEGDTLVNRI ELKGIDFKED GNILGHKLEY NYNSHNVYIT ADKQKNGIKA NFKIRHNIED
      601 GGVQLADHYQ QNTPIGDGPV LLPDNHYLSY QSALFKDPNE KRDHMVLLEF LTAAGITEGM
      661 NELYK
```

6.2.5 Sensor Tomato 1 x CNB-B Ultramarine

GenBank Header

```
LOCUS       Tomato_1xCNB-B_U        1965 bp    DNA             07-APR-2016
FEATURES             Location/Qualifiers
     misc_feature    49..81
                     /label=T7_gene10_leader
     misc_feature    85..123
                     /note="upstream homologous region"
     misc_feature    13..42
                     /note="10 x His"
     misc_feature    124..822
                     /note="Tomato"
     misc_feature    823..1272
                     /note=""
     misc_feature    1933..1965
                     /note="downstream homologous region"
     misc_feature    823..1269
                     /note="1 x CNB-B"
     gene            1270..1932
                     /note="Ultramarine"
     source          1..1965
                     /dnas_title="Tomato_1xCNB-B_Ultramarine"
```

DNA

```
ORIGIN
        1 ATGCGGGGTT CTCATCATCA TCATCATCAC CATCACCATC ACCGCGGTAT GGCTAGCATG
       61 ACTGGTGGAC AGCAAATGGG TCGGGATCTG TACGACGATG ACGATAAGGA TCCGGCCACC
      121 ATGGTGAGCA AGGGCGAGGA GGTCATCAAA GAGTTCATGC GCTTCAAGGT GCGCATGGAG
      181 GGCTCCATGA ACGGCCACGA GTTCGAGATC GAGGGCGAGG GCGAGGGCCG CCCCTACGAG
      241 GGCACCCAGA CCGCCAAGCT GAAGGTGACC AAGGGCGGCC CCCTGCCCTT CGCCTGGGAC
      301 ATCCTGTCCC CCCAGTTCAT GTACGGCTCC AAGGCGTACG TGAAGCACCC CGCCGACATC
      361 CCCGATTACA AGAAGCTGTC CTTCCCCGAG GGCTTCAAGT GGGAGCGCGT GATGAACTTC
      421 GAGGACGGCG GTCTGGTGAC CGTGACCCAG GACTCCTCCC TGCAGGACGG CACGCTGATC
      481 TACAAGGTGA AGATGCGCGG CACCAACTTC CCCCCCGACG GCCCCGTAAT GCAGAAGAAG
      541 ACCATGGGCT GGGAGGCCTC CACCGAGCGC CTGTACCCCC GCGACGGCGT GCTGAAGGGC
      601 GAGATCCACC AGGCCCTGAA GCTGAAGGAC GGCGGCCACT ACCTGGTGGA GTTCAAGACC
      661 ATCTACATGG CCAAGAAGCC CGTGCAACTG CCCGGCTACT ACTACGTGGA CACCAAGCTG
      721 GACATCACCT CCCACAACGA GGACTACACC ATCGTGGAAC AGTACGAGCG CTCCGAGGGC
      781 CGCCACCACC TGTTCCTGTA CGGCATGGAC GAGCTGTACA AGCAGGCATT TCGCAAATTT
      841 ACCAAAAAGCG AACGTAGCAA AGATCTGATC AAAGAAGCAA TCCTGGATAA CGATTTTATG
      901 AAAAACCTGC CGGAAGAAAT TCTGAGCAAA CTGGCAGATG TTCTGGAAGA AACCCATTAT
      961 GAAAACGGCG AATATATCAT TCGTCAGGGT GCACGTGGTG ATACCTTTTT TATCATTAGC
     1021 AAAGGCAAAG TGAACGTGAC CCGTGAAGAT AGCCCGAATG AAGATCCGGT TTTTCTGCGT
     1081 ACCCTGGGTA AAGGTGATTG GTTTGGTGAA AAAGCACTGC AGGGTGAAGA TGTTCGTACC
     1141 GCCAATGTTA TTGCAGCAGA AGCAGTTACC TGTCTGGTTA TTGATCGTGA TAGCTTTAAA
     1201 CATCTGATTG GTGGTCTGGA TGATGTGAGC AATAAAGCAT ATGAAGATGC AGAAGCCAAA
     1261 GCCAAATATA TGTCTGTTAT CGCCACCCAG ATGACCTACA AAGTTTACAT GTCTGGCACC
     1321 GTGAACGGCC ACTACTTTGA AGTGGAAGGC GACGGCAAAG GCCGCCCTTA CGAAGGCGAA
     1381 CAGACCGCCA AACTGACCGT GACCAAAGGC GGCCCTCTGC CTTTCGCCTG GGACATCCTG
     1441 TCTCCTCAGT GCCAGTACGG CTCTATCCCT TTCACCAAAT ACCCTGAAGA CATCCCTGAC
     1501 TACGTGAAAC AGTCTTTCCC TGAAGGCTTC ACCTGGGAAC GCATCATGAA CTTTGAAGAC
     1561 GGCGCCGTGT GCACCGTGTC TAACGACTCT TCTATCCAGG GCAACTGCTT CACCTACCAC
     1621 GTGAAATTCC GCGGCACCAA CTTCCCTCCT AACGGCCCTG TGATGCAGAA AAAAACCCAG
     1681 GGCTGGGAAC CTAACTCTGA ACGCCTGTTC GCCCGCGGCG GCATGTTGAT CGGCAACAAC
     1741 CGCATGGCCC TGAAACTGGA AGGCGGCGGC CACTACCTGT GCGAATTTAA AACCACCTAC
     1801 AAAGCCAAAA AACCTGTGAA AATGCCTGGC TACCACTACG TGGACCGCAA ACTGGACGTG
     1861 ACCAACCACA ACAAAGACTA CACCTCTGTG GAACAGTGCG AAATCTCTAT CGCCCGCAAA
     1921 CCTGTGGTGG CCTGATAATA GGAATTCCTA TAGTGTCACC TAAAT
//
```

Protein Sequence Sensor Tomato 1 x CNB-B Ultramarine

Leader Peptide
Donor FP
Sensor Domain
Acceptor CP

```
        1 MRGSHHHHHH HHHHRGMASM TGGQQMGRDL YDDDDKDPAT MVSKGEEVIK EFMRFKVRME
       61 GSMNGHEFEI EGEGEGRPYE GTQTAKLKVT KGGPLPFAWD ILSPQFMYGS KAYVKHPADI
      121 PDYKKLSFPE GFKWERVMNF EDGGLVTVTQ DSSLQDGTLI YKVKMRGTNF PPDGPVMQKK
      181 TMGWEASTER LYPRDGVLKG EIHQALKLKD GGHYLVEFKT IYMAKKPVQL PGYYYVDTKL
      241 DITSHNEDYT IVEQYERSEG RHHLFLYGMD ELYKQAFRKF TKSERSKDLI KEAILDNDFM
      301 KNLPEEILSK LADVLEETHY ENGEYIIRQG ARGDTFFIIS KGKVNVTRED SPNEDPVFLR
      361 TLGKGDWFGE KALQGEDVRT ANVIAAEAVT CLVIDRDSFK HLIGGLDDVS NKAYEDAEAK
      421 AKYMSVIATQ MTYKVYMSGT VNGHYFEVEG DGKGRPYEGE QTAKLTVTKG GPLPFAWDIL
      481 SPQCQYGSIP FTKYPEDIPD YVKQSFPEGF TWERIMNFED GAVCTVSNDS SIQGNCFTYH
      541 VKFRGTNFPP NGPVMQKKTQ GWEPNSERLF ARGGMLIGNN RMALKLEGGG HYLCEFKTTY
      601 KAKKPVKMPG YHYVDRKLDV TNHNKDYTSV EQCEISIARK PVVA
```

7 Acknowledgements

This thesis was not created in a vacuum and many people deserve recognition and thanks for making it possible.

First and foremost, I want to thank Oliver Griesbeck for giving me the opportunity to be part of his lab and for supporting my research throughout the years. Furthermore, I want to thank our collaborators in France, Pierre Vincent and Dahdjim Beltolngar for testing my sensors in the real world and giving helpful input.

I have to highlight the importance of the IMPRS program for my doctoral studies; it opened many doors in my life and enriched my experience greatly. My special thanks goes to the IMPRS coordinators Hans-Joerg Schaeffer, Ingrid Wolf and Maximiliane Reif, who make it all happen and supported me in countless ways. Hans-Joerg deserves recognition for being the possibly best organized person on the face of the planet.

Special thanks go to all current and former members of “se Griesbecks” for creating a positive and supportive environment, which is the envy of many groups. Without you it would not have been possible. Specifically, I want to thank Anselm Geiger, Thomas Thestrup, Julia Litzlbauer and David NG - Anselm, for being the greatest communicator and diplomat and the source various forms of music. Thomas used to say: “Take your time young man”, something that stayed in my head, quite literally. Together with Julia I busted many myths and spent way too much time in the cold room.

I want to thank David Ng for the many years of collaboration and all the lively discussions including, which dangerous animals are best kept as pets and the

world's most absurd sources of food. I hope you stay awesome and can focus on your new job.

After all of you left I tried my best to uphold the spirit you created and pass it on to the next generation in the hope they can enjoy the same support I received through you.

Furthermore, I want to thank Tobias Kruse for being a good friend and collaborator and for being part of the style-over-substance-initiative.

I have to thank Daniel Hornburg, who has been my closest friend for more than a decade now; a bit scary how time flies. I enjoyed all of it and I thank you for your friendship; it carried me through many depths in my life. I value your input on scientific and personal matters, which keeps things in perspective and provides constructive insights few people are able to give. I consider myself lucky for also being able to work on a project together with you.

I want to thank my fiancée Radhika, for being my time vortex and the most special person in my life. Together, we made it through both of our PhDs, sharing all of its highs and lows. I'm looking forward to many games of Dominion and spending the rest of our lives together.

Last but not least I want to thank my family, without them I would not be the person I am today. My mom and dad always encouraged and supported me. Everything I accomplished I owe to them. They were always there in times of trouble and kept me fed through two decades of education. I wish my mom could have seen the end of it.