Master Thesis

In Nanoscience

Multi-Color approach to track *Salmonella* during infection

Raphael Urbani

Supervised by Prof. Dirk Bumann
Department of Infection Biology
Biozentrum
University of Basel

October 15, 2009

Abstract

The aim of this work was to track various *Salmonella strains* during competitive mouse infection, using different fluorescent proteins.

To do so, sequences coding for fluorescent proteins GFP, SYFP, GFPuv, tagCFP and dsRED were fused with a kanamycin resistance cassette using PCR. The fusion products were integrated at the *sifB* locus of *Salmonella enterica* serovar Typhimurium using lambda red mediated homologous recombination followed by phage transduction. The resulting constructs were validated by PCR, restriction digest, sequencing and *in vitro* fluorescence analysis.

In vivo validation in one single mouse with seven different fluorescent strains was performed.

The data revealed successfully constructed strains carrying GFP, GFPuv, CFP and SYFP. Strains carrying dsRED did not show detectable fluorescence.

FACS with appropriate filter settings permitted individual detection of each fluorescent strain in mixtures *in vitro* and *in vivo*.

Contents

1 Abbreviations and glossary 4

2 Introduction 6

3 Materials and methods 7

3.1 Media and cultures 7

3.2 Polymerase chain reaction (PCR) and gel electrophoresis 7

3.3 Competent cells and electroporation 10

3.4 Phage transduction 10

3.5 Sequencing 11

3.6 Induction and analysis (FACS) 12

3.7 Mouse infection and sample preparation 13

3.8 *Salmonella* strains 14

4 Results 15

4.1 Strategy 15

4.2 Molecular biology 17

4.3 *In vitro* validation 20

4.4 *In vivo* validation 22

5 Discussion 27

6 Conclusion 28

7 Outlook 28

8 References 29

9 Acknowledgment 31

10 Appendix 32

10.1 Strain sequences 32

1. Abbreviations and glossary

CI:	Competitive index
DMSO:	Dimethylsulfoxide
Daughter gate:	A gate created on a population that is already constrained by another gate (events need to fulfill both gate criteria)
dsRED:	Red fluorescent protein with excitation and emission maxima at 550nm and 585nm respectively
EBU:	Evans Blue-Uranine
EGTA:	Ethylene glycol tetraacetic acid
FACS:	Fluorescence Activated Cell Sorting
FP635:	Red fluorescent protein with excitation and emission maxima at 589nm and 636nm respectively
FSC:	Forward/small angle scatter
GFP:	Green fluorescent protein
GFPuv:	Green fluorescent protein with excitation maxima at 405nm and emission maximum at 505nm
KAN:	Kanamycin resistance cassette
LA:	LB Agar, for growth on solid media
LB:	Luria Bertani broth
M9/M9*:	Minimal media used for induction of *sifB* promoter
OD_{600}:	Optical density measured at 600nm
PBS:	Phosphate Buffered Saline
PCR:	Polymerase Chain Reaction
PFA:	Paraformaldehyde, dissolved in PBS as formaldehyde
pKD46:	Plasmid carrying the λ-red recombinase system
RBS:	Ribosomal binding site
RFLP:	Restriction Fragment Length Polymorphism
SSC:	Side/right-angle scatter
SPI2:	*Salmonella* pathogenicity island 2

S.Tm:	*Salmonella enterica* serovar Typhimurium
S.Ty:	*Salmonella enteric*a serovar Typhi
SYFP:	Super yellow fluorescent protein with excitation and emission maxima at 515nm and 530nm respectively
Syto60:	DNA staining dye with excitation and emission maxima at 640nm and 657nm respectively
tagCFP:	Cyan fluorescent protein with excitation and emission maxima at 455nm and 480nm respectively
T3SS:	Type three secretion system
ToPro-3:	DNA staining dye with excitation and emission maxima at 640nm and 656nm respectively

2. Introduction

Salmonellae are flagellated Gram-negative bacteria that infect a wide range of animals and cause different types of diseases. They are typically taken up orally with contaminated food or water. They enter the small intestine and can traverse the intestinal mucous layer [1].

In humans, *Salmonella enterica* serovar Typhimurium infection can cause diarrhea, fever and vomiting. In contrast, serovar Typhi disseminates systemically causing life threatening typhoid fever, which includes high fever, diarrhea, gastroenteritis and profuse sweating.

Wildtype mice are resistant against both serovars but serovar Typhimurium can infect immunodeficient BALB/c mice systemically and causes a typhoid fever like disease that mimics important aspects of human disease caused by S.Ty [2].

To investigate the role of a particular *Salmonella* gene in virulence, the gene is inactivated. The resulting mutant is inoculated in mice as a mixture with wildtype *Salmonella*. After several days of infection, *Salmonella* are recovered from infected organs. The output ration of mutant and wild type is compared to the input ratio to assess the virulence phenotype of that mutant.

Compared to single-strain infections, competitive infections have less interindividual variation and thus provide more accurate results with fever experimental animals.

In competitive infections, wildtype and mutant *Salmonella* are usually distinguished based on antibiotic resistance markers. Only few such markers seem to be suitable, limiting the number of strains that can be analyzed in a single experiment. Furthermore, replica-plating of several dilutions and colony counting is time- and labor-intensive.

As an alternative to antibiotic resistance markers, fluorescent proteins could be used for differential labeling. The large variety of available fluorescence colors could enable simultaneous analysis of multiple strains. Flow cytometry could provide rapid and accurate enumeration of individual strains in such mixtures [3, 4].

The goal of this thesis was to explore this new approach by constructing differentially fluorescing *Salmonella* strains and *in vitro*/*in vivo* analysis of mixtures.

3. Materials and methods

3.1 Media and cultures

Liquid cultures were done in Luria Bertani low salt medium (LB) if not otherwise mentioned, plates were done with LA low salt medium.

For induction of the *sifB* promoter regulation expression of the fluorescent proteins, two different media were used: M9 and M9*. M9 media containing 12,8g/L Na_2HPO_4 x $7H_2O$, 3g/L KH_2PO_4, 0.5g/L NaCl, 1g/L NH_4Cl, 0.4% glucose, 100µM $MgSO_4$, 0.004% histidine. After mixing, the pH value was set to 6 by adding HCl, the medium was then sterile filtered. M9* medium is the same as M9 medium, except that it lacks $MgSO_4$ completely.

For selection, antibiotics were added; 100µg/ml for ampicillin, 50µg/ml for kanamycin and 90µg/ml for streptomycin.

"LB broth" used for phage transduction donor preparation was prepared from LB medium supplemented with 0.2% glucose, 5g/l K_2HPO_4, 0.2g/l $MgSO_4x7H_2O$, 2g/l citric acid and 3.5g/l $NaHNH_4PO_4xH_2O$. EBU plates for selection of phage-free clones consist of LA low salt with 0.00125% Evans-Blue, 0.0025% Fluoresceine, 0.5% K_2HPO_4 and 0.25% glucose.

3.2 Polymerase chain reaction (PCR) and gel electrophoresis

Two different PCR protocols were used, one for the Promega GoTaq® Flexi polymerase and another one for the Platinum® Pfx polymerase from Invitrogen.

For amplification with the GoTaq polymerase, 2µl template (single colony dissolved in 7% DMSO or plasmid/DNA solution in water) for a total volume of 25µl were used. Green buffer (final concentration 1x), $MgCl_2$ (2mM), dNTPs (200µM), primers (0.3µM

each) and polymerase (0.2µl per 25µl total volume) were added and filled up with nuclease free water.

For PCR with the Platinum pfx polymerase, $MgSO_4$ (1mM final concentration) instead of $MgCl_2$, 300µM dNTPs, 0.7µM primers, 1x of the provided buffer, 0.25µl polymerase (for 25µl total volume) and 2µl template were added.

The cycles for both programs are shown in table 1.

Table 1: PCR cycles

Cycles	Temperature	Time
1x	94°C	120s
10x	94°C	15s
	65°C (decreasing each cycle by 1°C)	30s
	72°C (GoTaq)/68°C (Pfx)	150s
25x	94°C	15s
	58°C	30s
	72°C (GoTaq)/68°C (Pfx)	120s
1x	72°C (GoTaq)/68°C (Pfx)	300s
	8°C	∞

For fusion PCR, the same basic protocol was used, except that two different DNA sequences instead of one single were used as template.

For both, amplification and fusion PCR, concentrations of Mg salts, template and polymerase as well as program properties were changed for optimization of production.

Plasmid preparations for the template plasmids of the protein sequence and the KAN-cassette were done using the Wizard® *Plus* SV Miniprep DNA Purification System from Promega.

The primers used for the PCRs are listed in Table 2. The forward primers for the fluorescent protein coding sequence and the reverse primer for the kanamycin cassette each contain a 40bp long homologous sequence for insertion into the *sifB* locus.

Table 2: Primers used for amplification and fusion PCR

Amplified sequence (direction)	Primer sequence
Kanamycin cassette (forward)	5'-GTGTAGGCTGGAGCTGCTTC-3'
Kanamycin cassette (reverse)	5'-TCTCACTCTTTAAAAATCCTCTCCCGATAG-TAATTGGCATGTGTAGGCTGGAGCTGCTTC-3'
SYFP, dsRED and tagCFP (forward)	5'-CCAGTAATGAAGTATCATATAATCACTTGT-GGTCTACATTGATCCTCTAGATTTAAGAAG-3'
SYFP (reverse)	5'-GGTCGACGGATCCCCGGAATCTTGGCTGCA-GTTACTTGTA-3'
tagCFP (reverse)	5'-GGTCGACGGATCCCCGGAATCTCATCCGCC-AAAACAGCCA-3'
dsRED (reverse)	5'-GGTCGACGGATCCCCGGAATCAGAAGCTT-CTACAGGAACA-3'
GFP 20% and SYFP 20% (forward)	5'-CCAGTAATGAAGTATCATATAATCACTTGT-GGTCTACATTGATCCTCTAGATTTAAGAACGA-GAT-3'
sifB test primer (forward)	5'-GTTGACGGTTCAGGACA-3'
sifB test primer (reverse)	5'-CACTTTGAGAAAACATC-3'
Test primer for phage DNA(forward)	5'-GGCAGGCTGAACTGGCTA-3'
Test primer for phage DNA (reverse)	5'-GGACATCATCGAAGCCAG-3'
Control primer for bacterial DNA (forward)	5'-GCAGTTGTAACCCGTGGA-3'
Control primer for bacterial DNA (reverse)	5'-CGTAACACGAACTTCGAAG-3'

The purification of the PCR products for further downstream applications was done with the Promega Wizard® SV Gel and PCR Purification System, according to its protocol.
To check the fusion PCR products, a Restriction Fragment Length Polymorphism (RFLP) analysis was performed. All PCR products were digested with Tsp509I in a total volume of 10µl, with 8µl PCR product and 1µl enzyme.

The gels for checking the PCR product were all done using the agar for routine use (Sigma, Buchs), at a concentration of 1%, in 1x TBE and the voltage was set on 80V/cm. The amount of loaded marker and DNA was 2-5µl, depending on the pocket size.
For analysis of restriction products, 2% gels were used.

3.3 Competent cells and electroporation

An LB culture of the virulent S.Tm strain 12023 containing the pKD46 plasmid was inoculated with a colony picked from a fresh plate and shaked at 28°C overnight. On the next day, a new culture, containing 0.2% Arabinose (to induce the lambda red recombinase on the plasmid) and ampicillin, was inoculated 1:50 with the overnight culture. At OD_{600nm} 0.9 (exponential growth phase), the culture was aliquoted into 50ml Falcon tubes (45ml culture each) and incubated on ice for 30 minutes.

After incubation on ice, the cells were centrifuged for 15mins at 4000rpm and at 4°C. The supernatant was poured off and the cells were resuspended and pooled in 4°C cold water, then the cells were centrifuged again for 15mins at 4000rpm and 4°C. This washing cycle was repeated 3 times in total.

After the last centrifugation step, the cells were resuspended in approximately 1ml per tube of original culture of water and equally distributed into 1.5ml Eppendorf tubes (ca. one tube per 45ml of original culture). After centrifugation of the Eppendorf tubes, the supernatant was taken away and 15µl of purified PCR product were mixed into the cells. The cells were then transferred into pre-cooled electroporation cuvettes.

After electroporation, the cuvette was washed twice with 0.5ml LB to take up the cells and the LB was then incubated at 37°C for 1h. Cells were plated afterwards in three different amounts (0.1ml, 0.2ml and 0.5ml) out on kanamycin plates and incubated over night at 37°C.

On the next day, colonies were picked, dissolved in 7% DMSO, checked by PCR and RFLP and for fluorescence in the FACS or stored at -20°C.

3.4 Phage transduction

After testing the electroporation products (S.Tm 12023 with fusion PCR product in *sifB* locus) for fluorescence in the FACS, a standing overnight culture at 37°C was made for donor preparation of phage transduction. On day two, 0.5ml of the overnight culture was mixed with 5ml of “LB broth” and incubated (standing) at 37°C for 1-2 hours. After

that, 10µl of P22 phage were added and the culture was incubated standing overnight at 37°C. On day 3, 1ml of the overnight culture was mixed with 200µl chloroform, vortexed vigorously and centrifuged at 12000rpm for 10 minutes. The supernatant containing phages with bacterial DNA was transferred into glass tubes and stored at 4°C.

A culture of the recipient (S.Tm SL1344) was incubated at 28°C overnight. 100µl of overnight culture were mixed with the same amount of LB with 10mM $CaCl_2$ and 20mM $MgSO_4$ added and with three different amounts of the prepared phage lysate (1, 5 and 20µl). Controls without phage and without cells were done as well.
The mixtures were incubated at room temperature for 15 minutes, then 1ml LB with 10mM EGTA was added and the cells were incubated at 37°C for one hour.
The cells were spun down at maximum speed and the supernatant was poured off. 100µl LB were added to resuspend the pellet and plated out on plates with kanamycin and streptomycin. Plates were incubated at 37°C overnight.

To check for phage free clones, colonies were picked from the plates and streaked out on EBU plates with kanamycin and streptomycin. White colonies, that have no more phage DNA in their cells, were picked and checked by phage test PCR, which contained primers for phage DNA and for chromosomal DNA of the bacteria (table 2). If the PCR did not result in a band around 2kb for the phage DNA but only the band for bacterial DNA around 800bp, which means that there is no phage DNA left in the bacteria cell, an overnight culture of that clone was inoculated. On the next day 1ml of the overnight culture was centrifuged at maximum speed for 2 minutes, resuspended in PBS and stored at -80°C.

3.5 Sequencing

The sequencing was done using the Barcode Economy Service from Microsynth Laboratory, Balgach. Results were analyzed with Chromas version 1.41.

3.6 Induction and analysis (FACS)

After streaking out the respective strain on a plate, an overnight culture was inoculated with 1-5 colonies in LB. On day two, 1ml of the culture was centrifuged and the pellet was resuspended in M9 medium. A new overnight culture in M9 medium was inoculated 1:50 with the washed bacteria. On day three, 1ml of the M9 culture was centrifuged and the pellet was resuspended in M9* medium. Again a new overnight culture in M9* medium was inoculated 1:50 with these resuspended bacteria. On day four, 1-2ml of the M9* overnight culture were centrifuged and resuspended in PBS containing 1% formaldehyde for fixation. To have accurate concentration for FACS measurements, the fixed cells were diluted 1:1000 before analysis.

The FACS analyses were done with a BD FACS Canto II. The laser and filter setups are listed in table 3.

Table 3: Laser and filter setup for FACS measurement

Excitation 405nm		Excitation 488nm		Excitation 633nm	
long pass	band pass	long pass	band pass	long pass	band pass
556	585/42	650	670 (LP)	735	780/60
502	510/50	556	585/42	-	660/20
		525	542/30		
		495	514/27		
		-	488/10		

A negative sample (wild type) was measured and its emission mean value was set to 1000 in every channel, flow rate was adjusted to an appropriate event number per second (not above 10’000). For *in vitro* validation 10’000 events were recorded, for ex vivo measurements, a fixed amount of cell solution was recorded (100µl to 1ml, depending on the concentration). The FACS Diva software was used for recording and the FCS Express V3 software for analysis of the data.

3.7 Mouse infection and sample preparation

After growing S.Tm SL1344 strains with different fluorescent proteins up to an OD_{600} of 1.5, 3ml of each strain was washed four times with PBS (1ml final volume). The OD_{600} was measured again, and the solutions were diluted to get the same concentration for every strain. To evaluate the correct number of injected bacteria, 100µl and 50µl of each strain was plated out on plates with kanamycin and streptomycin. The different solutions were mixed together to get a final cfu of 2000/ml and per strain.

An 8 week old female BALB/c mouse was infected intravenously with 100µl of the mixture, which gives a final infection with 200cfu per strain.

After four days, the mouse was sacrificed and spleen and liver were collected. Both organs were cut into pieces, half of the spleen was collected in 1ml 10% PBS, 0.1% Triton X-100 solution for FACS analysis and plating. The other pieces were collected in 4% formaldehyde (PFA) for microscopy.

The spleen collected for FACS analysis was mashed with the rear part of the plunger of a syringe. Afterwards, the solution was taken, diluted 1:3 and centrifuged at 500 x g for 5 minutes, the supernatant was collected and centrifuged again. 100µl of the supernatant were diluted 1:100, 1:1000 and 1:10'000 and plated out on plates with kanamycin and streptomycin. The rest of the supernatant (ca 2.5ml) was fixed by adding 2.5ml PBS with 2% formaldehyde. After dilution of 1:5 of the solution, ToPro-3 dye from Molecular Probes was added to a final concentration of 0.25µM, after that, the sample was analyzed by FACS.

The spleen and liver pieces for microscopy were kept in 4% formaldehyde (PFA) for 4h after sacrifice of the mouse. Then, the PFA was removed and the pieces were washed with PBS four times, leaving them in PBS for 5 minutes during every washing cycle. After washing, they were put in 10% sucrose for at least 5 minutes until they sedimented, after sedimentation the 10% sucrose was replaced by 20% sucrose. Again, after sedimentation or 15 minutes, the 20% sucrose was replaced by 30% sucrose, and then, after additional 15 minutes by 40% sucrose. After adding the 40% sucrose, the pieces were stored at 4°C over night.

The next day, the pieces were embedded in Sakura Tissue-Tek Oct Compound (Gentaur, Aachen). 10µm sections were cut on a cryostat at -15°C and adhered to glass slides. Sections were allowed to dry for 30min at room temperature and stored frozen at -20°C.

Bacteria (Alexa Fluor 647) and macrophages (Alexa Fluor 700) were stained for microscopic analysis.

3.8 *Salmonella* strains

All *Salmonella enterica* serovar Typhimurium strains used or produced during this work are listed in table 4.

Table 4: All *Salmonella Typhimurium* strains used or produced in this work

strain name	genotype	phenotype
12023	wild type	virulent lab strain, no resistances
SL1344	wild type	virulent lab strain, streptomycin resistant
SDFR3	SL1344, GFP and KAN at *sifB*	strong GFP expression
SKS3	SL1344, FP635 and KAN at *sifB*	no fluorescence by FACS measurable
SKS6	SL1344, GFPuv and KAN at *sifB*	strong GFPuv expression
SRU1	SL1344, GFP with 20% RBS and KAN at *sifB*	weak GFP expression
SRU2	SL1344, GFP with 20% RBS and KAN at *sifB*	weak GFP expression, but brighter than SRU1
SRU3	SL1344, SYFP and KAN at *sifB*	strong YFP expression
SRU4	SL1344, SYFP with 20% RBS and KAN at *sifB*	weak YFP expression
SRU5	SL1344, tagCFP and KAN at *sifB*	strong CFP expression, no ideal FACS filters available
SRU6	SL1344, dsRED and KAN at *sifB*	no fluorescence detectable by FACS

4. Results

Virulence studies of mutant bacteria strains in competitive mouse infection require hours of plating and counting colonies. Due to limited choices for antibiotic resistance markers, only two to three mutants can be tested simultaneously. Replacement of resistance markers with fluorescent proteins could minimize efforts and the amount of mice as more mutants might be tested at the same tame.

To develop such a method, genes coding for fluorescent proteins were inserted at appropriate sites in the *Salmonella* genome.

Prerequisites for this technique are different fluorescent proteins that can be well distinguished with available FACS laser and filter settings and a locus for insertion, which does not affect virulence of the bacterial pathogen, but enable sufficient *in vivo* expression

It has been shown, that two-color flow cytometry is an adequate method to separate different fluorescently labeled cells from each other or from tissue autofluorescence [4, 5]. In this work, six strains producing different fluorescent proteins were constructed and it was shown, that they were separable from tissue and from each other by FACS.

4.1 Strategy

As a first step, appropriate fluorescent proteins with suitable excitation and emission spectra had to be selected. The excitation lasers of our FACS system (BD FACS Canto II with 405nm, 488nm and 630nm lasers) limit the number of adequate fluorescent proteins, as for instance many orange or red fluorescent proteins (like FP635) are not well enough excited at either 488nm or 633nm.

The spectra for the chosen proteins (GFP [6], tagCFP [7], SYFP [8], dsRED [9] and FP635 [10]) are shown in figure 1. In addition, another green fluorescent protein (GFPuv [11]), that has an emission spectrum similar to the GFP shown below, but an excitation

spectrum with a global maximum around 405nm and a local maximum (25%-30%) around 480nm, was chosen.

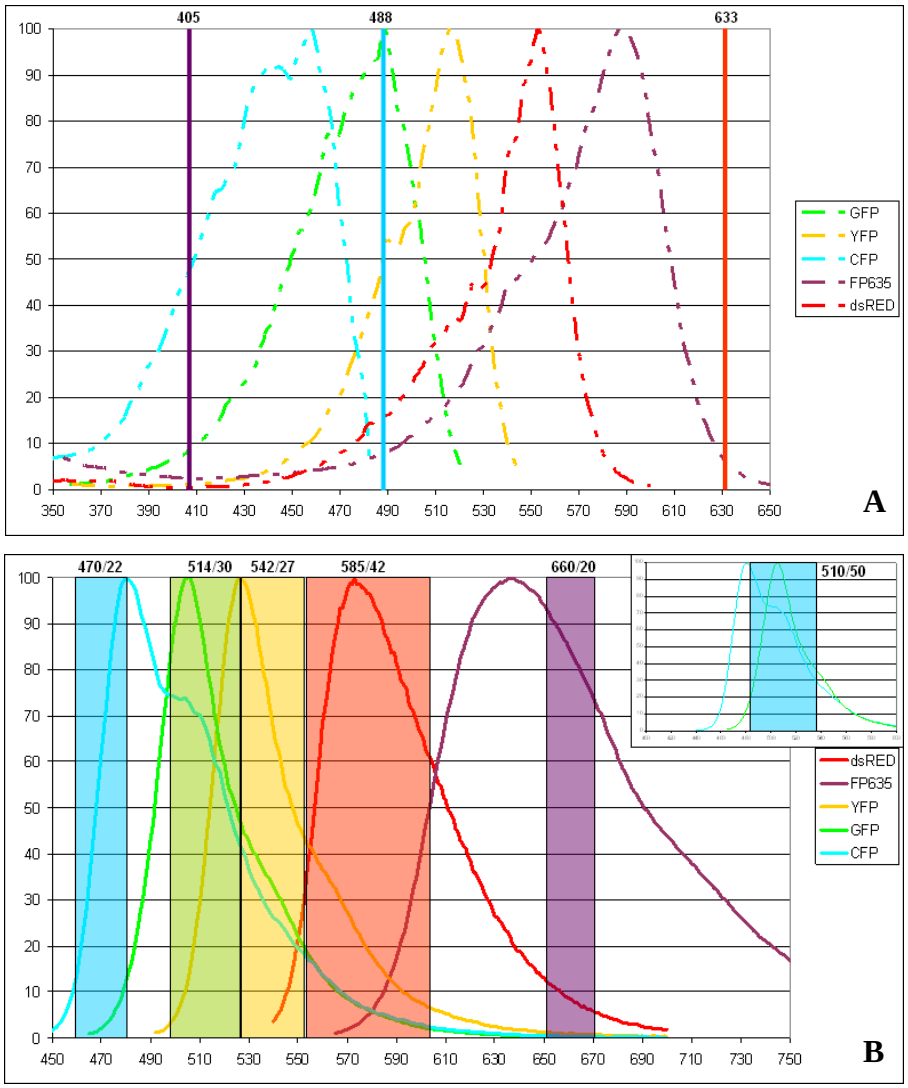

Fig. 1: Excitation (A) and emission (B) spectra for the different fluorescent proteins, the box in the upper right corner of B shows the emission spectra of CFP and GFP/GFPuv that are detected using a 510/50 bandpass filter

Fig. 1A shows the excitation spectra for the proteins. Except for FP635 and dsRED, all proteins should be at least 40% excitable with the 405nm laser or the 488nm laser. Since dsRED encoded on plasmid shows bright fluorescence in the red channel (585/42), it was assumed that suboptimal excitation at 488 or 633nm might be sufficient.

Fig. 1B explains the choice of the filter settings in the FACS.

Strains containing GFP (SDFR3), FP635 (SKS3) and GFPuv (SKS6) were already available in the lab.
In this thesis, strains with SYFP, tagCFP (the sequence for tagCFP was taken from the Evrogen webpage, a vector with this sequence, optimized for expression in *E.coli* was ordered from GeneArt) and dsRED containing strains were constructed. In addition, strains with diminished GFP or SYFP expression were constructed by mutating the ribosomal binding site. The ribosomal binding site (ACGAG instead of AGGAG) was changed by using a modified forward primer for amplification of the PCR product.

It has been shown, that plasmid-encoded fluorescent proteins can result in strains with a broad range of fluorescence intensities interfering with spectral separation from other strains [12]. Strains carrying the fluorescent protein sequence in their chromosome have a much narrower fluorescence distribution. Therefore the fluorescent protein sequence was inserted together with a kanamycin resistance cassette (for transformant selection) in single copy into the chromosome.
The locus chosen for the insertion was the *sifB* locus. The PCR product was inserted downstream of the promoter region, just before the start codon of the *sifB* gene. The *sifB* gene is an effector protein of the *Salmonella* type three secretion system (T3SS) on pathogenicity island 2 (SPI2) [13] and is constitutively expressed during infection [12]. Earlier experiments have shown that an insertion at the *sifB* locus or the deletion of the gene have no significant effect on the virulence of the bacterium [14].

4.2 Molecular biology

The amplification of the two sequences (kanamycin cassette and fluorescent protein sequence) from the plasmids was performed with the GoTaq polymerase and used as template for a PCR with pfx polymerase and a total volume of 100µl or more. After gel purification, the concentration of the eluted DNA, which was between 100ng/µl and 200ng/µl for all products, was measured in analytical gels using ImageJ version 1.38I.

The experiments showed that for amplification from the plasmid, the GoTaq polymerase worked better than the proofreading pfx polymerase. The amplification with the total volume of 100µl was done with the proofreading polymerase to decrease the possibility of mutations.

The forward and reverse primers used for the amplification of the fluorescent protein encoding sequence carried a 40bp overlap homologous to the *sifB* promoter region and a 20bp overlap homologous to the first 20bp of the kanamycin resistance cassette, respectively. The forward primer used for amplification of the kanamycin resistance cassette had no overlap, because previously better results had been observed without the overlap. The reverse primer for the kanamycin resistance cassette had a 40bp overlap homologue to the first 40bp of the *sifB* gene.

For fusion PCR, salt concentrations, template concentrations and ratios, annealing temperatures, annealing and extension times and polymerases were varied to optimize efficiency. There was no optimal protocol found to perform all fusion PCRs, therefore, the settings and concentrations had to be optimized for each fusion PCR.

Control gels were made after each step, including a restriction digest for the fusion PCR product (Fig. 2).

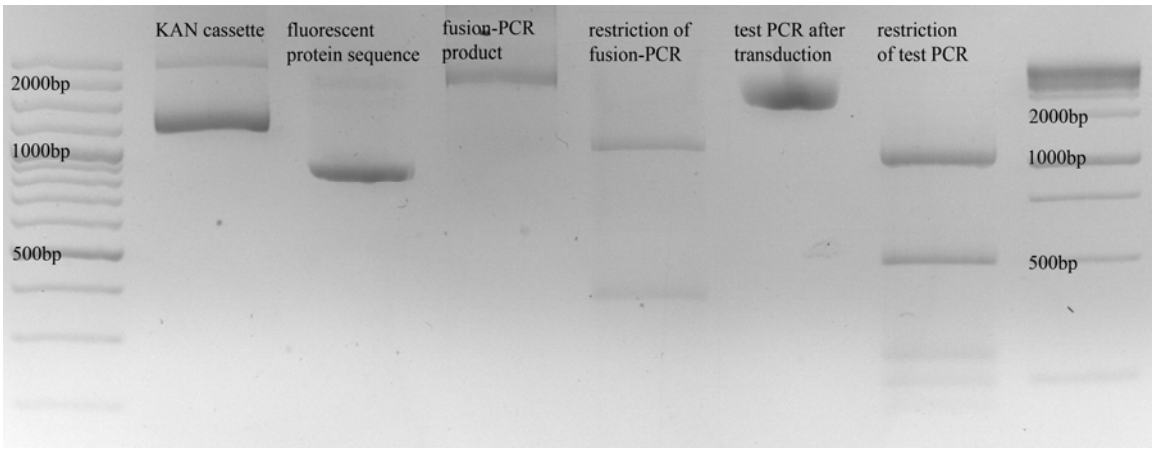

Fig. 2: Gel picture of each step during the construction of a new fluorescent mutant: amplification of the single DNA parts, purified fusion PCR product, test PCR after phage transduction and restrictions from the fusion product and the test PCR with Tsp509I (here for tagCFP as representative)

After purification and restriction digest of the fusion PCR product, the bacteria were transformed with the PCR product by electroporation.

The voltage varying from 2.5kV to 5kV was applied for 2.5ms to 50ms. The transformation efficiency for the different mutants differed from 0.01 to 0.3 colony forming units (cfu) per ng DNA.

Transformants were tested using PCR with primers which bind about 100bp outside of the insert. Correct clones were analyzed for fluorescence by FACS analysis.

In case of positive results, the locus of interest was transduced with phage P22 into wildtype S.Tm SL1344. After transduction, clones were tested for residual phage by PCR (Fig. 3) and PCR-RFLP for the fluorescent protein locus (Fig. 2).

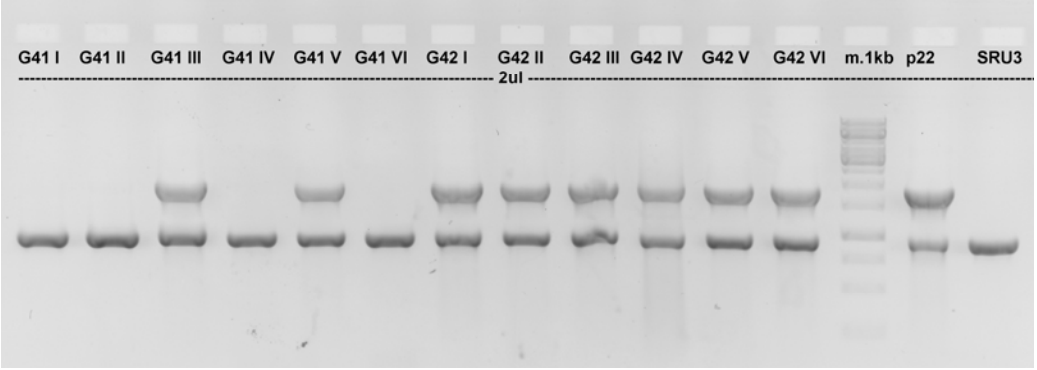

Fig. 3: PCR to check for phage DNA after transduction, clones I, II, IV and VI are phage-free

Finally the inserts were sequenced, to detect mutations, which were potentially introduced during PCR.

The two YFP variants (SRU3 with 100% and SRU4 with 20% intensity) did not have any mutations, one of the two weak (20%) GFP clones (SRU1) and the CFP clone (SRU5) showed only silent mutations that did not affect the amino acid sequence. Another weak GFP clone (SRU2), that had slightly different fluorescence properties (Fig. 4A), and the dsRED clone (SRU6) had non-synonymous mutations (L236P for SRU2, V122M and E212G for SRU6). The V122M mutation in the dsRED sequence was in the middle of the protein and might interfere with folding fluorophore formation. SRU6 was not further investigated.

4.3 *In vitro* validation

In the various constructs, fluorescent protein expression is driven by the P_{sifB} promoter that is highly active during infection but not during standard *in vitro* broth culture. To induce *in vitro* expression of the fluorescent proteins, the strains were grown sequentially in M9 and M9* media. In M9 medium, the bacteria grew slightly slower than in LB, but already showed some production of the fluorescent proteins. In M9* medium, the bacteria grew very slowly and showed a maximal fluorescence intensity.

First, the different strains were analyzed separately, each one with its appropriate laser and filter settings. The histograms are shown in figure 4.

Figure 4A shows three different green fluorescent strains excited by the 488nm laser. The two different weak GFP strains have both almost no overlap with either the non fluorescent wild type or the strong GFP in the SDFR3 strain. The difference between these two strains was consistently observed and might be due to the mutation of one amino acid (L236P), but this is not validated yet.

The two YFP producing strains are shown in figure 4B, the strong YFP was well separated from the wild type whereas the weak YFP strain (SRU4) substantially overlapped with the non fluorescent SL1344, impairing its use.

In figure 4C, the fluorescence histogram of the CFP- (SRU5) and the GFPuv-producing strains (SKS6) are shown. Both fluorescent proteins were well separated from the wild type and each other. Beside dsRED, which had a mutation that could affect fluorescence, also the FP635 (SKS3) did not show any fluorescence. For the FP635 this was most likely due to inefficient excitation.

Multi-color analysis is more informative than single parameter histograms for detection of individual strains in complex mixtures. Figure 5 shows that all strains except the SRU4 (weak YFP) were well distinguishable in two dimensional dot plots and did not overlap with any other strain including the wild type.

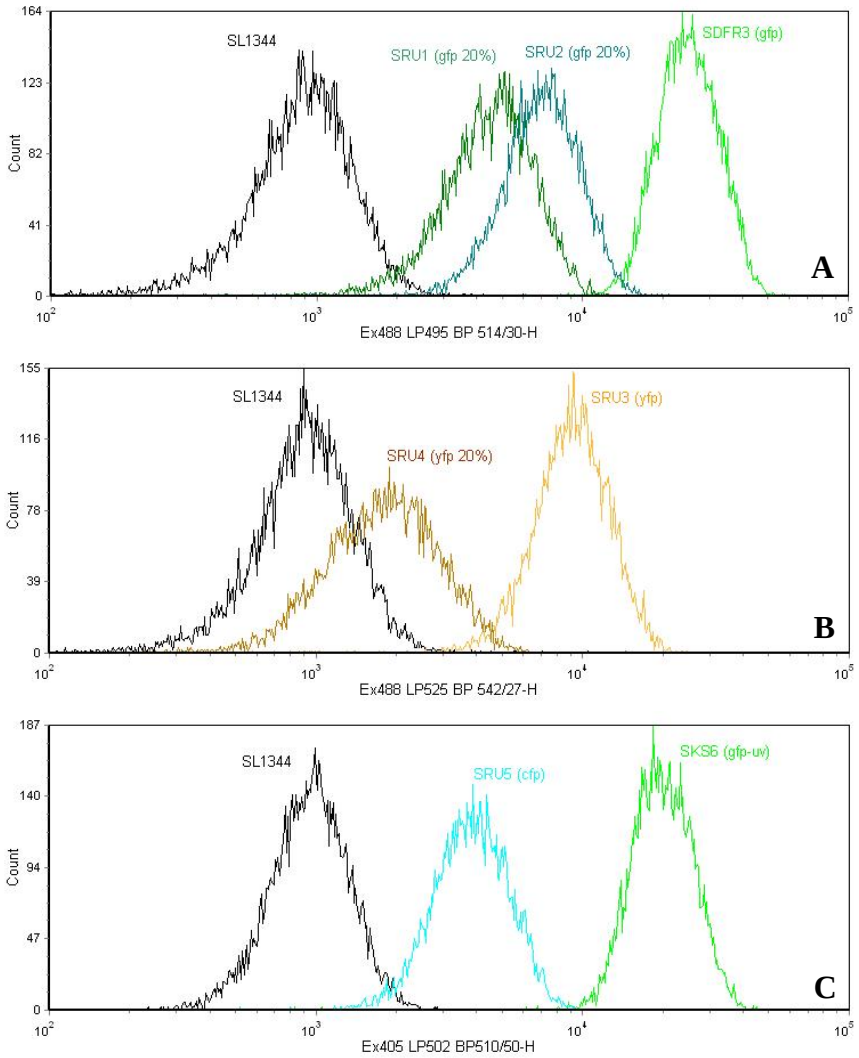

Fig. 4: Histograms of strains encoding various fluorescent proteins. A for GFP mutants, B for YFP mutants and C for CFP and GFPuv mutants

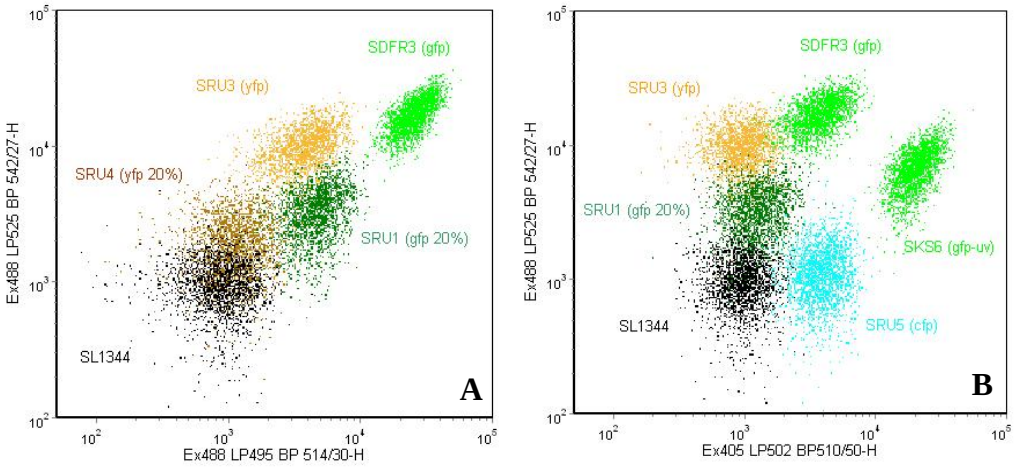

Fig. 5: Two color dot plots of several fluorescent strains. A shows strong and weak strains of GFP and YFP in a YFP-channel vs. GFP-channel plot. B shows a YFP-channel vs. UV-channel plot with both GFP strains, the YFP, CFP and GFPuv strain

4.4 *In vivo* validation

For *in vivo* evaluation, seven different strains were used. SKS3 (FP635) was chosen despite its undetectable fluorescence in FACS as a potential red fluorescent strain for microscopy analysis. The other six strains were SDFR3 (GFP), SRU1 (weak GFP), SRU3 (YFP), SRU4 (weak YFP), SRU5 (CFP) and SKS6 (GFPuv). 100 to 600 bacteria of each strain were mixed. The total number of injected cells was about 1880 as determined by plating each strain individually.

One BALB/c mouse was infected intravenously and sacrificed four days after infection. Due to the high infection amount and the long infection time (four days), the mouse showed clinical signs of disease including scrubby fur but was still climbing on the cage. Liver and spleen were collected. The numbers of colonies on the plates gave a total bacteria number in the spleen in the expected range of 10^7 cfu.

Previous experiments by another lab member (Petra Spröte) indicated that use of red fluorescent DNA stains can facilitate separation of fluorescent *Salmonella* from autofluorescent host tissue fragments. Possibly this reflects a higher DNA content of bacteria as compared to host fragments of similar size. DNA staining with ToPro-3 and Syto60 was compared which are both excited at 633nm and can be analyzed with a

660/20 bandpass filter but have no contribution to other colors in the shorter wavelength range. The results of the ToPro-3 staining showed a better distinction between tissue and bacteria than the Syto60 dye (Fig. 6). Even better distinction between bacteria and tissue background was obtained by diluting the centrifuged spleen sample first with Triton/PBS solution before fixing with formaldehyde (done by Petra Spröte, data not shown).

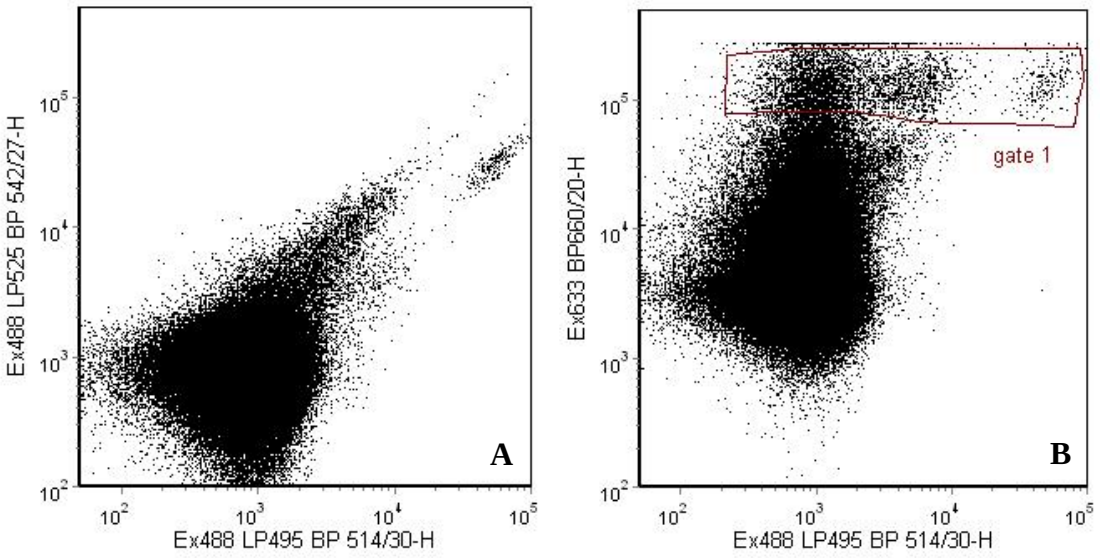

Fig. 6: Prepared *in vivo* sample without ToPro-3 staining (A) and with staining (B)

Gate 1 was created around ToPro-3^{hi} populations to suppress most of the tissue background. In a second step, a dot plot only showing bacteria within this gate was displayed (Fig. 7A), plotting the forward scatter (FSC) versus the side scatter (SSC). This scatter plot revealed a dense cluster of particles with typical properties of *Salmonella* (gate 2) as well as larger host fragments. This assignment was verified using specific analysis of the clearly distinguishable GFP^{hi} population (not shown).
Combination of ToPro-3^{hi} (gate 1) and scatter gate (gate 2) revealed separate populations in a multi-color dot plot (Fig. 7B).

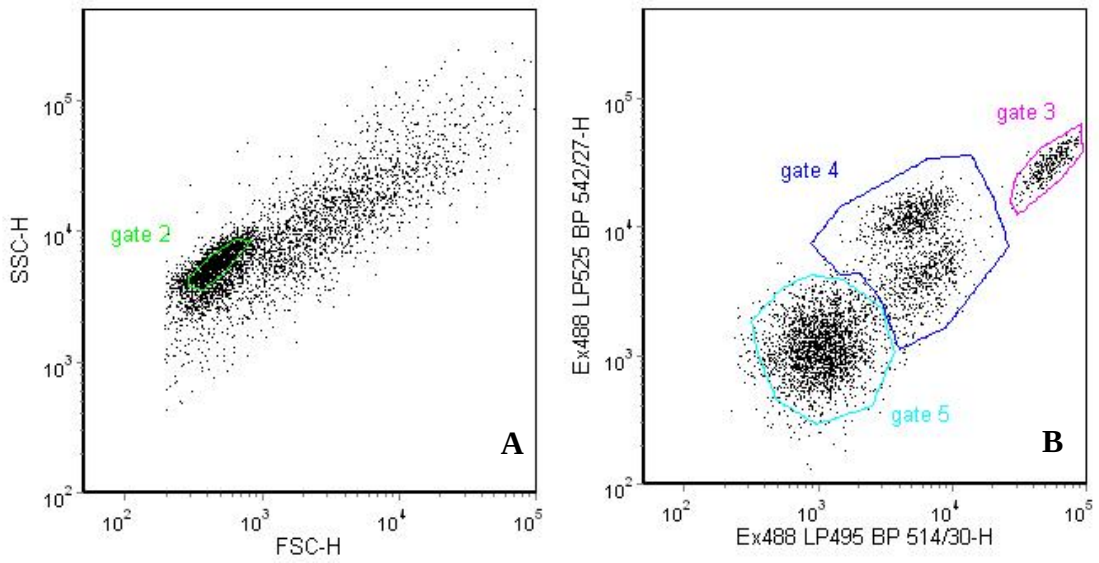

Fig. 7: Dot plot of gate 1, with gate 2 for the size (A) and (gate 1) AND (gate 2) (B), gate 3 is only for the GFP population, gate 4 separates the rest of green or yellow emitting particles from the non emitting ones (gate 5)

Gate 3 in figure 7B separated the GFP containing bacteria (SDFR3). Gate 4 distinguished the particles that showed fluorescence in either of the two channels ("YFP"-channel on Y axis and "GFP"-channel on X axis) from the ones that did not have any emission in these two channels (gate 5). All these gates were created as daughter gates of (gate 1) AND (gate 2) shown in figures 6B and 7A.

Gate 4 and 5 were then plotted in other fluorescence channels. Gate 4 was analyzed in "YFP-channel versus "UV"-channel (Fig. 8A) clearly separating GFPuv expressing SKS6 strain (gate 6) from strains expressing YFP and the weak GFP (SRU4 and SRU1, gate 7). This gate 7 was plotted again to separate SRU4 from SRU1 (Fig 9A).

In figure 8B, separation of the CFP expressing strain SRU5 (gate 8) and the non emitting strains SKS3 (FP35) and SRU4 (weak YFP) (gate 9) is shown.

Gate 10 in figure 9A separates the YFP expressing bacteria (SRU3) from the strain expressing weak GFP (SRU1) (gate 11).

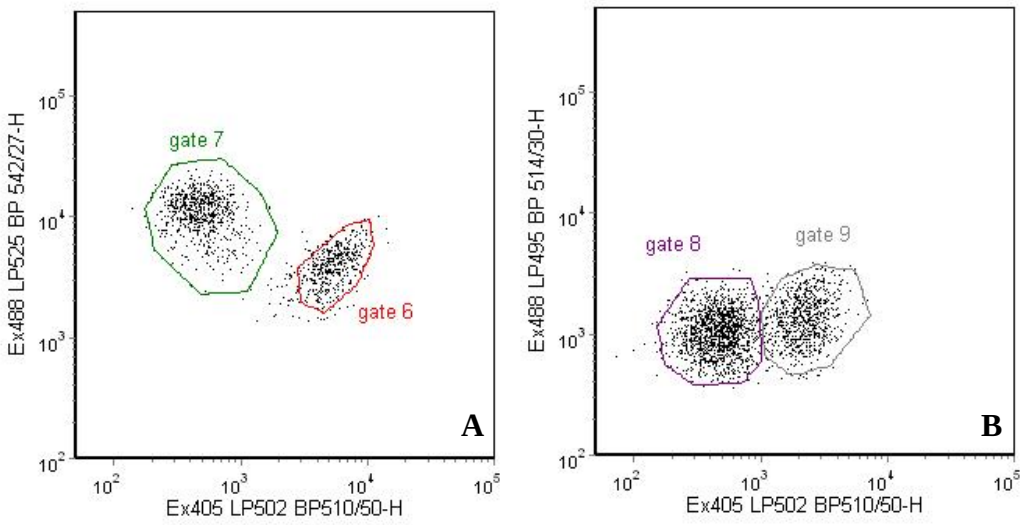

Fig. 8: A shows (gate 1) AND (gate 2) AND (gate 4) with SKS6 (gate 6), SRU3 and SRU1 (both gate 7), B shows (gate 1) AND (gate 2) AND (gate 5) with SRU 5 (gate 9) and non emitting SKS3 and SRU4 (gate 8)

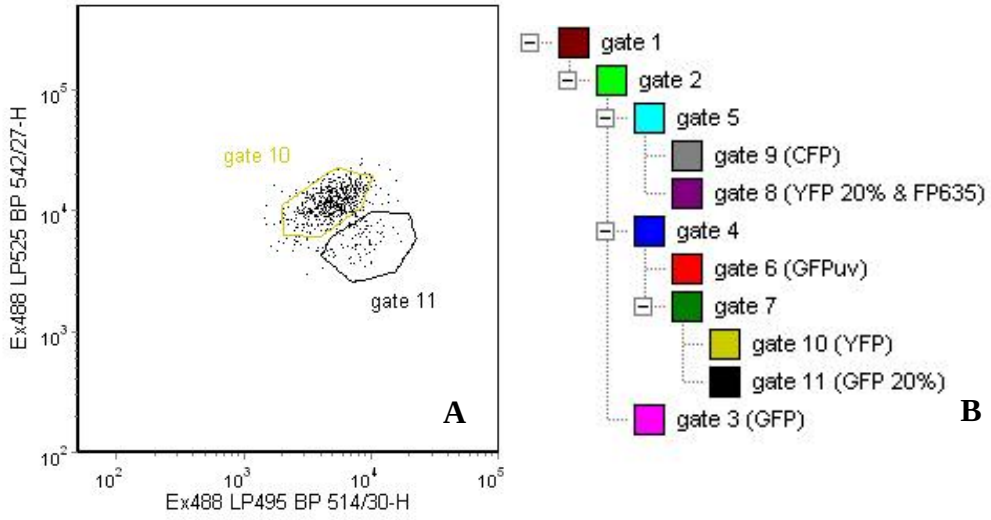

Fig. 9: A shows (gate 1) AND (gate 2) AND (gate 4) AND (gate 7) with the weak GFP strain (SRU1, gate 11) and the YFP strain (SRU3, gate 10), B shows an overview over all gates created during the analysis

By creating specific gates/gate combinations, the numbers of each fluorescent strain could be determined. Table 6 shows the number of gated cells for GFP (gate 3), YFP (gate 10), weak GFP (gate 11), CFP (gate 9), GFPuv (gate 6) and the non emitting cells (weak YFP and FP635, gate 8).

Table 5: Absolute numbers and ratios of total number of bacteria for each strain

	GFP	weak GFP	GFPuv	YFP	CFP	non em.
abs. num.	1577	446	2451	3186	3506	8416
rel. num.	0.081	0.023	0.125	0.163	0.179	0.43

Since the absolute number of the gated cells varied with the size of the gates, it was important to create gates that were similarly narrow compared to the density of the population.

The competitive index was calculated to analyze the effect on virulence of the different fluorescent proteins. To do this, the relative numbers (number of events per strain divided by total number of all strains) of gated cells (table 6) were divided by the relative numbers of each strain in the infection inoculum. Numbers and ratio of non-fluorescent SKS3 and SRU4 were summed up to the total number of non emitting events. Since mutation of the *sifB* locus does not affect virulence [12], the competitive index was expected to be 1.0, if each fluorescent protein variant was equally well tolerated.

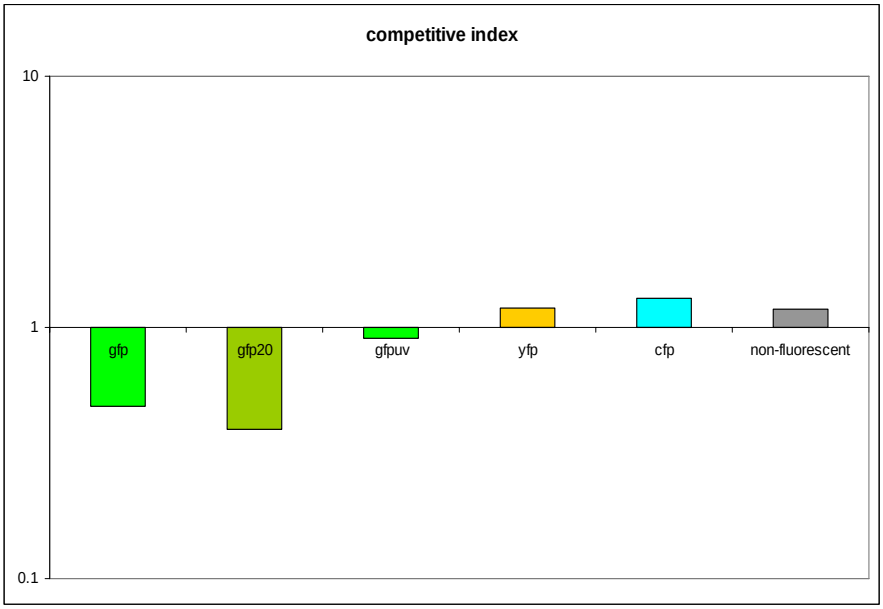

Fig. 10: Competitive indices for the different strains, logarithmic scale

The calculated competitive indices (CI) are shown in figure 10. For the strains producing GFPuv, YFP or CFP and for the non emitting cells, a CI close to 1 was obtained, showing that they did not have an effect on virulence.
The strains expressing high and low levels of GFP ("gfp" and "gfp20" respectively) were underrepresented. Previous experiments indicated that GFP at high level did not impair virulence and the deviations from the theoretical value 1 were still small and comparable to analogous results for replica plating [5]. More experimental data will be required for definite evaluation of relative *in vivo* fitness of the various strains.

5. Discussion

The goal of this work was to replace antibiotic resistance markers with fluorescent proteins as markers for competitive *in vivo* infections.
There was no universal protocol for fusion PCR that could be used for all constructs. However, all desired constructs could be generated. Strain SRU6 contained mutations that might have caused its undetectable fluorescence.
The difference between SRU1 and SRU2, (weak GFP strains), might be due to one mutation (L236P), although its location at the very C-terminal end of the coding sequence is far from other known mutations affecting GFP fluorescence [15]. Mutations in the promoter region (has not been sequenced) might have occurred.
Construction and application of the other strains was successful (SRU3 and SRU5 without mutations).
In vivo and the preliminary *in vitro* data suggest that it should be possible to track up to five different fluorescent S.Tm strains during competitive infection in the same host environment when analyzed using multi-color flow cytometry with appropriate filter settings. Preliminary CI data suggested negligible effects of the various constructs on virulence. The CI data also showed that the non fluorescent bacteria were well separated from the autofluorescent tissue by DNA staining.

6. Conclusion

Nine S.Tm SL1344 strains producing different fluorescent proteins were analyzed, out of which six strains were constructed in this work. The fluorescent protein encoding sequence was fused to a kanamycin resistance cassette by PCR. Electroporation of the fusion product was followed by phage transduction into *S. Typhimurium* SL1344. *In vitro* validation by PCR, restriction digest, sequencing and FACS measurements showed that at least five strains with different fluorescent proteins could be constructed and detected in mixtures *in vitro* and *in vivo* using multi-color flow cytometry (GFP, weak GFP, SYFP, tagCFP and GFPuv).

In vivo validation in a mouse infection showed that staining of DNA facilitates separation of bacterial fluorescence from host tissue autofluorescence. At least five different strains (including the wild type or SKS3 it would be six) were well separated from each other and host cell fragments using flow cytometry with appropriate filters.

7. Outlook

To further enlarge the set of fluorescent *Salmonella* strains, additional fluorescent protein variants could be tested. A yellow fluorescent protein with an emission maximum at 542nm (turboYFP) could result in higher fluorescence detection values and permit detection of strains with high and low levels (similar to the GFP strains). A mutation-free dsRED containing mutant could be constructed. An additional green laser (561nm) will provide efficient excitation for several red-shifted variants. A combination of two different colors in the same strain could also increase the number of detectable strains. In particular CFP does not interfere with any of the other colors. For this purpose a second locus would be beneficial. Previous data suggest that the *virK* locus could be suitable [14].

Finally, the infection inoculum could be induced *in vitro* for accurate and efficient determination of each strain by flow cytometry also for the input pool.

8. References

[1] Haraga A, Ohlson MB, Miller SI. Salmonellae interplay with host cells. Nat Rev Microbiol 2008;6:53-66

[2] Mastroeni P, Grant A, Restif O, Maskell D. A dynamic view of the spread and intracellular distribution of Salmonella enterica. Nat Rev Microbiol 2009;7:73-80

[3] Hewitt CJ, Nebe-Von-Caron G. The application of multi-parameter flow cytometry to monitor individual microbial cell physiological state. Adv Biochem Eng Biotechnol 2004;89:197-223

[4] Nebe-von-Caron G, Stephens PJ, Hewitt CJ, Powell JR, Badley RA. Analysis of bacterial function by multi-colour fluorescence flow cytometry and single cell sorting. J Microbiol Methods 2000;42:97-114

[5] Wendland M, Bumann D. Optimization of GFP levels for analyzing Salmonella gene expression during an infection. FEBS Lett 2002;521:105-108

[6] Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. Green fluorescent protein as a marker for gene expression. Science 1994;263:802-805

[7] Xia NS, Luo WX, Zhang J, et al. Bioluminescence of Aequorea macrodactyla, a common jellyfish species in the East China Sea. Mar Biotechnol (NY) 2002;4:155-162

[8] Kremers GJ, Goedhart J, van Munster EB, Gadella TW, Jr. Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Forster radius. Biochemistry 2006;45:6570-6580

[9] Wall MA, Socolich M, Ranganathan R. The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. Nat Struct Biol 2000;7:1133-1138

[10] Shcherbo D, Merzlyak EM, Chepurnykh TV, et al. Bright far-red fluorescent protein for whole-body imaging. Nat Methods 2007;4:741-746

[11] Crameri A, Whitehorn EA, Tate E, Stemmer WP. Improved green fluorescent protein by molecular evolution using DNA shuffling. Nat Biotechnol 1996;14:315-319

[12] Thone F, Schwanhausser B, Becker D, Ballmaier M, Bumann D. FACS-isolation

of Salmonella-infected cells with defined bacterial load from mouse spleen. J Microbiol Methods 2007;71:220-224

[13] Waterman SR, Holden DW. Functions and effectors of the Salmonella pathogenicity island 2 type III secretion system. Cell Microbiol 2003;5:501-511

[14] Rollenhagen C, Sorensen M, Rizos K, Hurvitz R, Bumann D. Antigen selection based on expression levels during infection facilitates vaccine development for an intracellular pathogen. Proc Natl Acad Sci U S A 2004;101:8739-8744

[15] Heim R, Tsien RY. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. Curr Biol 1996;6:178-182

9. Acknowledgment

I would like to thank:

Prof. Dirk Bumann for the opportunity to work on this interesting project and for a lot of fruitful discussions

Anne Kathrin Schemmer for introducing me into microbiology methods and for revising this thesis

Dr. Petra Spröte for revising this thesis and for helping with DNA staining and microscopy sample preparation

Beatrice Müller for a lot of support in the lab and at the FACS machine

Dr. Alain Mazé for a lot of useful advices and for helping with the mouse infection

Benjamin Steeb for support and help in the lab

All the rest of the Bumann working group for helping me with bigger or smaller problems and for a nice working atmosphere

Kasper Renggli for regular coffee breaks and inspiring discussions

10. Appendix

10.1 Strain sequences

The sequences shown are the results from the sequencing analysis, including the overlaps for *sifB* promoter region (5'-end) and kanamycin cassette (3'-end).
Start and stop codon are highlighted in red.

SRU1 (20% GFP)

GATCCTCTAGATTAAAGAACGAGATATACATATGAGTAAAGGAGAGGAACT
TTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGC
ACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAAC
TTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACA
CTTGTCACTACTTTCGCGTATGGTCTTCAATGCTTTGCGAGATACCCAGATCA
TATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAG
GAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAA
GTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTG
ATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATA
ACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAG
TTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAG
ACCATTATCAACAAAATACTCCGATTGGCGATGGCCCTGTCCTTTTACCAGA
CAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAG
AGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATG
GCATGGATGAACTATACAAATAAATGTCCAGACCTGCAGCCAAG

SRU2 (20% GFP)

GATCCTCAGATTTAAGAACGAGATATACATATGAGTAAAGGAGAAGAACTT
TTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGC
ACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAAC
TTACACTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAAC
ACTTGTCACTACTTTCGCGTATGGTCTTCAATGCTTTGCGAGATACCCAGATC
ATATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGA
AGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATT
GATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTAT
AACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAA
GTTAACTTCAAAATTAGACACAACATTGAAGACGGAAGCGTTCAACTAGCA
GACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAG
ACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAA
GAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACAT
GGCATGGATGAACCATACAAATAAATGTCCAGACCTGCAGCCAAG

SRU3 (SYFP)

GATCCTCTAGATTTAAGAAGGAGATATACATATGGTGAGCAAGGGCGAGGA
GCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAAC
GGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGC
AAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGC
CCACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCC
CGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTAC
GTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGC
GCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAG
GGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTAC
AACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGC
ATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGC
TGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCA
ACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGA
TCACTCTCGGCATGGACGAGCTGTACAAGTAACTGCAGCCAAG

SRU4 (20% SYFP)

GATCCTCTAGATTTAAGAACGAGATATACATATGGTGAGCAAGGGCGAGGA
GCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAAC
GGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGC
AAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGC
CCACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCC
CGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTAC
GTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGC
GCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAG
GGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTAC
AACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGC
ATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGC
TGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCA
ACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGA
TCACTCTCGGCATGGACGAGCTGTACAAGTAACTGCAGCCAAG

SRU5 (tagCFP)

GATCCTCTAGATTTAAGAAGGAGATATACATATGAGCGGTGGTGAAGAACT
GTTTGCAGGTATTGTTCCGGTTCTGATTGAACTGGATGGTGATGTGCATGGT
CATAAATTTAGCGTTCGTGGTGAAGGCGAAGGTGATGCAGATTATGGCAAA
CTGGAAATTAAATTTATTTGCACCACCGGAAAACTGCCGGTTCCGTGGCCGA
CCCTGGTTACCACCCTGGCATGGGGTATTCAGTGTTTTGCACGTTATCCGGA
ACACATGAAAATGAATGATTTTTTTAAAAGCGCCATGCCGGAAGGTTATATT
CAGGAACGCACCATCCATTTTCAGGATGATGGCAAGTATAAAACCCGTGGC
GAAGTTAAATTTGAAGGTGATACCCTGGTTAATCGTGTTGAACTGAAAGGCG
AAGGTTTTAAAGAAGATGGCAACATTCTGGGCCATAAACTGGAATATAGCG
CCATTAGCGATAATGTGTATATCATGCCGGATAAAGCCAATAATGGCCTGGA
AGCCAACTTTAAAATCCGCCATAATATTGAAGGTGGTGGTGTTCAGCTGGCA
GATCATTATCAGACCAATGTTCCACTGGGTGATGGTCCGGTGCTGATTCCGA
TTAATCATTATCTGAGCTGTCAGAGCGCAATTAGCAAAGATCGTAACGAAGC
ACGTGATCACATGGTTCTGCTGGAATCTTTTAGCGCATATTGCCATACCCAT
GGCATGGATGAACTGTATCGTTAAATGTCCAGACCTGCAGCCAAGCTTGGCT
GTTTTGGCGGATGAG

SRU6 (dsRED)

GATCCTCTAGATTTAAGAAGGAGATATACATATGGCATCCACCGAGGACGTC
ATCAAGGAGTTCATGCGCTTCAAAGTGCGCATGGAGGGCTCCGTGAACGGC
CACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACC
CAGACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCCCTTCGCCTGG
GACATCCTGTCCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACC
CCGCCGACATCCCCGACTACAAGAAGCTGTCCTTCCCCGAGGGCTTCAAGTG
GGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGA
CTCCTCCCTGCAGGACGGCTGCTTCATCTACAAG**A**TGAAGTTCATCGGCGTG
AACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACTATGGGCTGGGAG
CCCTCCACCGAGCGCCTGTACCCCCGCGACGGCGTGCTGAAGGGCGAGATC
CACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAG
TCCATCTACATGGCCAAGAAGCCCGTGCAACTGCCCGGCTACTACTACGTGG
ACTCCAAGCTGGACATCACCTCCCACAACGAGGACTACACCATCGTGG**G**GC
AGTACGAACGCACCGAGGGCCGCTACCACCTGTTCCTGTAGAAGCTTCTG