Technische Universität München
TUM Campus Straubing für Biotechnologie und Nachhaltigkeit

Establishment of a microfluidic sorting system for the identification of improved enzyme variants in ultrahigh-throughput screenings

Robert Genth

Vollständiger Abdruck der von dem TUM Campus Straubing für Biotechnologie und Nachhaltigkeit der Technischen Universität München zur Erlangung eines

Doktors der Ingenieurwissenschaften (Dr. Ing.)

genehmigten Dissertation.

Vorsitz: Prof. Dr.-Ing. Jakob Burger

Prüfende der Dissertation:

1. Prof. Dr. rer. nat. Volker Sieber
2. Prof. Dr.-Ing. Michael Zavrel

Die Dissertation wurde am 14.04.2025 bei der Technischen Universität München eingereicht und durch den TUM Campus Straubing für Biotechnologie und Nachhaltigkeit am 17.07.2025 angenommen.

Table of Contents

Abstract

Enzyme engineering is a powerful tool to tailor the properties of natural enzymes. By substitution of certain amino acids within the polypeptide chain, the 3D structure can be adapted towards the requirements of the latter industrial application process. However, it is not facile to estimate the impact of one or more amino acid substitutions and one of the easiest ways to assess differences is the measurement of the enzymatic turnover. By alternating several positions within a protein, the number of variants easily rises to several tenths of thousand to millions of variants, which exceeds the capacities of traditional screening platforms. Therefore, screening systems with even higher sample throughputs are needed. In order to improve the throughput and screen bigger enzyme libraries a droplet-based microfluidic sorting system was established over the course of this thesis. For that, microfluidic chips for droplet production and measurement were manufactured. The corresponding droplets were measured and an assay for the determination of enzyme turnover was characterized at different temperatures and pH. A self-customed LabVIEW program for the absorbance-based microfluidic sorting system was established and connected with the hardware components that allow isolation of interesting cells within droplets. To prove the functionality of the established system, three different enzymes were engineered:

First, the substrate specificity of a glucose dehydrogenase from *Bacillus subtilis* (*Bs*GDH) was improved towards xylose. In the first screening step the specific activity was improved 3-fold by a single mutation. In a second screening round three mutants with different mutations could be discovered, showing up to 197% higher enzyme turnover in the microtiter plate rescreening. In a final characterization the best variant *Bs*GDH$_{M168L}$ did not show an improved substrate affinity, but a 15 % improved catalytic efficiency compared to the first round.

For the second enzyme, a thermophilic glucose dehydrogenase from *Saccharolobus solfataricus*, the cofactor affinity was engineered towards the artificial cofactor cNADP (carba-nicotinamide adenine dinucleotide phosphate). For the screening process, droplets were incubated for several hours at elevated temperatures to simulate the later application process. In several screening steps, 12 variants with improved activity could be identified. The best variant *Ss*GDH$_{V63S\ T189L\ S280N\ R362T}$ showed a quadruple substitution and improved the catalytic efficiency by 10-fold.

In the last part of the thesis, two aldehyde dehydrogenases from *Herbaspirillum seropedicae* (*Hs*ALDH) and *Variovorax paradoxus* (*Vp*ALDH) were screened towards improved thermostability. For this purpose, the incubation temperature for the enzyme libraries was chosen close to the melting point of the template enzymes. With this setting, enzymes with an improved thermostability are favored to show an increased total turnover number. In a first screening round the *Hs*ALDH sorting yielded 30 mutants with improved thermostability.

Acknowledgements

First and foremost, I would like to thank my doctoral advisorProf. Volker Sieber for his tireless support, invaluable advices and his always-open door for scientific exchange. Under his guidance, I was able to grow throughout the course of this dissertation and further develop my analytical skills.

I thank Prof. Michael Zavrel for reviewing and examining my thesis. I would also like to thank Prof. Jakob Burger for taking the chair of my thesis committee and ensuring that my defense went smoothly.

My sincere gratitude goes to my mentor Prof. Florian Hollfelder for giving me the opportunity to visit his group in Cambridge. This stay opened the door for me to the until then unknown world of microfluidics. I am also grateful to Prof. Lieleg and Prof. Simmel for the collaborations between our research groups and the ongoing scientific exchange.

My heartfelt thanks goes to Ulrike Obst, who introduced me to the lab at the beginning of my doctoral studies. I thank Ioannis Zachos and Samuel Sutiono for their collaboration and their ideas for our joint project. Our mutual exchange made a significant contribution to my research and motivated me even after frustrating days. I would also like to thank Mariko Teshima for the exciting project work on aldehyde dehydrogenases.

A special thanks goes to all of my colleagues from the microfluidics group, especially Makis, Matthias Steiger and Dennis Romeis, for their support and creative exchange throughout our collaborative work.

I would also like to express my sincere gratitude to all the staff at the CBR, especially Magdalena Haslbeck, Petra Lommes and Lena Würstl, for their help and tremendous support. A special thank you goes to Elisabeth Aichner, who guided me through all organizational and administrative challenges.

I am very grateful for our joint running sessions with Torben Hüsing, Samed Güner and Korbinian Sinzinger and for participating together in the Straubing City Run.

I am deeply grateful to my family and friends for their love, patience, and constant support throughout this intense and challenging journey. Your trust in me has always given me the strength to continue this work.

Last but not least, I thank my wife and daughter for their unwavering support, understanding, and patience. You are my greatest source of strength and inspiration. Without you, this journey would not have been possible.

Abbreviations

2ACy-2	1-Methyl-3-[(E)-2-(1,3,3-trimethyl-3H-indolium-2-yl) ethenyl] quinolinium iodide trifluoromethanesulfonate
AADS	Absorbance activated droplet sorter/-sorting
AcAADS	Acoustic absorbance activated droplet sorter
ABTS	2,2'-azino-bis (3-ethylbenzothiazoline-6-sulfonic acid)
A/D	Analog to digital
ADH	Alcohol dehydrogenase
ALDH	Aldehyde dehydrogenase
BAW	Bulk acoustic wave
cNADP	Carba NADP
CV	Coefficient of variation
D2D	Droplet to digital
DA-64	N-(carboxymethylamino-carbonyl)-4.4'-bis-(dimethylamino)-diphenylamine
DE	Directed evolution
DMP	2,6-dimethoxyphenol
DMSO	Dimethyl sulfoxid
DNA	Deoxyribonucleic acid
dNTP	Deoxyribose nucleoside triphosphate
epPCR	Error-prone PCR
E. coli	*Escherichia Coli*
FACS	Fluorescent activated cell sorter/-sorting
FAD	Flavin adenine dinucleotide
FADS	Fluorescent activated droplet sorter/-sorting
FPGA	Field progammable gate array
GDH	Glucose dehydrogenase
GFP	Green fluorescent protein
HEPES	4-(2-hydroxyethyl)-1-piperazineethane sulfonic acid
INT	2-p-iodophenyl-3-p-nitrophenyl-5-phenyl tetrazolium chloride
IVTT	In vitro transcription translation
IMAC	Immobilized ion affinity chromatography
k_{cat}	Turnover number
k_{cat}/K_M	Catalytic efficiency
K_M	Michaelis-Menten constant
LED	Light emitting diode

LHS	Liquid handling system
mPMS	1-methoxy-5-methylphenazinium methyl sulfate
MTP	Microtiter plate
MTS	MTS(5-(3-carboxymethoxyphenyl)-2-(4,5,-dimethylthiazolyl)-3-(4-sulphophenyl) tetrazolium
MTT	3-(4',5'-dirnethylthiazol"i-y1)-2,5-diphenyltetrazolium bromide
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NBT	Nitroblue tetrazolium
NOX	NADH oxidase
oePCR	Overlap extension PCR
PCR	Polymerase chain reaction
PDMS	(Poly-)dimethylsiloxane
PFO	1H,1H,2H,2H-perfluorooctanol
PMT	Photo multiplier tube
QC	Quikchange
rpm	Revolutions per minute
SAW	Surface acoustic wave
SDM	Site-directed mutagenesis
TSAW	Traveling surface acoustic wave
VI	Virtual instrument (LabVIEW program)
WST	Water soluble tetrazolium salt
XTT	2,3-bis(2-methoxy-4-nitro-5-sulphopheny1)-5-[(phenylamino)carbonyl]-2-H-tetrazolium hydroxide, sodium salt

List of Tables

List of Figures

1 Introduction

1.1 Enzymes

1.1.1 Enzymes as natural catalysts

Enzymes are proteins, which catalyze biochemical reactions. They reduce the activation energy and thereby increase the velocity of the catalyzed reaction. They exist in every organism on our planet and their main function is to control the metabolic pathways. In 1836, Theodor Schwann proved the existence of enzymes by extracting pepsin from gastric juice of animals. He could show that pepsin specifically breaks down proteins into smaller peptides. While this seems to be an occasional finding, microorganisms and their enzymes have been studied more thoroughly in the last centuries. Modern techniques like light microscopy and chromatographic methods allowed not only the observation of the enzyme holding cells, but also the purification of the desired enzymes. In today´s life, enzymes established a central role. They are used as additives in cleaning detergents[2], for food processing like the production of cheese[3] and clearing up of fruit juices[4], production of pharmaceuticals[5] and the production of biofuel[6]. The reasons for this are simple:

Mild conditions

Many chemical reactions need extreme conditions. For example, the Haber-Bosch process for the production of ammonia is conducted at 400 °C – 500 °C and a pressure of up to 350 bar. The equivalent process in nature catalyzed by nitrogenase can be conducted at ambient temperature and pressure. Since most enzymes are derived from microorganisms, they also catalyze reactions at neutral pH conditions (pH 6 - 8). This makes these reactions easy to control and cheap in terms of energy consumption. The previously mentioned enzymes are summarized in the group of mesophiles. In contrast to that there are also enzymes from microorganisms, which prefer extreme conditions, e.g. higher temperatures. A very famous example is the *Taq*-polymerase, which was derived by the eubacterial microorganism *Thermus aquaticus*[7]. Naturally, living in hot springs, the bacterium adapted to the prevalent conditions and the enzymes catalyzing its metabolism can withstand elevated temperatures. The optimal working temperature of *Taq*-polymerase was found to be between 75 °C – 80 °C. This enzyme therefore belongs to the group of thermophilic enzymes. The ability to withstand higher temperatures can be very useful. The *Taq-polymerase* for example enabled the controlled polymerase chain reaction. Previously performed with DNA-polymerase derived from the mesophilic microorganism *Escherichia coli* (*E. coli*), a high concentration of polymerase was needed because the enzyme denatured at higher temperatures and therefore quickly lost its activity. By taking the enzyme from a thermophile, this bottleneck could be circumvented. Even

those thermophilic microorganisms require elevated temperatures, the requirements for an enzymatically catalyzed reaction can still be categorized as milder, compared to many chemically catalyzed reactions.

Selectivity

Enzymes can achieve very high selectivity towards the substrate they catalyze. Unlike chemical catalysts, they do not catalyze a broad range of reactions, but can be tuned to mainly catalyze the desired reaction. This can be achieved by alternating the genetic code, which later decides the selection of amino acids at a particular site within the polypeptide chain. The selection of amino acid also defines the three dimensional structure of the enzyme and especially of the catalytic site, where the substrate reaction takes place. Depending on the amino acid size and polarity, the attributes of the catalytic site can be changed towards higher regio- and enantioselectivity[8]. This adaptation process occurs naturally as a part of evolution, but it can be also artificially forced by mutagenesis of the DNA encoding the enzyme. The latter approach is called directed evolution and is applied for enzyme engineering within this work, which is further described in the upcoming chapters (1.1.3).

Availability

Many chemical catalysts are metals like gold, paladium or titanium[9]. Certain metals (or their derivatives) can only be found in particular areas on the world and the main way to obtain them is mining with expensive equipment[10]. In contrast, enzymes can be produced by cells, which can be cultured in any biological laboratory regardless of its location. The ability of cells to naturally divide, allows theoretically unlimited production of the desired enzyme and makes them ubiquitously available.

1.1.2 Dehydrogenases

Dehydrogenases belong to the enzyme class of oxidoreductases. They require cofactors like nicotinamide adenine dinucleotide (NAD) or nicotinamide adenine dinucleotide phosphate (NADP) to store their redox potential. Under alkaline conditions the substrate is preferably oxidized, while the cofactor is reduced to NAD(P)H. This is accomplished by binding both, the substrate and cofactor, at the catalytic center. A hydride group is then transferred from the substrate to the cofactor (Figure 1).

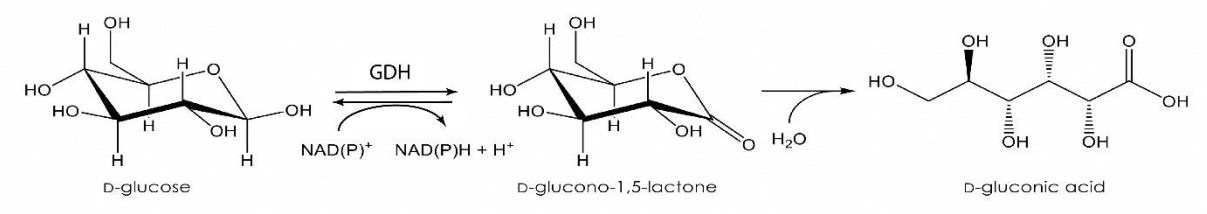


Figure 1: Dehydrogenase reaction illustrated by Glucose dehydrogenase as an example. In the first step D-glucose is oxidized by the glucose dehydrogenase to D-glucono-1,5-lactone under reduction of $NAD(P)^+$. In the second step the ring structure is opened by addition of water and D-gluconic acid is obtained.

Dehydrogenases can separate and upgrade monosaccharides from biomass liquors[11]. They are taken for the production of high-value compounds from biomass like gluconic acid and xylonic acid. These acids can be applied in various applications in food industry[12], pharmaceutical synthesis[13] and as precursors for the synthesis of higher alcohols like 1,2,4-butanetriol[14] and are consequently classified among the top 30 potential high value compounds from biomass by the US Department of Energy[15]. Apart from that, dehydrogenases are also used for cofactor recycling, because they require inexpensive substrates and the conversion is almost irreversible[16]. Cofactor recycling is especially interesting for industrial approaches, where the equimolar consumption of cofactor would be too expensive for technical applications[17]. In these reactions, dehydrogenases are added as a supportive enzyme, which recycles the cofactor for the main reaction. For example, Zhou et al. introduced a glucose dehydrogenase (GDH) for the chiral synthesis pathway of (R)-carbonyl reductase in a coupling system[18]. They could show that by addition of GDH the biotransformation efficiency of 2-hydroxyacetophenone to (R)-1-phenyl-1,2-ethanediol is increased by 64% and the reaction time is reduced two-fold.

1.2 Enzyme engineering

Natural enzymes usually show a decent activity towards their catalyzed reaction(s) under the predominant environmental conditions. These environmental parameters are mainly defined by the temperature, pH, the concentration of substrates, but also the concentration of other beneficial or harmful molecules, which help or interfere with the catalyzed reaction. Since the beginning of life, every (micro-) organism and its enzyme repertoire evolutionary adapted towards these conditions and only the ones who sufficiently adapted were able to survive. Obviously, for industrial applications this natural evolution would be too time consuming and the outcome would be uncertain and prone to failure. Besides, it is not always possible to perform industrial enzymatic reactions under the accustomed environmental conditions. Especially, if several enzyme reactions are performed in parallel in a one-pot reaction (enzyme cascade), concessions for temperature and pH will be necessary. However, if one of the above-mentioned parameters changes, the efficiency of the enzyme towards the catalyzed reaction usually decreases as well. Enzyme engineering can be utilized to change the polypeptide sequence of the enzyme and thereby adapting the enzyme towards these new parameters. A good enzyme catalyst is desired to have high activity and high stability within the given process parameters[19]. It is also important that the enzyme shows high specificity and selectivity towards the given substrate to avoid side reactions and have a high efficiency for the given reaction[20]. Especially, side reactions might inhibit the main enzyme reaction and decrease the overall conversion rate. Another important economical aspect that is often neglected is the protein yield[21]. A typical scenario would be that the expressed enzyme shows a slightly increased activity, but most of the expressed enzyme is not correctly folded and as a result unfunctional. The correctly folded enzyme variant itself would be improved, but since the yield is drastically decreased, the discovered enzyme mutant might not be an overall improvement for the process. For a successful enzyme engineering campaign, it is important to examine the enzymes under the conditions of their later application process. All the above-mentioned factors should be included in the screening process. As pointed out by Frances Arnold: “You get what you screen for”, the applied conditions have a direct influence on the screening result[22]. The idea of directed evolution is that variants, which show a higher fitness towards the varying parameters (e.g. higher temperature), will outperform the wildtype enzymes activity and/or stability, since their structure is more suitable to catalyze the enzyme reaction under the given conditions. The upcoming chapters will explain different methods to artificially create gene-expressing enzyme libraries and describe the isolation process of improved variants.

1.2.1 Generation of genetic diversity

1.2.1.1 Random Mutagenesis

Artificial evolution can be enforced in several ways. First described in 1927, Hermann Muller discovered that X-ray radiation could cause genetic mutations in fruit flies. He could show that the dose of X-ray radiation is proportional to the amount of mutations within the chromosomes[23]. Another way to enforce random mutations is the chemically induced mutagenesis. Here, chemicals like ethylmethane sulphonate[24] or hydrazine[25] are brought in direct contact with DNA. Most of the chemicals then directly bind to the nucleic acids of the DNA and interfere with the polymerase chain reaction. Both methods aim to change the DNA and different repair mechanisms try to reestablish the original sequence of nucleotides. Depending on the dose or concentration, this is not always successful and a different nucleotide is incorporated in the DNA. This eventually leads to mutations within the genetic code and can later change the nucleotide triplet (codon), which is translated into a different amino acid. However, the main disadvantage of the before mentioned methods is that they are indirect and therefore all available DNA is mutated and not only the plasmid that encodes the protein of interest. This decreases the selectivity of the approach and increases the chances of malfunctions within the expressing microorganism.

Error prone - Polymerase Chain Reaction (epPCR)

Several methods for random mutagenesis of defined plasmids/genes were developed by different groups around 1990 and later summarized under the name error prone – PCR[26,27]. Here, only DNA fragments, which are actively amplified, can include mutations. The main idea of this approach is to control the polymerase chain reaction by a low fidelity DNA polymerase and intentionally allow mutations within the amplified DNA. Therefor, the polymerase I from the thermophilic microorganism *Thermus aquaticus* was originally taken. It lacks 3´ to 5´exonuclease proofreading activity and has an error rate reported to be 1 in 18.000 nucleotides[28]. This error rate and/or the number of resulting mutations can be further increased by:

1) Addition of manganese ions as a substitution for magnesium
2) Increasing the concentration of magnesium ions
3) Using unequal dNTP (deoxynucleoside triphosphat) concentrations
4) Extending the elongation time
5) Increasing the iterations of PCR cycles

Once a mutation occurs in a daughter strand, it will be taken to the next iteration of PCR. This copy acts as parental strand and passes the mutation to all upcoming daughter strands, which can also include new mutations[29]. As a result, it would be theoretically possible to add more and more mutations to the template DNA and explore a higher number of variants. However,

more mutations hold also the risk to include stop codons in the middle of the DNA-sequence, which will yield only a part of the enzyme, once the translation is conducted. These truncated parts are usually not active and therefore not interesting for screening campaigns. Another problem is, that the chances to include mutations at highly conserved regions also increases. Consequently, the enzyme might not fold correctly and the resulting variant could be unfunctional as well. Therefore, it is crucial to test different epPCR conditions and evaluate the number of mutations and their impact on the functionality of the enzyme reaction. With the number of mutations **k** and the sequence length **n**, the number of possible variants **N(k)** within the library can be estimated as follows[30]:

$$N(k) = \frac{n!}{(n-k)! * k!} * 3^k$$

Since mutations are randomly distributed within the enzymes DNA, epPCR is usually applied for enzymes, where the structure is not well explored for example via protein crystallization. As a drawback, epPCR performed with *Taq* polymerase shows a bias towards A and T base pair mutations[31]. Other modern random mutagenesis kits (e.g. GeneMorph II, Agilent Technologies) reduce that bias, by utilizing a mixture of different polymerases[32]. Another bottleneck of epPCR is the low chance of subsequent mutations. A single nucleotide substitution within a codon triplet only results in an average substitution of 5.7 out of 19 possible amino acids[33]. Therefore, approaches, which specifically aim to change amino acids at a defined position, can be an interesting alternative.

1.2.1.2 Targeted mutagenesis

For proteins, where previous knowledge was aquired, site directed mutagenesis (SDM) or site-saturated mutagenesis (SSM) are usually the methods of choice for enzyme engineering. Compared to random mutagenesis, where mutations are arbitrarily distributed over the amplified DNA strand, SDM and SSM can incorporate mutations, deletions or insertions at predetermined points within the DNA strand. While SDM targets specific, predefined mutations, SSM can introduce a handful to all possible mutations at a specific site, depending on the primer motif. This allows changing specific amino acids within a protein, but only generating a maximum of 20 different variants per mutation. Subsequently, the two most prominent approaches to alter nucleotide sequences are shortly explained.

QuikChange (QC)

The Quikchange™ Site-Directed Mutagenesis System was developed by the company Stratagene (La Jolla, CA)[34]. To introduce mutations, complementary forward and reverse primers, which contain the desired mutations in the nucleotide sequence, are taken. Subsequently, the whole plasmid is amplified. Since all PCR amplified plasmids contain a

strand break (nick), they cannot be used for further rounds of amplification. Hence, the number of plasmid copies is not increasing exponentially (like in a normal PCR), but linearly and a rather high starting concentration of template DNA is necessary. In order to avoid transformation of the template DNA from the original microorganism, the endonuclease R from *Diplococcus pneumoniae* (DpnI) is taken as a restriction enzyme for parental plasmids. DpnI was found to digest only methylated DNA, which was produced by microorganisms[35]. In contrast to that the plasmid DNA obtained by PCR is unmethylated and is not digested by DpnI. After the digestion, the amplified PCR plasmid can be directly taken for transformation of competent cells.

Overlap extension PCR (oePCR)

Introduced by *Ho* et al. in 1989, this method allows the precise incorporation of several mutations at different positions within a plasmid[36]. The main idea is to split the gene of interest at the position where a mutation should occur. In the simplest version, the gene of interest is separated in two different fragments. To achieve this, the gene of interest is amplified with two different sets of primers in two separated reactions. Each of these reactions uses one flanking primer, which binds at the end of the target sequence and one internal primer, which binds partially at the site of the mutation. The two internal primers are designed in such way, that the parts, which do not hybridize with the sequence of interest, are complementary to each other. After purification of both PCR products, they are mixed together without addition of primers and form a fusion product by denaturation and annealing. The overlap introduced by the internal primers allows one strand from each fragment to act as a primer for the complementary fragment and introduce the mutation to the newly formed PCR product. After elongation, the flanking primers are added and the fused PCR product is amplified containing the new mutations. As the final step, the PCR product can be cloned into an expression vector and taken for transformation of competent cells.

1.2.2 Isolation of improved variants

The created libraries can then be expressed and the corresponding enzyme variants tested for their activity under the desired conditions.

Methods for monitoring product formation

Many of the products of enzymatic reactions cannot be detected directly, the formation must be coupled to a detectable signal. One of the most common ways for monitoring product formation is by measurement of a signal change from the redox cofactor: NAD(P). Depending on the nature of the reaction, $NAD(P)^+$ is reduced for oxidation reaction or NADPH is oxidized for reduction reaction in equimolar amounts. This conversion can be photometrically measured at 340 nm. Another way for monitoring product formation is given by coupling the enzymatic

reaction to conversion of a chromogenic dye. Depending on the enzyme reaction, the dye concentration is either increased or reduced in a equimolar ratio, allowing to quantitatively monitor the reaction progress with spectrophotometric measurements. The advantage of this approach is that chromogenic dyes are visible by eye and usually have a higher extinction coefficient than the cofactor based NAD(P)-measurements (>6220 $L \cdot mol^{-1} \cdot cm^{-1}$). This allows the detection of even lower product concentrations, which is particularly important if the pathlength of the measurement or the number of cells contributing to the product formation within the assay are reduced.

Selection vs. Screening

There are two approaches to verify the activity towards a substrate and eventually isolate improved variants. The first method is a selection of all variants. This approach gives a qualitative answer (yes or no) if an enzyme activity can be coupled to a cell growth. All cells bearing the desired enzyme will grow and can be taken to the next round of evolution. All other inactive variants are eliminated and cannot be studied subsequently. Therefore, selections are particularly applied to very large libraries (up to 10^{11} variants) in order to significantly reduce the number of variants[37], or when a variant must meet a specific selection criterion to be considered useful. Typical examples are selections for antibiotic resistance, substrate limitation or growth under acidic/alkaline conditions. The second method to isolate improved variants is given by screenings. Compared to selections, screenings give a quantitative answer. This means that the activity for every variant is determined and can be directly compared to other variants (e.g. template variant). Depending on the degree of adaption to the predefined conditions, the desired reaction can be performed faster and more product formation can be observed. Quantitative measurements have the advantage, that the screening criterion can be adapted during the measurements, while it is predefined and unchangeable for selection methods. However, quantitative screening approaches make the measurement of all variants of an enzyme library obligatory. This restricts the library size and make high-throughput approaches extremely desirable.

High-throughput methods

Traditionally, agar plates are a powerful method for high-throughput selection campaigns. They only require cheap disposables and agar medium, which contains the selection criterion. Especially for huge libraries (>10^9 variants) with a strong selection criterion (e.g. antibiotic substances), this selection method can help to drastically reduce the number of variants within an enzyme library. However, for screening campaigns only a few cells per plate can be investigated, as there is a high risk that colonies overgrow and the genotype of interesting colonies is mixed with colonies in close proximity. This fact along with the preparation time of several hours reduces the possible library size and throughput for screening campaigns (Table 1). A more advanced method is screening in microtiter plates. Here, cells bearing the enzyme

of interest are typically separated in up to 1536 different wells within the microtiter plate, avoiding genotype mixtures. Micro titer plates are (cheap) disposables, which are ready to use. Compared to agar plates, the reaction volume for each sample is reduced to the range of several microliters, which benefits the relative cost per sample. Another advantage is that microtiter plate screenings can be automated by liquid handling stations allowing the investigation of libraries with up to 10^5 variants. The throughput is mainly restricted by the readout capacity, which is traditionally performed by spectrophotometers. Additionally, measurements are prone to exterior effects (e.g. evaporation or difference in cell growth), which increase the deviation between measurements. Hence, replicates for measurements have to be made to reduce the deviation, which in turn reduces the actual throughput. To further reduce the reaction volume and avoid deviations caused by difference in cell growth, novel screening approaches focus on single cell measurements.

Table 1: Comparison of different high-throughput screening methods.

	Agar plates	**Microtiter plates**	**FACS**	**Microfluidic compartments**
Library Size [n]	10^4 (38)	10^5 (39)	$<10^8$ (38)	Up to 10^8 (40)
Measurement Throughput [per second]	<10	$<10^2$	$<5x10^4$ (41)	$<3x10^4$ (42)
Investment costs	€	€ - €€€	€€ - €€€	€ - €€
Relative costs per sample (10^7 variants)	€€€	€€ - €€€	€ - €€	€ - €€
Single Cell Analysis	No	No	Yes	Yes
Advantages	Simple Visual selection Cheap, disposable material	Separation of variants in different wells Automation Quantitative measurement	High-throughput Single cell analysis High sensitivity	High Throughput Single cell analysis Flexible assay conditions
Disadvantages	Long preparation time Restricted to colorimetric activity assays Diffusion problems	Prone to exterior effects (cell growth,evaporation) Expensive Equipment for automation No single cell analysis	High invest costs Requires trained user Assays restricted by cell wall	Requires trained user Necessary adaptions of equipment and software

One novel screening method is given by fluorescent activated cell sorting (FACS)[43]. Here, single cells are measured in an aqueous continuous phase and guided through a narrow measurement channel. Each cell is measured with light from a specific wavelength and the

light signals emitted by the cells are collected by forward and side scatter detectors, giving information about the size and granularity of the measured cells. Since most cells are only a few micrometers in size, the throughput for FACS can be tuned up to $5x10^4$ cells per second[41]. This allows the exploration of libraries with up to 10^9 variants, making it an extremely powerful tool for high-throughput screenings. Downsides of this technique are the high investment costs for the sorter and the restrictions for the assay predetermined by the composition of the cell wall (see chapter 1.2.3). To overcome these obstacles but keep most of the advantages of a FACS, microfluidic compartments were introduced. For this technique, single cells are encapsulated into small compartments, which act as reaction vessels. Depending on the composition of the compartments, molecules of specific size or charge can be kept within the compartment. This makes applications extremely versatile, but also requires specific adaptions of software and equipment for different screening approaches. The upcoming chapter gives a brief overview about microfluidic compartments for enzymatic measurements.

1.2.3 Compartments for enzymatic turnover measurements on single cell level

For a successful screening campaign, it is crucial to maintain the linkage between phenotype (referring to the assay signal caused by the improved enzyme) and genotype (referring to the corresponding DNA of the enzyme) over the whole course of the screening process. Only when the assay signal can be retraced to the genetic information within the DNA of the enzyme, the corresponding substitutions can be identified and used for production of the improved enzyme or further rounds of evolution. Therefore, the compartment in which the enzyme reaction proceeds plays an important role as it decides about the success of the screening campaign.

Cells

Designed by nature, cells with their cellular membrane can be directly taken as a powerful and selective reaction compartment. One major advantage is that cells can be directly measured in a commercially available FACS (fluorescent activated cell sorter), which allows cell sorting frequencies up to 50 kHz[41]. However, measuring the enzymatic activity within a cell comes with certain restrictions. First, if the enzyme is not presented on the cell surface it will be necessary to transport the substrate through the cell membrane into the cytoplasm to start the enzymatic reaction. This confines the size of the acceptable substrates and restricts the selection of molecules to the membrane-bound channels. Second, to measure the enzymatic turnover it is necessary that the product is not transported out of the cell. This can be achieved by increasing the size[44] or changing the charge of the product molecule[45]. Lastly, neither the substrates nor the products of the assay reaction should be catabolized by the cell. Therefore, it might be necessary to use engineered strains to conduct the screening. For example, Aharoni et al.[44] used the *E. coli* strain JM107 $NanA^-$ for engineering a sialyltransferase. In order

to avoid catabolization of the substrates, this strain lacked β-galactosidase (Δ*lacZ*) and Neu5Ac aldolase (Δ*nanA*) activity and allowed the engineering of sialyltransferase.

Water-in-oil emulsions

To avoid those restrictions water-in-oil emulsions, also called droplets, were invented[46]. For the production of droplets, the cells are within the dispersed aqueous phase. The continuous phase consists of fluorinated oil and separates the dispersed phase (chapters 1.3.2.1, 1.3.2.2). The ratio between cells and droplet volume is selected to ensure that ideally, each droplet contains only a single cell. The expected cell occupation is described by the Poisson distribution. To stabilize the droplets, surfactant molecules are needed[47]. These amphiphilic molecules consist of a polar head group and a long unpolar tail. This allows the surfactant molecules to settle on the interface between water and oil phase[48]. By that, bigger molecules (e.g. DNA or dye molecules) cannot enter or leave the droplet and droplet merging is strongly hindered. The droplet acts as a compartment around the cell and thereby expands the possibilities for assays. Substrates do not need to enter the cell anymore, but can be co-encapsulated with the cell in a droplet. Ostafe *et al*. exploited that fact to establish a cellulase assay[49]. Instead of protein production within the cytoplasm, they used a yeast display system to present the cellulase on the surface of the cell. Like this, the cellulase had direct access to the co-encapsulated polysaccharides and with the help of an enzymatic cascade could cleave the fluorescein moiety from the co-encapsulated substrate without entering the cell. Another way to use the droplet compartments is given by protein secretion. For example, Beneyton *et al*.[50] employed the yeast *Yarrowia lipolytica* and expressed five recombinant genes from *Aspergillus niger* (CBS513.88 strain). The yeast cells were incubated in droplets and after secretion of the proteins, the fluorogenic substrate was added. Depending on the substrate and the expressed enzyme either the xylanase, cellulase or protease activity could be measured and the droplets were sorted accordingly. Since not all enzymes can be easily displayed on the surface of the cell or secreted, another method is to disrupt the cell wall. First established by Kintses *et al*.[40], this method co-encapsulates lysis buffer together with the cells and the substrate. To avoid cell lysis prior to encapsulation, cell lysis buffer and cells are injected from separate syringes and just meet right before encapsulation. Within a short time, the cell membrane is disrupted and all proteins, which were in the cytoplasm before are released to the droplet sphere, where they can react with the co-encapsulated substrate. Since the droplet itself is not permeable for DNA or enzyme molecules, the phenotype-genotype linkage is retained. However, when producing the droplets one needs to keep in mind, that the protein concentration within the cell is strongly diluted upon releasing the cells to the droplet sphere. Therefore, the time required to generate a strong readout is also prolonged.

Double-emulsions

Water-in-oil-in-water emulsions are another way of compartmentalization. The main idea of this approach is to make droplets accessible for fluorescent-activated cell sorting (FACS). Since FACS machines only operate with continuous aqueous phases, it is necessary to change the continuous phase from oil to water again. This is accomplished by taking water-in-oil emulsions and emulsify them with another aqueous phase. This can either be conducted on two separated microfluidic chips[51] or by combination of two junctions right after each other[52,53]. Especially in the beginnings of droplet-based microfluidics, double emulsion screenings were very popular since FACS machines were commercially available and only very few people had access to FADS systems. Mastrobattista *et al*. used double emulsion droplets for in vitro transcription translation (IVTT) to evolve *Ebg* (a protein with unknown function) towards beta-galactosidase activity[54]. With their FACS approach, they could sort 10^7 variants per day and found several mutations of *Ebg*, which led to lactose consumption. Another example of high-throughput sorting with double emulsion was given by Zinchenko *et al*.[51]: The group showed that it is possible to isolate a single hit out of 100.000 droplets and retrieve the genetic information by direct transformation of *E.coli* cells. By applying a lambda of ten for the negative population an enrichment of one to one million could be reached, which implies that it is possible to retrieve the genetic information of every sorted droplet hit. A very innovative approach was presented by Adams *et al*.[55]: Instead of one single core, the group showed that it is possible to encapsulate several different cores in a double emulsion. By changing the continuous phase from oil to a low melting temperature wax, they induced a controlled coalescence of all inner cores by applying heat to the incubated droplets. This approach allows a precise control over the reactions within one droplet. The resulting inner core is still protected by the outside shell and keeps the genotype-phenotype linkage.

Hydrogel beads

One major difference between a cell and emulsion droplets is the selectivity of the outer barrier. While cells are semipermeable and allow transport of distinct molecules, emulsion droplets aim for complete separation of inner and outer droplet space. However, as mentioned before one bottleneck for cell-based reactions is that the enzymes within a cell can degrade the product or transports it out of the reaction compartment. Therefore, Fischlechner *et al*.[56] conceived a biomimetic compartment made out of an alginate-agarose mixture. The cells, all assay compartments and a lysis buffer solution were co-encapsulated with this mixture. Once the enzyme reaction generated enough product for the fluorescent readout, all droplets were solidified by cooling them down to 4 °C. After breaking the emulsion and adding the polycation poly-(allylamine hydrochloride), which reacts with the alginate polyanion, a polyelectrolyte shell was formed. The resulting shell served as a template for more polyelectrolyte layer coatings. Depending on the choice of the different layers, a biomimetic cage with defined properties was

created. Like this, the group developed a hydrogel bead, which switched the molecular weight cut-off from 250 kD (pure agarose) to <2 kD. As a result of this approach, the enzyme, the plasmid and the reaction product stayed within the hydrogel bead and allowed a readout based on the enzymes activity. The group proved the functionality of their approach by engineering a phosphotriesterase. After solidification of the hydrogel beads, FACS sorting yielded a 20 times faster variant. One other aspect of hydrogel beads is that they cannot only be used in oil as a continuous phase, but also in aqueous continuous phases[57]. Once solidified they can be trapped on a chip and supplied with fresh medium. The hydrogel around the cell allows exchange of medium and the contained nutritions without destruction of the reaction compartment. This enables long-term incubation/storage of cells[58] and even allows growth studies[59]. For example, Kleine-Bruggeney *et al.* established a microfluidic chip, which was capable of trapping up to 2.000 hydrogel beads[60]. In combination with confocal microscopy, they were able to track the differentiation of up to 2000 pluripotent stem cells trapped in the hydrogel beads. Cells were cultured over the course of 68 hours and depending on the medium selection, the cells either lost (differentiated) or kept their GFP signal (not differentiated).

1.3 Droplet-based microfluidics

1.3.1 General introduction

The research field of microfluidics explores the behavior and manipulation of fluids within geometrical structures at micrometer scale. The focus of this work will be determined on the subcategory of **droplet-based** microfluidics. Here, two immiscible fluids are encountering each other. A continuous phase separates a dispersed phase into microliter-sized drops (droplets). By exploiting the fact that the liquids are immiscible, it is possible to create millions of droplets from small volumes, which can all act as independent reaction chambers. With this technique, it is possible to isolate single cells or even molecules into individual droplets and study their disparities. The idea to isolate single cells or molecules in microdroplets reaches back until the 1950's were e.g. Nossal and Lederberg describe a method to isolate lymph node cells[61]. For their publication, they use a spray method to produce water-in-oil droplets on a microscopy slide. The aqueous phase contains a mixture of lymph node cells and bacteria from different strains. From all produced droplets, they only selected those ones that contained one plasma cell. They investigated the influence of the antibodies produced by single lymph node cells on the motility of different bacterial strains. They could show that some cells were able to produce antibodies and were capable to inhibit the motility of one bacterial strain. More importantly, they could show that none of the selected cells, when mixed with a different bacteria strain, could produce antibodies to immobilize also the second strain. Like this, they could provide first evidence for the clonal selection theory (One cell - One antibody), which was proposed by Talmage[62] and Burnet[63] a year before. Soon after this in 1961 Rotman used a similar way to produce microdroplets to investigate the activity of a β-D-Galactosidase[64]. He determined the enzymes activity through the conversion of a fluorescent dye. With his approach, he was able to calculate the turnover number of single enzymes and showed the difference in activity of different enzyme variants. In contrast to these groundbreaking results, only minor progress was observed in the field of microfluidics for the next three decades (Figure 2). Reasons for this might be (I) The special equipment and chemicals that were required to produce microdroplets or microchips (II) The difference in droplet size, which hampers the comparison of different droplets among each other (III) The stability problems for microdroplets due to merging and evaporation, which did not allow long term observations (IV) The missing tools to manipulate droplets after their production (V) The need for devices allowing the analysis of vast amounts of droplets in short time to generate statistical relevant data.

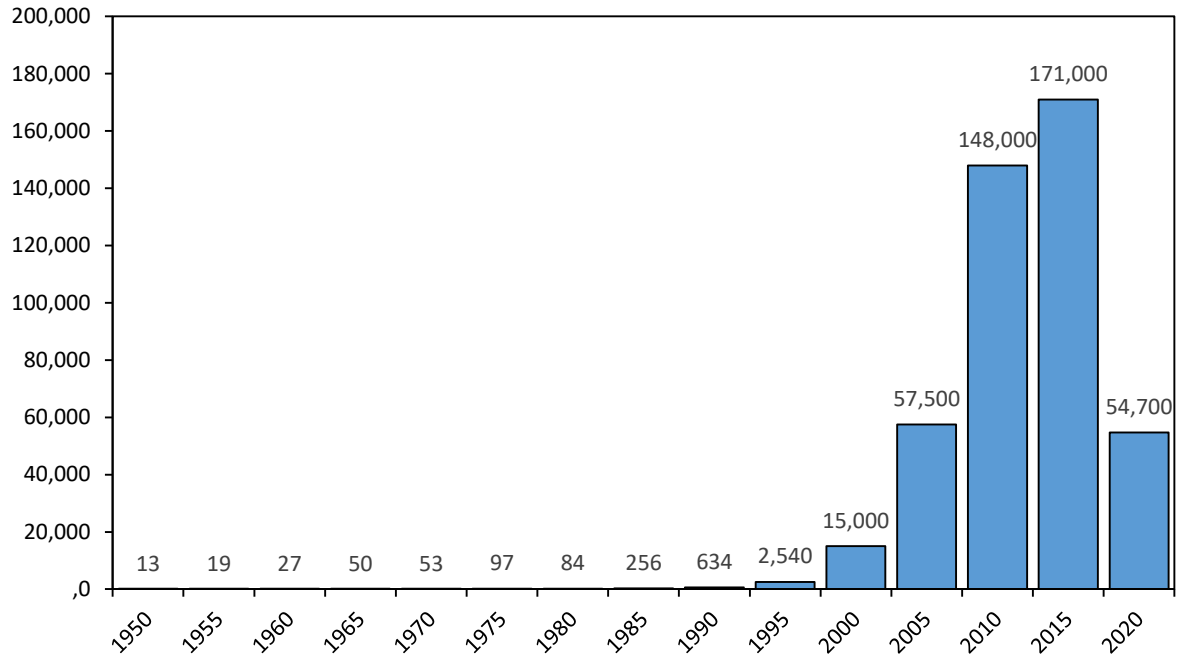


Figure 2: Number of articles published within 5 years (e.g. 1950: 1950-1954) with "microfluidic" as keyword in Google Scholar. Data was assessed in September 2022.

This resulted in a high investment, but poor flexibility and reproducibility for microfluidic experiments. Only when Whitesides et al.[65,66] introduced the production of microfluidic chips with PDMS ((poly-)dimethylsiloxane) to the microfluidic field in 1998, droplet-based microfluidics got more attention (Figure 2). It enabled the cheap and fast production of microfluidic structures. With PDMS, chip fabrication did not require etching chemicals anymore, but could simply be conducted with soft lithographic steps. Furthermore, PDMS permits a broad variety of analysis, since PDMS is translucent for 350 nm -1000 nm light[67]. Finally, its chemical structure allows oxygen transfer from the surrounding into the chip, which is necessary for many biological applications[68]. In the upcoming years different chip geometries were invented to produce monodisperse droplets and the addition of surfactants increased the stability an integrity of droplets[47,48]. With the production of stable and monodisperse droplets, the reactions within each produced droplet got comparable. Since the first droplet production chips could already produce droplets at 20-80 Hz[46] so that a lot of data could be generated within a short amount of time. The integration of valves[69] and electrodes[67-69] into the microfluidics chips allowed to control single droplets within a great population and manipulate them on demand. This expanded the possible applications. In combination with a laser-based measurement this led to the establishment of the first fluorescent activated droplet sorter (FADS)[73]. Baret *et al*. showed that the established sorting system could distinguish two populations of *E. coli* cells expressing either an active or an inactive variant of β-galactosidase. For that, they encapsulated the *E. coli* cells with a fluorogenic substrate, which released a fluorescent moiety up on conversion. This fluorescent signal was measured and the intensity of the emitting light signal was used as an argument to

decide if a droplet holds an active variant or not. Following up on this proof of concept experiment, Agresti et al. showed that the FADS can also be applied for directed evolution enzyme libraries[74]. In only 10 hours, they screened 100 million droplets and identified a horseradish peroxidase with 10 times higher catalytic efficiency. At this time, this meant a decrease in screening time by a factor of 1000 and a 1 million-fold reduction in cost compared to state-of-the-art robotic systems. Since then, several enzymes have been successfully engineered with FADS systems[75–77]. Most of them belong to the enzymatic group of hydrolases, as FADS assays rely on the secession of a fluorescent moiety from the native substrate. One major challenge of enzyme engineering within droplet-based microfluidics is to expand the application of the systems also to different enzyme classes. In this work a droplet-based microfluidic screening system, which relies on an absorbance readout, was established like reported before[78]. In comparison to the FADS system, it can be used to engineer dehydrogenases (chapter 1.1.2), which are expressed in the cyto- or periplasm and not secreted[79] nor presented on the cells surface[74]. Therefore, it is a valuable addition to the existing microfluidic screening technologies and expands the application of droplet-based microfluidics. The upcoming chapters will provide a brief overview about how the screening system is working and promising enzymes can be detected.

1.3.2 Droplets as biomimetic compartments

The production of droplets is a key step in droplet-based microfluidics. It allows the separation of single cells or molecules into independent reaction chambers and defines the reaction volume, where they can be studied by themselves. This is especially interesting for studying desparities within a population of cells that hold different variants e.g. of an enzyme. In this case, single cells can be encapsulated together with a substrate and assay components and the catalytic efficiency can be monitored for each enzyme variant within a droplet. The enzyme reaction is initiated by lysing the cells. The cells then release the enzymes into the droplet space and start the conversion of the substrate, which is monitored by a colorimetric assay. This allows studying the enzyme activity by comparing the corresponding signal intensities and eventually finding positions within the genetic code that can improve the desired attributes of an enzyme. Compared to already established techniques, e.g. assays in microtiter plates, there are several advantages of droplet-based microfluidics. At first, the sample volume is around 5 - 6 orders of magnitude lower[74]. Second, the high production and sorting rate of droplets reaching kilohertz speed allows investigating a gigantic number of reactions. This reduces the time and also the resources that are needed to investigate an enzyme library[80]. This ultrahigh-throughput (>100.000 components/day)[39] of droplet-based microfluidics also increases the overall capacities for investigating libraries and raises the chances to find synergetic effects between different mutations[81,82]. While traditional robotic liquid handling systems are restricted to several thousand enzyme variants per day, droplet-based microfluidics can investigate several million droplets within one day. Third, the droplets act as a second cell wall and can be used to retain the link between the improved enzyme variant and the corresponding DNA. This phenotype-genotype connection later enables the recovery of the DNA from the sorted droplet population and allows identifying interesting positions within the genetic sequence of the enzyme. To keep reactions within each droplet comparable, it is mandatory to produce monodisperse droplets. This means that all droplets have ideally the same volume and concentrations of substrate and assay components within the aqueous phase. For that, it is essential to have good control over the droplet production and the flow of the dispersed and continuous phase. Since the introduction of PDMS for manufacturing microfluidic chips, a lot of work has been contributed to produce monodisperse droplets with different chip designs. The next chapter will therefore introduce different chip geometries to form monodisperse droplets.

1.3.2.1 Chip geometries for production of monodisperse droplets

Microfluidic droplets are usually formed within continuous fluid phases. The most prominent way is to bring two immiscible fluids together in a microfluidic chip. By introducing one immiscible fluid (dispersed phase) into another (continuous phase) droplets are formed. The fluids are pumped into the chip either with syringe pumps[83] or pressure driven systems[84] applying a constant pressure. The two main parameters to characterize monodisperse droplets are: First, the coefficient of variation (CV), which defines the standard deviation of droplet diameters within a population. Second, the production rate, which decides how fast droplets can be produced and ultimately how much time difference between the first and the last droplet, can be expected. The chip designs for monodisperse droplet production can be divided into four groups: Cross-flow junctions (I), co-flow designs (II), flow focusing junctions (III) and step emulsification (IV) (Figure 3).

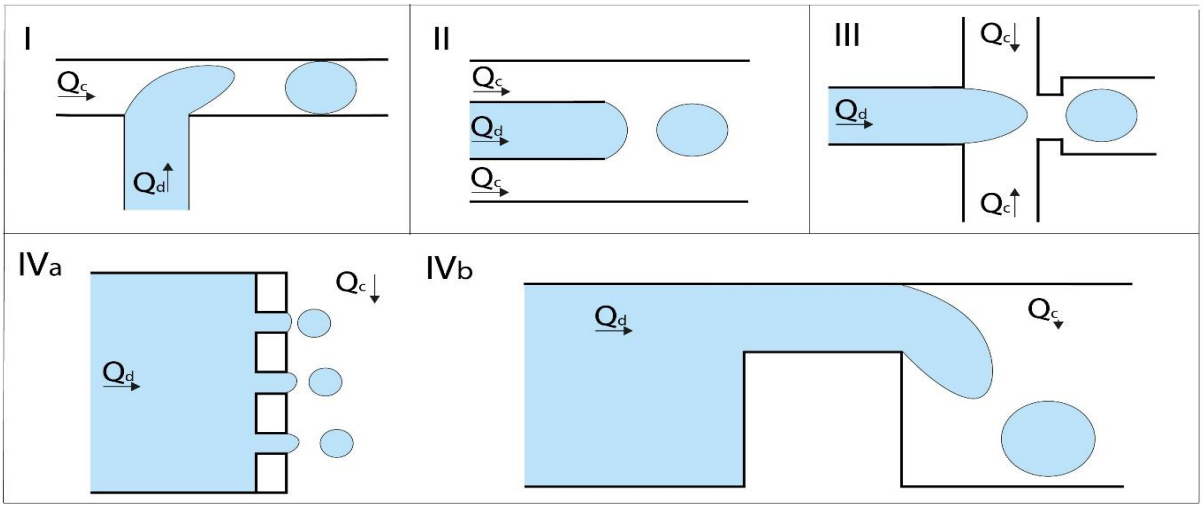


Figure 3: Schematic of different channel geometries for the production of monodisperse droplets: (I) Cross-flow junction (II) Co-flow junction (III) Flow-focusing junction and (IV) Step emulsification a) top view b) side view. Qc indicates the flow direction of the continuous phase and Qd the flow direction of the dispersed phase.

(I) First introduced by Thorsen et al.[46] the two immiscible fluids meet orthogonally in a 90° angle, which also gave this cross-flow design the byname “T-junction”. The oil phase applies shear forces on the surface of the aqueous phase, which compete with interfacial forces that impede the elongation of the fluid. For droplet production, the continuous oil phase is set in that way, that the shear forces applied by the continuous phase are higher than the interfacial forces between continuous and dispersed phase. This leads to the creation of a liquid neck within the dispersed phase, which eventually promotes the droplet formation. The droplet size is strongly depending on the channel dimensions of the continuous and dispersed phase, but also on the corresponding flowrates. Droplet production chips featuring a T-junction have been shown to produce monodisperse droplets with a very low coefficient of variation of about 2%[85]. The main advantage of this junction type is the ease of design and fabrication.

(II) In co-flow designs, the continuous and dispersed phase encounter each other in parallel streams. First introduced by Umbanhowar et al.[86] a tapered capillary holding the dispersed aqueous fluid is directly introduced into the continuous oil phase. Droplets detach from the capillary once the streamwise forces exceed the interfacial forces. The later droplet diameter is strongly depending on the diameter of the capillary and flowrate of the continuous phase. Similar like for cross-flow junctions monodisperse droplets can be produced with a low CV of 3% at kilohertz speed[86]. The main advantage of the co-flow design compared to the cross-flow junction is its ability to produce droplets with a diameter below 10 µm[46]. The droplet diameter can be scaled down to only several hundred nanometers[87], which makes them especially interesting for monitoring reaction of single molecules to increase signal intensity by reducing the droplet volume[88].

(III) For flow-focusing junctions, the continuous phase encounters the dispersed phase also in a 90° angle, like previously described for the cross-flow junction. The main difference is that for flow-focusing junctions the continuous phase approaches the dispersed fluid from two opposing sides, which hydrodynamically focuses the dispersed phase. The dispersed fluid is contracted until the capillary forces cause instability, which eventually leads to the droplet formation. First introduced by Gañán-Calvo for a gas-liquid droplet production[89], Anna *et al*. converted this idea into a chip-based flow-focusing junction for a liquid-liquid microfluidic system[90]. Flow-focusing junctions offer a great variability of droplet sizes, by adjusting the flow of dispersed and continuous phase. They have been used to produce monodisperse droplets at kilohertz speed[91] with a low CV of 4.1 %[92]. This comparably high CV can be tuned down to 0.2% by using pressure driven systems with feedback control[93].

(IV) In contrast to the before mentioned designs (I - III), step emulsification does not rely on the application of shear forces, but on capillary pressure. The dispersed phase is guided from a reservoir through a channel, which is narrowed in x- and z-dimension (step) and then released into an oil reservoir. From the transition between step and oil reservoir the droplets are pinched off by interfacial forces and transported towards the outlet of the chip. Producing monodisperse droplets with step emulsification results in several advantages. First of all, the droplet production is less prone to flow rate or pressure fluctuations allowing the production of monodisperse droplets with 1% CV[94]. The droplet size mainly depends on the channel dimensions and is barely influenced by the fluid properties. This results in the second and main advantage of step emulsification: Since the droplet production is mainly controlled by the flowrate of the dispersed fluid, it is easier to parallelize several step emulsification designs on one and the same chip. Multiple channels with the same properties can be created on one chip

and produce monodisperse droplets in parallel. For example, Xu et al. created a design, which featured 2.000 parallel nozzles. They could reach a production rate of 15 kHz with only 3% CV[95], which shows the great potential of these designs. However, one needs to keep in mind that these designs also have a bigger area of production, which requires an immaculate chip structure and no interference of particles or cell debris during the droplet production process as this might result in polydisperse droplets.

1.3.2.2 Droplet regimes for the production of monodisperse droplets

Droplet generation is always triggered by fluid instabilities. The kinetic energy introduced by syringe pumps or pressure driven systems is partially converted into interfacial energy, which enables the destabilization between the liquid-liquid interface and eventually leads to the discrete droplet shedding of the dispersed phase[96]. However, depending on the forces involved of dispersed and continuous phase, different mechanisms for droplet production can be distinguished referring to different microfluidic regimes. In general, these regimes can be divided in three different groups: squeezing[97], dripping[86] and jetting[98] (Figure 4 A, B and C respectively).

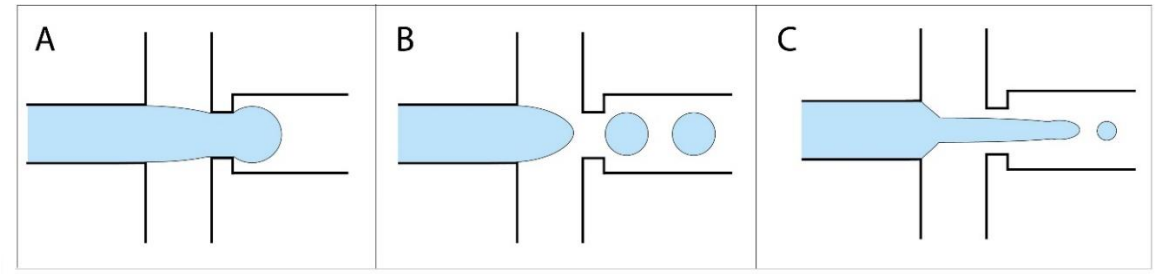


Figure 4: Different mechanisms of droplet production. A) Squeezing of droplets in flow-focusing channel. B) Dripping of droplets in flow-focusing channel. C) Jetting of droplets in flow-focusing channel

The squeezing regime is characterized by droplets, which are bigger than the channel diameter of the continuous phase. The dispersed phase occupies the entire space of the cross section in the junction preventing the elongation of the continuous phase, which in turn leads to an increase in pressure for the continuous phase. This increase in pressure allows the continuous phase to squeeze on the dispersed phase and eventually leads to the droplet formation behind the cross-section area (Figure 4 A). It is worth to mention, that the squeezing regime can just be observed if the width of the dispersed and the continuous channel have roughly an equal ratio. If the channel width of the continuous phase exceeds a ratio of 4:1 (continuous phase : dispersed phase) the dispersed phase is unable to fill the cross section of the junction[99]. The dispersed phase will be pinched off by the occurring shear forces from the continuous phase, which is the typical attribute for the dripping regime. The dripping regime occurs when the shear forces of the continuous phase overcome the interfacial tension between continuous and dispersed phase. For equal channel dimensions, this is usually the case when the flow

rate of the continuous phase is higher than the flowrate of the dispersed phase. Here, the droplets are pinched off within the cross section of the junction (Figure 4 B). The resulting droplets are smaller than the channel diameter of the continuous phase. Finally, the jetting regime can be only observed if both flowrates, the dispersed and continuous phase, are high. The dispersed fluid is squeezed by the continuous fluid within the cross section of the microfluidic junction. In comparison to the squeezing regime, the dispersed phase is not completely filling the cross section of the junction, but leaving space on the side for the continuous liquid to pass by. The result is that the pressure of the continuous phase is not increasing. In the middle of the cross section, the dispersed fluid appears as a stream, but the droplet pinch off can be observed behind the cross section of the junction (Figure 4 C). The resulting droplets are much smaller than the channel width of the continuous phase. From all three regimes, the jetting regime allows the highest production rate of droplets, but also needs to be carefully controlled, as pressure fluctuations might lead to a polydisperse population of droplets.

1.3.3 Droplet storage and manipulation

After droplet formation, cells and substrates are co-encapsulated and the product formation within the droplets can start. For fluorescent assays, detection of fluorescein product formation as low as 4 nM in 2 pL droplets is possible[75]. This translates into the conversion of 3.000 product molecules. Depending on the selected enzyme, this conversion can be achieved within seconds and the droplets can be immediately measured and sorted directly after production. However, if the desired enzyme has lower catalytic efficiency towards the selected substrate or the cell wall has to be lysed to start the assay reaction[40], it might be necessary to incubate the produced droplets or add an additional component.

Incubation

By incubation the time for product conversion prior to the measurement can be increased from some seconds to several days[75]. This incubation step can be performed either on- or offside the microfluidic chip. On-chip incubation can be achieved by expanding the microfluidic channel structure and add a delay line before the sorting junction[74]. If the delay line consists of straight channels, which have a small width the droplet order and also the spacing can be kept[100]. However, this approach is limited to a few minutes of incubation, due to the available space for the channel structure on the chip and the high flowrates for droplet production. To increase the incubation time, one possible approach is to widen the channel width. This allows the oil to pass by the droplets and therefore reduces the velocity of the droplets within the delay line. This effect can be increased by extracting the spacing oil at the beginning of the delay line and space the droplet population right before sorting. The main disadvantage of this

approach is caused by the dispersion effect. Droplets that are in the middle of the channel move much faster than droplets, which are close to the channel wall. The result of this approach is a great variety of incubation times of up to 90 %[101]. In order to reduce this effect, Frenz *et al.*[101] added constrictions to their channel design. Every three centimeters the channel width was reduced from 1000 µm to 75 µm (diameter of one droplet). This allowed the droplets to be mixed repeatedly and prevented droplets to stay in the same faster or slower flow line. With this adjustment, the group was able to reduce variety of incubation times from 90% down to 10%. They showed that droplets could be incubated up to one hour on the chip. To further extend the incubation time several methods for off-chip incubation were reported. For that droplets will leave the chip for a defined time and are reinjected after the incubation. An easy method to extend the delay time is the incubation of droplets within a tubing[102,103]. After droplet production the droplets are guided offside the chip into a tubing where they are continuously pushed towards the end of the tubing. The incubation time can be determined by the flow rate and the length of the tubing. This approach is easy to set up, but limited by the pressure, which is increasing with the length of the tubing. Another method to incubate the droplets offside the chip is the storage within a syringe[75]. Therefore, droplets are guided offside the chip via a short tubing and collected in a tube. Once the production process is finished, all droplets within the tube are collected by withdrawing them into a syringe partially filled with oil. To avoid evaporation, another layer of mineral oil is withdrawn into the syringe. Like this the droplets are in the middle of both oil layers. A novel approach was reported by Neun *et al*.[104]. Here the droplets are collected in a droplet chamber. The chamber has two openings, one on top of the chamber and one at the bottom. Both openings are connected with tubings. For the droplet production, the upper tubing is taken to collect the produced droplets from the droplet production chip, while the lower opening allows the release of excess oil. Due to Buoyancy forces, the droplets gather on top of the droplet chamber and are not flushed out. After finishing the droplet production, a syringe filled with oil is connected to the lower opening prior to incubation. For long-term incubations at higher temperatures, the oil is partly released from the chamber and the upper tubing part is sealed to avoid evaporation. After storage of droplets for a defined time, the upper opening is opened and connected to a sorting chip. By applying pressure through a syringe pump, the droplets are pumped into the chip where they can be sorted.

Addition of reagents

Not for all assays, it is desirable to start the product formation immediately after droplet production. In some cases, it is favorable to delay the assay reaction to avoid background and increase the signal to noise ratio. Therefore, in droplet-based microfluidics it is necessary to add certain components of the assay reaction after droplet formation has been performed. One way to achieve this is the passive fusion of two droplets. Here, the chip structure is designed

in a way that two consecutive droplets are forced to collide with each other. For this purpose, Niu *et al*.[105] designed a narrow pillar structure. The pillars are assembled in a way that once the first droplet enters the structure, it is slowed down and the second droplet collides into it and both droplet fuse. Another way of passive fusion was shown by Mazutis *et. al.*[106]. In this paper, two droplet populations with different sizes were fused by a zig-zag channel structure. The smaller droplet population moves faster through the droplet channel and therefore collides with the bigger droplet population. By optimization of the surfactant concentration for the two droplet populations and the carrier oil, the group showed that it is possible to force a one-to-one droplet fusion independently of the ratio of both droplet populations. The key to the success of this concept is that the big droplet population is produced with a very low surfactant concentration (0.1%), while the small droplet population has a very high surfactant concentration (2.8%). After fusion of both droplet species, the surfactant concentration within the big fused droplet increases the surface tension. This prevents the fusion of more small droplets. However, this also describes the main bottleneck of passive fusion. It will only succeed with droplets that are not stabilized in advance by high surfactant concentrations. However, for long-term incubation of droplets it is necessary to reduce the space in between the droplets. If the droplets are not stabilized, unwanted droplet fusion (also called merging) or leakage (diffusion of substances from the droplet into the oil phase or into other droplets) can occur. Therefore, active droplet fusion is a valuable asset for droplet-based microfluidics. It allows droplet fusion, even at high surfactant concentrations. First introduced by Priest *et al*.[107], this technique relies on the short application of an electric field between two electrodes. By application of the electric field, a dynamic instability of the interface between two droplets is induced. By tuning the parameters of the electric field (intensity, frequency and time), controlled droplet coalescence can be achieved. Brouzes *et al.*[108] applied active droplet fusion for their cytotoxicity test with mammalian cells. Since the cytotoxin had no fluorescent readout, they used droplet fusion to code the cytotoxin concentration with a fluorescent dye. With increasing concentrations of cytotoxic substances, they also increased the concentration of fluorescent dye in the first droplet population. This allowed them to make the concentration of cytotoxic substances visible in their readouts. The second droplet population contained mammalian cells stained with another fluorescent dye indicating the viability of the cells. By active droplet fusion, they could draw direct conclusions on the influence of the cytotoxin concentration towards the cell viability. However, the main bottleneck of active droplet fusion is the pairing of different droplet populations. Usually, the droplet populations are produced with different sizes to guarantee droplet pairing prior to the fusion event. However, using droplets of different sizes makes it hard to control a one-to-one pairing of droplets. Therefore, it might occur that a big droplet will not pair with a small droplet or that a big droplet will pair with two small droplets. This leads to different starting concentrations for the assay within the

droplet and might shift the results of the product formation. To avoid big differences in assay component concentrations, Abate *et al*. introduced picoinjection to the microfluidic field[109]. Here, a channel filled with aqueous solution is placed orthogonally to the main microfluidic channel. On the opposite side of the channel, a set of electrodes creates an electrical field, which destabilizes all bypassing droplets upon activation and allows fusion of the aqueous solution into the bypassing droplets. The injected volume can be controlled by the spacing between the droplets and the pump rate of the aqueous solution. Beneyton *et al*. used picoinjection for enzyme engineering of several recombinant proteins with the yeast *Yarrowia lipolytica*[50]. After droplet formation with single cells, all droplets were incubated for 16 hours offside the chip. This allowed the yeast cells to grow, divide and secrete the recombinant protein. After incubation, the droplets were reinjected into a microfluidic chip and the substrate solution was added by picoinjection into each bypassing droplet. Because several cells secrete the same protein variant, the protein concentration within each occupied droplet was increased and the readout of the product formation could immediately be performed. With their approach, they could improve the enzyme activity of a recombinant xylanase by 4.7 fold.

1.3.4 Detection of product formation in microfluidic droplets

In the past, most of the work that has been contributed to measure product formation in droplet-based microfluidics, was performed with fluorescent activated droplet sorters (FADS[73]). Here, the concentration of a fluorogenic substrate within a droplet[77,103,110] or cell[111] is measured by excitation with a laser light of a defined wavelength. The emitted light is then separated with the help of dichromatic mirrors and band pass filters and allows detection of multiple fluorescent dyes. With a photo multiplier tube (PMT) the signal is amplified, allowing the detection of product concentrations as low as 4 nM[75]. However, the detection with FADS is mainly limited by the availability of substrates. The substrates mostly consist of a chemically synthesized combination of the native substrate and a fluorogenic moiety. The enzyme later cleaves that connection and thereby enables the fluorescent readout. Therefore, hydrolases were the enzyme class, which was mainly engineered[73,77,103]. Yet, it is interesting to expand the capabilities of microfluidic screening systems and allow screening of many different enzyme classes. Over the last decade, droplet-based microfluidics experienced a great increase in exploring new methods for determining product formation. For example Maceiczyk *et al*. used photothermal spectroscopy to quantify the concentration of the absorbance dye Erythrosin B within microfluidic droplets[112]. Moreover, the group could distinguish between droplets that were unoccupied or occupied with water-soluble tetrazolium salt 1 (WST-1) stained cells. Recently, Gielen *et al*. developed a microfluidic screening system, which implemented an absorbance-based readout[78]. The product formation is indirectly quantified by the reduction of the cofactor NAD^+ to NADH. Via the electron coupling agent 1-methoxy-5-methylphenazinium methyl sulfate (mPMS) the colorless WST-1 dye is reduced, while the

cofactor is recycled. The reduction cleaves the tetrazol structure in the middle of the molecule and turns the colorless WST-1 into the orange WST-1 formazan dye (Figure 5).

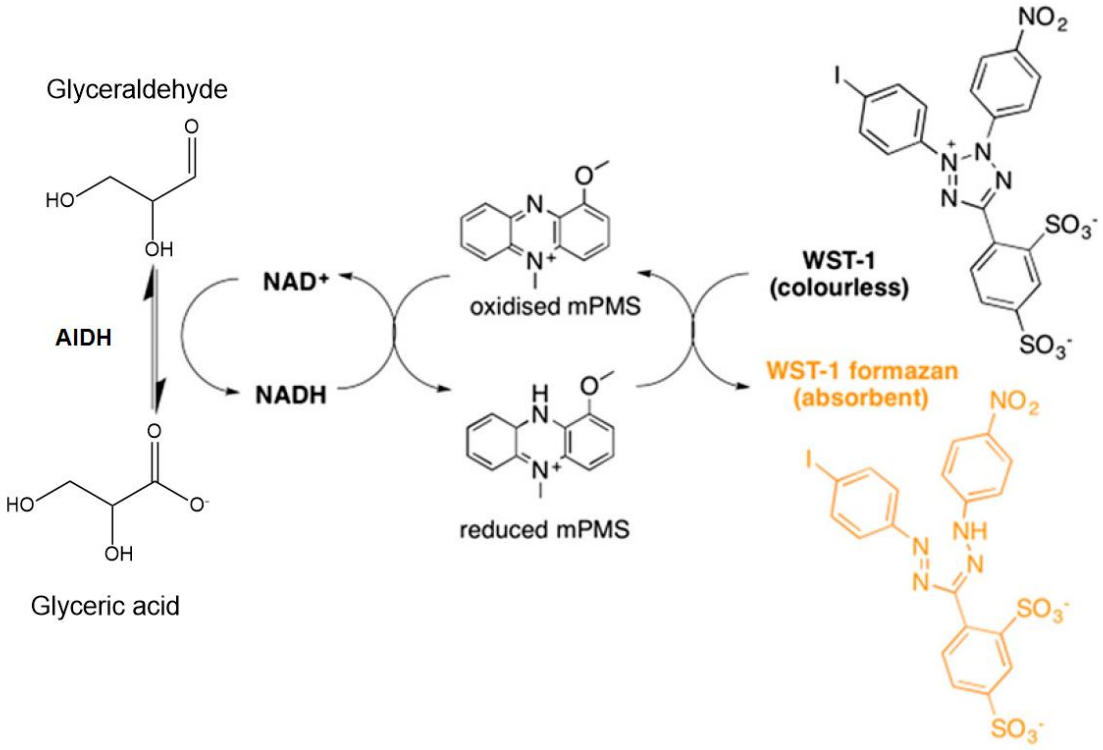


Figure 5: Electron transfer reaction for reduction of WST-1 (colorless) to WST-1 formazan (orange). The substrate glyceraldehyde is converted to glyceric acid with an aldehyde dehydrogenase. Thereby, the cofactor NAD is reduced to NADH. Via the electron coupling agent mPMS, the WST-1 dye is reduced and appears orange. The cofactor NADH is recycled to NAD (Figure adapted from Gielen and Hours et. al 2016).

This allows the indirect quantification of the enzymatic turnover by the cofactor consumption. The expected absorbance signal intensity can be described by the Beer-Lambert law:

$$A = \varepsilon * l * c$$

A= Absorbance
ε= Molar extinction coefficient
l= Length of the lightpath
c= Concentration of product

Compared to NADH (ε_{340nm}= 6220 M^{-1} cm^{-1}), which is used for quantification of many photometrical microtiter plate assays, WST-1 has an approximately six-fold higher extinction coefficient (ε_{433nm} =37000 M^{-1} cm^{-1}), which increases the sensitivity of the absorbance assay tremendously. For the microfluidic system established over the course of this work, the same absorbance-based readout is taken. The quantification of product formation is performed within a narrow channel with a width of 50 µm. The channel squeezes the droplets, which usually have a diameter of around 80 µm and allows normalization of the path length. Orthogonally to the analysis channel, a glass fibre conducts light from a light emitting diode (LED) with a wavelength of 450 nm (Figure 6). The light goes through all bypassing droplets and depending on the product concentration, the light intensity is reduced. The light signal is then captured by

another glass fibre on the other end of the microfluidic channel, where it is directed into a photodetector. The intensity of the light signal is converted into a voltage signal, before it is split in two directions. One direction activates the sorting if a certain threshold is decreased, which will be explained in the following chapter (1.3.5). The other direction enables the measurement of the data by sending the voltage signal to an analog to digital converter and makes the data available for PC analysis with LabVIEW (chapter 3.1.4).

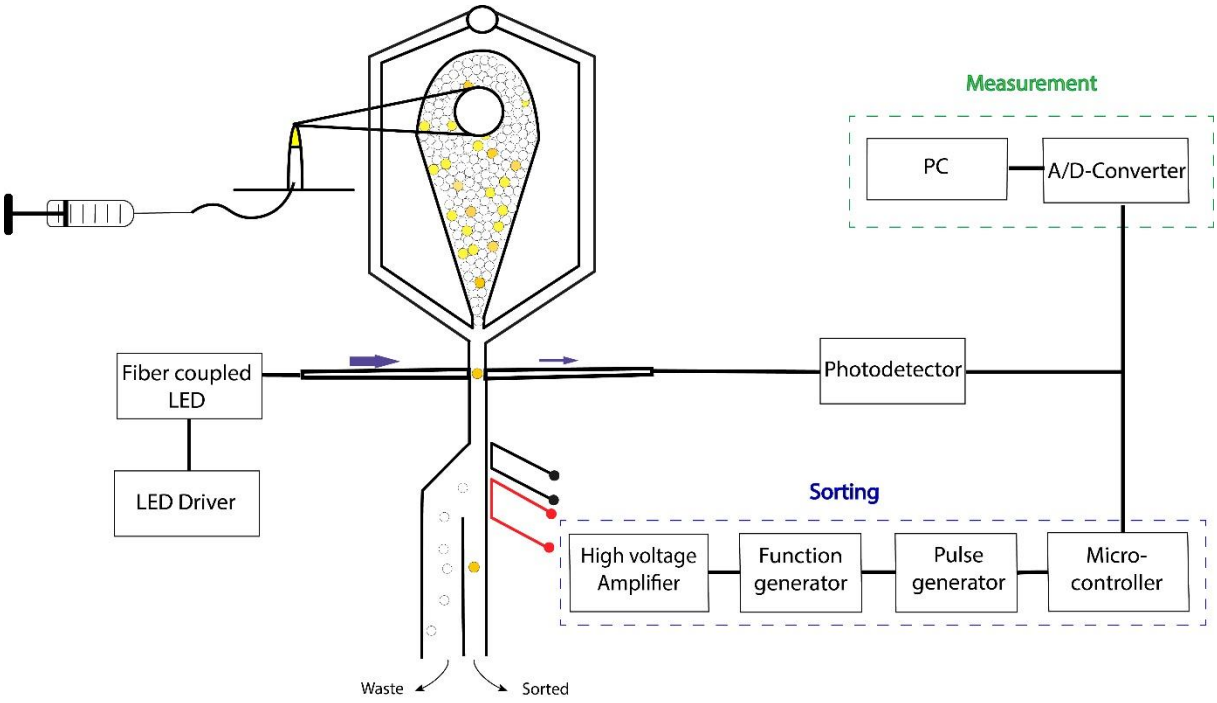


Figure 6: Absorbance activated droplet sorting scheme: Droplets are reinjected into the AADS chip after incubation. Droplets holding an enzyme variant will produce WST-1 formazan dye and therefore appear yellow/orange. Droplets are spaced by an oil junction prior to sorting. Blue light (450nm) from a fiber coupled LED is brought through a glass fiber into the microfluidic chip where it is guided through bypassing droplets. Product formation within the droplet reduces the light signal. A photodetector converts the light into an analog signal, allowing measurement and sorting on demand of all bypassing droplets.

Since absorbance activated droplet sorting (AADS) is based on the reduction of the incoming light signal, amplification with PMT's is not possible. Therefore, the minimum detection level is reported to be between 10 µM[78] and 100µm[113] and by that about three orders of magnitude higher than for fluorescent assays. The reduced sensitivity leads to increased incubation times, as the encapsulated enzymes need longer to convert enough product for the detection limit. However, in contrast to most fluorescent assays, the capacity of the WST-1 assay allows detection of up to 5 mM of product, which makes it extremely valuable for long-term experiments and highly active enzymes.

1.3.5 Droplet sorting

The determination of product formation within microfluidic droplets allows visualizing the difference in enzymatic activity for each enzyme variant within a droplet (chapter 1.2.4). However, for retrieving the genetic information, corresponding to the improved enzyme variant within a droplet, it either is necessary to separate the droplets showing improved enzyme activity from the great majority of droplets holding an enzyme with no improved activity or are vacant. There have been several ways reported to divide these two populations. In general, these methods can be divided into passive (i) and active (ii) droplet sorting.

(i) For passive droplet sorting no extra equipment is needed, as the sorting relies on the hydrodynamic properties of the droplet itself. For example, Chabert and Viovy[114] passively sorted droplets by utilizing the fact that occupied droplets show an increased diameter compared to vacant droplets with their pinjet production method. By optimizing the flows of aqueous and oil phase within their chip they were able to geometrically sort the droplets according to their diameter and by that separate occupied and vacant droplets. Another example can be given by Joensson *et al.*[115]. who exploited the fact, that droplets holding yeast cells are shrinking after almost eleven hours of incubation due to the medium consumption of the cells. The droplets were then reinjected after incubation into a chip with a pillar array and separated according to their size by deterministic lateral displacement. The described methods are inexpensive as they rely on the geometric abilities of the chip, but the parameters for the geometry are strongly depending on the droplet size. Thus, the efficacy is strongly decreased by different droplet sizes and cannot be simply applied for other applications without testing and adaption of the parameters. Furthermore, these sorting methods are vulnerable to droplet merging, which occurs especially for longer incubation periods and elevated temperatures. Finally, it is complicated to link reactions to the hydrodynamic properties of a droplet, which reduces the possible applications.

(ii) Active droplet sorting relies on a measurement, which in turn triggers a positive or negative effect, changing the flow direction of the droplet. In order to keep deviations for droplet movement neglectable, sorting junctions are usually implemented directly after the point of measurement. The channel structure of the chip splits up into at least two channels to divide the droplet population in droplets, which surpass the threshold (later also called “hits”) and droplets, which do not meet the expectations (later also called “waste”). Usually, the waste channel has a wider or shorter geometric structure in order to lower the hydrodynamic resistance. This ensures that all droplets are guided into the waste channel if no trigger event is initiated. Active droplet sorting has been conducted by using magnetic fields[116,117], pneumatic controls relying on hydrodynamic pressure[118] and valves[119] or by thermocapillary forces[120]. However, all the above-mentioned methods have clear limitations either in terms of application

or in terms of throughput. A more commonly used method for sorting in droplet microfluidics is acoustophoresis. It relies on the generation of acoustic waves. Piezoelectric actuators are used to create either surface acoustic waves (SAW)[121] or bulk acoustic waves(BAW)[122]. Both wave types are produced in megahertz frequency and deflect the droplet of interest into the sorting channel. Compared to the above-mentioned methods, the frequency of the acoustic waves can be changed to enable sorting of different droplet sizes. Additionally, the throughput for the sorting can be drastically increased. Schmid *et al*.[123] integrated a SAW sorter in a FADS and reached sorting rates of up to 3,000 Hz. The main bottleneck of sorting with acoustic waves is the complicated and costly manufacturing of the chips. Therefore, the most popular way of sorting microfluidic droplets is with dielectrophoresis. First introduced by Ahn *et al*.[71] it relies on the short creation of an electric field between two electrodes. Upon activation, the droplets are dragged in the direction of the electric field. The electric field is positioned in a way, that it deflects the droplets from the waste channel, with less hydrodynamic resistance, into the sorting channel. The sorted droplets are then collected in a separate tube and the DNA of the hits can be identified after hit recovery (chapter 1.2.6). Compared to acoustofluidic devices, the manufacturing process is simplified. The sorting electrodes can be added by either pressing molten alloy[124] in the designated channels or filling them with high concentrated salt water solution[125]. Connection with wires to a function generator and a high voltage amplifier allows precise control of frequency and intensity of the electric field on demand. This ensures a high control over droplets of different sizes and different velocities. However, the size of the droplets is also a limiting factor for dielectrophoretic sorting in terms of throughput. For small droplets (~25 µm) sorting rates of up to 30 kHz[42] have been reported. In this work, an absorbance-based detection was used, which required bigger droplets (70 - 80µm) to reach enough signal intensity in order to distinguish between the different droplet populations. Working with bigger droplets makes it necessary to increase the time of activation for the electric field to direct the droplets into the sorting channel. As a result the sorting throughput of the absorbance based system is one to two orders of magnitude lower at 100 - 300 Hz[78]. Recently, the same group showed that with several improvement steps, the capability for this sorting system can be raised to 1 kHz[113]. Another approach to improve the sorting speed of the AADS was given by Richter *et al*. which replaced the dielectropheretic sorting unit and used surface acoustic waves[126].

1.3.6 Recovery of genetic information from microfluidic droplets

After the sorting, the droplet hits are separated from the rest of the population for further analysis. The population of sorted droplets is directed along a separate channel offside the chip, mostly with a tubing at the end of the sorting channel. Droplets leaving the chip can either be collected altogether in a collection tube[79,127] or spatially separated to avoid mixing of the genetic information. This is especially important if the droplet holds more than one genetical information e.g. heavy and light chain of antibodies. Here, it is mandatory that the droplets are collected separately to retain the heavy and light chain linkage also after breakup of the emulsion. For the separation of droplets after the sorting, several mechanisms have been reported. Droplets can either be placed on different positions on an agar plate[128] or a droplet matrix[129,130], where either the tubing or the collection plate is moved by a motorized system. Every time a droplet is sorted, a new position on the collection plate is approached and allows spatial separation of each sorted droplet. In combination with computer software, the information about the droplet (e.g. signal intensity) and its position can be preserved and used for later analysis. Another way of droplet separation is provided by droplet to digital (D2D) microfluidics. Here, the water-in-oil droplets are transported along an electrode railway. D2D enables several microfluidic operations after the sorting process. The sorted droplets can be merged and mixed[131], they can be split to perform different analysis[132] or they can be simply parked on the chip for observation purposes and later analysis[133]. For most experiments, it is interesting to retrieve the genetic information of the sorted enzyme variants by recovering the DNA from the aqueous phase of the droplet. Therefore, it is necessary to break the emulsion of the droplets either chemically[106] or with a static gun[134] and collect the DNA, which was co-encapsulated with the enzyme. Depending on the purity of the aqueous phase within the droplet, it can be helpful to clean up the aqueous phase and remove RNA and cell debris. For this the aqueous phase can be either washed several times or cleaned up with a plasmid column[135]. In a final step the recovered plasmid DNA can either be directly used to transform competent *E. coli* cells[51] or amplified first[56] to increase the amount of DNA. The transformed *E. coli* cells can then be taken for rescreening and eventually sent for sequencing to retrieve the genetic modifications.

1.4 Aims of this work

1.4.1 Establishment of a microfluidic sorting system

Prior to this work, screening of enzyme libraries (at the Chair of Chemistry of Biogenic Resources) were conducted either by hand in microtiter plates or with the support of a liquid handling station. As indicated in chapter 1.3.2 these microtiter plates approaches are limited to sort only several thousand variants per week. To enhance screening capabilities and explore larger sequence spaces of the corresponding enzymes, this work aims to establish a microfluidic sorting system. The system should be capable of sorting several million enzyme variants per day, thereby increasing the likelihood of identifying variants with improved performance. To build such a system, microfluidic chips for droplet production, measurement, and sorting need to be designed, fabricated and tested. Additionally, the software must be developed to capture measurement signals and trigger droplet sorting on demand. Both hardware and software will then need to be fine-tuned to ensure a high level of control throughout the sorting process.

1.4.2 Proof of Principle: Discover improved variants from different enzyme libraries

In the second part of the thesis, the established system has to be verified for different enzyme libraries. In a first step, the functionality of the system should be proven. For that a small library has to be created and sorted with the newly established microfluidic sorting system. In parallel, the same library has to be sorted in microtiter plates and the results of both sortings should be compared to validate the functionality of the system. If the validation of the sorting system is successfull, bigger libraries (which are beyond the capacities of microtiter plate screenings) should be created and analyzed for improved abilities. To expand the applicability of the sorting system some of the investigated enzymes should be derived from thermophiles. For that, the stability of the whole workflow has to be adapted to elevated temperatures and should be stable for longer time periods. For an efficient procedure, a prescreening method should be established, which reduces the number of experiments necessary to find the optimal conditions (temperature, incubation time) for each produced library. Finally, it should be tested if the microfluidic sorting system can be used to screen for enzymes with improved thermostability.

2 Material and Methods

2.1 Materials

2.1.1 Devices and consumables

List of devices

Device	Manufacturer
Autoclave Varioklav 135S	Thermo Fischer Scientific
Camera Axiocam Erc5s	Zeiss
Camera Mako U29B	Allied Vision
Centrifuges:	
Sorvall RC 6+	Thermo Fischer Scientific
Rotor F10-6x500y	Piramoon Technologies Inc
Rotor F9-4x1000y	Piramoon Technologies Inc
Rotor SH-3000	Thermo Fischer Scientific
Rotor SS-34	Thermo Fischer Scientific
Mini Centrifuge	Fisherbrand
Fresco 21 Centrifuge	Thermo Fischer Scientific
Freezer -20 °C	Liebherr-Hausgeräte
Freezer -80 °C Forma 906 -86°C ULT	Thermo Fischer Scientific
Gas-tight Syringes (100µl, 1000µl, 2500µl)	SGE
Gel documentation: Gel iX Imager	Intas Science Imaging Instruments GmbH
Gel electrophoresis:	
Agarose electrophoresis apparatus Mini-sub cell GT system	Bio-Rad Laboratories GmbH Munich
SDS electrophoresis chambers	Bio-Rad Laboratories GmbH Munich
Mini-PROTEAN®-Tetra Cell	Bio-Rad Laboratories GmbH Munich
Mini-Protean®3Multi-Casting Chamber	Bio-Rad Laboratories GmbH Munich
Power supply PowerPacTM Basic	Bio-Rad Laboratories GmbH Munich
Heating oven Function Line T12	Thermo Fischer Scientific
High-pressure homogenizer Cell Disruption System Basic Z	Constant Systems
Homogenizer Ultra Turrax T18 basic	IKA-Werke GmbH & C. KG
Incubation cabinets:	
Climate chamber KBF 240 E5.1/C	BINDER GmbH, Tuttlingen
Incubator Function Line B12	Thermo Fischer Scientific, Heraeus
Incubation shakers:	
HAT Minitron	Infors AG
MaxQ 2000	Thermo Fischer Scientific
TiMix 5 control	Edmund Bühler GmbH
Liquid chromatography:	
ÄKTA™ purifier equipped with	GE Healthcare Europe GmbH
Pump P-900	GE Healthcare Europe GmbH
Sample pump UP-960	GE Healthcare Europe GmbH
Control unit UPC-900	GE Healthcare Europe GmbH
Columns:	

HisTrap FastFlow 5 mL	GE Healthcare Europe GmbH
HiTrap Desalting HiPrep 26/10 Desalting	GE Healthcare Europe GmbH
Magnetic stirrer:	
MR 3001 K	Heidolph Instruments GmbH & Co.KG
VMS-C7	VWR International GmbH
Variomag Telesystem	Thermo Fischer Scientific
Microfluidic System Equipment:	
Breadboard 400 Pins	Qita
Fiber-coupled LED M455F1	Thorlabs
Function Generator TG2000	Aim TTI
High Voltage Amplifier Model 601C	TREK
LED Driver LEDD1B	Thorlabs
Microcontroller Arduino Due	Arduino
Multifunction I/O Device USB 6009	National Instruments
Photodetector PDA36A	Thorlabs
Pulse Generator TGP110	Aim TTI
Microliter pipets, electric:	
Transferpette® S -8 electronic 10-200 µL	BRAND GmbH & Co. KG
Research pro 8x 1200 µL	Eppendorf AG, Hamburg
Research pro 12x 300 µL	Eppendorf AG, Hamburg
Microwave MH 25 ED	ECG
Microliter pipets, manual:	
Transferpette® S : 0.1-10000 µL	BRAND GmbH & Co. KG
Microscope	
Stemi 508	Carl Zeiss AG
Axio LabA.1	Carl Zeiss AG
Axio ObserverZ.1	Carl Zeiss AG
PCR equipment:	
MJ MiniTM Personal Thermo Cycler	Bio-Rad Laboratories GmbH
MyCyclerTM Thermal Cycler	Bio-Rad Laboratories GmbH
CFX96 Touch	Bio-Rad Laboratories GmbH
Ultrapure water system PURELAB Classic	ELGA LabWater, Celle
Ultrasonic Cleaner	VWR International GmbH
Ultrasonic homogeniser UIS250v	Hielscher Ultrasonics GmbH
Sonotrode VialTweeter	
Sonotrode LS24d10	
pH meter and electrodes	
FiveGoTM	Mettler-Toledo GmbH
FiveEasyTM	Mettler-Toledo GmbH
InLab® Expert Pro pH 0-14 ; 0-100 °C	Mettler-Toledo GmbH
InLab® Micro Pro pH 0-14 ; 0-100 °C	Mettler-Toledo GmbH
Pressure reducer Oxyclass 3 (10bar)	Kayser
Refridgerator +4 °C (KG39EVL4A)	Siemens
Rocking platform	VWR International GmbH
Scales	
Microscales PioneerTM	Ohaus Europe GmbH
TE6101	Sartorius AG
TE1502S	Sartorius AG

Syringe pump: Flow-Start Evo	Future Chemistry
Thermoblock Tmix	Analytik Jena Aktiengesellschaft
UV-Vis spectrophotometer	
Multiskan Spectrum	Thermo Fischer Scientific
Varioskan	Thermo Fischer Scientific
Nanophotometer P300	Implen GmbH
Vortex Genie 2	Scientific Industries Inc
Waterbath ED-33	JULABO Labortechnik GmbH

List of consumables

Consumable	Manufacturer
96-well plate (polystrol, F-ground)	Greiner Bio-One GmbH
96-well plate (polystrol, U-ground)	
Biopsy Punch (1mm)	Kai Medical
Fine Bore Polythene Tubing (ID=0.38, OD=1.08)	Smith Medical
Fotomask	Zitzmann GmbH
Jump wire cable (male to male)	Qita
Resistors 10kOhm	Yageo corp.
SMA to SMA Fiber Patch Cable (50µm)	Thorlabs
Disposable Scalpels Cutfix (Bladetype: 15)	B. Braun Melsungen AG

2.1.2 Software and Databases

Product	Manufacturer	Application
Advanced Plotting Toolkit	Heliosphere Research	Add-on Labview (2D-Histogramm)
Arduino	Arduino	Programming Microcontroller
Chem Draw	Cambridge Soft (Cambridge)	Drawing chemical structures
DraftSight 2017 x64	Dassault Systèmes	Microfluidic Chip Design
ImageJ win64 1.51w	Wayne Rasband (1997)	Droplet picture analysis
MATLAB R2017A	The MathWorks, Inc.	Statistical Analysis
NI LabVIEW 2017	National Instruments	Microfluidic Statistics
Plasma Ant	Plasma technology GmbH	Plasma oven
Snapgene v2.3.2	GSL Biotech LLC	DNA Sequence Visualization
Vimba Viewer v3.1	Allied Vision	Camera Microscopy
ZEN 2.3 (blue edition)	Carl Zeiss Microscopy GmbH	Camera Microscopy

2.1.3 Chemicals

All common chemicals used in this study are presented below. List adapted from thesis Samuel Sutiono (2020).

Chemicals	Manufacturer	Catalog number
BSA (Bovine Serum Albumin)	Amresco	332
Polymixin-B-sulfate	Applichem	A0890
Acrylamide/Bis-solution 30 % (37.5:1)	Carl Roth Gmbh & Co. Kg	3029
Agar	Carl Roth Gmbh & Co. Kg	5210
$CaCl_2.2H_2O$	Carl Roth Gmbh & Co. Kg	52395
Carbenicillin	Carl Roth Gmbh & Co. Kg	6344.3

cNADP (carba nicotinamide adenine dinucleotide phosphate)	Roche Diagnostics GmbH	
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth Gmbh & Co. Kg	8040
Glycerol 99.5%	Carl Roth Gmbh & Co. Kg	3783
H_2SO_4 96 w/v%	Carl Roth Gmbh & Co. Kg	4623
HEPES (4-(2-hydroxyethyl)-1-piperazineethane sulfonic acid)	Carl Roth Gmbh & Co. Kg	HN78
IPTG (Isopropyl β-d-1-thiogalactopyranoside)	Carl Roth Gmbh & Co. Kg	CN08
K_2HPO_4	Carl Roth Gmbh & Co. Kg	P749
Kanamycin	Carl Roth Gmbh & Co. Kg	T832
KH_2PO_4	Carl Roth Gmbh & Co. Kg	3904
$MgCl_2.6H_2O$	Carl Roth Gmbh & Co. Kg	21891
Na_2PO_4	Carl Roth Gmbh & Co. Kg	P030
NaCl	Carl Roth Gmbh & Co. Kg	P029
NADH (Nicotinamide Adenine Dinucleotide reduced)	Carl Roth Gmbh & Co. Kg	AE12
Peptone	Carl Roth Gmbh & Co. Kg	8952
Yeast extract	Carl Roth Gmbh & Co. Kg	2363
Novec 7500	Iolitec GmbH	fl-0004-hp
APS (Ammonium Persulfate)	Merck Millipore	1.01201
Bromophenol blue	Merck Millipore	1.08122
HCl 37 w/v%	Merck Millipore	30721
β-mercaptoethanol	Merck Millipore	8.0574
Agarose	SERVA Electrophoresis GmbH	11404
Coomassie Brilliant Blue G250	SERVA Electrophoresis GmbH	17524
SDS (Sodium Dodecyl Sulfate)	SERVA Electrophoresis GmbH	20760
α-D-glucose monohydrate	SERVA Electrophoresis GmbH	22720
$(NH_4)HCO_3$	Sigma Aldrich	9830
D-(+)-Glyceraldehyde	Sigma Aldrich	49800
TRIS (Tris(hydroxymethyl)aminomethane)	Sigma Aldrich	T1503
Pico-Surf™ (5% (w/w) in Novec™ 7500)	Sphere Fluidics	C022
Ethanol absolute	VWR Chemicals	20821.321

2.1.4 Cultivation media

For *Escherichia coli* (*E. coli*) cultivation either LB medium (lysogeny broth) or TB medium (terrific broth) were used. Both media were autoclaved at 121 °C and 2 bar for 20 min and stored at room temperature. Solid media for preparation of agar plates were obtained by adding agar-agar (1.5 % (w/v)) before autoclaving. For gene expression, autoinduction medium was prepared. Heat stable components were autoclaved (121 °C, 2bar, 20 min), while heat labile components were sterile filtered (0.2 µm filter). All components were mixed after autoclavation and antibiotics were added if necessary.

	Component	Ingredients	Concentration (w/v) in Component / in final solution
LB medium		Yeast extract	0.5 %
		Sodium cloride	1.0 %
		Tryptone	1.0 %
TB medium		Yeast extract	2.4%
		Peptone	1.2 %
		Dipotassium phosphate	1.25 %
		Monopotassium phosphate	0.25 %
Autoinduction medium	**ZY -medium**	Yeast extract	0.5 %
		Tryptone	1.0 %
	50x5052-solution	Glycerin	25 % (0.50 %)
		D-glucose monohydrate	2.61 % (0.05 %)
		Lactose monohydrate	10.47 % (0.21 %)
	20xNPS-solution	Ammoniumsulfate	6.6% (0.33 %)
		Na_2HPO_4	13.6 % (0.68 %)
		KH_2PO_4	14.2 % (0.71 %)
	1M Magnesium sulfate		1mM

All listed Antibiotics were dissolved in water and sterile filtered (0.2µm filter) prior to storage at -20 °C. Stock solutions were pipetted into the medium to reach final concentration.

Solution	Stock concentration	Final concentration
Carbenicillin	100 mg/mL	100 µg/mL
Kanamycin	100 mg/mL	100 µg/mL
		50 µg/mL

2.1.5 Plasmids

Plasmid	Description	Reference
pET24a	A vector that encodes for a C-terminal Histag. Operated under T7 promoter with kanamycin as a selection marker	Novagen
pET28a	A vector that encodes for an N-terminal Histag and an optional C-terminal Histag. Operated under T7 promoter with kanamycin as a selection marker.	Novagen
pCBR	A pET28a plasmid with improved Multiple cloning site (for cloning of synthetic gene without tag)	CBR plasmid collection

pET28Nhis_*Bs*GDH	A pET28a plasmid containing a N-terminal histag fused to a gdh mutant gene from *Bacillus subtilis*	CBR plasmid collection
pCBR-GDH2	A pET28a plasmid containing a C-terminal histag fused to a gdh mutant gene from *Saccharolobus solfataricus*	CBR plasmid collection
pET24a_WP_013233459.1_Chis	A pET24a plasmid containing a C-terminal histag fused to a aldh mutant gene from *Herbaspirillum seropedicae z67*	CBR plasmid collection
pET24a_ADU38700.1_Chis	A pET24a plasmid containing a C-terminal histag fused to a aldh mutant gene from *Variovorax paradoxus EPS*	CBR plasmid collection

2.2 Microfluidic Methods

2.2.1 Chip production

The production of microfluidic chips is divided in two steps. First photolithography is used to create a cast template, later called master, defining the later structure and height of the chip. Secondly, soft lithography uses the master created in the first step to produce a negative polydimethylsiloxane (PDMS) stamp of the cast template. Lastly, the crafted stamp is bonded to a substrate, either glass or PDMS, forming the desired channel structure in between. The single steps are described more in detail in the upcoming paragraphs.

2.2.1.1 Photolithography

The master for microfluidic chips were crafted using the design files stated in Gielen *et al*. 2016 (https://openwetware.org/wiki/DropBase). For the droplet production chip, the later application required three inlets with a junction width of 50 µm. Several chip designs were assembled next to each other for increasing the number of PDMS stamps per wafer. Assembly was performed with DraftSight. Fotomasks were produced by Zitzmann GmbH. For the master production, photolithography was used. 3 - 5 ml photoresist EpoCore50 (Micro resist technology GmbH, Berlin) were poured in the centre of a 3-inch silicon wafer. The wafer was then placed in a spin coater and vaccum was applied. The wafer was spun 15 seconds at 300 rpm to spread the photoresist and afterwards 35 seconds at 1000 rpm to reach the desired thickness of 50 µm. The wafer with the photoresist layer was than prebaked on a hot plate at 50 °C for 5 minutes. After that the wafer was baked at 90 °C for 20 minutes. The wafer was cooled down for 10 minutes before using the photomask to avoid sticking effects. The fotomask was adjusted on top of the wafer and UV-light with 100 W was applied for 135 seconds. The wafer was baked another time at 50 °C for 5 minutes and afterwards at 85 °C for 20min. After cooling down for

10 minutes the wafer was developed with mr-DEV 600, removing the photoresist, which was not exposed to the UV-light. The wafer was hard baked a last time, first at 65 °C for 10 minutes, later at 120 °C for 30 minutes.

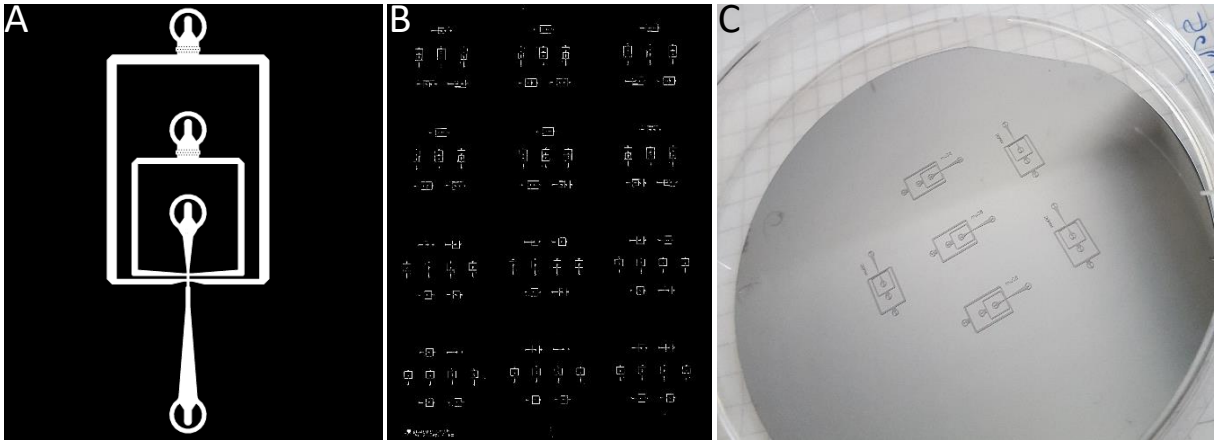


Figure 7: Photolithography steps for microfluidic chip production. (A) Design of droplet production chip taken from Gielen *et al*. 2016. (B) Assembly of several droplet production chips with three (top) or two (bottom) inlets for fotomask. (C) Resulting silicon wafer master for creating droplet production chips.

2.2.1.2 Softlithography and plasma bonding

The stamp was produced with Sylgard 184 (Dow Corning, USA). PDMS and curing agent (trimethylsilyl chloride) were mixed in a ratio of 10:1 (w/w) initiating the polymerization reaction. Air bubbles were removed by applying vaccum in a dessicator for approx. 30 minutes. The degassed PDMS was poured on the master and cured for 3 hours at 65 °C in an oven (Heating oven Function Line T12, Thermo Fischer Scientific). After that, the stamp was pealed off from the master by cutting the PDMS with a surgical scalpel (B. Braun Melsungen AG). Holes for inlets and outlets were stamped out with a 1 mm thick biopsy punch (Kai Medical, Japan). The stamp was cleaned with glue tape removing dust particles. For the bottom-layer, microscope slides (VWR Microscope Slides, Cut Edges) were used and cleaned with ethanol prior to bonding. The bonding was performed with a plasma oven (MiniFlecto®, plasma technology GmbH, Herrenberg-Gültstein), using 99,99 % oxygen. Stamps and microscope slides were placed upside down inside the oven. The program was set to 0.6 mbar, 80% and 40 seconds. After the treatment, the upper side of stamp and slide were thoroughly pressed together allowing both parts to bond. The resulting micro channels obtained between PDMS and glass were flushed with 2 % (v/v) Trichloro(1H,1H,2H,2H-perfluorooctyl)-silane diluted in HFE-7500 for hydrophobicity. Chips were dried and stored at 65 °C to prevent wetting.

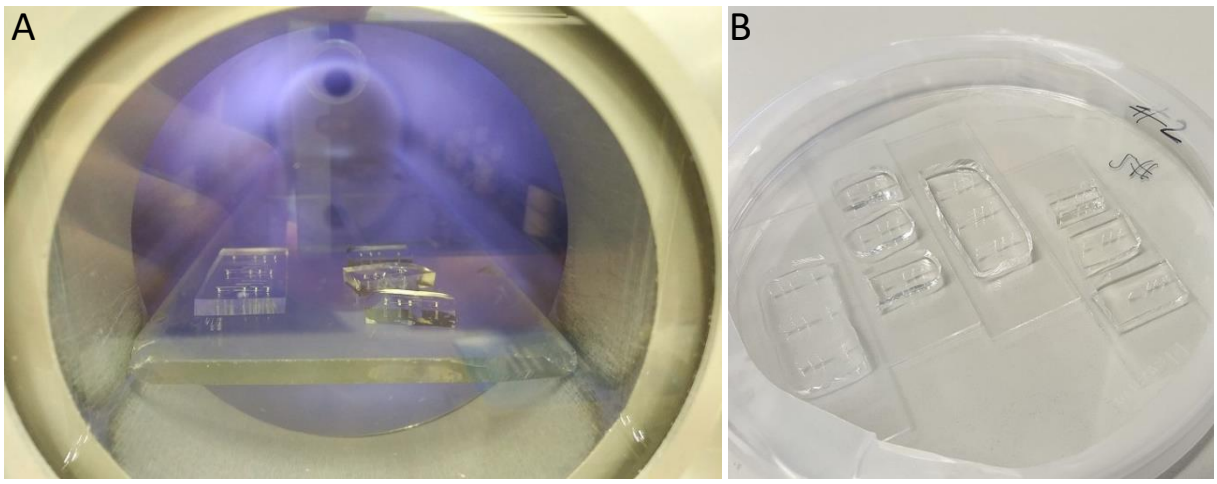


Figure 8: Softlithographic steps for microfluidic chip production. (A) Plasma bonding of PDMS stamps and glass slides. (B) Resulting microfluidic chips after plasma bonding.

Additionally, for the measurement chip a SMA to SMA fiber patch cable (M14L02, Thorlabs) was cut in the middle and approx. 1 cm of the outer and inner clading were removed at both inner ends with a scissor. The inlets for the for the glass fibre channels were filled with HFE 7500 oil. The chip was placed under a stereo microscope and the stripped fibres were smoothly inserted until the end of the channel close to the measurement channel. One end of the SMA Fibre was connected with the LED source and the other end with the photo detector. The alignment of both fibres was adjusted by tracking the light signal with a custom-made labview program. After adjustment, both fibres were fixated with epoxy glue.

2.2.2 Droplet production

For droplet production, a flow focusing junction chip with a junction width of 50 µm was used. The chip has three inlets (fluorinated oil, substrate solution and cell solution) and one outlet for collecting the produced droplets. The fluorinated oil syringe contained 1.5% (v/v) Picosurf1 (Sphere Fluidics, Cambridge, UK) diluted in HFE 7500 (3M Novec 7500). The syringe with the substrate solution contained 10-400 mM substrate, 1-5 mM cofactor, 10 mM WST-1, 20 µg/mL mPMS, 8% (v/v) CelLytic™ B Cell Lysis (10x), 1.5 mM WST-1 formazan and 100 mM buffer. Before droplet production, the solution was filtered and 0.1% rLysozyme (Merck KGaA, Darmstadt, Germany) was added afterwards. Lastly, the cell solution was prepared. Cell culture was photometrically measured and diluted to a lambda of 0.1 assuming that 1 OD of *E. coli* cells contain 3*10^8 cells. Cells were washed by spinning down 1 mL of harvested cell culture, removing the supernatant and dissolve the pellet with 950 µL of buffer. This step was repeated twice prior to encapsulation. Additionally, 25 % (v/v) of Percoll (VWR) and 100 mM buffer (HEPES pH 8 for *Bs*GDH and TrisHCl pH 9 for *Ss*GDH, *Vp*ALDH and *Hs*ALDH) were added to the cell syringe solution, aiming for a final OD of 0.06. Flowrates for oil, substrate and cell solution were 30 µL/min, 8 µL/min and 8 µL/min, respectively.

2.2.3 Droplet Incubation

For droplet incubation a self-made chamber was prepared by cutting the lid of a 1.5 mL Eppendorf tube and punching two holes (one at the bottom of the Eppendorf tube, the other on the side) with a 1 mm biopsy punch (Kai Medical, Japan) in the Eppendorf tube. The tube was turned upside down and bonded to a microscopy slide by using PDMS and curing agent. After two hours hardening in 65 °C oven, two tubings with a length of approx. 20 – 30 cm were cut and placed in the holes. Connections were fixed by applying instant glue on the interface between tubing and tube. For droplet production, the chamber was completely filled with HFE 7500. The upper tubing was connected with the droplet production chip. Droplets gathered on top of the tube due to buoyancy forces. After droplet production, both ends were sealed by lighting the tubing and pressing it together with a tweezer. For incubations at higher temperatures, oil was released from the bottom tubing, giving air space over the droplets and prevent bursting of the chamber by overpressure when heated. After incubation at the desired temperature, sealed tubings were cut off and connected with a syringe on the bottom tube. The upper tube was connected with the measurement chip to reinject droplets.

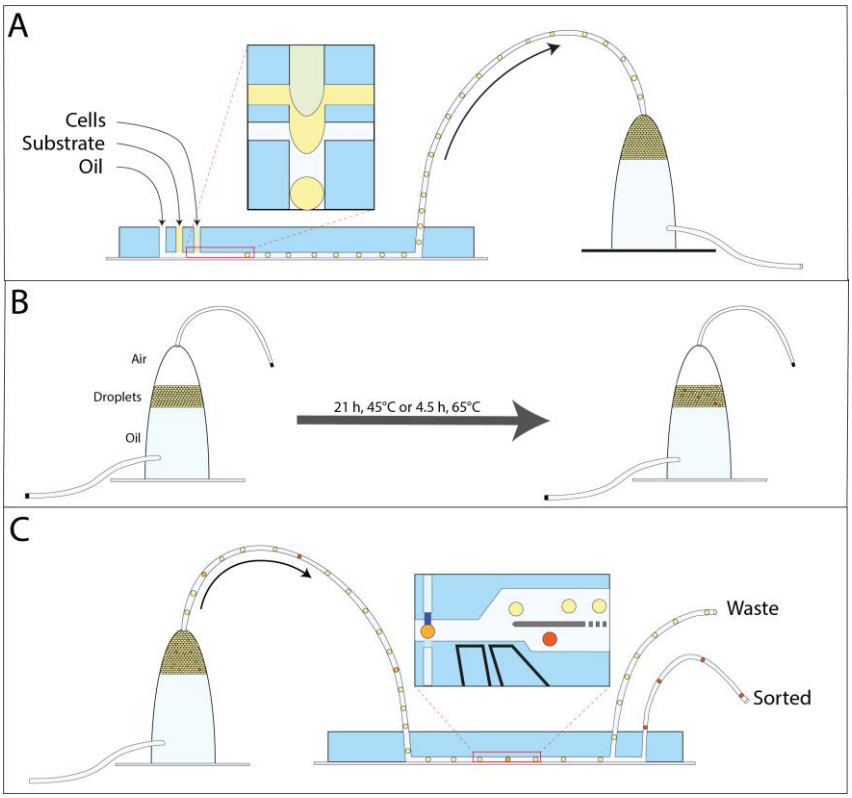


Figure 9: Droplet incubation in self-made chamber. (A) Production of microfluidic droplets with a flow-focusing chip. Cells with different enzyme variants are encapsulated together with substrate and assay components in monodisperse droplets and collected in a self-ustomed chamber. Droplets gather on top of the chamber due to buyoancy forces. Excess oil is released at the bottom outlet of the chamber. (B) Oil is partially released from the chamber. Droplets stay in the middle of oil and air phase. Tubing endings are sealed prior to incubation at 45 °C or 65 °C. Droplets with improved enzyme variants show a higher conversion of WST-1 and appear dark orange. (C) Tubing endings are opened again after incubation at elevated temperatures. A gas-tight syringe filled with oil is connected to the bottom inlet tubing and the chamber is filled with oil. Droplets are reinjected into a measurement chip and sorted according to their photometric value.

2.2.4 Droplet measurement

Droplets were pumped into measuring chips with 2 µL/min and spaced with HFE 7500 (30 µL/min). For estimation of product turnover, the conversion of WST-1 Formazan was measured. Droplets within the measurement channel were squeezed to 50µm to unify the measured pathlength and light with a wavelength of 455 nm was guided through the droplet via a glass fiber. Readouts of the remaining light were guided through another glass fiber going out of the chip into a photodetector, which converted the light intensity into an analog voltage signal. Analog signals were converted to digital signals using an analog to digital converter (NI-USB 6009, National Instruments). Digital signals were gathered and analyzed with a custom Labview program identifying the size and product concentration of each droplet. The corresponding data was saved in files for later analysis. The theoretical background of the program is further explained in chapter 3.1.4 and Appendix 3.1.4.1.

2.2.5 Droplet sorting

Analog signals from photodetector were guided through a breed board with resistors, reducing the maximum voltage signal to 3.3 V matching the maximum voltage for the microcontroller. Attenuated signals were sent in a microcontroller comparing incoming signals with a threshold value set for sorting. Signals lower than the threshold triggered an onboard pin and set the signal high, activating a pulse generator (TGP110, Aim-TTi, United Kingdom). The signal was modified to a 5 V pulse with a length of 10 – 12 ms and 4 ms delay. The pulse generator triggered a function generator (20 MHz DDS Function Generator TG 2000, Aim-TTi) generating a 13 kHz square signal with 10 V_{pp}. The signal was amplified 100 times with a voltage amplifier (TREK 601c) to actuate the sorter. Sorting channels on the measurement chip were filled with 5 M NaCl.

2.2.6 Sorting hit recovery

Sorted droplets were collected in an Eppendorf tube. The tubing was rinsed inside and outside with buffer to collect all sorted droplets. Emulsion was terminated by adding 100 µL 1H,1H,2H,2H-perfluorooctanol (PFO) (Sigma Aldrich). Droplets were shaken and centrifuged to reach complete phase separation. The top layer containing the aqueous phase with DNA of the sorted droplets was carefully removed and pipetted in another tube. 200 µL of distilled water were added for second extraction. After shaking and centrifuging the aqueous phase was collected again. DNA was purified through a spin column (Macherey-Nagel) eluting with 50 µL elution buffer from the kit. PCR was performed increasing the amount of DNA for transformation.

2.2.7 Rescreening and hit identification

Recovered plasmid DNA was used to transform *E. coli* BL21(DE3) according to the protocol described in 2.4.2. Afterwards, cells were grown on agar plates. Single colonies were picked and incubated in 96 deep well plates with each well containing 1 mL autoinduction media. Cultivation was performed on a plate shaker over night at 37 °C, 1.000 rpm. 100 µL from each well was transferred into a 96 - F-Bottom plate (Greiner Bio one) and mixed with 100 µL of 50% glycerol before storing as backup plate at -80 °C. Another 100 µL culture were taken and transferred into a 96 - U-Bottom plate (Greiner Bio one). U-Bottom-plates were centrifuged and supernatant was discarded. Cell pellets were incubated with 100 µL Lysis buffer (1.5 mg/mL Lysozyme, 0.5 mg/mL Polymyxin B Sulfate dissolved in buffer) for 1 hour at 37 °C, 1000 rpm. After lysis, plates were centrifuged again and supernatant was taken for measuring enzyme activity (2.6.2). Variants that showed increased activity were taken from the backup plate and incubated overnight. Plasmid DNA was isolated and purified by mini-prep and sent for sequencing (Eurofins Genomics Germany GmbH, Ebersberg, Germany).

2.3 Molecular biology methods

2.3.1 Isolation and purification of plasmid DNA from *E. coli*

For the production of plasmid DNA *E. coli* DH5α or NEB Turbo were used. Overnight cell cultures containing the desired plasmid were centrifuged at 4.000 xg for 10 minutes and the supernatant was discarded. The plasmid DNA was isolated and purified using the GeneJET Plasmid Miniprep Kit (Thermo Fischer Scientific) following the instructions of the manual. Plasmids were eluted with 30 – 50 µL of elution buffer. Plasmids were stored at -20 °C.

2.3.2 Determination of DNA concentrations

The concentrations of plasmid DNA were measured photometrically with a NanoDrop Spectrophotometer at 260nm. For measurements, 1 µL of sample was transferred on the light beam and measured with a lid factor of 50. Purity was assured checking A260/A280 – ratio for protein contamination and A260/A230-ratio for other contaminants (e.g. EDTA). Both ratios were at least 1.8 and 2.0, respectively.

2.3.3 Sequencing of plasmid DNA from *E. coli*

For verification of DNA sequences and identification of mutations, plasmid samples were sent to Eurofins Genomics Germany GmbH, Ebersberg, Germany. According to companies' guideline, each sample contained at least 15 µL of dsDNA with a concentration of 75 ng/µL. Samples with significantly higher concentration were diluted with dH_2O. In order to identify mutations, sequences were aligned using Snapgene v2.3.2 (GSL Biotech LLC).

2.4 Microbiology methods

2.4.1 Preparation of competent *E. coli* cells for chemical transformation

A single colony of *E. coli* BL21(DE3) was picked from an agar plate and transferred to a 50 mL baffled Erlenmeyer shake flask containing 10 ml LB medium. The culture was incubated at 37 °C and 150 rpm overnight. All culture was harvested and transferred to a 2 L baffled Erlenmeyer flask with 500 ml LB media under sterile conditions. The culture was incubated at 37 °C, 120 rpm until an OD of 0.4 was reached. The culture centrifuged at 2000 xg for 10 min, at 4 °C. All subsequent steps were performed on ice. After centrifugation, the supernatant was poured off and the pellet was resuspended with 50 mL of sterile and ice-cold 100 mM $CaCl_2$ solution. The solution was then transferred into a sterile 50 mL Falcon tube and incubated on ice for one hour. After that, centrifugation at 4 °C, 800 xg was performed for 10 min and the supernatant was discarded. The pellet was resuspended with 5 mL of sterile and ice-cold 85 mM $CaCl_2$-solution containing 15 w/v% glycerol. Cells were incubated for 1 hour on ice. Finally, the cells were aliquoted as 100 µl solution in a 1.5 ml Eppendorf tubes. Directly after aliquotation all tubes were flash frozen in liquid nitrogen and stored at -80 °C until further use.

2.4.2 Transformation of chemically competent cells

Chemically competent cells were taken from -80 °C freezer and thawed for 10 min on ice. Afterwards 1 to 5 µL plasmid DNA were added and cells were incubated another 30 min on ice. In the meantime, the water bath was heated up to 42 °C. Cells were heat shocked for 45 seconds and directly put on ice for 10 min. 900 µL prewarmed SOC medium was added and cells were incubated at 37 °C, 250 rpm, 1 hour for recovery. After recovery 100 µL were spread on an agar plate containing appropriate antibiotics. The remaining cells were spun down, supernatant was decanted and cells were resuspended. Another 100 µL of concentrated cell solution was plated and both plates were incubated at 37 °C overnight.

2.4.3 Cultivation of *E. coli* strains

In preparation for sequencing, single colonies were picked or cells from a single well in a backup plate were taken and cultivated in 5 mL LB-medium at 37 °C, 300 rpm in culture tubes overnight. Antibiotics were added for the appropriate selection markers of the strains (2.1.4). For precultures, *E. coli* strains were cultivated in 20 – 30 mL LB medium at 30 ° or in 100 mL baffled shake flasks at 160 rpm (MaxQ 2000) overnight. After harvesting, cell growth was determined measuring optical density (2.4.5) and main cultures were inoculated at an OD of 0.1.

2.4.4 Expression of recombinant genes in *E. coli* strains

For expression of recombinant genes, the same procedure as described in chapter 2.4.3 was performed with auto induction medium. Expression was demonstrated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) (2.5.1). Therefore, cells were harvested and transferred into an Eppendorf tube. After centrifugation (12.000 xg, 5 min) cells were resuspended in binding buffer (2.5.2) and mechanically disrupted by sonication (Sonotrode VialTweeter) with 100 % amplitude, 0.5 ms duration, and 3x 45 s pulse/30 s cooling on ice. After centrifugation (12.000 xg, 10 min), soluble fraction and pellet were separated and analyzed via SDS-page (2.5.1) for protein content. For microfluidic measurements, expression cultures were aliquoted in Eppendorf tubes. Each tube contained 1 mL of cultivation broth and was spun down for 5 min, 12.000 xg. Supernatant was carefully removed and discarded. Cells were frozen and stored at -20 °C. For protein characterization, *E. coli* strains were cultivated as described in chapter 2.4.3. Cell growth was determined by OD-measurement (2.4.5) and the main culture consisting of 1 L auto induction medium was inoculated with an OD of 0.1. Cells were grown at 16 °C or 30 °C overnight, depending on the solubility of the protein. Cells were harvested and transferred to 1 L centrifuge tube prior to centrifugation (Sorvall, RC 6+ Centrifuge with SH-3000 rotor, 4.000 g, 15 min, 20 °C). The pellet was resuspended in buffer and the cell solution was moved into a falcon tube. The cell solution was than directly taken for protein purification or stored as pellet at -20 °C after centrifugation (4,000 xg, 15 min, 20 °C).

2.4.5 Optical density measurement for *E. coli* cell solutions

The cell growth was determined measuring the optical density at 600 nm (Implen Nanophotometer P-class). Plastic cuvettes with a path length of 10 mm were filled with 1 mL of cell solution. Samples were diluted if OD exceeded 0.8. An OD_{600} of 1 corresponds to 4.4x 10^8 cells/mL (3.1.2).

2.5 Biochemical Methods

2.5.1 SDS-Page

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed to identify protein expression in the supernatant (soluble fraction) or in the pellet (insoluble fraction). According to the protocol of Laemmli[136] the following solutions were prepared as stated in the internal chair protocol:

Ammonium persulfate (APS) stock solution, 10%

10% (w/v) APS were solved in ddH2O and stored at 4 °C.

Coomassie Brilliant Blue staining solution

0.2% (w/v) Coomassie Brilliant Blue G250 und R250, 50% (v/v) ethanol, 10% (v/v) pure acetic acid. The solution was stirred for 3 h before filtration and stored, protected from light, at room temperature.

Acrylamide/bisacrylamide solution

Acrylamide and bisacrylamide were mixed in a ration of 37.5:1 and mixed with water to a final concentration of 40 % (w/v).

4x SDS-PAGE separating gel buffer

0.8% (w/v) SDS, 1.5 M TRIS HCl pH 8.8. The pH was adjusted with 37% (v/v) HCl.

4x SDS-PAGE stacking gel buffer

0.8% (w/v) SDS, 0.5 M TRIS/HCl, pH 6.8. The pH was adjusted with 37% (v/v) HCl.

10x SDS-PAGE buffer

1% (w/v) SDS, 0.25 M TRIS, 1.92 M glycine. The resulting pH of 8.5 was not changed.

5x SDS loading dye

50% (v/v) glycerol; 12.5% (v/v) β-mercaptoethanol; 7.5% (w/v) SDS; 0.25 M TRIS, 0.25 g/L Bromophenol blue, pH 6.8. The pH was adjusted with 37% (v/v) HCl. Gels were prepared using the Multi-Casting Chamber (BioRad Laboratories GmbH). For 14 gels, 80 ml of separation gel (12 %) and 40 ml of stacking gel (5 %) were prepared. Separation gel consisted of 24 ml acrylamide/bis-acrylamide solution, 20 ml separation gel buffer (4x), 34.32 ml ddH_2O, 800 µl SDS (10 % (w/v)). The stacking gel consisted of 5.2 ml acrylamide/bisacrylamide solution, 10 ml stacking buffer (4x), 23.96 ml ddH_2O and 400 µl SDS (10 % (w/v)). The polymerization was initiated with 0.1 % Tetramethylethylenediamine (TEMED) and 1% APS stock-solution. If necessary samples were diluted with water and mixed with 5x SDS loading dye adjusting a final concentration of 1x. After heat denaturation at 95 °C for 5 min, samples were spun down (Thermo Fischer Scientific Heraeus, Fresco 21 Centrifuge) 17.000 g, 1 min and either divided in pellet and supernatant fraction or they were stored at 4 °C for further usage. In that case, the samples were heated and centrifuged again. Depending on the expected concentration 5 µl to 15 µl were pipetted into the gel well and the electrophoresis was started at a constant current of 30 mA per gel for about 50 – 60 min (Bio-Rad Laboratories GmbH, Mini-Protean Tetra Cell). For detection of proteins, the gels were first stained with Coomassie Brilliant Blue for 15 min. For a first analysis, the gels were destained in boiling water. If necessary gels were further destained by putting them on a nutating shaker in water at room temperature overnight.

2.5.2 Cell disruption and protein purification

Cell lysis by sonication

First cells were disrupted by sonication to release the protein within the cells. Therefore, cells were harvested and centrifuged for 15 min at 5000 xg, 4 °C (Sorvall RC 6 Plus, F10-6x500y). The supernatant was discarded and the pellet was resuspended in binding buffer (50 mM potassium phosphate (KPi) pH 8.0, 500 mM NaCl, 10% glycerol, 20 mM imidazole). Cells were transferred into a 50 ml falcon tube and 4 µL DNase (10 mg/ml) were added before cooling the sample on ice. Disruption was performed by sonication (80% amplitude, 0.5 s pulse, 3x 10 min (Ultra Turrax IKA T18 basic). After sonication, the samples were again centrifuged for 30 min at 40,000 xg, 4 °C. The supernatant was filtered (0.45 µm filter) and in case of glucose dehydrogenase from *Saccharolobus solfataricus* (*Ss*GDH) a heat shock (30 min, 70 °C water bath) was applied to reduce the amount of *E. coli* protein within the sample and to heat-activate the protein. His-tagged proteins were purified by immobilized ion affinity chromatography (IMAC).

Immobilized ion affinity chromatography (IMAC)

A HisTrap FF column (GE Healthcare) with a column volume (CV) of 5 ml was used to perform IMAC. The His-tagged proteins are bound onto the column and washed with binding buffer (50 mM KPi (pH 8.0), 500 mM NaCl, 10 % glycerol, 20 mM imidazole). Elution is performed with elution buffer (50 mM KPi (pH 8.0), 500 mM NaCl, 10 % glycerol, 500 mM imidazole) containing high concentrations of imidazole, leading to detachment of the His-tagged proteins. For the purification procedure, the column was mounted in an ÄKTA purifier system (GE Healthcare) and the proteins were purified as shown in Table 2.

Table 2: Steps for His-tagged protein purification with ÄKTA purifier system

N°	**Step**	**Volume**	**Solution**
1	Equilibration	5 CV	Binding buffer
2	Loading	5-8 mL/min	Filtered Supernatant
3	Washing	5 CV	Binding Buffer
4	Elution	5 CV	Elution buffer (gradient 0 -100 %)
5	Cleaning	10 CV	Elution buffer (Until UV signal is at baseline)
	If more than one Protein was purified steps 1 -5 were repeated		
6	Remove buffer	5 CV	ddH_2O
7	Storage	5 CV	20 % Ethanol

The complete purification process was monitored via UV-measurement at 280 nm allowing to detect the eluted protein (UV increases) and indicating the status of the cleaning process (UV decreases). Step 6 and 7 were only performed, if the purification process was finished and no desalting was performed afterwards to avoid contamination of the column. Protein desalting was performed to remove salts and imidazole and exchange the buffer to the buffer in which

enzymes were characterized (50 mM HEPES pH 8 for glucose dehydrogenase from *Bacillus subtilis* (*Bs*GDH) and 50 mM TrisHCl pH 8.9 for *Ss*GDH). For desalting the HiPrep 26/10 desalting column (1 CV = 53 ml) was mounted to the ÄKTA purifier system. Table 3 shows the procedure.

Table 3: Steps for desalting protein solutions with ÄKTA purifier system

N°	Step	Volume	Solution
1	Equilibration	5 CV	Characterization buffer
2	Loading	up to 15 mL	Protein solution
3	Elution	2 CV	Characterization buffer
4	Cleaning	3 CV	Characterization buffer (UV signal needs to be at baseline)
	If more than one Protein was desalted steps 1 -4 were repeated		
5	Remove buffer	3 CV	ddH_2O
6	Storage	2 CV	20 % Ethanol

After desalting the protein solution was measured for its protein content (2.5.3) and then dripped with a Pasteur pipette into liquid nitrogen resulting in small protein beads. The beads were stored at -80 °C until further use.

2.5.3 Determination of protein concentration

Purified protein solutions (2.5.2) were measured photometrically at 280 nm with a nanophotometer (Implen P330). For the measurement 1.5 µL of protein solution were pipetted on the measuring window and closed with Lid 10 (10 mm pathlength). Extinction coefficients and molecular weight of the proteins were derived using the ExPASy ProtParam tool. The tool calculates the extinction coefficient based on the sum of all extinction coefficients from the aromatic amino acids tryptophan, tyrosine and cysteine[137]. The corresponding molar extinction coefficient and molecular mass for each protein were entered in the nanophotometer and allowed quantification of the protein concentration according to Beer-Lambert law. For protein variants that either gained or lost the aromatic amino acids tryptophan, tyrosine and phenylalanine due to mutation, the molar extinction coefficient was adapted prior to measurement.

2.6 Determination of enzymatic activity

2.6.1 Quantification of enzymatic stability with gradient incubation

Single colonies expressing the desired protein were picked from agar plate and incubated in shake flask or deep well plate overnight (2.4.4). Cells were harvested and the OD_{600} was determined to obtain cell density. Cultures were diluted with lysis agent (1.5 mg/mL Lysozyme, 0.5 mg/mL Polymyxin B dissolved in buffer) to a final OD of 0.1 and 8x 50µl of each sample

were pipetted in PCR tube strips. Strips were incubated for 1 hour at room temperature to allow lysis. The strips were put in a thermocycler in the direction of the applied gradient. Samples were taken after defined times and centrifuged with a mini centrifuge to remove precipitated sample parts. Supernatant (10 - 20 µL) was pipetted in a microtiter plate and residual enzymatic activity was determined (2.2.6).

2.6.2 Quantification of enzymatic activity in microfluidic droplets

Single cells were encapsulated together with substrate, cofactor and lysis buffer as described in chapter 2.2.2. By lysing the cells, enzymes within the cell were released and the enzyme reaction started. By oxidation of the substrate, the cofactor NAD^+ was reduced to NADH. The electron coupling agent mPMS recycles the cofactor and transfers the electrons to the colorless WST-1, which is reduced into the colorful WST-1 formazan salt. Enzyme activity was measured by guiding the droplets through a constant LED light signal, which was connected to a photodetector on the other side. Every time a droplet passes through the light beam, the light intensity is reduced proportionally to the amount of produced WST-1. Droplets with increased enzymatic activity were sorted and further investigated (2.4 ff.)

2.6.3 Cell lysate rescreening

For the rescreening of protein variants derived from microfluidic screening, single colonies were picked and cultivated in microtiter plates as described before (2.2.6). The increase of NADH was photometrically measured at 340 nm. Measurements were performed with the same concentration of substrate and cofactor as before in the microfluidic screening. Measurements up to 45 °C were performed in the Multiskan Spectrum plate reader (Thermo Fischer Scientific Inc.), while measurements above 45 °C were performed with the microplate spectrophotometer epoch 2 (BioTek Instruments Inc.). Prior to the measurement, all solutions were prewarmed up to screening temperature. Reactions were started by pipetting 180 µL of master mix (containing substrate, cofactor and buffer with defined concentrations) on top of 20 µL of the protein cell lysate, which was pipetted in each microtiter plate well in advance. Starting the measurement, the plate was initially shaken for 10 seconds at 1000 rpm to assure good mixing before starting the measurement. Kinetics of the measurement were exported in Excel for further analysis. In order to avoid inaccuracies in the calculation of the slope for the enzyme activity due to cooling down of solutions at room temperature, the maximal slope of 4 consecutive measurement points was taken.

2.6.4 Enzyme characterization

For enzyme characterization, purified enzymes (2.5.2) were thawed from -80 °C freezer and diluted in the corresponding buffer. Since either cofactor or substrate concentrations alternates for the kinetic measurements, aliquots for master mixes of each concentration were pipetted

in a deep well plate. Microtiter plates containing 20 µL of enzyme solution were prepared and the reaction was started by adding 180 µL mastermix (containing substrate and cofactor) from the deep well plate with a multichannel pipette on top of the enzyme solution. If necessary, solutions were prewarmed to assure optimal conditions for the kinetic measurements. The measurement and analysis were conducted in the same way as described before (2.6.1). Enzymatic activity was calculated by Beer-Lambert law.

3 Results

3.1 Establishment of a droplet-based microfluidic screening system

Over the course of this thesis a droplet-based microfluidic screening system was established, which is capable of measuring differences in enzymatic activities on the single cell level. For that, single cells and assay components were encapsulated within microfluidic droplets, allowing measurement of hundreds of enzyme variants per second. Interesting enzyme variants yielding high product concentrations can be sorted on demand and the corresponding DNA retrieved to identify mutational changes. Figure 10 illustrates the workflow of a microfluidic screening campaign.

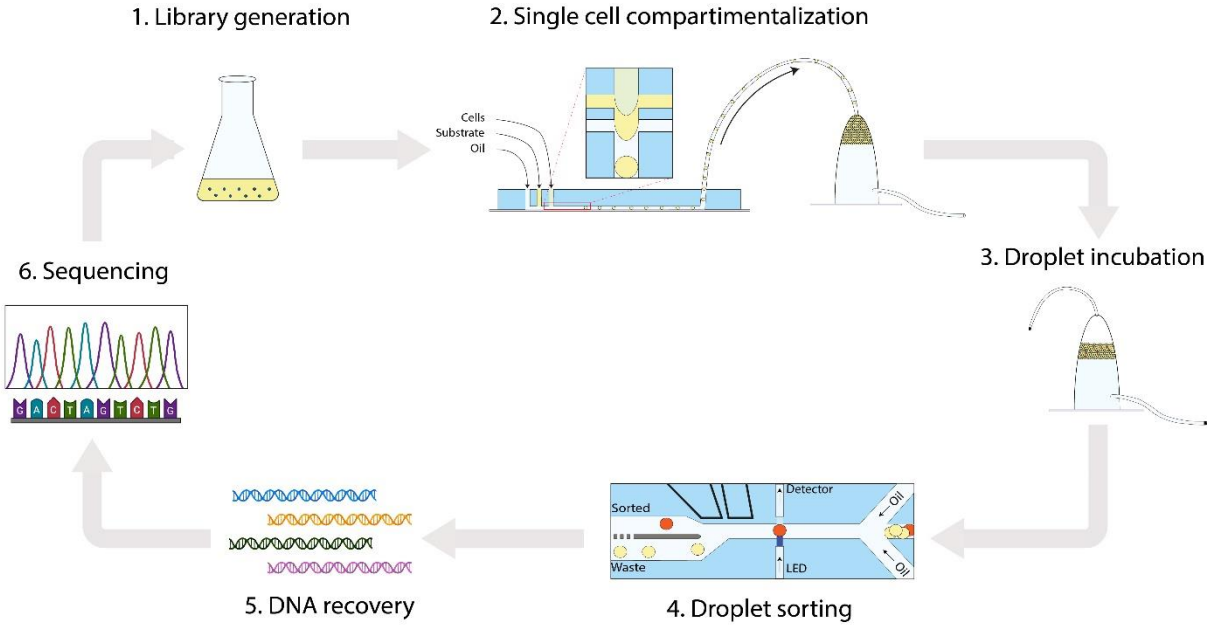


Figure 10: Workflow for droplet-based microfluidic sorting. (1) Plasmids for recombinant proteins were expressed with competent *E. coli* cells. (2) Single cells holding expressed proteins were encapsulated together with assay components and lysis buffer. Droplets were captured in a self-made droplet chamber. (3) After cell lyses, enzymes were released into droplet space and start the enzyme reaction resulting in WST-1 dye formation. Droplets were incubated at desired temperature up to several days. Droplet chamber prevented evaporation and merging of droplets. (4) After incubation, droplets were reinjected in a second sorting chip. The dye formation was measured and desired droplets were separated by dielectrophoretic sorting. (5) Captured cells within droplets were de-emulsified and DNA was recovered. (6) DNA of improved variants was sequenced and analyzed in-silico to identify benefitial mutations. These mutations can be included to the new template variant and a new screening round can be started.

In the upcoming chapters, the process of producing monodisperse droplets of defined size will be explained and droplet sizes with different flowrates and chip designs are verified (chapter 3.1.1). Next, droplet size is validated by counting the cells within droplet populations and comparing it with the theoretical Poisson distribution (chapter 3.1.2). The assay for the subsequent screening process is tested for different pH values and temperatures in microtiter plates to determine stable conditions (chapter 3.1.3). The theoretical background for establishing the LabVIEW measurement program is explained in chapter 3.1.4. First, a

program with a predefined VI (virtual instrument) is assembled and tested. The limitations of this approach are discussed and as a result, a self-customed program is developed to overcome the identified issues. Since enzyme engineering in this work is performed with thermophilic enzymes, the droplet stability and possible leakage at elevated temperatures is investigated (chapter 3.1.5). Finally, the dieletrophoretic sorting is tested for different flowrates and pulse parameters (intensity, width and delay) and verified by a sorting experiment (chapter 3.1.6). These steps pave the way for the enzyme engineering approaches presented in the later chapters of this work (chapter 3.2, 3.3 and 3.5)

3.1.1 Droplet production

The size of microfluidic droplets decides about different factors: First and most importantly, the droplet size decides about the parameters for the assay reaction. If one cell is encapsulated in a small droplet, the reaction will be quickly conducted since the number of enzyme molecules is comparably high to the substrate molecules within the droplet. However, this will also reduce the capacity of the assay and consequently the observation time, since the solubility of WST-1 formazan dye is restricted. Second, the enzymatic turnover is measured by the reduction of light absorbed by WST-1 formazan. Therefore, the diameter of the droplet decides about the path length for subsequent measurements. The Beer-Lambert law defines the dependencies between the path length, the concentration and the extinction coefficient of the dye (1.2.4). The smaller a droplet and consequently the path length, the higher the difference in WST-1 formazan concentration needs to be to distinguish two different populations. Lastly, the size of the droplet also influences the throughput for the sorting. The bigger a droplet, the more likely it is that it disrupts when applying high flowrates. Additionally, the pulse intensity and sorting time also increases with the droplet size, which directly defines the maximal throughput of a sorting experiment.

For the droplet production in this thesis a flow-focusing chip with a junction width of 50 µm was chosen (chapter 2.2.1.1, Figure 7). The corresponding chip was manufactured, and droplet production was tested with different flowrates for oil, substrate and cell syringes to determine droplet sizes and standard deviations. For simplification purposes, substrate and cell syringes were filled with buffer solution (50 mM HEPES, pH 8). The oil syringe was filled with HFE 7500 containing 1.5 % Picosurf1 as surfactant, which stabilized the droplet surface for long-term storage of droplets. Syringes were connected to the chip with tubing and pumps were started simultaneously. After pressure equilibration of all three flows was attuned, monodisperse droplets were produced and collected in 1.5 mL reaction tubes. Droplets were transferred on a glass slide by a pipette and investigated under the microscope. Additional HFE 7500 oil was pipetted on the glass slide to surround the droplets and decelerate evaporation caused by the

microscopic illumination source. This retained the round shape of the droplets and allowed measurement of droplet diameters. Pictures of each droplet population were taken for each flow rate separately by a stereo microscope (Zeiss Axiocam). After size calibration, diameters for 20 droplets per flowrate combination were determined manually by using the diameter function of ZEISS Blue software. The initial measurements with a chip depth of 80 µm resulted in droplets with a diameter of 90 to 137 µm, depending on the respective flowrates (Figure 11). The results showed that the water to oil ratio mainly determined the size of the droplets. For the subsequent setup, smaller droplets were preferred, as larger droplets would decrease the number of droplets that can be produced from a given amount of substrate and cell solution. Moreover, the strength and duration of the dielectrophoretical sorting signal should be increased to sort droplets successfully. This would decrease the throughput and consequently the sorting speed. Hence, the droplet diameter should be reduced to avoid these problems. To address this issue, the height of the master photoresist was reduced from 80 to 50 µm. Droplets produced with flow rates of 30 µL/min and 8 µL/min for oil and aqueous solutions respectively, resulted in a diameter of roughly 80 µm. By decreasing the aqueous flowrates to 4 µL/min or even 2 µL/min, the diameter could be further decreased to 76 µm. However, using double or even four times the amount of oil to form droplets with a 4 µm reduced diameter was not a sustainable tradeoff. Interestingly, increasing the oil flow rate to 37 µL/min and keeping the aqueous solutions at 8 µL/min resulted in bigger droplets and bigger standard deviation.

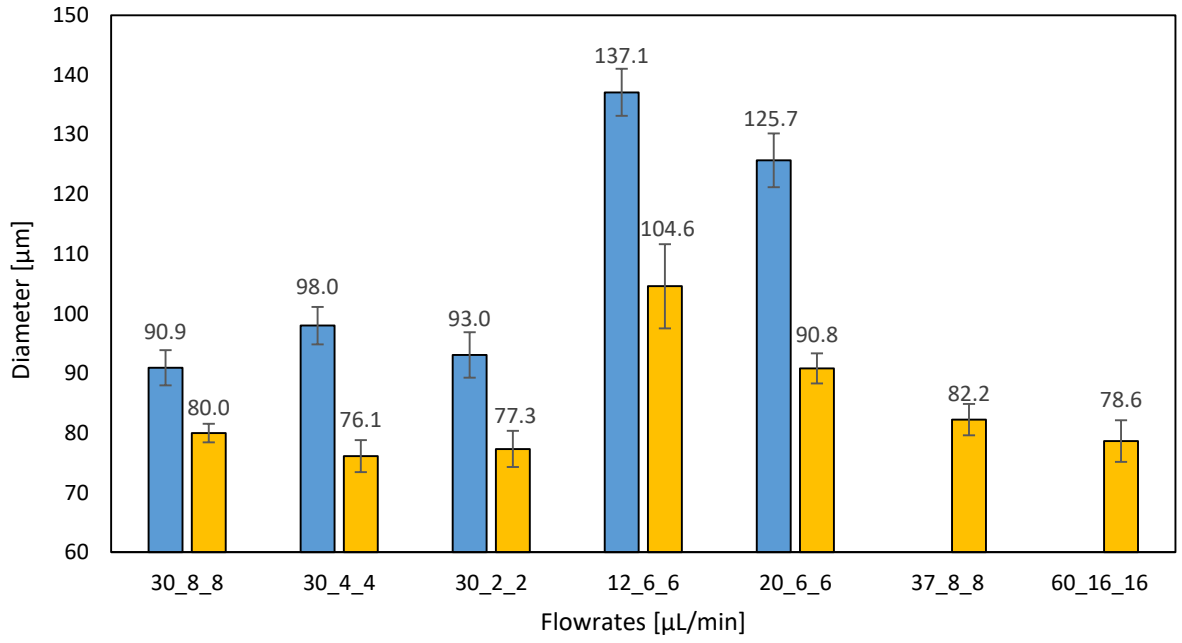


Figure 11: Droplet diameter for the production with different flowrates with 80 µm (blue) and 50 µm (orange) channel height. Flowrates are indicated by the numbers below the corresponding bars. First number defines the oil flowrate, while second and third number describe the substrate/cell syringe flowrate in µL/min.

Finally, the flow rates for the oil and aqueous solutions were increased twofold (60:16:16) to determine if this could increase the droplet production speed. The measured diameter of the

droplet population was nearly identical (79 µm), but increasing the system pressure led to a higher standard deviation of 3.5 µm compared to 1.5 µm for the droplets made with the default flowrates (30:8:8). Since the volume of the droplets increases cubically with respect to the diameter, the flow rates for producing droplets were set to 30:8:8 [µL/min] for oil, substrate, and cell solutions, respectively.

3.1.2 Droplet occupation

The next step was to assess the occupation of droplets. For the subsequent screening approaches, it was crucial that single cells were encapsulated, as each cell produced a different variant of the enzyme. If more than one cell was encapsulated, the resulting product formation may be higher, not because the enzymes within the droplet were improved, but simply due to a higher enzyme count, if encapsulation occured randomly and independently. The probability for droplet occupation by a certain number of cells is defined by the Poisson distribution. The distribution describes the likelihood how many cells will be encapsulated per droplet. The ratio for the average number of cells per droplet is defined as lambda. If lambda is 0.1, the number of droplets is 10 times higher than the number of cells. Lowering lambda is advantageous as its decreasing the chance of double occupations. On the other hand, it also decreases the overall percentage of droplets with single encapsulations and rises the number of unoccupied droplets. This in turn decreases the actual throughput of enzyme variants investigated per second (also referred to as biological throughput). To adjust lambda, the total amount of cells per volume needs to be controlled. Therefore, *E. coli* BL21(DE3) was grown over night to a final OD of 0.592. Three different dilutions (1:100, 1:200, 1:500) were prepared and counted with a thoma chamber. According to that, an OD of 1 corresponds to 4.4×10^8 cells. Knowing that an average droplet has a diameter of 80 µm and assuming that a droplet has a perfect spherical shape, the volume of one droplet could be calculated 268 pL. With 100 µL aqueous solution (50 µL substrate and 50 µL cell solution) a total amount of 373.134 droplets could be produced. A lambda of 0.1 was adjusted, resulting in 37.314 cells within in the cell solution. This required an optical density (OD_{600}) of 0.006, including dilution factors for cell solution and mixture of substrate and cell solution. As a proof of principle, cell solutions for lambda 0.1 and lambda 0.2 were produced and the corresponding droplets were investigated under the microscope for cell occupation.

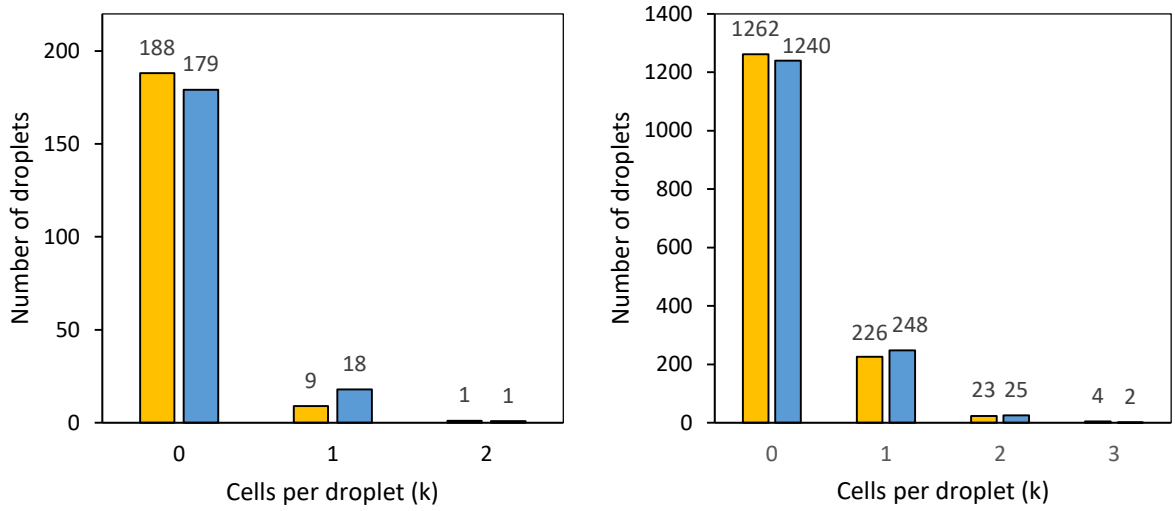


Figure 12: Cell distribution for lambda 0.1 (left) and lambda 0.2 (right). Cells within droplets were counted (orange) and compared to the expected distribution according to Poisson (blue).

In total 198 droplets were investigated for lambda 0.1 and 1,515 droplets for lambda 0.2 (Figure 12). Both droplet samples showed a distribution, which was below, but still very close to the expected numbers of single and multiple occupied droplets. The results confirmed that the calculated conversion factor from optical density to total cell number was accurate. Additionally, by verifying the droplet occupation, the size of the droplets was also validated.

3.1.3 Determination of assay stability

For monitoring the enzymatic activity, the cofactor coupled WST-1 assay was used (1.3.4). Since the activity of enzymes from one single cell was investigated with an absorbance assay, the incubation process for the enzyme turnover took several hours or even days at room temperature. During this time, WST-1 formazan needed to remain stable in order to identify improved variants. To achieve this, a spectrum of WST-1 was measured to determine the wavelength with the highest extinction. The highest extinction was found at 440nm. For the subsequent assay application, different buffers were tested to observe variations in the maximal turnover of the assay. The reaction was started by mixing 5 mM WST-1 with 0.3 mM $NADH+H^+$ in a microtiter plate, measuring the conversion of WST-1 formazan at 440 nm (Figure 13). Multiple measurements were performed at different time points to detect any potential decay of the dye.

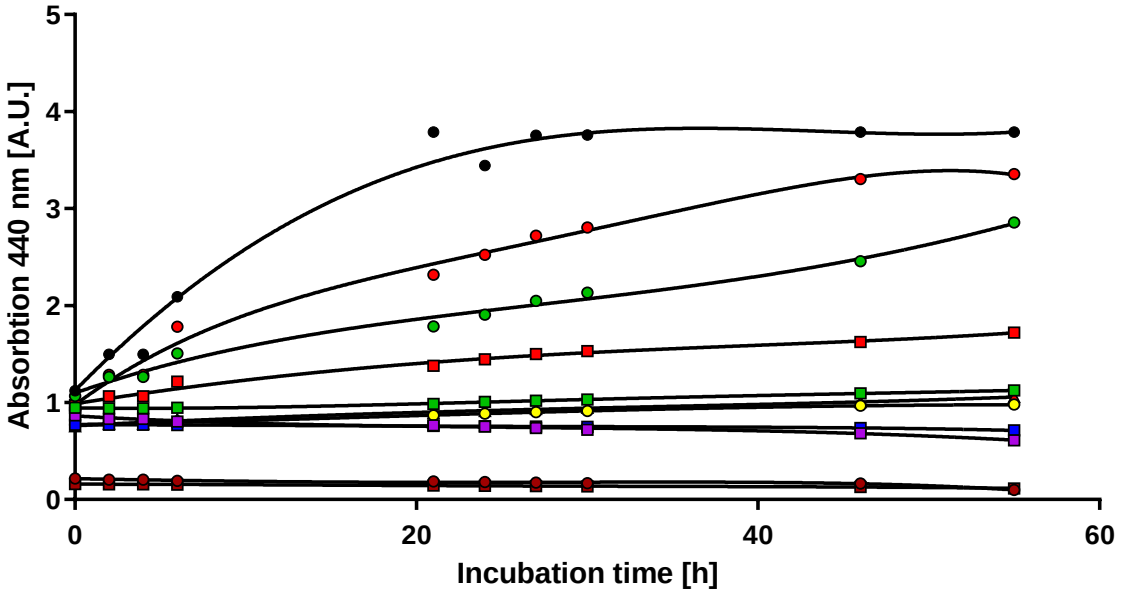


Figure 13: Stability test of WST-1 formazan at 440nm in different buffers over time at room temperature. Buffers were 500 mM glycine sodium hydroxide buffer pH 10.0 (black circle), 500 mM HEPES buffer pH 7.9 (red circle), 1M Tris HCl pH 9.0 (green circle), 50 mM HEPES buffer pH 7.9 (red square), 50 mM Tris HCl pH 8.0 (green square), 50 mM HEPES buffer pH 7.0 (red triangle), 100 mM PIPES buffer pH 7.0 (yellow circle), deionised water (blue square), 500 mM potassium phosphate buffer pH 8.0 (purple square),1 M potassium phosphate buffer pH 6.0 (brown circle) and 50 mM phosphate buffer pH 5.5 (brown square).

As shown in Figure 13, only 500 mM potassium phosphate buffer pH 8.0 (purple squares) showed a minor decay in signal intensity over the time course of 55 hours. For 500 mM glycine sodium hydroxide pH 10, 50 mM and 500 mM HEPES pH 7.9 and 1M Tris HCl pH 9 the absorption even increased over the course of the measurement. This indicates that absorption values increased for buffers with higher pH. The experiment also revealed, that subsequent measurements have to be performed in basic milieu. No conversion of WST-1 formazan dye was observed at pH 7.0 or below. Throughout this thesis, droplets were incubated at elevated temperatures (e.g. 45 °C) to reduce the incubation time and to assess the temperature stability of different enzyme variants and ultimately using the temperature as a selection criterion. Consequently, it was essential to verify the stability of the WST-1 assay at these higher temperatures. Therefore, the assay was performed with 5 mM WST-1 and 10 µM mPMS mixed in different buffers. As before, the reaction was started by addition of 0.3 mM NADH (Figure 14).

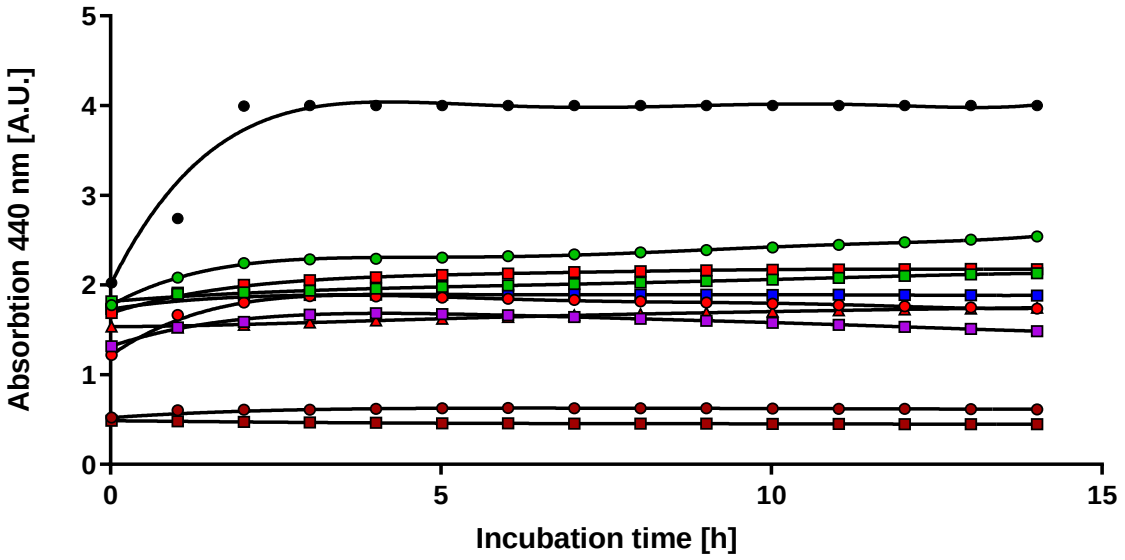


Figure 14: Stability test of WST-1 formazan measured at 440nm in different buffers over time at 45 °C. Buffers were 500 mM glycine sodium hydroxide buffer pH 10.0 (black circle), 1M Tris HCl pH 9.0 (green circle), 50 mM HEPES buffer pH 7.9 (red square), 50 mM Tris HCl pH 8.0 (green square), deionised water (blue square), 500 mM HEPES buffer pH 7.9 (red circle), 50 mM HEPES buffer pH 7.0 (red triangle), 500 mM potassium phosphate buffer pH 8.0 (purple square), 1 M potassium phosphate buffer pH 6.0 (brown circle) and 50 mM phosphate buffer pH 5.5 (brown square).

As previously observed at room temperature, the signal intensity strongly depended on the pH value of the buffer. Due to the increased temperature, the reactions were catalyzed much faster. The highest conversion rate was found with 500 mM glycine sodium hydroxide pH 10. It reached the maximal absorption value of the photometer after 2 hours of incubation. All other tested buffers also reached their maximum conversion within the first 5 hours. The only exception was 1 M Tris HCl pH 9.0 (green circle), which slowly increased over time. The only reaction that showed a decay was 500 mM potassium phosphate buffer pH 8.0 (purple squares). The stability of the WST-1 assay could not be tested for 60 °C in the plate assay since the evaporation distorted the measurement. Sealing the plate with a cover resulted in non-reproducible results, since evaporated water drops forming under the sealing and interfered with the readout. Therefore, the assay stability was directly tested in droplets (3.1.5).

3.1.4 Measurement of product formation in microfluidic droplets

For the measurement, a microfluidic chip with an optical fiber was produced (2.2.1.2). After production, droplets were captured in a self-made chamber and incubated for a defined time to allow cell lysis and release of enzymes into the droplet sphere initiating the enzyme reaction. The enzymatic turnover was monitored with the WST-1 cofactor coupled assay (1.3.4). Droplets were reinjected into the measurement chip after incubation. The droplets were spaced allowing clear separation of measurement signals and guided through the measurement point. Reduction of the light signal was continuously measured by a photo detector and converted into an analog signal (2.2.4.1). The analog signal was forwarded to an A/D - converter

(NI USB 6009) and processed on the computer with a LabVIEW program. Within the LabVIEW program, the signal was initially gathered with a frequency of 10,000 Hz, meaning that every 0.1 ms the signal intensity of the photo detector was captured. This high frequency allowed a precise reproduction of the light intensity for every droplet that passed through the optical measurement and ensured that the highest concentration, resulting in the lowest signal intensity (from now on called “Peak”), can be found. To achieve this, it was important to develop a program that can automatically identify the peaks and collect the data for analysis.

Establishment of a peakdetector program from predefined LabVIEW building blocks

First, a progam based on a predefined building block (subVI) from LabVIEW, called “Peakdetector.vi” was employed to analyze the raw data. For basic peak identification, the subVI needed a value for the peak threshold (1), the argument whether a peak or a valley should be identified (2) and the expected width of the peak signal (3). With the given subVI, peaks for droplet measurements could be identified. However, all three parameters had to be chosen manually and might change for each screening, depending on the throughput and the signal intensity of the different droplet populations. In this case, expert knowledge would be required to adjust the parameters properly and once the parameters were set for a measurement, only manual changes can be applied during the screening process. Therefore, the following paragraph will describe how the subVI was gradually improved and eventually replaced by a self-customed program (Appendix 3.1.4.1). This program improved the screening process, by automatically adjusting the three points mentioned above and enabled screening for non-expert users.

(1) The main challenge in defining the peak threshold, was the inconsistent baseline intensity. Depending on the relative position of the optical fibers to each other, the signal intensity showed a certain variance. The peak threshold served as a criterion to determine which measurement points were considered part of the peak signal and which were excluded from the peak analysis. Therefore, the threshold was chosen as low as possible to identify all droplets and only exclude baseline noise. If the baseline intensity differed from the default value, for example the intensity was set to 8 V in the program, but the photodetector measured 7 V, the subVI would identify peaks within the baseline, as the subVI was designed to find the highest (peaks) or lowest (valley) value within the peak width. This led to random point identification on the baseline. In order to identify the peaks reliably and independently from the baseline intensity, the variable baseline was subtracted from all values allowing normalization of the raw data signal. Therefore, the mean of all 10,000 values captured within one second was calculated. As most of the points belonged to the baseline and therefore had the baseline value, the mean of all values was similar to the baseline and subtraction led to values close to

0 V (Figure 15 B). The normalized data was then forwarded to the “Peakdetector.vi” for peak identification.

(2) The predefined LabVIEW subprogram “Peakdetector.vi” could identify either peaks or valleys by default. Since WST-1 formazan formation reduced the light intensity at 455 nm, droplets with higher product formation resulted in valleys within the baseline signal. Interestingly, droplets with low WST-1 formazan concentration resulted in higher light transmission, which in turn increased the signal intensity, creating peaks. Gielen et al.[78] found that liquids with higher refractive indices result in higher light signals. As a result, droplets filled with water (or buffer) passing the optical measurement, resulted in an increased light signal. This increase could be measured as a peak formation (Figure 15 A). To count all droplets (with high and low WST-1 formazan concentration) it was at first necessary to include two separate peakdetector subprograms and sum the number of droplets counted by both peakdetectors. Even this approach was able to correctly differentiate droplets with high and low WST-1 formazan concentrations, it required fine tuning of the corresponding thresholds. As shown in Figure 15 A, both peak signals consisted of several smaller valleys and peaks, which masked the signal peak and made it difficult to set thresholds precisely to avoid double counting. Additionally, if thresholds were chosen too low, the subprogram which should identify valleys, would find the valleys produced by the droplets with low formazan (Figure 15 A, [around measurement point 30]) as a potential valley of a solitary droplet. This would result in counting the droplet twice. On the other hand, if the threshold would be set too high, droplets which produced less than 1.5 mM of WST-1 formazan would not be counted. Both events would distort the later statistical analysis. The best solution for this problem was the reduction of the two subprograms to one subprogram, which can identify both populations at once. This was achieved by calculating the absolute value (modulus) of the light intensity after subtracting the baseline value. Figure 15 C illustrates the converted light intensity signal after threshold subtraction and application of the modulus based on the raw data (Figure 15 A). Now, both signals appeared as peak values and could be recognized by the peakdetector.vi in “peak”-mode. However, this solution brought up a new problem. The concentration of WST-1 was proportional to the reverse light intensity in the middle of a droplet, indicated by the red dots in Figure 15 C. For the droplet with low WST-1 formazan concentration (Figure 15 C, blue dots), the middle of the droplet signal was a valley, masked by two peaks (between measurement point 15 and 30) caused by refraction. Moreover, the middle of the droplet with high WST-1 formazan concentration (red) was next to an even higher peak (measurement point 69). In order to reliably identifying the middle of the droplet, the peak width needed to be defined.

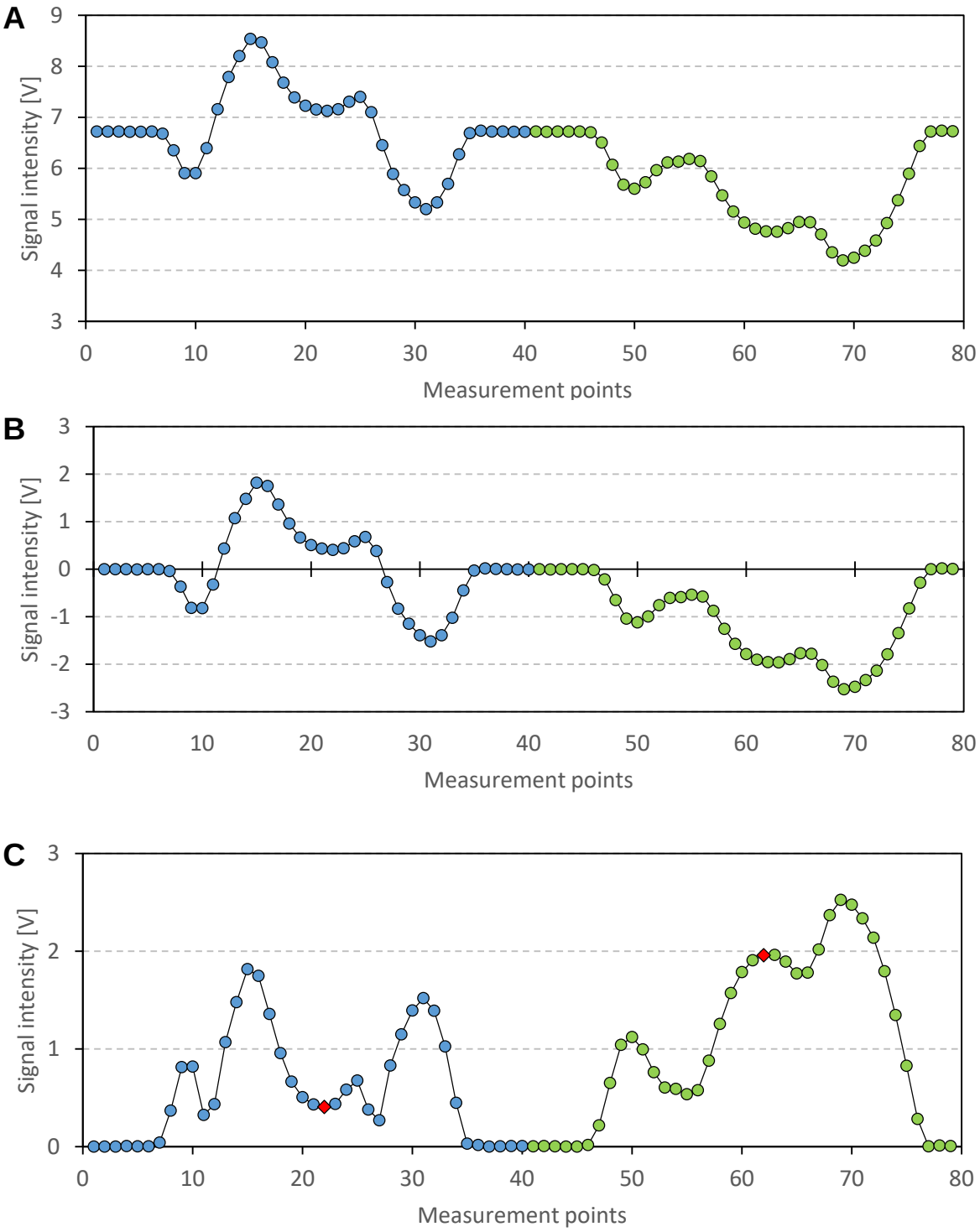


Figure 15: Peak signal of droplets without WST-1 formazan (blue) and with 1.5 mM WST-1 formazan (green). Signals intensities for droplet peaks are taken from raw data and fused together for comparison. (A) Raw data as received from photo detector. (B) Raw data with subtracted baseline. (C) Raw data with subtracted baseline and modulus. Points indicated by red square represent the middle point of a droplet and correspond to the concentration.

(3) A typical peak signal with the given flowrates consisted of 28 – 30 measurement points at a measurement rate of 10,000 Hz. This meant that a droplet usually took around 3 ms to completely pass the optical measurement point. The recommendation for the “peak width” -value was defined as half the number of measurement points that described the peak. The peak was then found by applying a quadratic function and by the method of least squares to

find its maximum. For the given case, this would result in a “peak width” - value of 15. However, this did not work reliably as the droplet peak itself consisted of several smaller peaks. Instead of reducing the peak width points, the number was increased to 50, forcing the peak finder to include 10 measurement points of the baseline before and after the actual peak signal. The result of this increase was that the least squares fitting performed better for peak values located in the middle of the signal. By increasing the peak width value even higher (e.g. 80), the peakfinder always found the middle point even it was a valley. However, the drawback of this approach was that the throughput of the subsequent screening process would be restricted to approx. 100 droplets per second. The higher the throughput, the more likely it was for two droplet signals to be close enough to fall within the range of the neighboring peak signal. This led to the identification of peak values at the baseline level, which again caused distortion in the droplet signal analysis. Therefore, it was necessary to create an own peakdetector subprogram, which could analyze the droplet signals more precisely and allow higher throughputs for the later screening process.

Establishment of a self-customed subprogram for identification of droplet peaks

As pointed out before, the main difficulty for the “peakdetector.vi” subprogram was to recognize the middle of the droplet signal. Since the peak in the middle was masked by other surrounding peaks, the subprogram needed to include many measurement points outside of the droplet signal to find the middle of the peak reliably. Therefore, the self-customed program should not identify peaks or valleys, but rather recognize which measurement points of the light intensity signal belong to one droplet signal and which belong to separate droplets. To efficiently find the middle of the peak, the light intensity signal was reduced by the baseline intensity and all values of the signal were converted to absolute values like shown before (Figure 15 B, C). The “peakdetector.vi” was replaced by three successive for-loops, which allowed the separation of peak signals and the reliable identification of the peak middle point. In the first for-loop, all 10,000 light intensity values of 1 second were checked whether they exceed a user-defined threshold (e.g. 0.5 V). All values exceeding the threshold were collected as indices. Since some of the values within a droplet signal might fall below the defined threshold (Figure 15 C, measurement point 27) a second for-loop was needed preventing the recognition of two droplets and recognizing that the two separated signals belong to one droplet. Therefore, the second for-loop exploited the fact that droplet signals were rather short compared to the baseline signal in between two successive droplets. The loop checked whether the distance between two peak signals was higher than the user defined threshold (e.g.10 measurement points). If the distance between the collected indices was higher than 1 ms (or 10 measurement points) it was very likely that the detected values, which exceed the threshold, belong to two different peak signals. If the distance between two peak signals was lower than one ms, the program recognized that the collected signals belong to one droplet signal and filled up the

missing indices in between the two droplet signals. Finally, in the third for-loop the length of each droplet signal was counted. This was achieved by subtracting the successive indices from each other. If the successive indices were serial numbers, the count for the length of the droplet signal was increased by one. If the successive indices were not serial numbers, the for-loop counting was stopped and the length of the droplet signal was forwarded before resetting to zero for the next droplet signal. By taking the median of the peak length and adding it to the indices of the first measurement point that belongs to the peak signal, the middle indices of the peak could precisely be determined. In case the peak length was an odd number, the number was rounded up, as this represented the middle point of the peak. After finding the index of the droplet peak, which was the relative position of the peak within 10,000 measurements per second, the indices were collected and taken back to the raw data signal to determine the corresponding light intensity value. By reducing the number of measurement points necessary to identify a droplet peak the throughput could be increased to 600 droplets/s.

Integration of online data analysis and export functions

The establishment of the self-customed peak detector increased the capacity of possible measurements per second approximately 6-fold. However, creating big amounts of data always goes along with the ability to manage the created amounts of data and being capable of handling it. LabVIEW was taken as the program of choice, because it was able to process data online and present it to the user in a lucid way. Peak data points found by the peak detector were counted and allowed to determine the throughput (droplets/s) and the overall measured droplets. Furthermore, all peaks were collected in a 2D-histogram and arranged according to their intensity. This allowed the user to get an overview of the droplet population and distinguish droplets with enhanced enzyme variants from droplets holding variants with lower activity. It allowed the user to monitor the progress of the enzyme reaction within the droplets and evaluate the separation between different enzyme variants. The 2D-histogram laid the foundation for deciding whether the parameters (e.g. reaction time, temperature) were well chosen or need to be improved to reach better separation. Once separation was achieved, a threshold could be defined within the program, allowing to count only droplets deceeding this threshold. By division, the percentage of droplets with increased activity could be determined, helping to project the total number of droplet hits expected in a sorting campaign. Besides the light intensity reflecting the productivity of the encapsulated enzymes, the size of each droplet was an important factor. Droplets should have a monodisperse shape, meaning they should be of roughly the same size to ensure the results from each droplet were comparable. Due to merging of two droplets, the size of a droplet can change substantially and by that the concentration of WST-1 formazan within the droplet, leading to false negative droplet hits. For long-term incubations, the risk of merging increased and it was therefore interesting to monitor how many merging events happened over the duration of the incubation period. Within the

program, the size of the droplet could be estimated by the number of measurement points, which reflected the peak signal of a droplet. For that all values above a user defined baseline threshold were counted. If two droplets merged, the number of measurement points for a droplet increased. For observation of the droplet size with respect to the light intensity, a 100 x 110 matrix was added to the LabVIEW program. At the start of each measurement, all entries of the matrix were set to 0. Each time the peak detector recognized a droplet, the number of measurement points (size) and the signal intensity (volt) were captured as a pair of values and used as coordinates for the matrix. Since coordinates for the matrix needed to be integers, the light intensity signal (which usually has one decimal place) was multiplied by 10 to allow correct positioning. In this way, the matrix added always one on top of the droplet count at the measured position, leading to a 3D-histogram. The data obtained for the 2D and 3D-histogram were exported into Excel-files, allowing later analysis of the signal intensities and size of the measured droplets (Figure 16). The total amount of droplets within the histogram was normalized. Therefore, the number of droplets with a specific light intensity was divided by the total number of droplets, giving the percentage of each subpopulation. This allowed the comparison of different datasets consisting of various sample sizes.

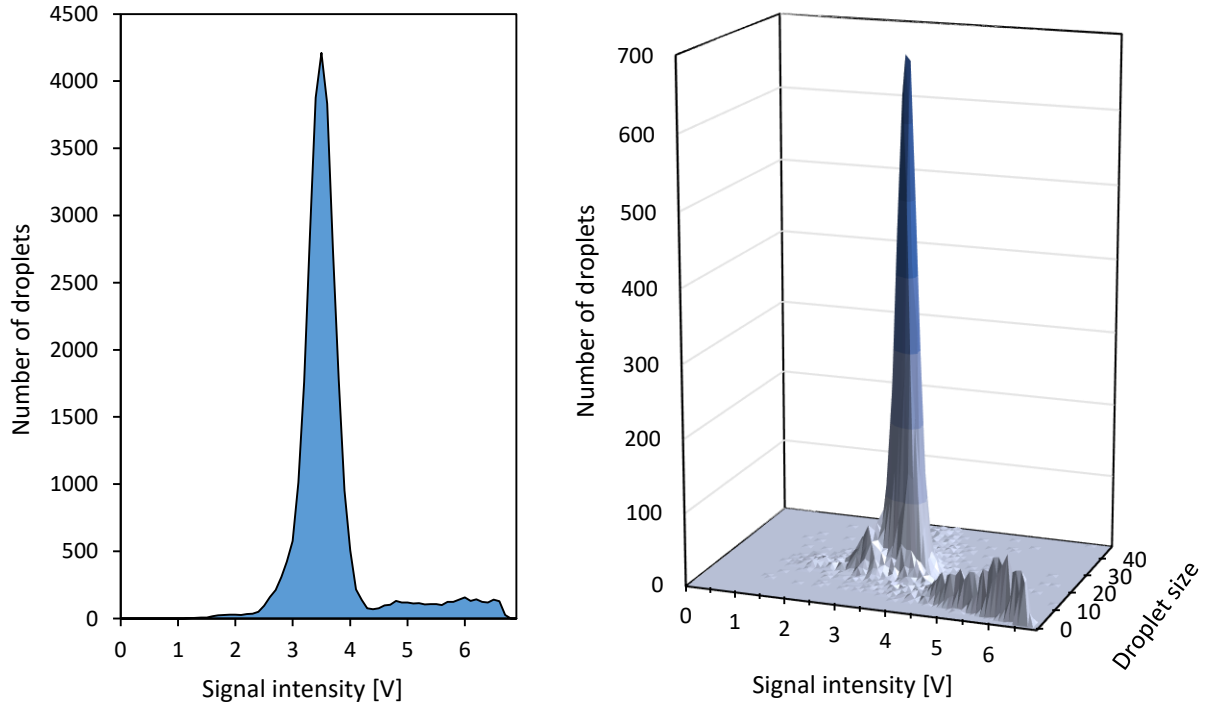


Figure 16: Example for exported 2D- (left) and 3D-histogram (right). The 3D-histogram allows identification of different droplet populations and can be taken to validate monodispersity of sorted droplets.

Calibration curve

The functionality of the program was proven by showing the identification of various dye concentrations. Droplets with 6 different tartrazine concentrations, ranging from 0 mM to 5 mM tartrazine, were produced. The baseline of the photodetector was set to 6 V and all droplet

populations were measured, one after another. Droplets with same tartrazine concentrations showed similar signal intensities and formed a peak in the 2D-histogram (App. 3.1.2.1). The average signal intensity of each population, representing one tartrazine concentration, was calculated and a standard curve was plotted (Figure 17).

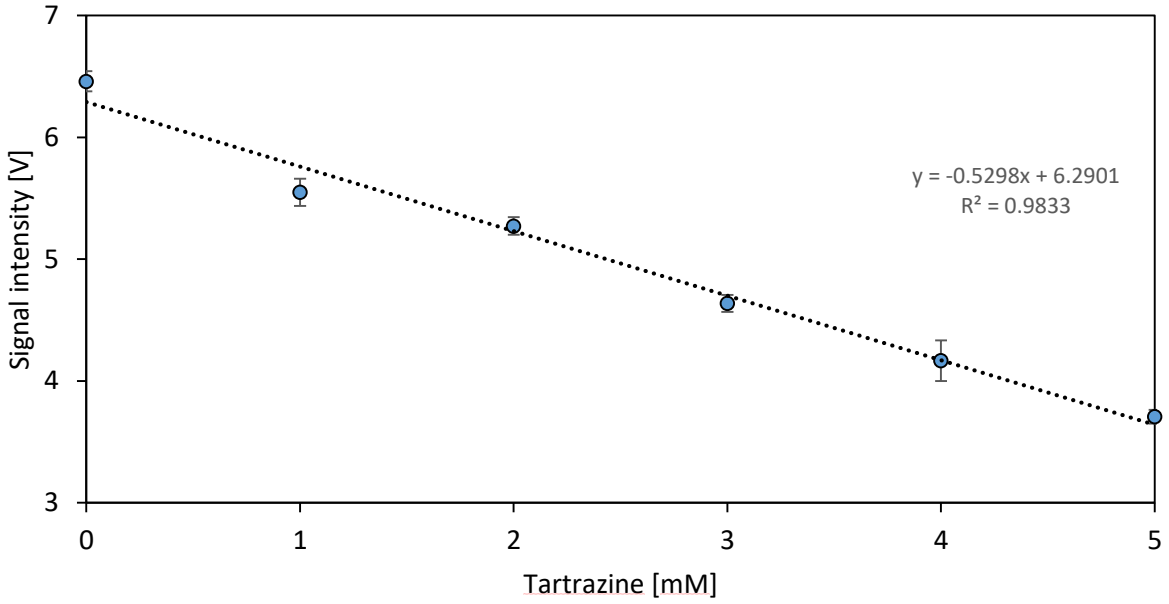


Figure 17: Standard curve for product formation against signal intensity in microfluidic droplets measured with the self-customed LabVIEW program.

As described by Beer-Lambert law, a linear standard curve was derived with a correlation index of 98.3%. For a baseline of 6 V, the signal intensity was reduced by 0.53 V for each millimolar of tartrazine produced by the cofactor coupled assay. Looking closer at the standard curve, it can be recognized that the point for 0 mM was above the standard curve and the point for 1 mM was under the average curve. This could be explained by the user-defined threshold, which was used to avoid peak recognition within the baseline noise.

3.1.5 Droplet stability

Leakage (chapter 1.3.3) is a phenomenon in microfluidics, which describes the process of unwanted leaving of a substance from the droplet sphere into the oil phase and possibly reentering in other droplets. Compared to merging (chapter 1.3.3), that can always be recognized by the droplet size, leakage cannot be detected in a subsequent sorting campaign. However, it was crucial to exclude or minimize leakage as it hinders the later evaluation of the enzyme turnover for each specific variant. In order to stabilize the droplets and prevent leakage, the amphilic surfactant molecule Picosurf1 was used. The surfactant stabilizes the droplet surface and hampers the assay dye to leave the droplets. To test if droplet leakage of WST-1 dye was prevented with the surfactant, droplets with 2 mM and 4 mM tartrazine were

produced. Both droplet populations were produced with 1.5% Picosurf1 and incubated in the same chamber, allowing close contact between the two droplet species like in subsequent sorting campaigns. If tartrazine leaks from one droplet to another, the droplet population with 4 mM tartrazine should reduce its tartrazine concentration and show higher signal intensity over time. If the dye can also reenter through the droplet surface, droplets with 2 mM tartrazine should show reduced signal intensity, absorbing more light. After a certain amount of time the separate populations would form a joint peak, making it impossible to distinguish the two populations.

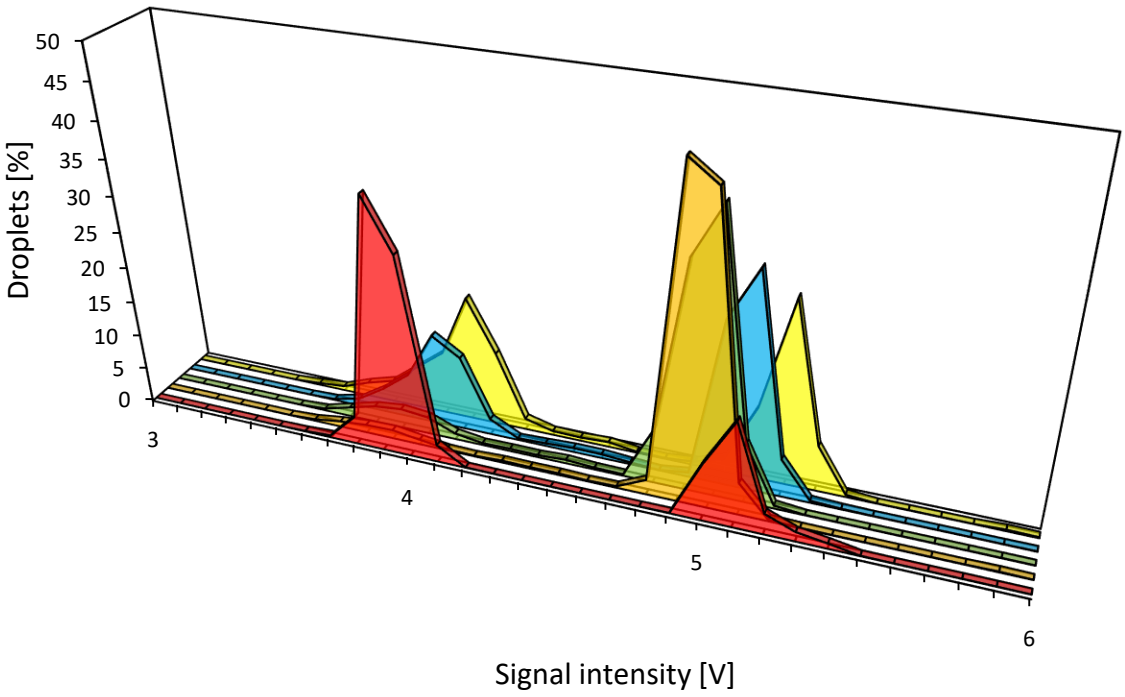


Figure 18: Droplet leakage at room temperature. Droplets with 2 mM and 4 mM tartrazine were produced and incubated at room temperature. Samples were measured after 0 hours (red), 2 hours (orange), 4 hours (green), 6 hours (light blue) and 24 hours (yellow) to investigate droplet leakage.

At first, droplets were incubated at room temperature. Samples were initially taken every two hours and after 24 hours to investigate long-term effects (Figure 18). For each measurement time the average peak signal for the two populations was calculated and checked whether the distance between the two populations decreases. For the time of investigation, the signal intensity difference between the two populations stays at a constant level of approx. 1 volt. The distance between the two peaks was alternating within 0.98 – 1.16 volt, but no clear tendency was indicated. Consequently, for the observed time no leakage occurred. However, looking closely at Figure 18, it can be observed, that some droplets appeared around 4.5 V, right in the middle of the two populations. These droplets were most likely merged droplets. In

comparison to the main two populations, only 2% of all measured droplets were in between the two main populations, meaning that 98% of the droplets were still monodisperse. Throughout this work, droplets were also incubated at higher temperatures, enabling the investigation of temperature stability and the measurement of enzyme activities for thermophilic enzymes. Therefore, the droplet stability was investigated under these conditions. For better distinction of the droplet populations, the baseline signal intensity was increased to 8 V and the product concentration was adapted to 0 mM and 5 mM tartrazine.

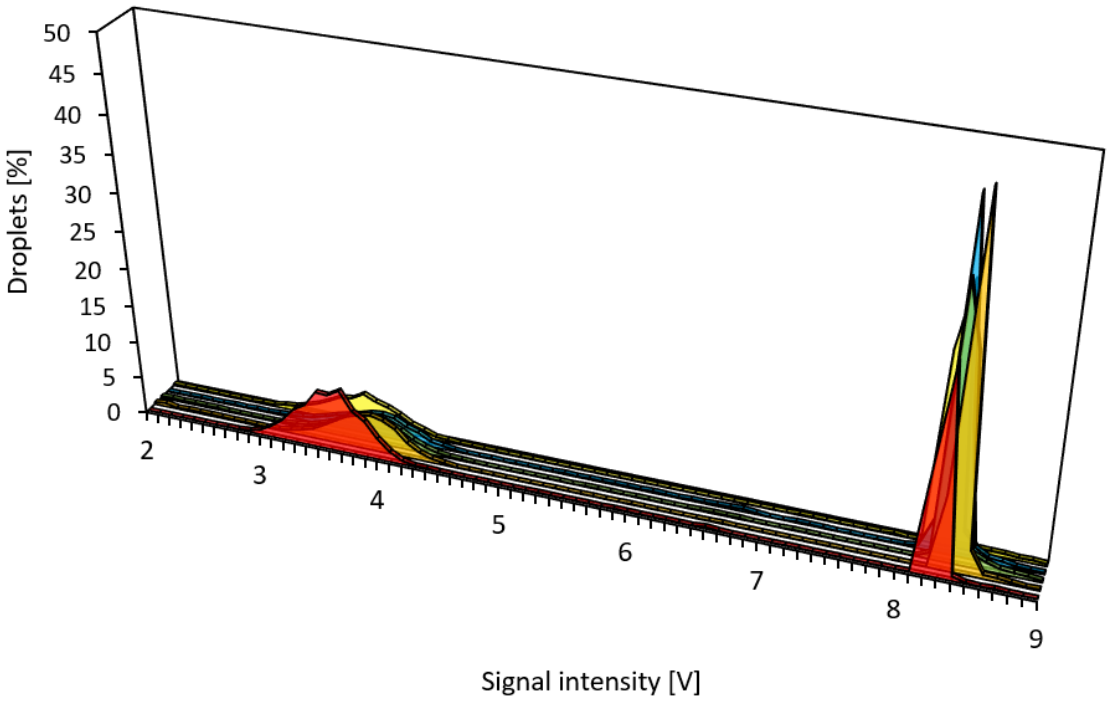


Figure 19: Droplet leakage after incubation at 60 °C. Droplets with 0 mM and 5 mM of tartrazine were produced and incubated at 60 °C. Samples were measured after 0 hours (red), 2 hours (orange), 4 hours (green), 20 hours (light blue) and 27 hours (yellow) to investigate droplet leakage.

As shown in Figure 19, the signal intensity stayed on a constant level even after 27 hours of incubation at 60 °C. Looking closer at Figure 19 it can be observed that the droplet population at 4 V for first measurement point (red) showed a higher percentage at the droplet population. This can be explained by the fact that this population was produced second and some of these droplets remained within the tubing of the chamber when mixing the droplets. After reinjection and measuring of the droplets, the overall percentage of this population was increased.

3.1.6 Droplet sorting

Once the software for the droplet assay readout was established, it enabled the distinction of droplet populations based on their product formation. This knowledge was required to

successfully separate droplets according to their signal intensity by sorting. Droplets, which showed a product formation higher than the user-defined threshold, were isolated by a dielectrophoretic pulse. This pulse dragged the droplet towards the electrodes and allowed changing the flow direction into the sorting channel of the microfluidic chip. At the end of this sorting channel only droplets, which fulfil the sorting criteria were collected. For the sorting, the light intensity of the droplets was measured with a photodetector, which converts the light signal into an analog signal. The signal was then split and guided to the A/D-converter for measurement (3.1.4) and a microcontroller (Arduino due) for sorting. The microcontroller was programmed in such a way, that a short signal pulse was released every time the light signal drops under a user-defined threshold. From the microcontroller the signal pulse was forwarded to a pulse generator, which modified the pulse width and the amplitude of the incoming signal pulse. The signal was then further transferred to a function generator, which converted the pulse to a sine wave with a defined frequency and amplitude. At last, the function generator was connected to high voltage amplifier (HVA), which amplified the signal 100 times and was therefore powerful enough to sort droplets dielectrophoretically. An additional connection between the HVA and the A/D-converter (NI USB-6009) was established. In contrast to the sorting signal, this connection sent an unamplified voltage signal to the A/D-converter, which was connected to the computer with the measurement program. This allowed monitoring of the sorting signal and later also signal synchronization and adaption of the electric field. For the sorting, electrode channels were filled with salt solution (5 M sodium chloride). Jumper wires were inserted in the positive and the ground electrode and connected with wire clamps to the HVA. Initially, the parameters published for the sorting chip were tested (Table 4).

Table 4: Sorting parameters for different droplet sizes and flowrates. Pulse delay not mentioned for Gielen *et al*.[78]

Parameter	Gielen et al. 2016	Set1	Set2	Set3
Droplet diameter	70 µm	82 µm	82 µm	82 µm
Flowrate spacing oil/ droplet inlet [µl/min]	30/2	25 / 0.4	30 / 0.8	30 / 2
Pulse width [ms]	5	25	25	12
Pulse delay [ms]	0 (?)	5	4	4
Sorting signal [V]	1000V	500V	500V	1000V
Droplet throughput [droplets/s]	100/s (300/s)	~10/s	30-40/s	100/s

However, these parameters only showed a low sorting efficiency. The reason for this might be the differences in the droplet diameter. For the measurement, the channel width was reduced to 50 µm. This led to a defined path length and standard deviations in droplet diameter did not interfere with the absorbance measurement according to Beer-Lambert law. However, since the droplet diameter for droplets produced in this study was approximately 12 µm larger, the droplet velocity was further reduced due to squeezing along the walls of the measurement

channel. Therefore, it was necessary to delay the pulse signal and initially increase the duration of the pulse. For observation purposes, the throughput of the droplets passing the measurement channel was drastically reduced to 10 Hz to assure that only the desired droplet was sorted. The voltage intensity was reduced since the droplets have more time to be dragged in the sorting channel. Once the parameters for the sorting at 10 Hz were found, the throughput was increased by doubling the reinjection flowrate of the droplets. This led to a 4-fold higher throughput and a reduced pulse delay. The throughput of the droplets was further increased until a throughput of 100 droplets per second was reached. For a successful sorting the pulse duration was shortened to avoid sorting of several droplets. The voltage intensity of the electric field was increased to 1000 V. All attempts to further decrease the pulse duration, resulted in lower sorting efficiency. To proof that functional parameters were obtained, the sorting efficiency was investigated with the parameters summarized in Table 4, set 3. As a test experiment, three different variants of *Bs*GDH strain 168 (template variant (Q252L, E170K), template variant with additional 258N substitution, and template variant with additional 258D substitution) were mixed, encapsulated and incubated at 50 °C for 24 hours to test also the heat stability of droplets for the designated parameters. Due to their different turnover rate and heat stability (Appendix 3.1.6) a broad range of activities could be measured and enabled a very stringent sorting (Figure 20, orange).

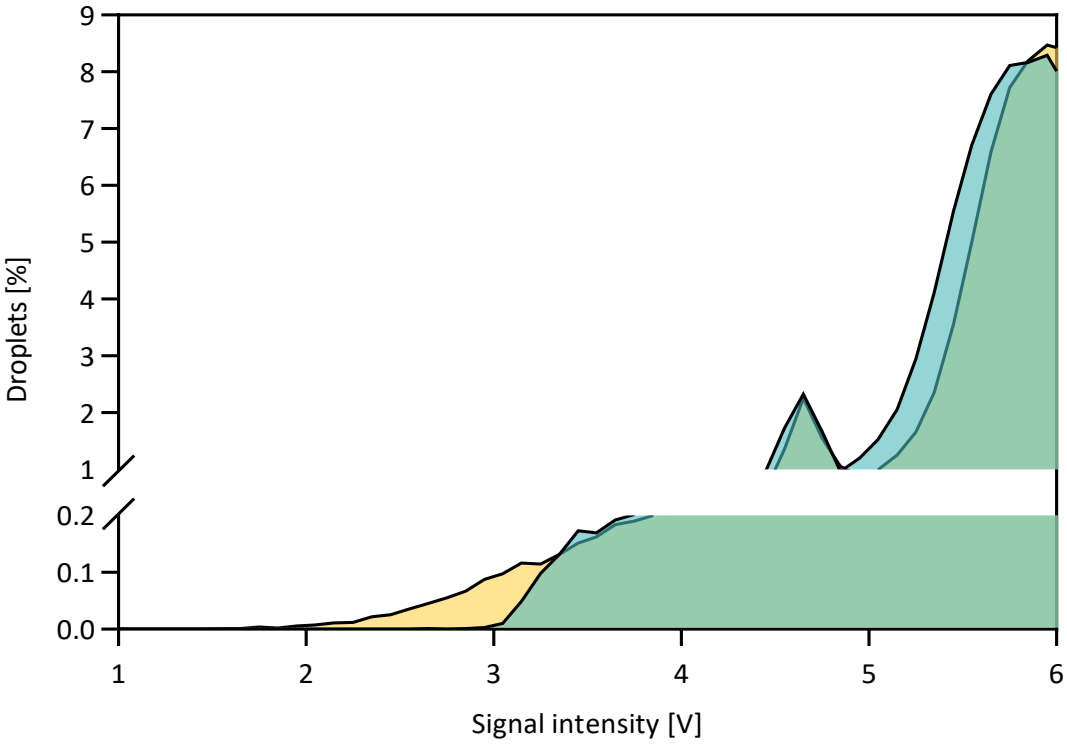


Figure 20: Testing sorting efficiency: Variant mix (orange) was initially sorted with a threshold of 3 V. Droplets that normally go to waste were recaptured in a new chamber and measured subsequently to identify sorting efficiency (blue).

For the sorting a threshold of 3 V was defined. In total 317.999 droplets were measured and every droplet which showed a light signal below the threshold was sorted. This way, 1236 droplets (~0.4% of all droplets) were sorted. In order to control if all droplets which fulfilled the criterium were successfully sorted, the droplets which usually were discarded were captured in another chamber and reinjected into the microfluidic sorting chip for a second measurement round (Figure 20, blue). From the 316.763 (317.999 – 1236) captured droplets, exactly 240.000 droplets could be measured. The remaining droplets in the tubing of the chamber could not be reinjected in the measurement chip due to buoyancy forces. From 240.000 reinjected droplets, only 22 droplets showed a light intensity below 3 V. Out of this 22 droplets, 8 droplets showed a light intensity of 2.95 V and were therefore probably just increased the product formation during the time they were measured for the first time until the moment they were reinjected for the second round. This shift in product formation can also be observed by comparing the empty droplet peak in the histogram, which appear a little bit more to the left for the second measurement (blue) compared to the first measurement (orange). Taking this into account a false negative rate of 1.13% (14 out of 1236) and a true negative rate of 99.99% (239.986 out of 240.000) was reached with the sorting experiment. The true positive and false positive rate could not be investigated, as the number of droplets sorted was

too low for reinjection. However, a strong difference in the color of sorted and unsorted population could be observed by eye (Figure 21).

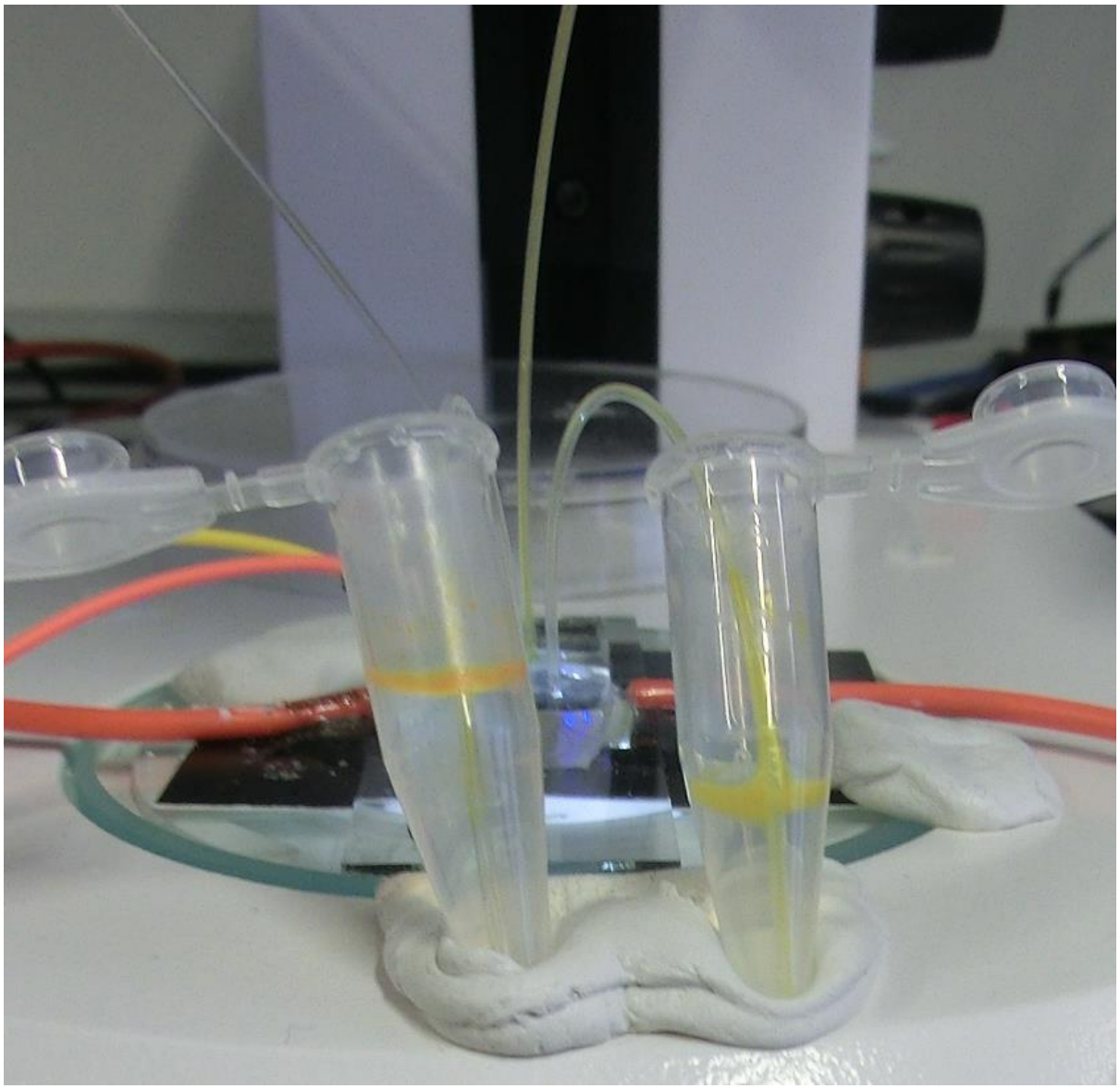


Figure 21: Visible separation of droplets with improved enzyme activity appearing in dark orange (left tube) and droplets with low enzyme activity appearing in yellow (right tube).

3.2 Engineering *Bs*GDH towards altered substrate specificity

After building up the microfluidic system and establishing the software for the online supervision and the subsequent analysis of the encapsulated enzyme variants (3.1), the next key step was testing the functionality of the established system. Therefore, enzymes with the following three requirements were needed: First, they need to catalyze a cofactor dependent reaction, which in turn allows WST-1 formazan production by the coupled assay. Second, the substrate needs to be oxidized and the cofactor needs to be reduced to allow cofactor recycling via the electron coupling agent mPMS (1-methoxy-5-methylphenazinium methyl sulfate). Lastly, the enzymes should have a stable expression and a sufficiently high activity, as one cell needs to express enough enzyme to monitor the differences in product formation over time. The glucose dehydrogenase from *Bacillus subtilis* strain 168 (*Bs*GDH template variant)[138] was chosen for this task as it fulfills all above mentioned requirements. With a V_{max} of 86 U/mg, it shows a good activity and affinity towards the substrate glucose (Figure 22). In this study, engineering the substrate specificity of the glucose dehydrogenase towards xylose conversion was chosen as an interesting application for engineering the *Bs*GDH. Initial tests revealed that *Bs*GDH was active on xylose, but the substrate affinity for xylose (K_M^{xylose}= 1099 mM, k_{cat}^{xylose} = 75 s^{-1}) was around 220 times lower compared to that for the native substrate glucose ($K_M^{glucose}$ = 5 mM, $k_{cat}^{glucose}$ = 47 s^{-1}). Consequently, high concentrations of xylose were necessary to reach the same activity with *Bs*GDH (Figure 22). Therefore, increasing the substrate affinity towards xylose was an interesting aim for enzyme engineering.

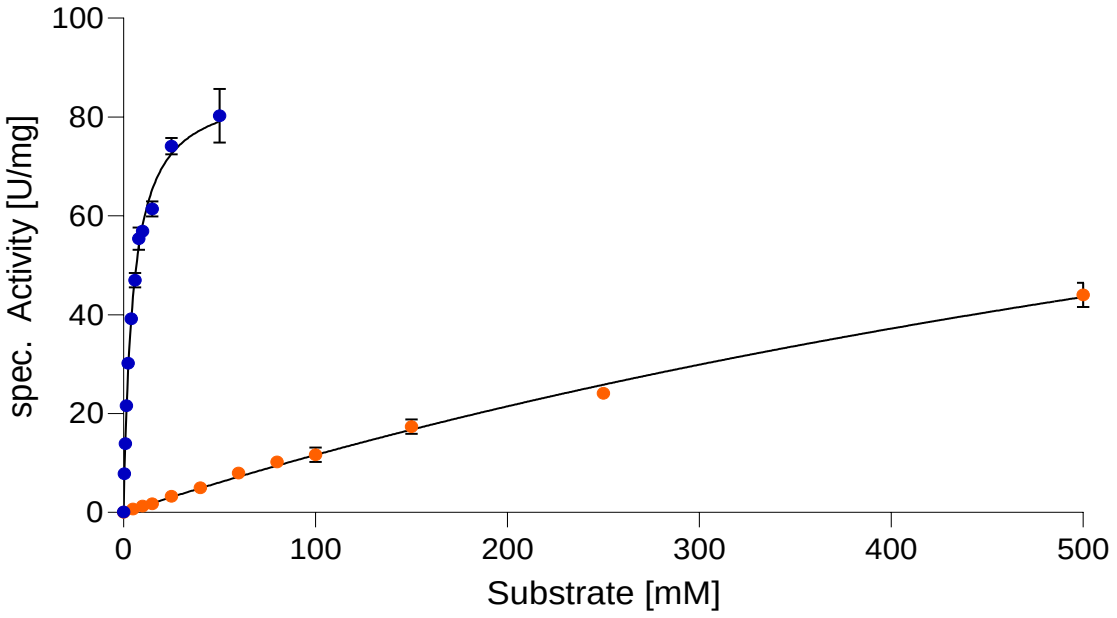


Figure 22: Michaelis Menten Kinetics of *Bs*GDH template variant with glucose (blue) and xylose (orange). Kinetics were performed with 1 mM NAD and the reaction velocity was determined by photometrical measurement at 340 nm, 25 °C.

After the performing initial tests in microtiter plates with purified enzyme, the feasibility in microfluidic droplets was tested. For that, microfluidic droplets were produced with a high xylose concentration of 400 mM, allowing an initial reaction velocity of up to 55 U/mg. This allows *Bs*GDH to convert nearly at the same reaction velocity as it would with glucose, even at lower substrate concentrations. Instead of purified enzyme, single cells expressing *Bs*GDH template variant were encapsulated into droplets. Initial incubations were performed with a droplet occupation of $\lambda = 0.1$ (meaning that 10 times more droplets than cells were used for encapsulation). According to Poisson distribution this results in 9.1 % single occupied droplets, but also 90 % unoccupied droplets. The occupation rate was chosen comparably low to keep the amount of multiple occupied droplets low (~1%). Droplets were produced and incubated at room temperature allowing cell lysis. Enzymes were released from the cell into the droplet sphere and the enzyme reaction was initiated. A portion of the whole droplet population was measured at different time points to monitor overall product formation over time. The first measurement showed that right after the production a signal intensity of around 6 V can be observed in the histogram (Figure 23, orange).

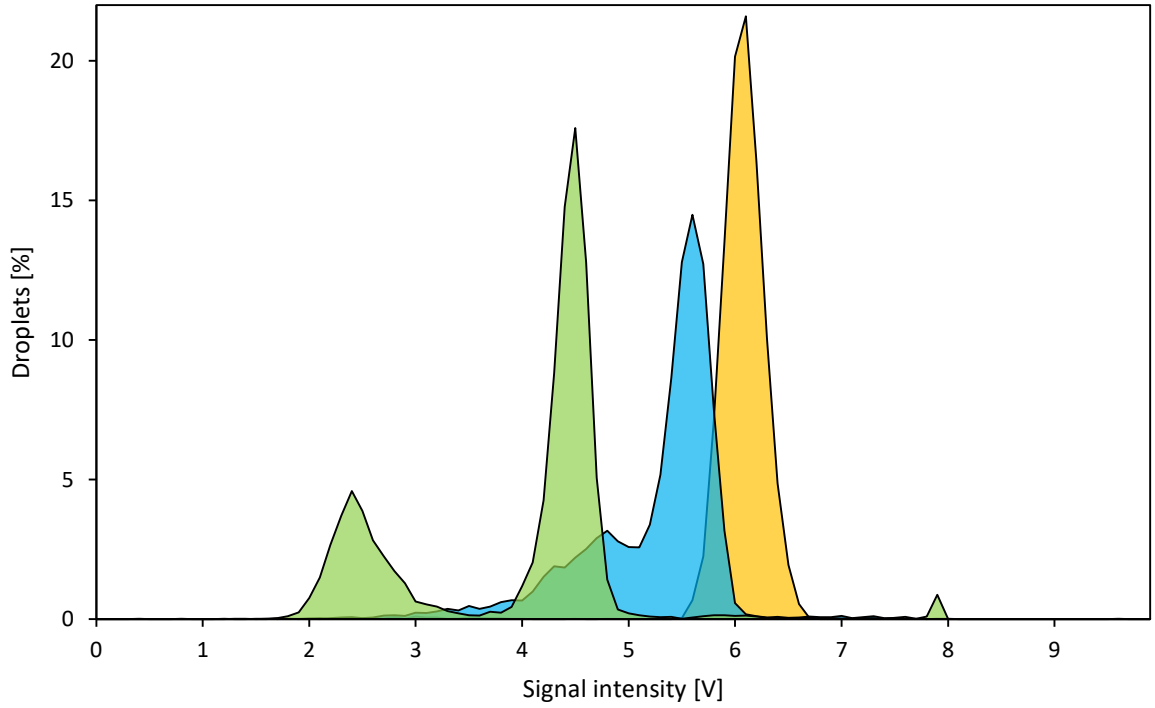


Figure 23: Activity of *Bs*GDH template variant with 1 mM NAD and 400 mM xylose in microfluidic droplets. Measurements were taken directly after production (orange), 7.5 hours after production (blue) and 25 hours (green) at 25 °C to observe product formation.

At this point, the two droplet populations could not be distinguished directly after production, since no (or not enough) product was formed yet. The droplets were measured again after 7.5 hours of incubation at 25 °C. It could be observed that the single peak moved to the left side and a shoulder has formed on the left side representing the occupied droplets, which contain more WST-1 formazan than the background of the vacant droplets (Figure 23, blue).

Over the course of the next measurements, this shoulder was moving further to the left, indicating more product formation (data not shown). After 25 hours, the shoulder was clearly separated from the bigger peak at the right (Figure 23, green). For the measurement after 25 hours, it can be observed that the wide left shoulder from the measurement at 7.5 hours was separating from the peak of unoccupied droplets and formed another distinct and more defined peak. This could be explained by the full conversion of WST-1 to its formazan dye. All droplets with full conversion of WST-1 formazan were gathering at a signal intensity of ~ 2.3 V. This first experiment showed that the enzymatic turnover of the *Bs*GDH template variant could be monitored with the WST-1 assay. It was possible to distinguish occupied and vacant droplets. Over the course of 25 hours, a background conversion of 0.13 mM/h was observed for vacant droplets. In contrast to that, the occupied droplets showed an average conversion rate of 0.4 mM/h, which was around 3 times higher than the background reaction. After showing that differentiation between occupied and vacant droplets was possible with a high substrate concentration of 400 mM xylose, the substrate concentration was reduced to 100 mM xylose. This adjustment was made since the ultimate goal of the project was to engineer the enzyme towards higher substrate affinity. For the purified enzyme this substrate reduction would cause a decreased reaction velocity of approx. 27 U/mg. Droplets were produced in the way like described above, just with reduced xylose concentration.

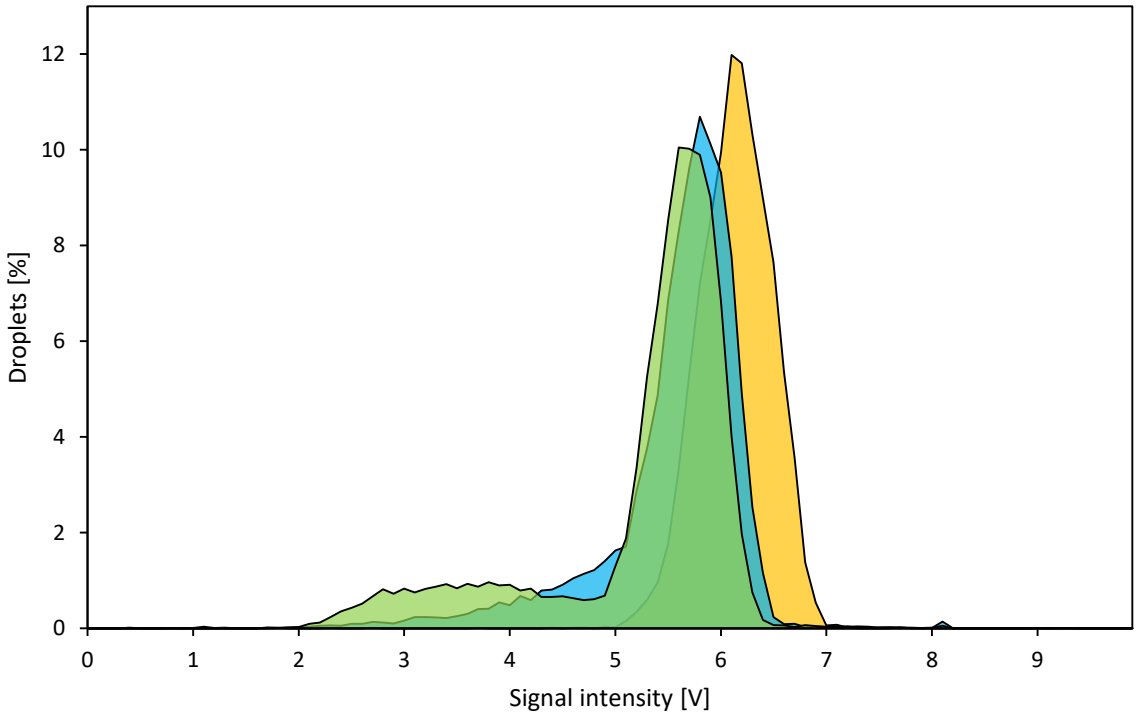


Figure 24: Activity of *Bs*GDH template variant with 1 mM NAD and 100 mM xylose in microfluidic droplets. Measurements were taken directly after production (orange), 21 hours after production (blue) and 44.5 hours (green) to observe product formation at 25 °C.

As expected, product formation in the droplets was slowed down. Within the first 21 hours, no clear separation of the occupied and vacant droplets was possible. However, a clear shoulder could be detected, which indicated substrate conversion over time (Figure 24, blue). After 44.5 hours, the shoulder moved more to the left and a notch could be recognized around 4.8 V, stating that less droplets were counted at 4.8 V than at 3.5 V (Figure 24, green). As an interesting side effect, it could be observed that even the experiment was running for almost 2 days, the background reaction of the assay for 100 mM xylose was lower than for 400 mM xylose substrate concentration. This would allow even longer experiment times, without possible concerns that the background reaction and the enzyme reaction in the droplets cannot be distinguished. Examination of the droplets after 45 hours through the light microscope showed that droplets were stable over the whole time of the experiment.

Summarizing the first two experiments, it could be proven that droplet separation was possible with high xylose concentrations. By decreasing the substrate concentration, the reaction velocity was drastically reduced and cannot be compared to the enzyme kinetics from the microtiter plate anymore. Occupied and vacant droplet populations could not be clearly distinguished anymore since the total enzyme turnover after 45 hours in some droplets was comparable to droplets with high background formation. However, the tested parameters gave a good starting point to explore improved variants in enzyme libraries. These variants would be expected to have a higher turnover rate and would therefore produce more formazan dye in shorter time, showing up further at the left side of the signal intensity histogram. The microfluidic screening setup proved to be stable over the whole experiment time of 45 hours.

3.2.1 Proof of concept with QuikChange library

As a proof of concept a small library was created, which allowed testing of the microfluidic setup for its efficiency to sort improved variants. To control the sorting results, the library was designed small enough to be screened using microtiter plates. This allowed verification whether all improved variants were found with the microfluidic system. For that purpose, a QuikChange library from *Bs*GDH was created by Samuel Sutiono. The QuikChange library allowed the alternation of a nucleotide triplet, which resulted in different amino acids at one desired position. For *Bs*GDH, position 258 was chosen as it was previously reported by Li *et al*.[139] to change substrate affinity of glucose dehydrogenase from Bacillus XY-1 sp towards xylose. A NNS-motif was chosen for the QuikChange approach leading to a potential library size of 32 variants. According to nBLAST, both proteins share 79 % of their nucleotide sequence, increasing the likelihood that the amino acid exchange for *Bs*GDH strain 168 will also result in an enhanced xylose affinity. At first colonies from the library were randomly picked and investigated in a microtiter plate screening, allowing a total sample size of 84 variants and

12 controls. Investigating 84 variants resulted in a theoretical oversampling of 2.6 times. Within the microtiter plate screening 13 out of 84 variants (15.5%) showed significantly improved activity, compared to the template variant (Figure 25 top). Sequencing of the variants revealed that substitution from alanine at position 258 towards histidine (258 H) and glutamic acid (258 E) resulted in two times higher enzymatic activity under the given assay conditions. Substitution with aspartic acid (258 D), glycine (258 G) and asparagine (258 N) increased the turnover rate to approximately three times higher activity. After the initial screening, all variants were exposed to heat challenge treatment (30 min, 70 °C) to investigate if thermostability could be retained (Figure 25 bottom). Interestingly, the variants with asparagine, glycine and histidine substitution retained the thermostability and showed 75 % - 90 % of their initial activity. Variants with aspartic acid and glutamic acid substitution could only retain 4 % - 16 % activity after the heat treatment and were therefore less efficient than the template enzyme.

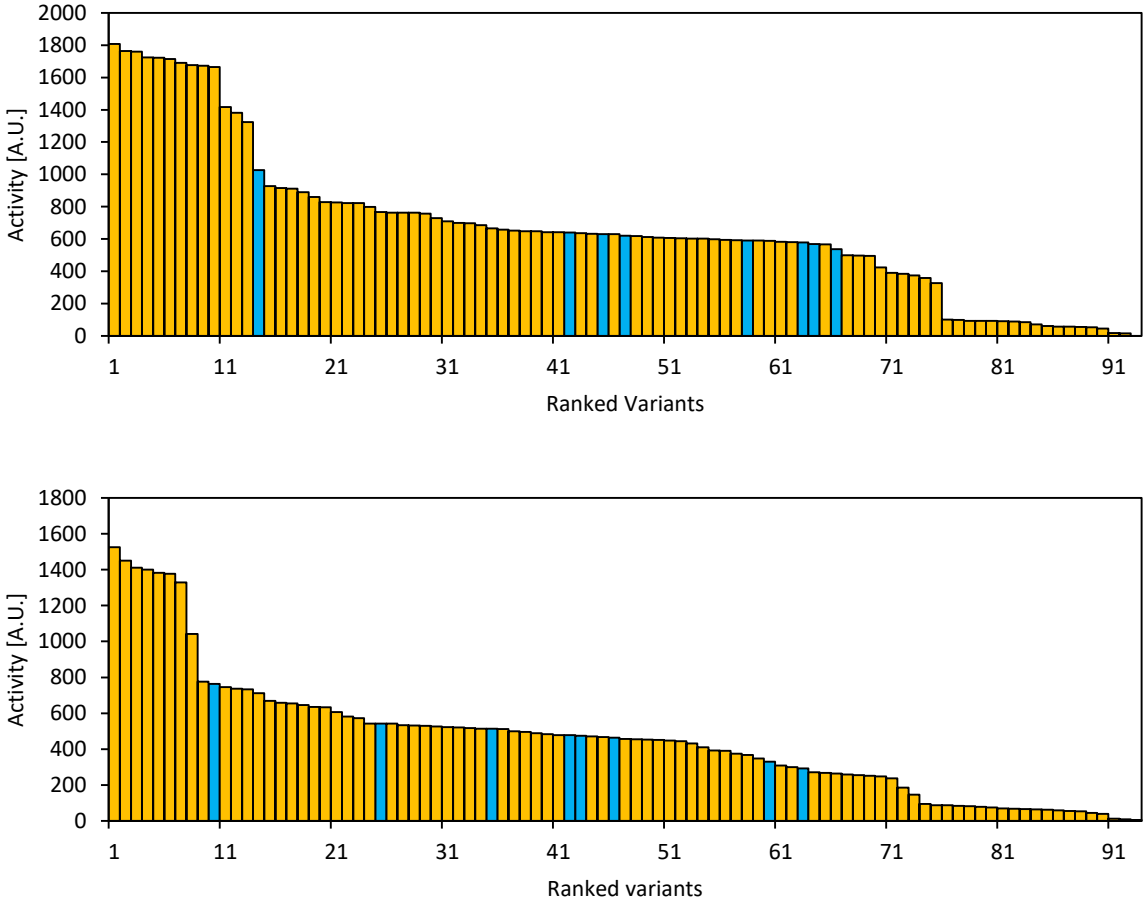


Figure 25: QuikChange library screening of *Bs*GDH in microtiter plates before (top) and after heat treatment (70 °C, 30min) (bottom). Variants (orange) and template controls (blue) were ranked according to their activity.

The microtiter plate screening proved that the library contains possible hits. The QuikChange library was encapsulated in droplets and prepared for sorting with the microfluidic system. Therefore, single cells of the library (λ=0.1) were encapsulated in droplets containing 100 mM xylose and 1 mM NAD^+ as previously performed with the parental template enzyme. As shown

before, incubation of wild type enzyme for 21 hours resulted in a decent shoulder formation. Expecting two to three times higher activity from the improved variants the droplets were incubated over night for 14 hours, prior to measurement. In comparison to the wild type measurement after 21 hours, the library measurement after 14 hours showed already an extended shoulder with some variants already reaching the maximal conversion rate at 2 V. Right after the measurement the sorting was initiated. The threshold for the sorting was set to 4 V, which corresponds to the best 3 % of all droplets measured after 14 hours (Figure 26).

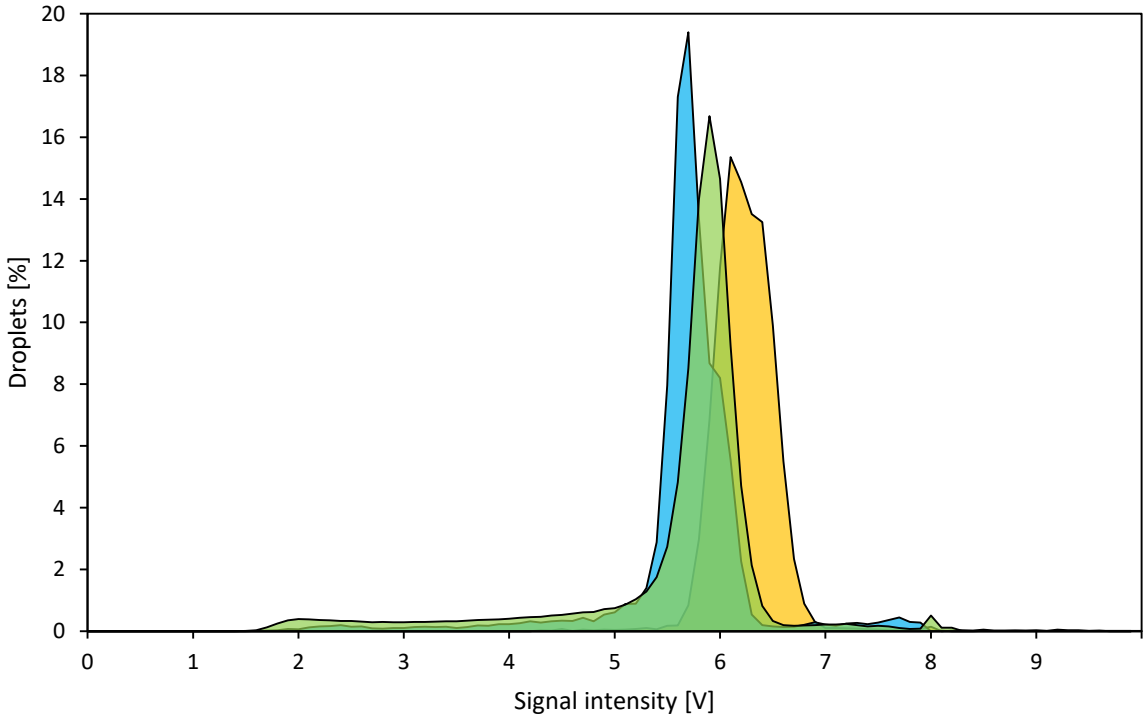


Figure 26: Normalized histogram of signal intensity distribution for *Bs*GDH QuikChange library after defined times. Measurements were taken right after droplet production (orange), after 14 hours of incubation (light blue) and for sorting after 17 hours (green).

Within the next 3 hours, 623.966 droplets were measured and 48.089 droplets were sorted. This corresponds to an overall sorting rate of 7.7 %. The rise in sorted droplets could be explained by the ongoing reaction within the droplets during the sorting. The DNA of all sorted variants was amplified via PCR and used to transform *E. coli* BL21 (DE3). After transformation, cells were plated and grown over night. 54 single colonies were picked to rescreen the sorted variants in microtiter plates.

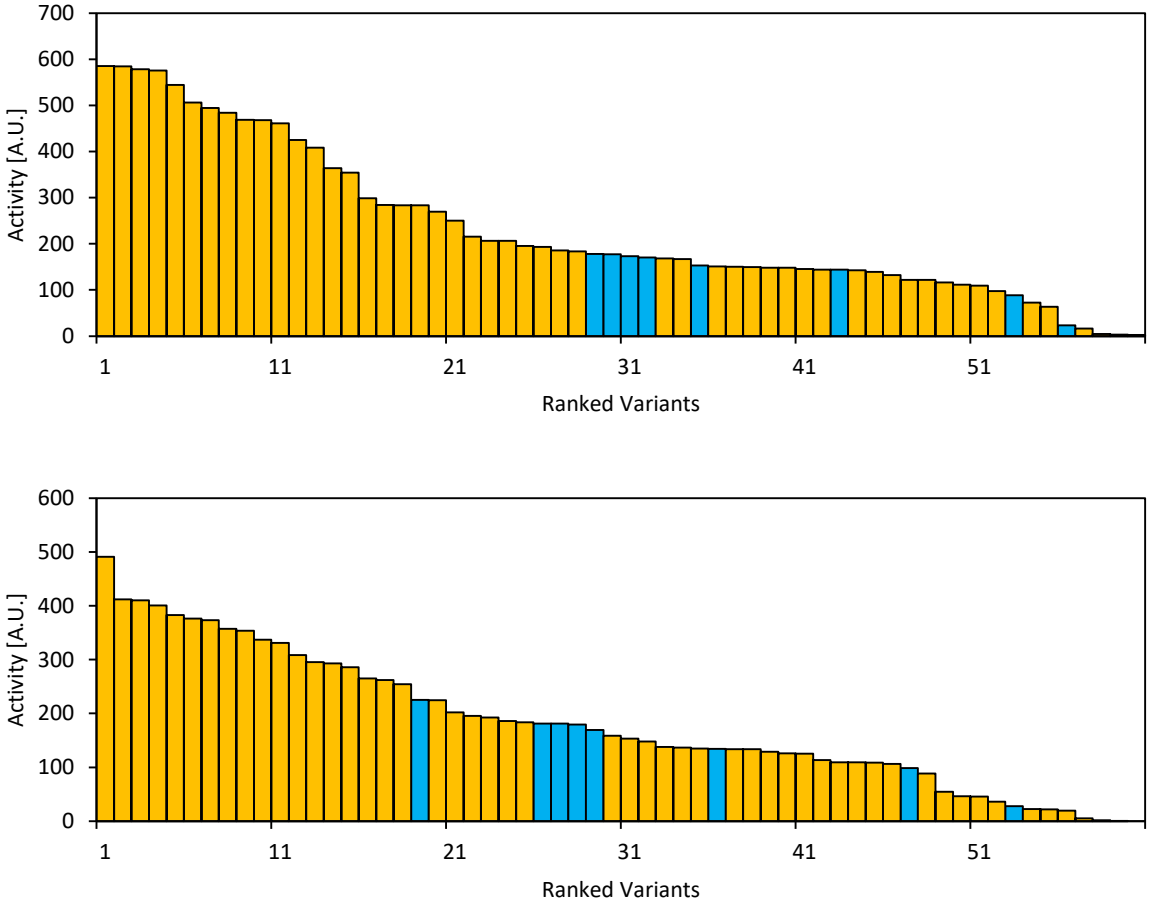


Figure 27: Activity measurement of *Bs*GDH QuikChange variants sorted by the microfluidic system before (top) and after heat treatment (70 °C, 30 min) (bottom). Variants (orange) and template controls (blue) were ranked according to their activity.

Rescreening of the microfluidic sorting led to 22 variants (40.7 %) with significant increased activity compared to the parental template (Figure 27 top). Compared to the random microtiter plate screening with 15.5% improved variants a 2.6-fold enrichment of improved variants could be achieved. However, in the second screening round the majority of variants is expected to show improved activity. This rather low percentage of hits for the rescreening might be explained by the rather conservative sorting threshold, which allowed 7.7% of all droplets to be sorted. With an overall occupation of around 15%, this resulted in the top 50 % of all occupied droplets to be sorted. Thereby, the chance to sort also template variants or enzyme variants with similar activity was rather high. To check for heat stability a heat treatment (30 min, 70 °C) was applied for all sorted variants from the microfluidic sorting (Figure 27 bottom). From 22 variants, 15 retained their temperature stability and showed significantly higher activity than the wild type. From these 15 variants, 6 hits were chosen for sequencing. The sequencing proved the findings from the microtiter plate screening. Three sequences showed a mutation at position 258 towards histidine and another three sequences showed the mutation towards glycine. As expected, the thermolabile variants with aspartic acid and

glutamic acid could not be found. Overall, four out of five improved variants found with microtiter plates could also be identified with the microfluidic sorting system. Only the thermostable variant A258 N could not be retrieved within the 6 sequences, which might be a consequence of the low sequence number in the rescreening.

3.2.2 Proof of concept with epPCR library

In the previous chapter, it was shown that the microfluidic system could identify and separate improved variants within a small library. Therefore, the next logical step was to increase the number of possible variants. For that an error prone PCR (epPCR) was performed with the glucose dehydrogenase from *Bacillus subtilis* by Samuel Sutiono. Here, the number of variants exceeded the capabilities for a microtiter-based screening system. One of the best performing variants from the first screening round, in this case A258N, was chosen as a template for the epPCR to further improve the catalytic efficiency of the enzyme towards xylose. In preparation for the epPCR library production, four different manganese chloride concentrations (0.125 mM, 0.25mM, 0.375 mM and 0.5 mM) were tested. Manganese chloride hinders the DNA polymerase from accurately replicating the DNA. As a result, the concentration of manganese chloride has a direct influence on the later mutation rate of the polymerase and such the number of mutations within the variants of a library. However, increasing mutation rates within an enzyme coding gene also increases the chance to express non-functional enzymes, which do not show any or lower activity than the parental template. For that, the average activity of the parental enzyme lysate was determined and compared with the lysate activity of randomly picked colonies from the four different epPCR approaches. Variants which showed at least 20% activity compared to the parental template were counted as active variants. With a manganese concentration of 0.375 mM, the average mutation rate was calculated from 10 sequenced variants and determined to be around 4 mutations per gene. The fraction of enzymes with remaining activity was determined to be around 12 %. To identify variants with an improved xylose affinity, the concentration of xylose was further lowered to 20 mM xylose for the epPCR screening. For comparison, the library and the template variant (A258N) were encapsulated in two populations of droplets and their activity over time was compared. After 26 hours at room temperature, the droplet population with the template variant clearly separated from the empty droplets (Figure 28, blue). As expected, the majority of droplets from the epPCR library population showed minor activity. Only a very low percentage of droplet encapsulated variants showed activity comparable or even higher than the template variant (Figure 28, bottom, orange). The sorting threshold was set rather conservative to 4 V since earlier experiments showed that recoveries with a low number of sorted droplets and therefore a low concentration of DNA might fail. As a results, a high false positive rate, with low enrichment was expected. In the microfluidic screening a total of 468.517 droplets was screened and 2.151 hits were sorted. This corresponds to the best 0.46 % of all droplets.

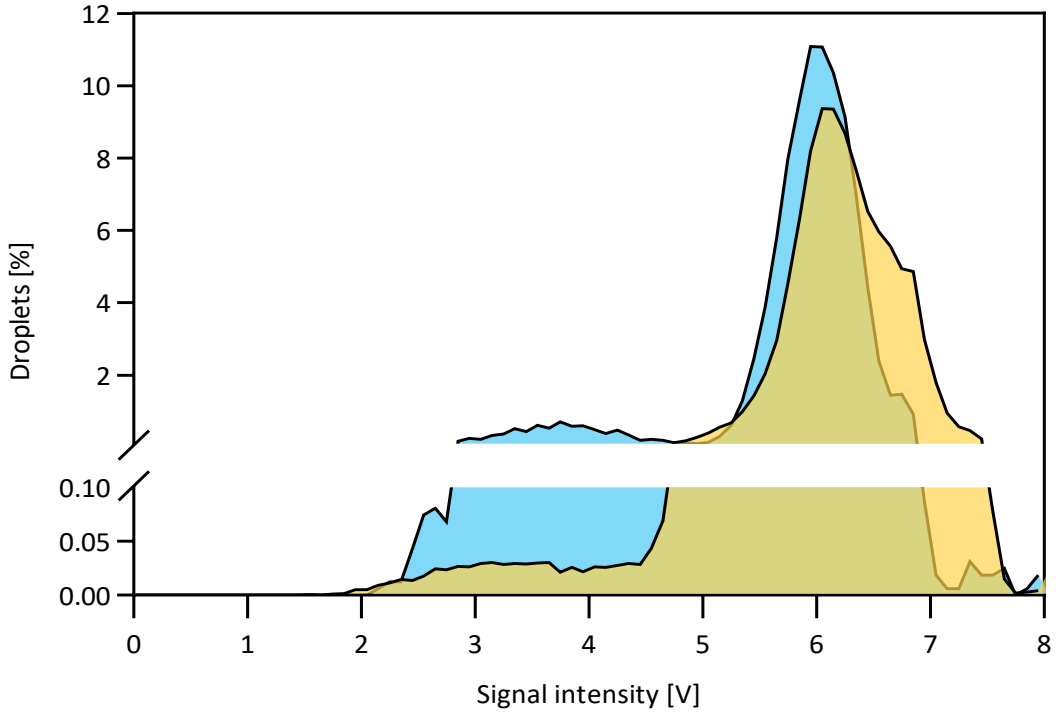


Figure 28: Comparison of *Bs*GDH template variant A258N (blue) and the corresponding epPCR library (orange). Activity was measured after 26 hours incubation at room temperature. Histogram was divided into two parts (y-axis) to show rare hits in epPCR library.

All variants were recovered from the droplets and the corresponding DNA was amplified and used to transform *E. coli* competent cells. After recovery process, all cells were plated and colonies were picked for rescreening of the obtained variants. In total 8 plates were cultured overnight and the activity of each variants supernatant was measured. Except for 8 wells, which did not show growth, all colonies showed a reasonable growth after incubation overnight.

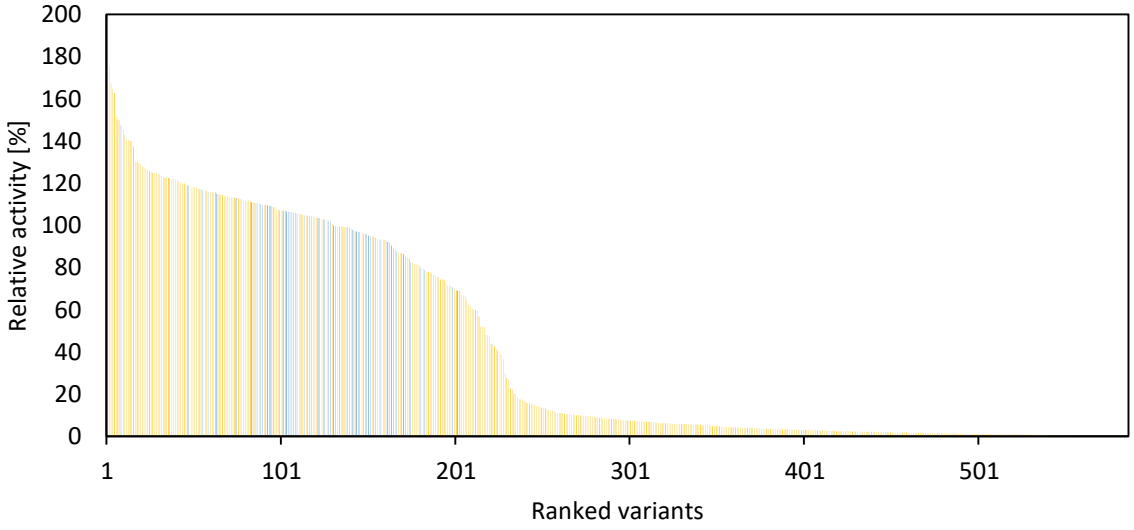


Figure 29: Recreening of amplified hits from microfluidic sorting. Variants (orange) were ranked for their activity. The activity was normalized to the mean activity obtained by the template variant (blue).

Measuring the supernatant of the lysed cultures approximately 30% of the cultured colonies showed activity comparable to the template variant in the photometric measurement. This corresponds to an enrichment factor of 2.5 and therefore confirmed the results from the first library screening. However, it also indicated that the chosen threshold, allowed the sorting of variants with lower activity and multiple occupied droplets. From the screened 632 variants, only 20 showed significantly higher activity than the parental template (Figure 29). The best variant in the microtiter plate rescreening reached 197% activity compared to the template. In the next step these 20 variants were inoculated from the backup plates and measured in triplicates to neglect the effect of the cell density within the wells (data not shown). The best three performing variants were sequenced and chosen for enzyme purification. For better comparison the template variant *Bs*GDH A258N was also purified and characterized (Figure 30, green).

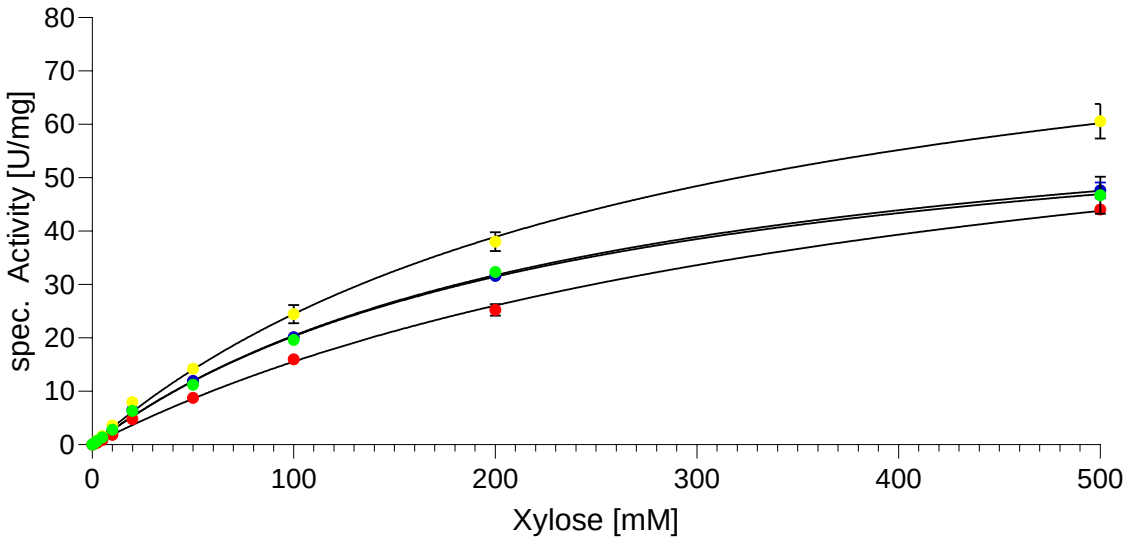


Figure 30: Substrate kinetic for the three best *Bs*GDH variants and the template variant A258N (green). The variants M168L (yellow), S255P (red) and the double mutant K170R N192Y (blue) were characterized with 8 mM xylose and 1 mM NAD at room temperature.

Interestingly, all three variants showed a slightly higher maximum velocity compared to the template mutant ($V_{max}^{Template}$= 70 ± 4 U/mg, $V_{max}^{K170R\ N192Y}$=71 ± 2 U/mg, V_{max}^{S255P}=80 ± 4 U/mg and V_{max}^{M168L}= 95 ± 6 U/mg). However, the affinity for xylose was decreased for all mutants ($K_M^{Template}$ = 244 ± 29 mM, $K_M^{K170R\ N192Y}$ = 248 ± 17 mM, K_M^{S255P} = 415 ± 38 mM and K_M^{M168L} = 287 ± 34 mM). The double mutant K170R N192Y showed similar reaction rates like the template mutant, while S255P had a slightly higher V_{max} but also an almost doubled K_M, making the catalytic efficiency worse than the template mutant. Only the mutant M168L showed a higher V_{max} and only a slightly increased K_M, resulting in a 15% higher catalytic efficiency. While all mutants showed significantly increased turnover rates in the lysis test, only the variant M168L showed slightly increased enzymatic reaction for all xylose concentrations tested. This difference might be explained by solubility of the enzyme variants. All enzyme

variants showed higher enzyme concentrations after the purification process (2.3 mg/mL - 2.6 mg/mL) compared to the template variant (1.5 mg/mL). Overall, the second screening round did not yield an improved enzyme in terms of xylose affinity, but it successfully demonstrated the functionality of the constructed microfluidic sorting system.

3.3 Increase catalytic efficiency for a *Saccharolobus solfataricus*-derived glucose dehydrogenase with the artificial cofactor carba-NADP$^+$

Oxidoreductases offer a wide variety of redox reactions. They need cofactors to store electrons as redox equivalents for efficient catalysis. In industry they play an important role for the biodegradation of lignin and can be utilized to produce polymer building blocks, sustainable materials from plant biomass[140] and fine chemicals[141]. After confirming that the microfluidic system could identify improved enzyme variants for engineering substrate specificity, the next goal was to engineer an enzyme towards acceptance of an artificial cofactor. Previous experiments revealed that the glucose dehydrogenase from *Saccharolobus solfataricus* was able to convert substrates with the artificial cofactor carba-NADP$^+$ [142]. This cofactor had a similar structure as the native NADP$^+$, but the oxygen in the ribose ring was replaced by a methylene group. This created a carbocyclic ribose analogue, which was less prone to hydrolysis at elevated temperatures compared to the native NADP$^+$ cofactor. With this change in the chemical structure, the lifetime of the cofactor was increased, making it interesting for industrial enzyme cascade approaches performed at elevated temperatures. To simulate these conditions and find variants with high activity at elevated temperatures, single cells expressing *Ss*GDH variants were encapsulated in droplets and incubated at 45 °C and 65 °C in the first and second screening round, respectively. One challenge of this approach was the storage of droplets at elevated temperatures for several hours, since droplet-based microfluidics was very prone to evaporation. The results of this chapter have been published[135].

First screening round: QuikChange combined with epPCR

To target the area around the cofactor binding site a pooled library consisting of 21 single saturated positions [G40, T41, E44, T189, G190, P191, I192, N211, R212, R213, P215, T254, G255, F277, G278, F279, S280, T281, L305, V306 and K354] was designed and produced by Ioannis Zachos. All selected positions were located in the 1st shell (≤ 4 Å) around the cofactor binding site. Each position was saturated with a NNK codon to obtain a library of 420 variants based on the amino acid level. As a template, a codon optimized *gdh* gene on a T7-based expression system was used. After harvesting cells from plates, libraries (2 x 10^4 pooled *E. coli* colonies) were incubated in auto-induction media for enzyme expression. Single cells holding the *Ss*GDH variant, were encapsulated with a lambda of 0.1 allowing a low percentage of multiple encapsulations. For the substrate reaction, 50 mM glucose/xylose and

1 mM cNADP$^+$/NADP$^+$ were added. As a feasibility experiment, the droplets were incubated at room temperature and the product conversion was monitored by measuring the product formation after defined time points. Since the activity of *Ss*GDH at room temperature was determined to be below 1 U/mg, even after 5 days of incubation hardly any product formation could be measured (data not shown). The reason for this low catalytic efficiency was due to the fact that *Saccharolobus solfataricus* is a thermophilic microorganism. Its glucose dehydrogenase has the highest efficiency at a temperature around 70 °C. Therefore, improving the catalytic efficiency by increasing the incubation temperature of the enzyme variants within the droplets would be beneficial. Unfortunately, microfluidic droplets proved to be very prone to higher temperatures, as the small volumes of liquid used could easily evaporate, making it impossible to monitor the reaction. To avoid evaporation a custom-made droplet chamber was manufactured (chapter 2.2.3). The droplets could be gathered and by sealing the tubing inlets after droplet production, a closed system was created, where evaporation was strongly reduced by the vapor pressure within the chamber. The chamber reduced shrinking and merging effects of the captured droplets and allowed comparable assay measurements. To determine if the droplets within the droplet chamber remain stable at elevated temperatures, the QuikChange library and the template variant of *Ss*GDH were incubated at 45 °C for 21 hours. The reaction was monitored for 50 mM xylose and 50 mM glucose with 1 mM cNADP (Figure 31).

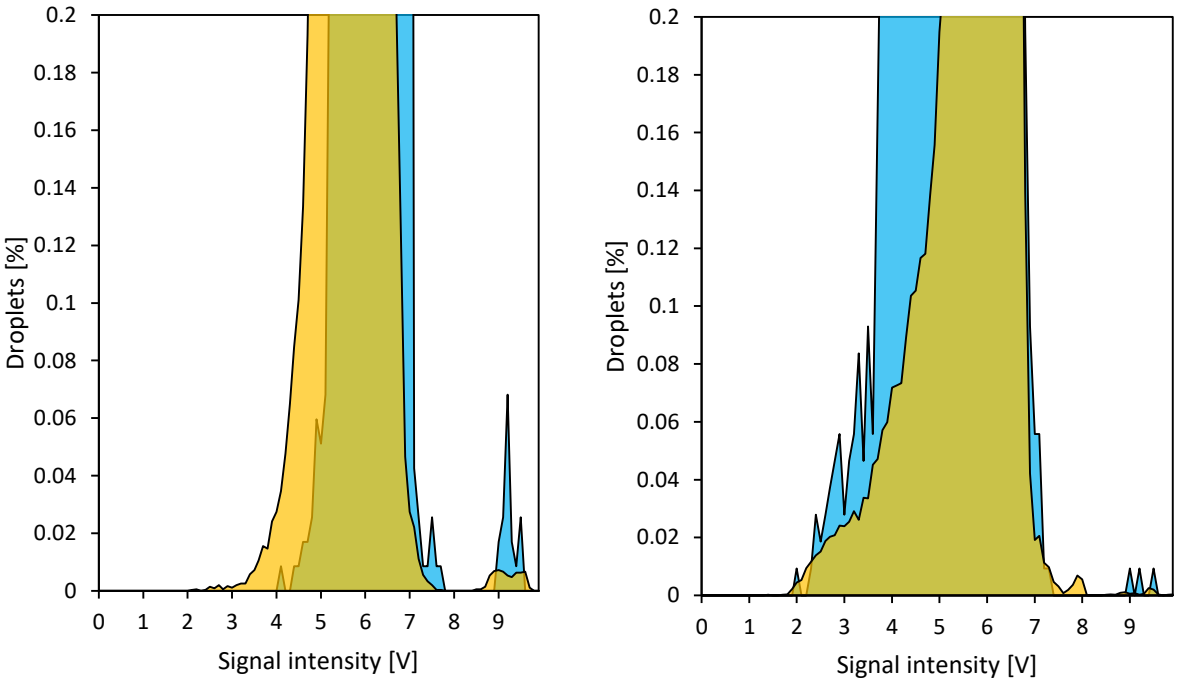


Figure 31: Microfluidic measurement of *Ss*GDH library (orange) against template variant (blue) activity for glucose (left) and xylose (right) after 21 hours. Number of droplets was normalized to allow direct comparison of both populations.

After 21 hours both droplet populations showed a reasonable turnover, which can be recognized as a small peak shoulder. The measurements suggested that both libraries could hold improved variants. This could be observed by comparing the histograms of both, wild type and library. The lower light intensities for the library in Figure 31 suggested that some variants showed higher activity than the template enzyme enzyme. In the screening experiment, a threshold of 4.5 V was applied for both libraries as a sorting criterion. This led to sorting of 1.274 hits (0.34%) and 5.094 hits (0.94%) for the glucose and xylose sorting, respectively. All sorted plasmids were recovered, amplified via PCR and used to transform competent *E. coli* cells. For the rescreening two micro titer plates full of colonies were picked for each sorting. The xylose rescreening yielded 32 improved variants, which showed up to 9 times higher activity for the lysis supernatant, compared to the template enzyme (Figure 32). This corresponded to an enrichment of 17% compared to the template enzyme activity (+3σ). The best 9 variants were send for sequencing and led to 7 unique sequences. The identified variants were single mutants T189L, N211L, N211F and double mutants K22E R213T, V63M N211L, E161K T189M, A210T N211L. The variants N211L and E161K T189M appeared twice in the sequences.

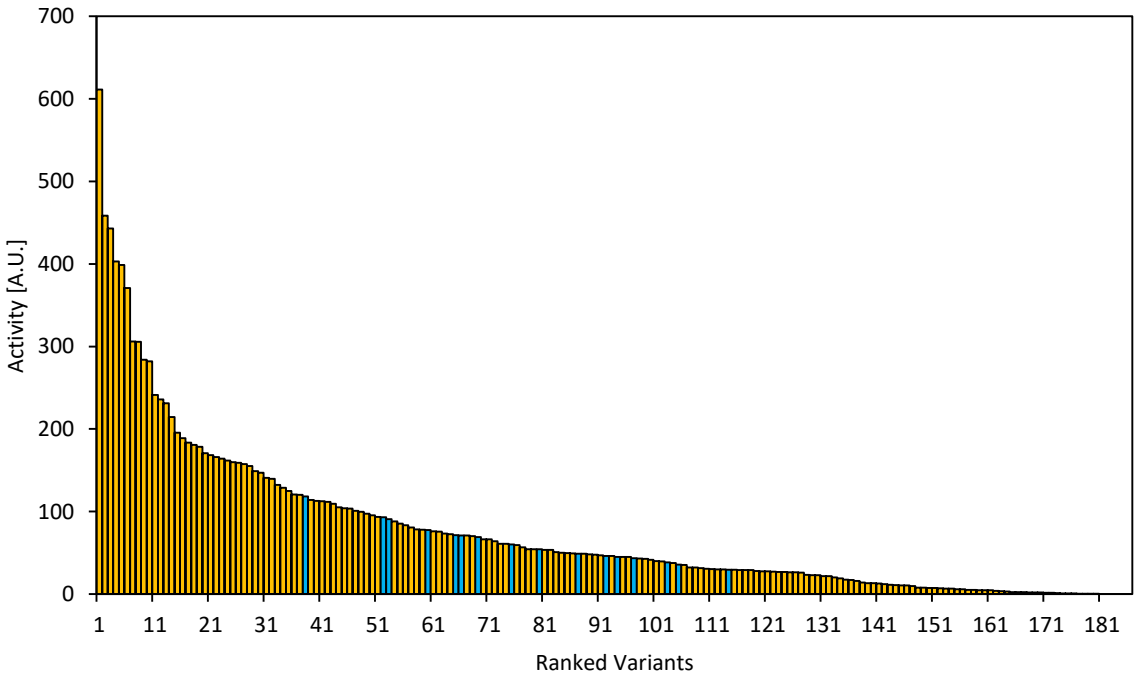


Figure 32: Microtiter plate recreening of sorted variants from QuikChange *Ss*GDH library, incubated with xylose. Activity of variants was determined from lysed cells supernatant with 8 mM xylose, 1 mM cNADP with TrisHCL pH 8 (65 °C). Variants (orange) and template controls (blue) were ranked according to their activity.

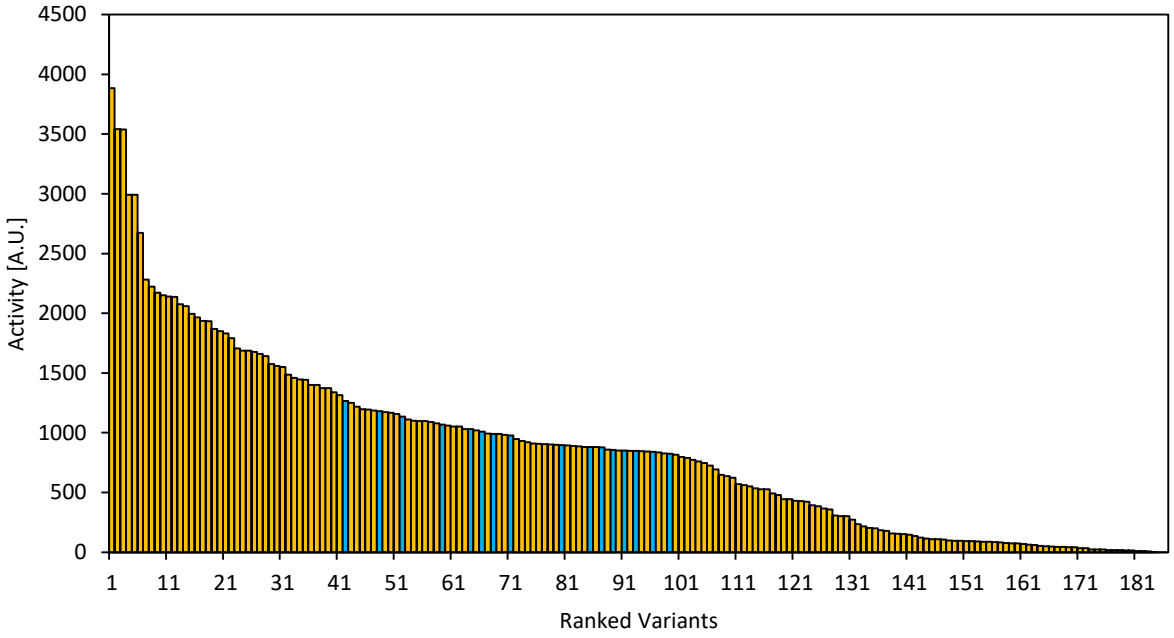


Figure 33: Microtiter plate recreening of sorted variants from QuikChange *Ss*GDH library incubated with glucose. Activity of variants was determined from lysed cells supernatant with 8 mM glucose, 1 mM cNADP with TrisHCL pH 8 (65 °C). Variants (orange) and template controls (blue) were ranked according to their activity.

The same procedure was conducted for the glucose rescreening (Figure 33). Here 39 from 186 variants (21%) showed a significantly higher activity (+3σ) compared to the template enzyme activity. The best 8 variants with up to 4 times higher activity for the lysis supernatant were sent for sequencing and led to 5 unique sequences. The identified variants were single mutants T189M, 189S, S280K and two double mutants S280K R362C and N211A S280K. Interestingly, not only the expected single mutants, but also double mutants were found in both rescreenings. This can be explained by the comparably high error rate and the lack of proof reading activity of the Taq polymerase, which was used for the recovery of the libraries. All sequenced hits were expressed and purified for kinetic analysis of the enzyme variants. For the purpose of simplification and with the aim to further increase the activity towards one specific substrate, the enzyme variants from both rescreenings were characterized only with glucose as a substrate. The results of this characterization were summarized in Table 5.

Table 5: Summary of Michaelis-Menten coefficients from substrate characterization of both screening runs for *Ss*GDH. All variants were characterized with varying glucose concentrations and 1 mM cNADP$^+$.

	*Ss*GDH variant	V_{max} [U/mg]	k_{cat} [s^{-1}]	K_M Substrate [mM]	k_{cat}/K_M [$mM^{-1}\ s^{-1}$]
Glucose Hits	**WT**	5 ± 0.2	4	22 ± 2	0.2
	S280K	11 ± 0.5	8	11 ± 1	0.7
	S280K R362C	15 ± 0.4	10	7 ± 1	2.0
	N211A S280K	12 ± 0.4	8	12 ± 1	0.7
	T189M	12 ± 0.5	8	23 ± 2	0.4
	T189S	9 ± 0.5	7	24 ± 3	0.3
Xylose Hits	**WT**	7 ± 0.2	5	46 ± 5	0.1
	N211F	5 ± 0.1	4	37 ± 3	0.1
	V63M N211L	11 ± 0.2	8	30 ± 2	0.3
	N211L	5 ± 0.2	4	36 ± 3	0.1
	E161K_T189M	10 ± 0.3	7	35 ± 4	0.2
	T189L	12 ± 0.3	8	39 ± 3	0.2

For the glucose screening, the substitution from the serine to lysine at position 280 appeared three times, either as a single substitution or combined with substitutions at position 211 and 362. All of the variants showed an at least two times higher V_{max} and additionally a two times lower K_M, underlining the importance of the mutation at position 280. The best variant for the glucose screening, was the double mutant S280K R362C, which showed a V_{max} of 15 U/mg and a K_M as low as 7 mM. Additionally, two variants showed a change at position 189 from threonine to either methionine or serine. Both of this single mutants showed no improvement in the substrate affinity, but doubled the maximal substrate velocity towards glucose. Analyzing both, the hits obtained in the xylose screening and glucose recreening, mutations at position 189 and the position 211 were found again as single (N211F and N211L) and double mutants (V63M N211L). Here, the single substitutions showed a lower maximal velocity compared to the wild type, but also a higher substrate affinity resulting in a lower K_M. The double mutant including a substitution from valine to methionine at position 63 showed a doubled maximal reaction velocity. All variants were also characterized towards the artificial cofactor cNADP$^+$. Therefore, the substrate concentration was set to 50 mM glucose, like in the screening. Here, the variants with substitutions at position 280 showed a three to five times higher maximal velocity compared to the template enzyme. On the other hand, the cofactor affinity for all variants was decreased resulting in lower substrate affinity for all variants. Therefore, the catalytic efficiency was doubled. The single substitution at position 189 showed a 2.5 times higher V_{max} but additionally a two times lower substrate affinity resulting in a slightly increased catalytic efficiency of 1.4 times. The double mutant with the additional mutation at position 161

showed even lower catalytic efficiency with only 1.15x increase. When directly comparing the variants N211L and V63M N211L it was interesting to observe that N211L did not seem to have an influence on the maximal reaction velocity, but halved he substrate affinity. Combined with the mutation at position 63 the variant showed an almost three times higher maximal reaction velocity, but also a four times lower substrate affinity.

Table 6: Summary of Michaelis-Menten coefficients from cofactor characterization of both screening runs for *Ss*GDH (glucose and xylose). All variants were characterized with varying cNADP$^+$- concentrations and 50mM glucose.

	***Ss*GDH variant**	**V_{max} [U/mg]**	**k_{cat} [s^{-1}]**	**K_M cNADP$^+$ [µM]**	**k_{cat}/K_M [mM^{-1} s^{-1}]**
	WT	3 ± 0.1	2	55 ± 5	43
Glucose	**S280K**	10 ± 0.4	7	100 ± 11	68
	S280K R362C	16 ± 0.7	11	132 ± 16	84
	N211A S280K	12 ± 0.2	8	96 ± 6	86
	T189M	8 ± 0.2	6	97 ± 8	58
Xylose	**E161K T189M**	6 ± 0.1	4	84. ± 4	49
	N211L	3 ± 0.1	2	109 ± 7	19
	V63M N211L	10 ± 0.4	7	194 ± 18	36

Shortly summarizing the first screening round: Five interesting substitutions at the positions 63, 189, 211, 280 and 362 could be identified and characterized. All positions either improved the maximal reaction velocity and/or increased the substrate affinity. None of the identified positions increased the affinity towards the artificial cofactor. The positions 189, 211 and 280 were part of the initially designed library with 21 positions in the proximity of the catalytic center of the enzyme. On the other hand, the mutations at position 63 and 362 were created by random mutagenesis and were located far away from the catalytic center. Therefore, it would be highly unlikely to identify these positions through in silico rational design.

Second screening round: Overlap extension PCR with defined codons at 5 positions

To find variants with even higher catalytic efficiency, all positions found in the first screening round were combined in a new library. The library was designed with the DYNAMCC tool by Ioannis Zachos and produced by overlap extension PCR (oePCR) by Samuel Sutiono. The design allowed not only the wild type and the substitutions found for the corresponding positions, but also amino acids with different properties at all positions (V63AKS T189NBT N211VNS S280VWS R362VNS). With these codons the library comprises a potential variant size of 112,640 proteins (331,776 variants on DNA level). The first production of the library yielded around 120,000 colonies after transformation resulting in only 36% coverage of the constructed library. Droplets were incubated in self-made droplet chambers for 43 hours at 45 °C and sorted below 4.2 V. Overall 492.514 droplets were measured and the best 0.3%

(2,514 droplets) were sorted. The amplified DNA was used to transform *E. coli* BL21(DE3) and a rescreening in microtiter plates was performed. In the plate rescreening, 16 variants out of 170 variants (9.4%) showed significantly higher activity (+3σ) compared to the template enzyme activity (Data not shown).

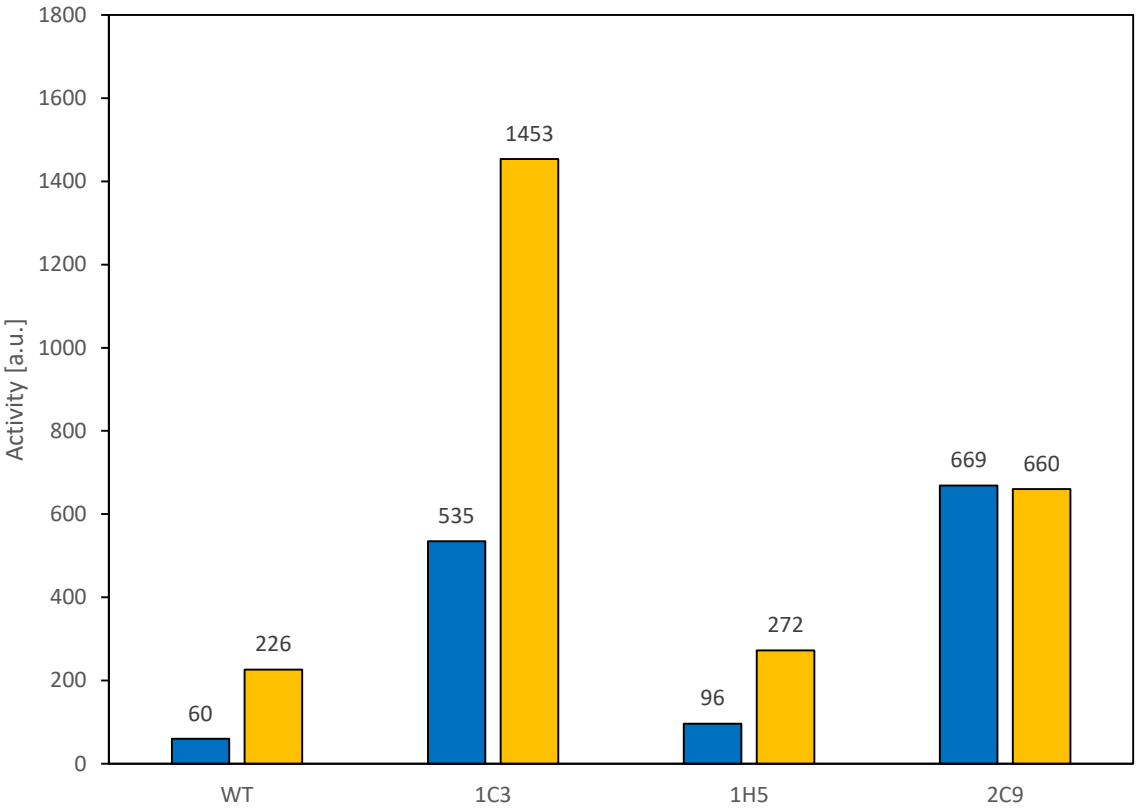


Figure 34: Plate rescreening of variants derived from microfluidic sorting of *Ss*GDH oePCR llibrary at 45 °C. Blue bars indicate activity after cell disruption with sonotrode. Orange bars show activity after heat activation (30 min, 70 °C).

The three most promising variants (designated as 1C3, 1H5 and 2C9) were further characterized (Figure 34). The variant 1H5 was found to perform only slightly better than the wild type. The variant 1C3 showed a ~9 times higher activity than the wild type and the variant 2C9 even 11 times higher activity. After the initial characterization, all variants were also exposed to a heat challenge (70 °C, 30 min) to exclude activity from other *E. coli* enzymes, which might convert glucose. Addtionallly, it was clarified whether the substitutions in the variants had an influence on the thermostability. Interestingly, the wild type and the variants 1C3 and 1H5 showed a heat activation. After the heat challenge, these variants demonstrated about three times higher activity than before. In contrast to that, the variant 2C9, which showed the highest activity before the heat challenge, did not improve its activity after the heat treatment. The variants 1C3 and 2C9 were purified for kinetic analysis and sent for sequencing. Variant 2C9 revealed 3 substitutions (V63M T189L N211V) and variant 1C3 4 substitutions (V63I M86I S280T R362T). Since the heat challenge at 70 °C showed a beneficial influence on the activity of *Ss*GDH, the screening temperature was raised to 65 °C for a second

screening round. Here, 15.6 million droplets were produced with an occupation rate of λ=0.1 to reach a theoretical coverage of over 95%[143]. Droplets were stored in 12 different microfluidic chambers, each comprising 1 – 1.5 million droplets, and incubated for 4.5 hours at 65 °C prior to measurement and sorting. In total 16,660 droplets (best 0.1 %) were sorted and hits were extracted for each microfluidic chamber separately and pooled for amplification. After the first amplification round, all products were mixed in equal amounts and amplified again prior to transformation with competent *E. coli* cells. The corresponding colonies grew poorly on agar plates and only 420 colonies were picked. The activity was measured in a plate rescreening. From the selected 420 variants only 1 variant showed a high improvement in activity and was send for sequencing. The variant showed quadruple substitutions (V63S T189L S280N R362T). Since the rescreening procedure only allowed screening of 420 out of 16,660 (2.5%) potential hits, the amplified hits from the first overlap extension screening round were plated again. The aim of this additional screening round was to rescreen all hits with 3x oversampling. For this enriched library, only 1 Mio. droplets were screened and 991 hits were sorted. The plate rescreening of the enriched library led to 20 hits out of 170 picked colonies (11.8%) with significantly improved activity (+3σ). Two hits demonstrated 26% and 53% increased activity compared to the template variant from the plate screening and were consequently sent for sequencing. Interestingly, one of these variants contained only a single mutation (V63M) emphasizing the importance of that position. The second variant contained 3 substitutions (V63S T189S R362I). All five variants of the overlap extension screening were purified for kinetic analysis. During the purification, a heat treatment (30 min, 70 °C) was applied like in the first screening round to observe potential heat activation. Interestingly, the yield of the different purified variants differed depending on their substitutions. Especially the variants containing a mutation at position 189, showed an 8 times lower yield compared to the wild type. In a separate incubation test, it was exemplarily shown that the enzymatic turnover rate from variant 2C9 (V63M, T189L, N211V) was decreased by 80% after 45 hours incubation at 70 °C. The enzymatic turnover of the wild type and variant 1C3 (V63I M86I S280T R362T) retained at 80-90% and did not show any loss of enzymatic activity with rising temperature. After purification, all variants were characterized with substrate (Appendix 3.3.1) and cofactor kinetics (Appendix 3.3.2). In addition to the artificial cofactor (cNADP), which was used for the sortings, the native cofactor NADP was also used for characterization. This comparison revealed if the substitutions found were only beneficial for the artificial cofactor or generally enhanced the catalytic efficiency of the enzyme. The results of the kinetic characterization were summarized in Table 7.

Table 7: Characterization of *Ss*GDH hits from oePCR. For all enzyme variants substrate and cofactor kinetics were measured for both cofactors (NADP and cNADP). Specific screening conditions were summarized in Appendix 3.3.1 and 3.3.2

		cNADP				NADP			
		Substrate-Kinetics		Cofactor-Kinetics		Substrate-Kinetics		Cofactor-Kinetics	
		V_{max} [U/mg]	K_M[mM]	V_{max} [U/mg]	K_M[µM]	V_{max} [U/mg]	K_M[mM]	V_{max} [U/mg]	K_M[µM]
	WT	6 ± 0.1	32 ± 1.6	5 ± 0.2	135 ± 13	7 ± 0.2	33 ± 5	9 ± 0.4	211 ± 24
1F1	V63S T189L S280N R362T	40 ± 0.6	19 ± 1.2	34 ± 0.5	186 ± 6	158 ± 3	2 ± 0.2	183 ± 12	543 ± 69
2C9	V63M T189L N211V	35 ± 0.6	35 ± 2.2	22 ± 0.6	93 ± 7	149 ± 6	3 ± 0.3	223 ± 11	630 ± 59
1C3	V63I M86I S280T R362T	22 ± 0.5	19 ± 1.7	20 ± 0.4	45 ± 3	35 ± 1	1 ± 0.1	28 ± 1	84 ± 12
1G9	V63S T189S R362I	15 ± 0.2	30 ± 1.3	16 ± 0.5	174 ± 14	25 ± 1	21 ± 5	31 ± 1	158 ± 12
1G8	V63M	12 ± 0.2	27 ± 1.3	8 ± 0.2	9 ± 2	12 ± 1	1 ± 0.2	8 ± 0.2	30 ± 3

In the second screening round, 5 improved variants were idientified. The variant with the highest catalytic efficiency (V63S T189L S280N R362T) exhibited a six times higher V_{max} and an almost halved K_M for glucose. This multiplied to a 10 times higher catalytic efficiency. The mutation at position 280 proved to benefit the substrate affinity, which also can be observed for variant 1C3 (V63I M86I S280T R362T). In contrast to the triple and quadruple mutation variants, the single mutant V63M was found in the second screening round. All variants contained a mutation at position 63, which emphasized the importance of that position. Judging from the obtained kinetic data it can be concluded that position 63 doubles the maximal reaction velocity and had a great impact on the cofactor affinity towards cNADP. The cofactor affinity for the single mutant V63M was increased almost 15 times by that single mutation. However, based on the kinetic data obtained for the native cofactor NADP, it was evident that the substitutions identified in the cNADP screening also enhance the catalytic efficiency for the native cofactor. Even some variants (V63S T189L S280N R362T and V63M T189L N211V) exhibited up to three 3 times lower cofactor affinity, all other parameters were improved. The best variant (V63S T189L S280N R362T) showed a 22 times higher maximal reaction velocity and a 15 times increased substrate specificity, giving an overall 350 times higher catalytic efficiency. It is interesting to point out that the single mutation at position 63 also increased the cofactor affinity for NADP, but showed a substrate inhibition for glucose concentrations over 10 mM. This substrate inhibition was abolished by using the artificial carba cofactor.

3.4 Establishment of a microfluidic screening method for selection of thermostable variants

In the previous chapter, a new method for the enzyme engineering of the thermophile *Ss*GDH towards higher enzymatic turnover at elevated temperatures was presented. It was shown that the established microfluidic screening method in self-made chambers, was capable to engineer thermophiles towards substrate affinity and/or maximum reaction velocity, while incubating the enzymes at elevated temperatures for several hours. However, not all enzymes have a high level of thermal stability naturally. In some cases, enzymes exhibit high catalytic efficiency but low thermostability. Therefore, we thought to investigate whether incubation temperature could serve as an additional selection criterion for the established microfluidic system. There were several challenges that had to be taken into account to choose the appropriate incubation temperature and incubation time. On the one hand, the assay was performed as an endpoint determination. This would mean that an enzyme with very high catalytic efficiency but low thermostability could also generate a comparatively strong assay signal if the incubation temperature was set too low or the incubation time was too short and the enzyme was not denaturated. On the other hand, an incubation temperature that was selected too high would directly lead to denaturation of all enzymes and the difference between enzymes with higher thermostability and enzymes with lower thermostability could not be measured. Finding the correct incubation time and temperature proved to be even more challenging, as every time the incubation chamber was removed from the incubator to measure a sample, the entire system was cooled down to room temperature before measurements could begin. It then needed to warm up again, when returned to the incubator. The challenge of this approach was given by choosing the highest possible incubation temperature to denature thermolabile enzymes, while at the same time incubating the enzymes long enough to measure differences between improved variants and template enzymes. In contrast to the previous chapter, the aim was not necessarily to increase the catalytic efficiency, but rather the thermostability (time that the enzymes activity can be measured at elevated temperatures). In chapter 3.2 the glucose dehydrogenase from *Bacillus subtilis* was engineered towards higher activity. The improved variants 258D and 258N were especially interesting as they both showed a doubled activity towards xylose containing a single substitution (Figure 35, blue bars). However, after incubating the lysed cells in an Eppendorf tube and checking the temperature stability after a heat challenge (30 min, 70 °C), the variant with aspartic acid lost almost all of its activity, while the variant with asparagine and the template variant retained 90% of their initial activity (Figure 35, orange bars).

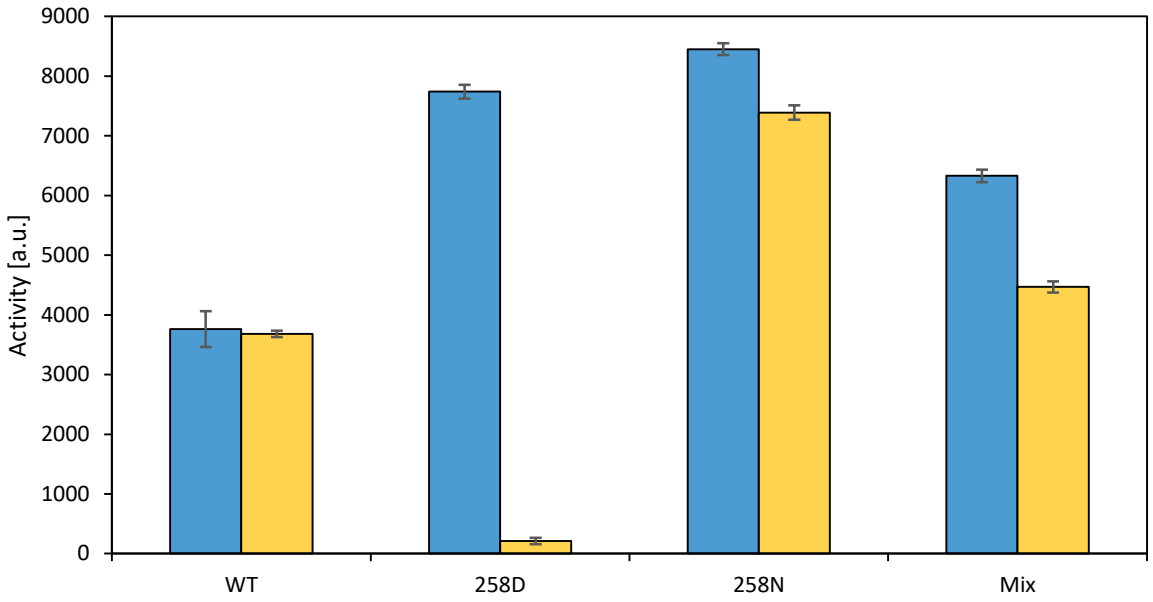


Figure 35: Activity of lysed *Bs*GDH variants before (blue) and after heat treatment (orange) in water bath for 30min at 70 °C. Activity was determined photometrically in 96 well plate at 340nm.

The mixture combined all three variants (*Bs*GDH template variant, variant 258D and variant 258N). The activity of the mix was determined before and after the heat treatment. As expected, the mixture of all three variants showed an activity higher than the template variant, but lower than the improved variants. After the heat treatment, the overall retained activity was 70%. As a proof of concept, the mixture of the three variants should be incubated, but only the thermostable variant 258N with improved activity should be identified and sorted by the system. Therefore, the optimal temperature for the experiment, where the activity of the template variant and the two variants differed the most, had to be identified to sort only thermostable variants. One major limitation of the microfluidic screening platform was its inability to test multiple reaction parameters e.g. temperature simultaneously. For each temperature, a droplet population had to be created in a separate incubation chamber and incubated in different incubators. This would require significant effort to determine the optimal screening conditions. To simplify this procedure, a gradient program in a thermocycler was performed with 100 µL of lysed cell culture expressing the different variants of the *Bs*GDH. Each culture was incubated at 8 different temperatures ranging from 45 °C – 70 °C. The activity of the variants was determined right after the lysis to obtain the maximal activity and later normalize the residual activity. Samples were taken after 18, 24 and 36 hours to evaluate the activity of the three different enzyme variants (Figure 36).

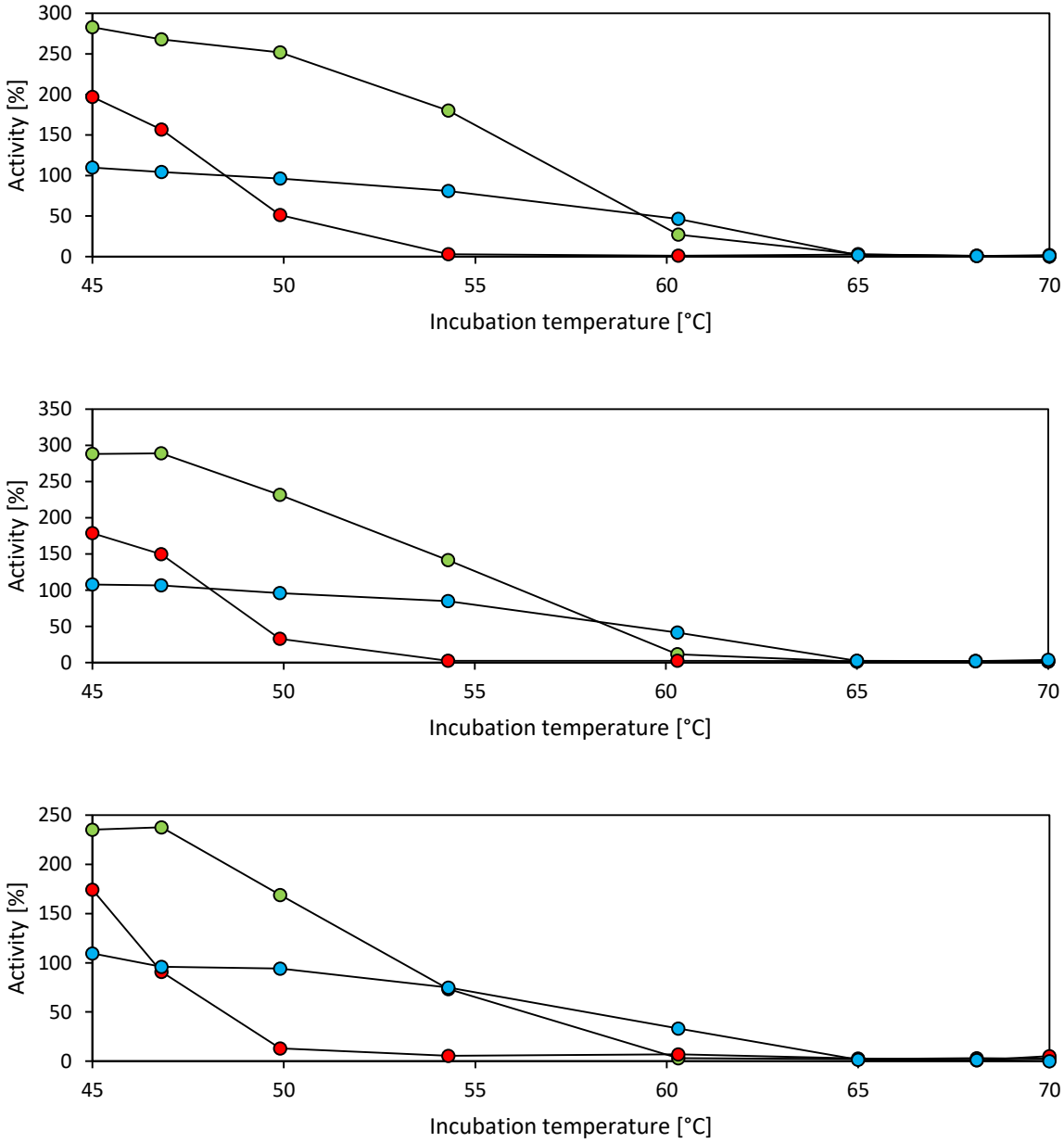


Figure 36: Residual activity after gradient incubation of lysed *Bs*GDH enzyme variants 258N (green), 258D (red) and template variant (blue). Samples were measured after 18 hours (top), 24 hours (middle) and 36 hours (bottom) incubation at different temperatures. Enzymatic activity was normalized to the initial activity of *Bs*GDH template variant directly after lysis. All activity measurements were performed at room temperature.

Similar to the measurements previously performed, the variant 258D lost its activity already at 55 °C after 18 hours of incubation, while the template variant and the variant 258N retained most of their activity. Over the course of the next 18 hours, the variant 258N also gradually lost its initial activity. At 60 °C and 36 hours incubation time the template variant retained 33% of its initial activity, while the two other variants lost almost all of their activity (Figure 36, bottom). The highest difference in enzymatic activity between the variant 258D and template enzyme compared to the heat stable variant 258N was observed after 18 hours at 50 °C (Figure 36, top). However, it had to be considered that the performed assay was an endpoint measurement, which did not reflect the enzyme's activity at that specific time point, but rather

its activity over the entire incubation period. Therefore, it should be considered to incubate the enzymes for longer time. This way, it could be assured that the temperature destabilization sufficiently reduced the enzyme activity of variant 258D and the overall turnover would be comparable or even lower than the template enzyme. Acknowledging this fact, the incubation at 50 °C for 24 hours was chosen to be the most promising set of parameters. Here, the template variant maintained 96% of its original activity, while the thermo labile variant 258D showed only 32% residual activity compared to the template variant. In contrast to that, the improved variant 258N showed 231% residual activity compared to the template variant. This implied that the improved variant 258N showed almost 2.5 times higher activity at 50 °C after 24 hours (Figure 36, middle). Next, the results obtained with the gradient program in the thermocycler were transferred and compared to microfluidics droplets. In a first measurement, all variants including the *Bs*GDH template variant and the substitutions 258D and 258N were solely encapsulated into droplets. The droplets were than incubated for 24 hours at 55 °C to denature temperature labile variants (Figure 37, top).

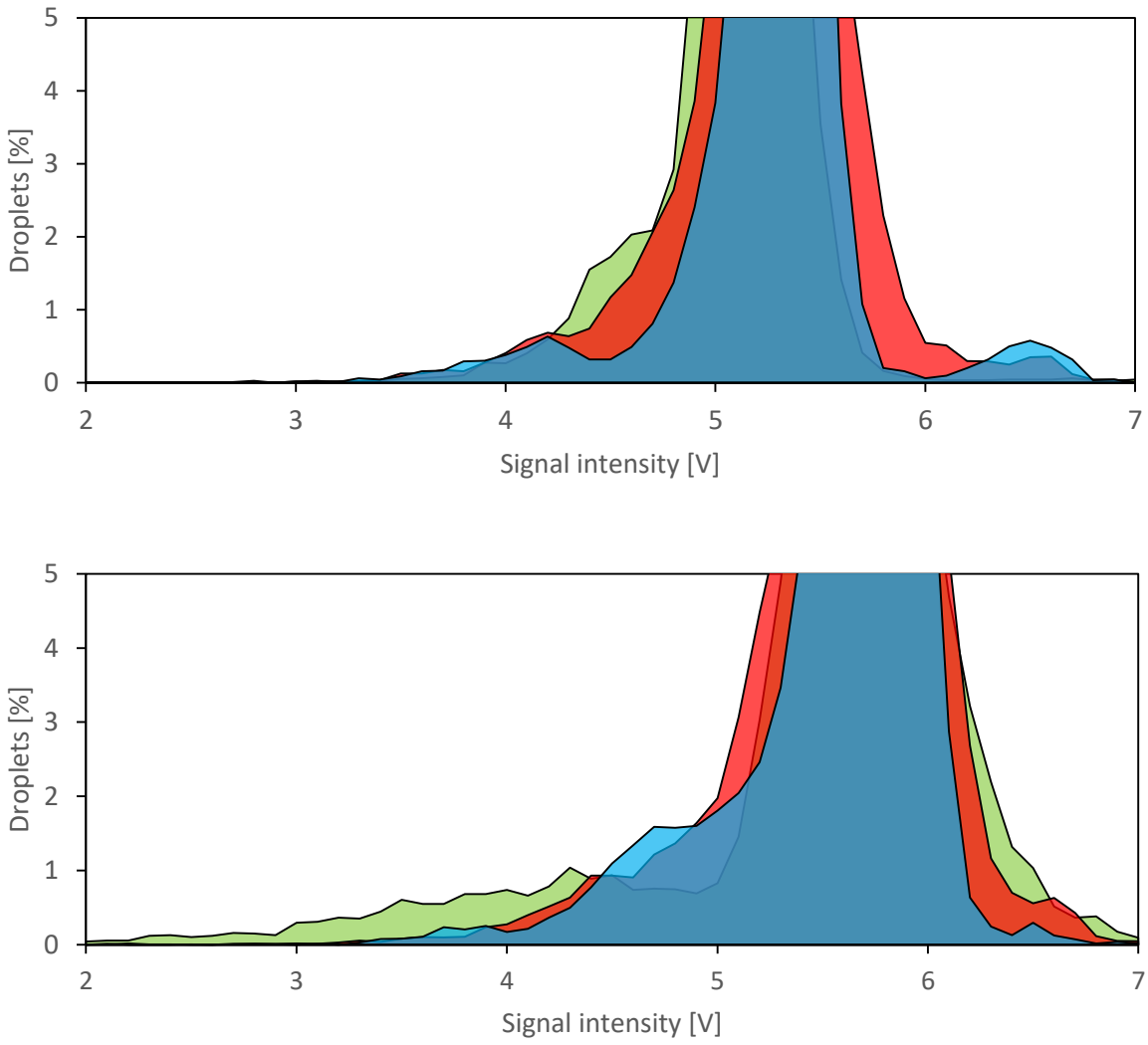


Figure 37: Microfluidic measurement of *Bs*GDH template variant (blue), variant 258D (red) and variant 258N (green) after 24 hours incubation at 55 °C (top) and 20 hours at 50 °C (bottom).

In contrast to the results obtained by the gradient incubation, the variant 258N did not show more activity than the template variant. By overlaying all obtained histograms it can be observed, that the droplets containing the template variant, showed more activity than the two mutants. This indicated that the enzyme variants 258D and 258N were deactivated earlier in the droplets than in the gradient incubation experiment. In a second run, the incubation temperature was reduced to 50 °C and the incubation time was shortened to only 20 hours (Figure 37, bottom). Here, the overlay of the histograms showed a clear green shoulder, which indicated that the variant 258N showed more enzymatic turnover in the observed period of time than the template variant. Additionally, the heat labile variant 258D was discriminated and showed comparable activity to the template enzyme. The histograms suggested, that the improved variant 258N would be predominantly found if the mixture would be sorted below 3 V. In the second step, all variants were incubated individually and additionally an equally mixed cell culture of all variants (template variant, 258D and 258N) was prepared and incubated in microfluidic droplets at 50 °C for 20 hours (Figure 38).

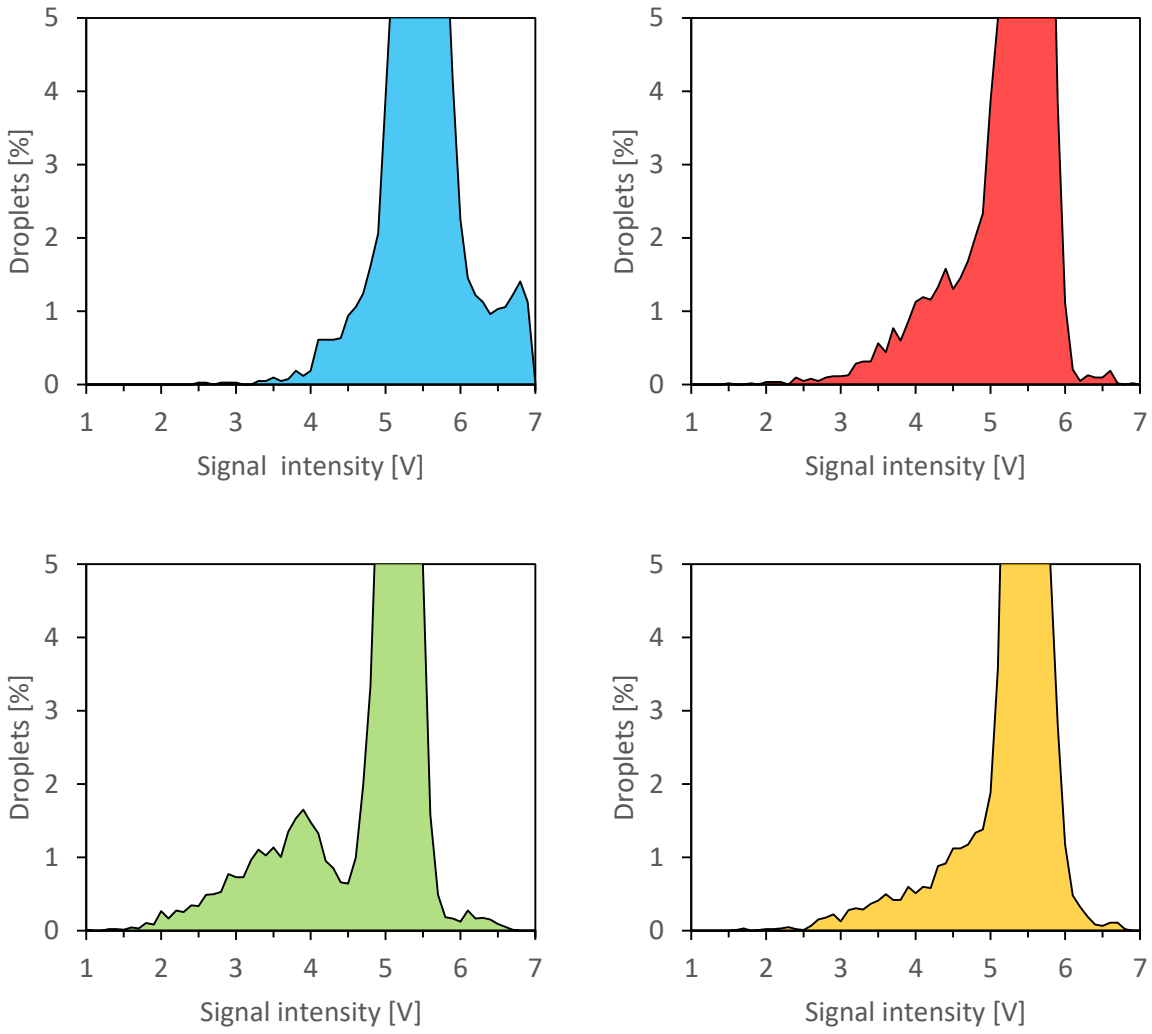


Figure 38: Microfluidic measurement of *Bs*GDH template variant (blue), variant 258D (red), variant 258N (green) and the mix of all 3 before mentioned variants (orange) after 20 hours incubation at 50 °C

At first, all single variants were measured. The results obtained in this experiment were reproducible with the histograms obtained in the prior experiment. The template variant (blue) and the heat labile variant 258D (red) showed a decent shoulder in the histograms, but the transition of the population with vacant and occupied droplets could not be distinguished. Only a few droplets showed activity below the signal intensity of 3 V, which might correspond to multiple occupied droplets. Looking at the heat stable variant 258N (green) an edge between both populations can be observed around 4.5 V, indicating the higher overall activity of this variant. Additionally, a reasonable number of droplets showed activity below 3 V as in the initial experiment. The mix of all 3 variants showed a decent shoulder with some variants reaching the intended threshold of 3 V. To show that the microfluidic method could be used to separate heat stable and heat labile variants the *Bs*GDH mix was sorted at 3 V to obtain only heat stable variants. In this first run 320.000 droplets with *Bs*GDH mix (λ= 0.1) were screened and 1.236 exhibited an activity below 3 V and were therefore sorted (fraction 1). All droplets, which were not sorted in the first run, were recaptured in a second microfluidic chamber and reinjected for subsequent sorting round. In this second screening round, 240.000 droplets were measured. Only 14 (0.006%) of the measured droplets showed an activity below 3 V indicating that the sorting of the first round was successful. For the second screening round the threshold was set to 3.5 V to sort not only for the heat stable variant 258N, but also allow to find the template variant and heat labile variant 258D to be captured. In this run another 1.130 droplets were sorted (fraction 2). The emulsion of both sorted droplet fractions was broken in separate tubes and the DNA was purified prior to amplification. After transformation of competent *E. coli* cells with the DNA, both fractions were plated. Single colonies were picked and incubated in fresh medium overnight to measure their enzymatic activity after lysis and heat challenge treatment (30 min, 70 °C). Overall, 49 colonies were picked from fraction 1 and 39 colonies for fraction 2. Only a few cultures on the plate showed reasonable activity, indicating that the expression efficiency was rather low. From the first fraction, five colonies showed a reasonable activity. The initial activity test after lysis and heat challenge revealed, that all variants showed a high activity and thermostability, indicating that the thermolabile variant 258D could be excluded. The sequencing then proved that all picked colonies were holding the improved and thermostable variant 258N. For the second fraction, six colonies showed activity. All of the measured colonies showed a template variant like activity after lysis comparable to the template controls. The heat challenge did not reduce the initial activity. The only exception was the culture in well F6 with a 60% lower activity than the template variant even before the heatshock. Four of the six active colonies were sent for sequencing.

The sequencing proved that no colonies with template-like activity showed any mutation at position 258. The enzyme variant with reduced activity (F6) contained only a shorter DNA fragment, which did not include position 258. This could explain the lower activity of the enzyme. All results were summarized in Table 8.

Table 8: Activity of sorted *Bs*GDH variants from microfluidic sorting after incubation at elevated temperatures (20 hours, 50 °C). Variants were sorted from BsGDH mix holding variants 258N, 258D and template variant. Activity of the variants was determined after lyses of the culture and after an additional heatshock (30 min, 70 °C) to determine thermostability. Fraction 1 was initially sorted at a threshold of 3 V. Fraction 2 represents the resorted droplets of Fraction 1, excluding all sorted droplets. Fraction 2 was sorted at 3.5 V.

	ID	Lysed	Heat treatment	Remaining activity [%]	Variant
Fraction 1	A3	7302	7305	100	258N
	A9	6830	6865	101	258N
	A12	6691	6505	97	258N
	C10	6842	6978	102	258N
	E4	7718	7698	100	258N
Fraction 2	F6	1141	1329	117	[1]
	F9	3425	3853	113	WT
	G5	3786	4123	109	WT
	H3	3314	3387	102	[2]
	H5	3487	3919	112	WT
	H8	3387	3531	104	[2]

[1] variant was send for sequencing, but the obtained sequence did not include position 258. [2] Variants were measured in plate assay but not send for sequencing.

This experiment proved that separation between the template variant and improved variants with either thermostable or thermolabile properties was possible. The key point of this experiment, was to identify the incubation temperature and incubation time where the heat labile variants were inactivated and the template variant enzyme was outperformed. This made the difference in product formation detectable and allowed the sorting of only heat stable variants with improved activity. After showing that this approach worked for a simple library consisting of only 3 variants, the next aim was to expand this approach and apply it to a larger library. Therefore, the QuikChange library of *Bs*GDH with substitutions at position 258 was screened after incubation for 20 hours at 50 °C. From 803.192 screened droplets, 1.193 were sorted and rescreened in micro titer plates. 84 colonies were picked and their activity after lysis and heat shock (30 min, 70 °C) was measured. Out of the 84 variants, 26 showed an improved activity compared to the template variant. The best five performing variants were send for sequencing. Additionally, the variant D1 was picked, as it showed lower activity than the template variant and a loss in activity after the heatshock (Table 9).

Table 9: Activity of sorted *Bs*GDH variants from microfluidic resorting after incubation at elevated temperatures (20 hours, 50 °C). Variants were sorted from *Bs*GDH QuikChange library allowing all amino acids at position 258. Activity of the variants was determined after lyses of the culture and after an additional heatshock (30min, 70 °C) to determine thermostability. All sequences were aligned and corresponding amino acids at position 258 investigated.

ID	Lysed activity [a.u.]	Activity after heatshock [a.u.]	Remaining activity [%]	Mutation at 258	Additional substitutions
C1	6783	8388	124	Glycine	-
D1	1935	663	**34**	Alanine	M104V
E2	6540	7885	121	Glycine	-
E4	8327	8851	106	Asparagine	V112I, N97Y
E5	6166	740	**12**	Glycine	-
G7	5367	9218	172	Glycine	-

The sequencing revealed that the five best variants all showed either a glycine or an asparagine mutation at position 258. These amino acids were found to improve the activity of *Bs*GDH while maintaining its thermostability (3.2). The variant in well D1 showed an alanine mutation at postion 258 (parental amino acid), but also contained an additional mutation at position 104, which could be the reason for the decreased activity. The variant E4 contained additional substitutions at positions 97 and 112, which did not affect the activity. The reason for these additional substitutions could be the amplification with Taq polymerase. Previous experiments showed that the Taq polymerase was able to amplify minor amounts of DNA. Therefore, it was chosen over other polymerases, which would provide a proofreading ability, but were not able to amplify the low amounts of DNA retrieved from a normal sorting. Interestingly, the variant E5 exhibited a decrease in activity after the heat challenge even it contained a thermostable glycine mutation at position 258 and no additional substitutions within the sequenced region. Overall, it could be shown that the established method for finding improved variants, retaining the thermostability, was also working for a larger QuikChange library with and expected size of 20 variants. However, when applying the method to the prepared epPCR library of *Bs*GDH, no improved variants could be found (data not shown).

3.5 Exploitation of incubation temperature for discovery of thermostable variants

The aldehyde dehydrogenase of *Thermoplasma acidophilum* (*Ta*ALDH) was previously reported as an important enzyme in the bio-based ethanol and isobutanol production[144]. It was reported to catalyze the oxidation of glyceraldehyde to glycerate, but importantly it showed no activity towards the other aldehyde intermediates of the cascade, e.g. acetaldehyde and isobutyl aldehyde. Therefore, it could be used as a precise biocatalyst for the bioethanol and isobutanol production. However, with only ~0.4 U/mg specific activity at 50 °C, it was a rather slow enzyme compared to the other cascade enzymes and therefore acted as a bottleneck of the cascade. Enzyme engineering of six different positions could increase the specific activity up to 1.5 U/mg[145] but further improvement was not achieved. As a new approach, two aldehyde dehydrogenases from *Herbaspirillum seropedicae* (*Hs*ALDH) and *Variovorax paradoxus* (*Vp*ALDH) were chosen to replace the *Ta*ALDH within the cascade. Both, *Hs*ALDH and *Vp*ALDH exhibited higher enzymatic activity towards glyceraldehyde: 25 U/mg and 15 U/mg, respectively[146]. However, the temperature stability for both enzymes was lower compared to the variant of *Thermoplasma acidophilum*, which was a problem for prolonged application in the cascade reactions performed at 50 °C. While *Hs*ALDH lost all its activity at 50 °C after 1 hour incubation, the *Vp*ALDH retained 50% of its activity even after 10 hours of incubation at 50 °C. For both enzymes, it would be beneficial to increase the temperature stability and retain or even increase the enzymatic activity. For that reason, two epPCR libraries were created by Mariko Teshima with 63.400 and 79.800 variants for *Hs*ALDH and *Vp*ALDH, respectively. For the first feasibility test, both enzyme expressing strains were encapsulated in microfluidic droplets (λ=0.1) and their overall activity measured at 30 °C, 40 °C and 50 °C. The temperature stability of the different enzymes in droplets should be assessed and an initial understanding of the differences between occupied and vacant droplets gained (Figure 39).

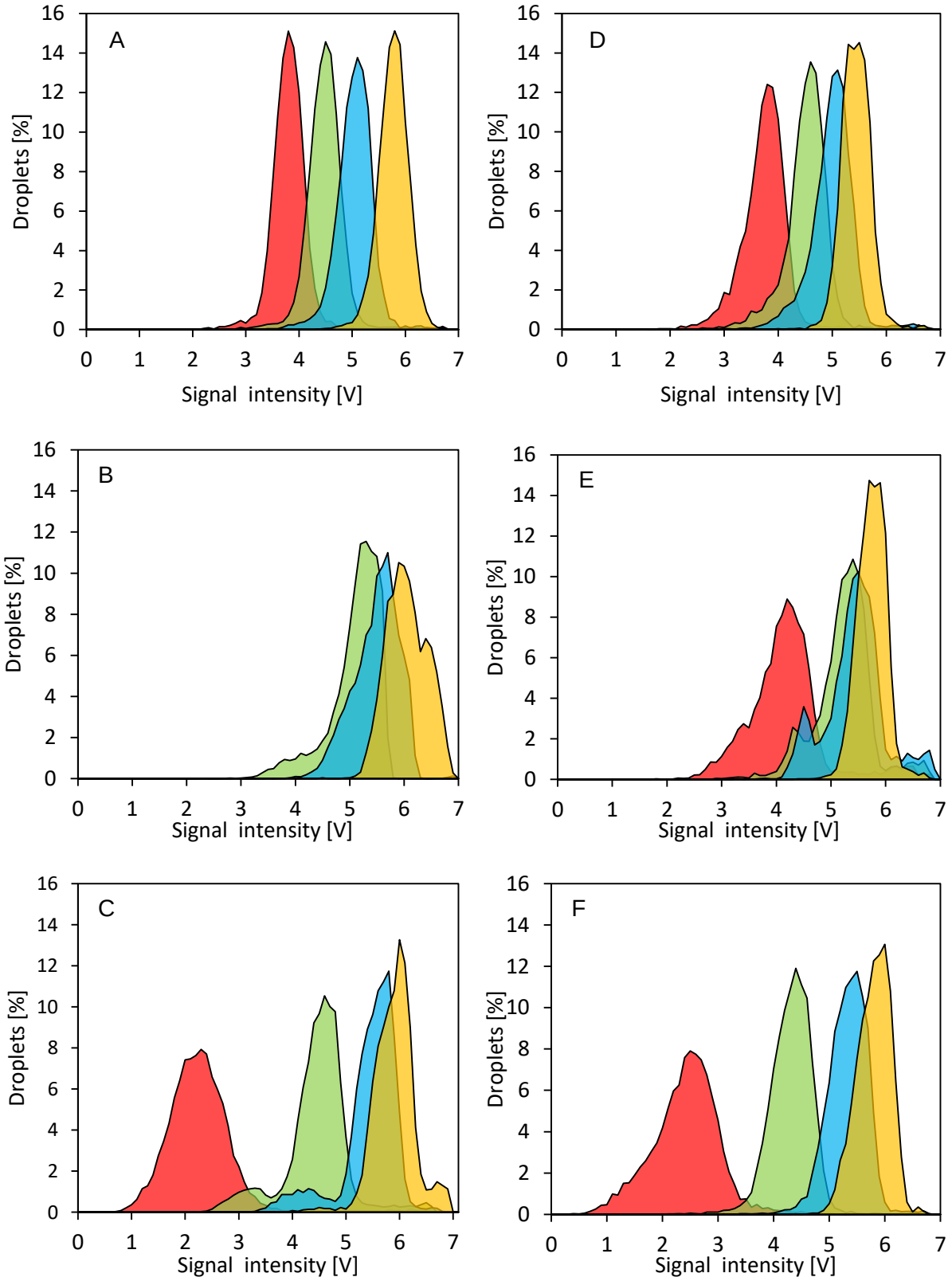


Figure 39: Activity measurement of *Vp*ALDH template (A-C) and *Hs*ALDH template (D-F) with 5 mM glyceraldehyde in microfluidic droplets. Droplets were incubated at 30 °C (A and D), 40 °C (B and E) and 50 °C (C and F). Samples were taken right after encapsulation (orange), 1.5 hours (blue), 4 hours (green) and 23 hours (red) after incubation. The chamber for the stored droplets of VpALDH broke before the last measurement. The measurment for 23 hours could not be obtained.

As expected, both enzymes exhibited only minor activity at 30 °C. The histogram of both populations showed a decent shoulder, which increased over time, indicating that the enzymes

were active over the complete time of investigation. The size of this shoulder was increased when the enzymes were incubated at 40 °C (Figure 39, B and E). For *Hs*ALDH the shoulder even turned into formation of another peak between 1.5 and 4 hours after incubation (Figure 39 E, blue and green). A similar observation could be made for *Vp*ALDH incubation at 50 °C. It showed a decent enzyme turnover after incubation at 1.5 and 4 hours, which could be recognized as a smaller peak left from the empty droplet population peak (Figure 39 C, blue and green). However, when measuring both populations again after 23 hours, the difference between the occupied and vacant population were not visible anymore, indicating that the enzymes were not stable over the whole time of observation. Additionally, it should be mentioned that the substrate glyceraldehyde showed an increased background in previous plate tests, when using higher substrate concentrations or incubation temperatures. Therefore, it was crucial to focus on optimizing the incubation temperature and time to determine the best parameters for identifying thermostable variants. In the previous chapter, it could be shown that gradient incubation in a thermocycler was a powerful technique to identify incubation temperatures and times at which enzymes or enzyme variants show different activities with respect to their temperature stability. Therefore, the cells of several enzyme variants were lysed after different incubation times at various temperatures and their residual activity was measured. By comparing these measured activities over the different temperatures and times, the biggest differences in enzymatic activity between those variants could be found and chances to separate thermostable from thermolabile enzymes were increased. The main advantage advantage of this approach compared to the microfluidic setup was the high parallelization. Eight different temperatures and times could be assessed in one experiment, which would be impossible for a microfluidic droplet experiment. Moreover, it was shown that the obtained temperatures and times for the different enzyme variants could be translated from the plate assay to the droplet-based microfluidic approach with minor adaptions. In this chapter, the described method was applied for engineering the thermostability of the cascade enzyme aldehyde dehydrogenase. Instead of single variants, the whole libraries of *Hs*ALDH and *Vp*ALDH were incubated and their average activity at different temperatures was compared to the activity of the template enzyme. Obviously, the measurement of residual activity for a library at different temperatures could not depict the activity of single improved variants within the library. However, it could provide a brief overview about the likelihood, which temperature and incubation time could be suitable to find improved variants. The final aim of the enzyme engineering approach was to increase the temperature stability to at least 50 °C. However, an engineering approach in several selection rounds would also be possible to increase the temperature stability stepwise. For that reason, eight incubation temperatures for the gradient incubation were chosen between 40 °C to 55 °C. With these small steps between the incubation temperatures, a better understanding of the new enzymes could be obtained

and the temperature stability assessed. As a comparison, both, templates and corresponding libraries were investigated (Figure 40).

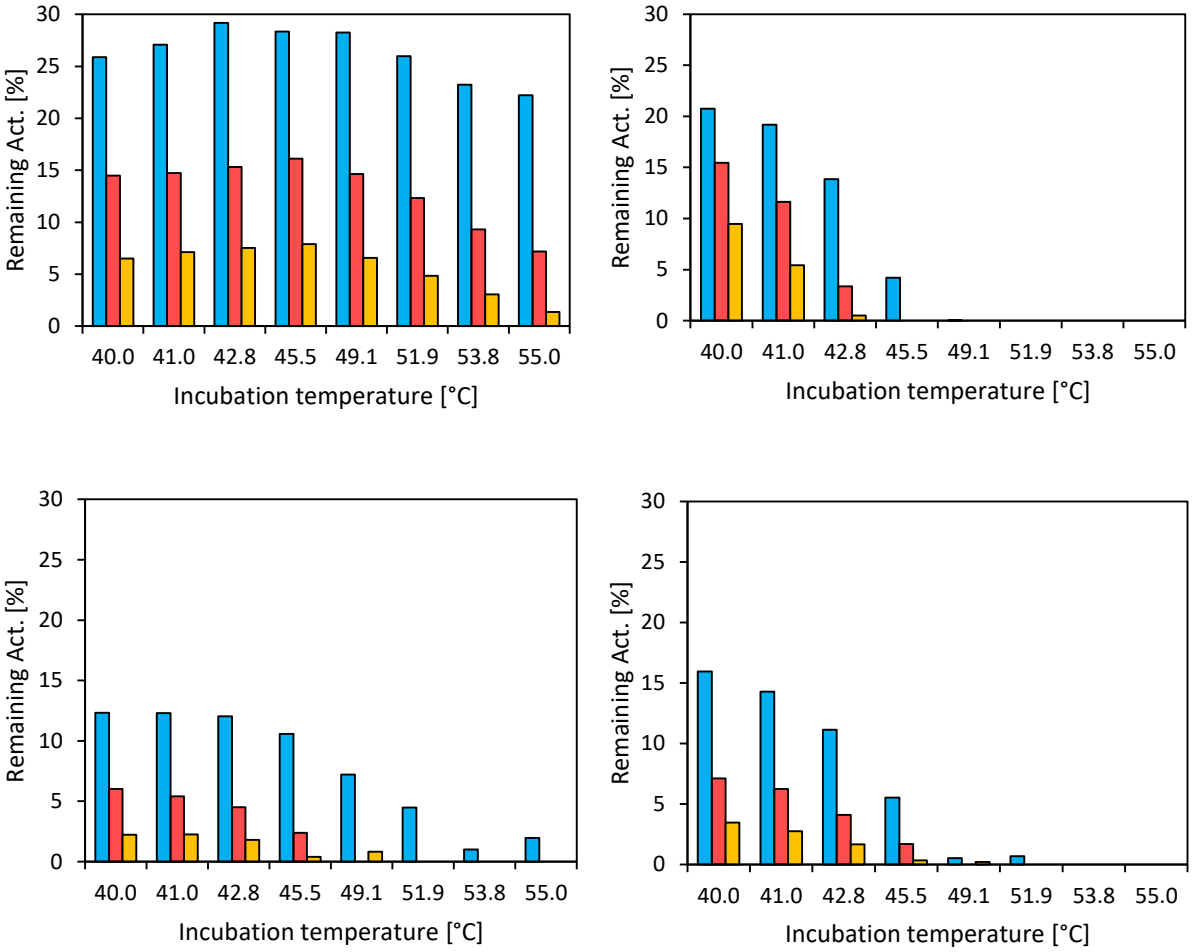


Figure 40: Remaining activity of ALDH after gradient incubation in thermocycler of *Variovorax paradoxus* (left) and *Herbaspirillum seropedicae* (right). Results for template variants are shown on top and library variants at the bottom. Cells were incubated for 1 hour (blue bar), 4 hours (red bar) and 6 hours (orange bar) and the residual activity of the enzyme supernatant was measured at 340 nm and compared with the initial activity (activity before heat treatment). Gradient was chosen between 40 °C and 55 °C. Steps were set by the thermo cycler.

As expected, both epPCR libraries lost some of their average activity, since (additional) mutations tend to destablize enzymes. For *Vp*ALDH a clear decrease of temperature stability for the majority of the variants can be observed. While the template enzyme showed activity after 6 hours of incubation at 55 °C, the library showed barely any activity after 6 hours incubation at 49.1 °C. Of course, it could not be excluded, that some of the variants within the library showed a strong increase in temperature stability. However, the likelihood that there were several enzymes with improved abilities, was rather low. The epPCR of *Hs*ALDH indicated some improvements in temperature stability for several variants. While the template enzyme did not show any residual activity after incubation at 45.5 °C for 4 hours, the library retained 1.7% of its initial activity. Even after 6 hours 0.3% of residual activity could be observed. This hypothesis was reinforced by the data obtained at 42.8 °C. After 1 hour of

incubation, the template enzyme showed 13.8% activity, while the library had only 11.1% residual activity. This ratio was changing after 4 hours of incubation. Here, the template enzyme had only 3.4% residual activity, while the library maintained 4.1% residual activity. After 6 hours there was only 0.5% of residual activity for the template enzyme and 1.7% for the library. Combined, the data suggested that there was a high likelihood, that the library contained variants with improved thermostability. The prescreening proved that the initial observations made with microfluidic droplets were comparable with the results obtained by the gradient incubation. The turnover for the *Vp*ALDH template enzyme was higher at 50 °C than at 40 °C. On the other hand, the *Hs*ALDH template was completely deactivated at that temperature and showed a higher activity around 40 °C. By comparing the obtained results from template enzyme and libraries, the course of the enzyme stability for different incubation temperatures could be analyzed in more detail. This knowledge helped to obtain the right incubation temperature and time for the corresponding libraries. For *Vp*ALDH library, an incubation temperature of 50 °C for 4 hours was chosen. According to the obtained data from the gradient incubation screening, some residual activity could be measured at 49.1 °C after 6 hours of incubation. Therefore, the incubation time was initially set to 4 hours at 50 °C.

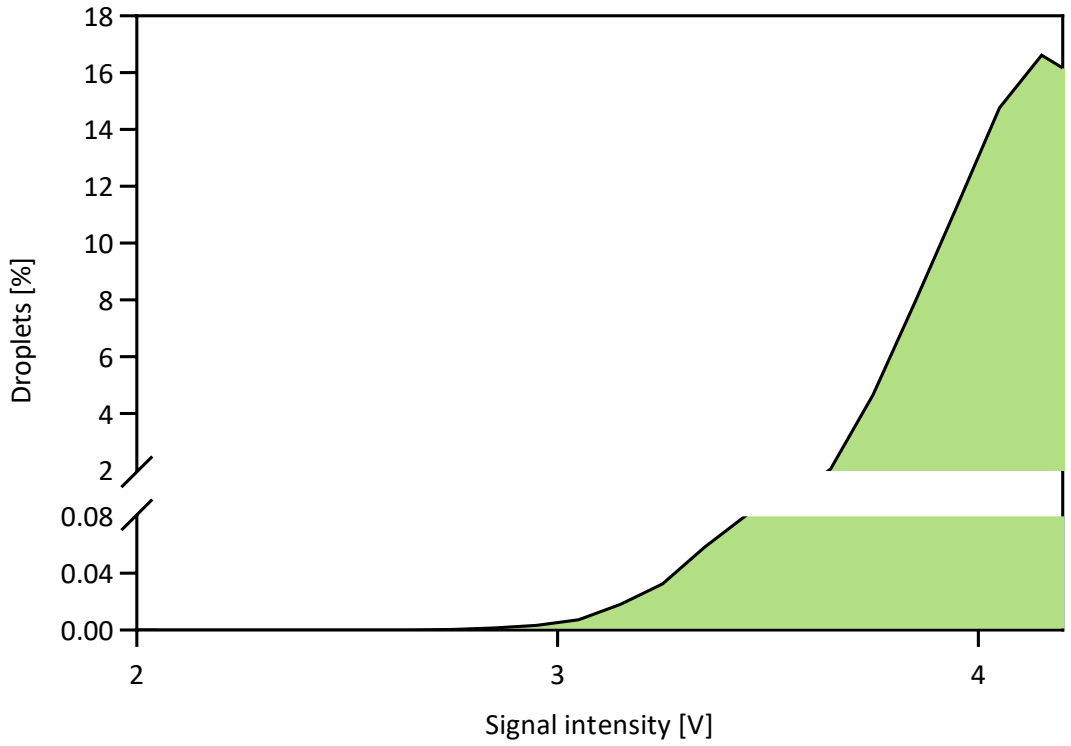


Figure 41: Histogram from microfluidic sorting of *Vp*ALDH library with 5 mM glyceraldehyde and 5 mM NAD after incubation at 50 °C for 4 hours. Histogram axis was divided in two parts to recognize the low amount of hits.

As predicted by the gradient incubation results, the majority of variants showed no activity at 50 °C and could not be distinguished from the empty droplet population (Figure 41). Only very few variants exhibited at least some enzymatic turnover and reached voltages down to 1.8 V.

For the sorting, the threshold was set to 3.3 V. Because only 0.02% of all measured droplets exhibited a higher activity, a total of 1,246,580 droplets needed to be screened (Figure 41) to reach a reasonable number of hits for the recovery. Previous experiments showed that the success rate for DNA recovery was drastically reduced, if the number of hits was too low. The chosen limit allowed double occupations, but was slightly above the theoretical limit for multiple occupations (cells ≥3; probability = 0.015%). This sorting strategy resulted in 246 hits, which were prepared for recovery. Unfortunately, the amplification PCR did not yield any DNA after gel extraction, which could be a result of the very low amount of starting DNA. For the sorting of *Hs*ALDH an incubation temperature of 45 °C for 6 hours was chosen. At this temperature and incubation time the data obtained from the gradient incubation prescreening showed a residual activity of 0.3% for the *Hs*ALDH library, while the template enzyme did not show any activity after 4 hours already (Figure 42).

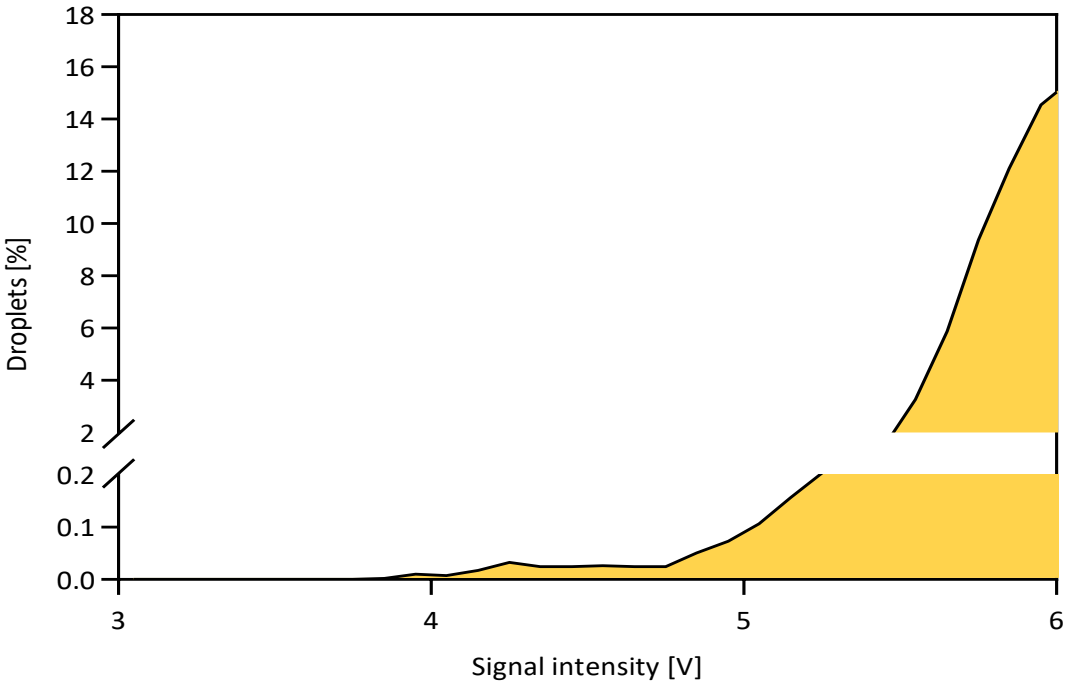


Figure 42: Histogram for microfluidic sorting of *Hs*ALDH library with 5 mM glyceraldehyde and 5 mM NAD after incubation at 45 °C for 6 hours. Histogram axis is divided in two parts to recognize the low amount of hits.

In contrast to the *Vp*ALDH library a clear shoulder could be recognized for the *Hs*ALDH library. For the sorting, the threshold was set to 4.5V. Overall, 1.3 Mio droplets were screened and 1072 hits were sorted. This corresponded to the best 0.08% of the droplet population. In this case, the amplification of the sorted DNA was successful and several plates with colonies were obtained. The colonies were picked by Mariko Teshima for rescreening and their activity was confirmed with a plate assay. From this plate rescreening Mariko Teshima could obtain 30 variants, which showed an improved temperature stability. The continuation of this enzyme engineering project will be conducted by Mariko Teshima.

4 Discussion and Outlook

4.1. Classification of the established microfluidic screening platform for discovery of improved enzyme variants

The functionality and specificity of an enzyme towards a substrate is defined by the order and assembly of the corresponding amino acids within a polypeptide chain. Naturally, for each position 20 different amino acids can be selected. For a dipeptide, this means already 400 different variants, but for an average enzyme consisting of 300 amino acids, it would mean a potential number of $2*10^{390}$ variants. Obviously, many variants can be excluded, as several repetitions of the same amino acid will not yield a functional enzyme Additionally, modern computing approaches[147–149] can help to predict the 3D-structure of enzymes and identify amino acids which are likely to improve the desired enzyme abilities, but until now these calculations are very time consuming. Moreover, *in silico* predictions can identify interesting positions, but not for all mutations predict the precise outcome of an amino acid substitution [150]. Consequently, enzyme engineering remains a numbers game and therefore the sample throughput in a screening is an essential parameter to explore the sequence space of an enzyme-encoding gene in a reasonable timespan. A higher sample throughput enables the investigation of bigger libraries with several mutations within each sequence and eventually increases the chances to find synergetic mutations[81,82]. Additionally, increased sample throughput allows the analysis of more samples in the same time period, which can be crucial for the survival of encapsulated cells[151] or product crosstalk between different droplet compartments over time[57].

Traditionally, enzyme engineering campaigns are conducted with 96 microtiter plates (MTP). They are comparably cheap, ready-to-use and allow enzyme screenings for 10^3 - 10^4 variants per day[37]. Supported by Liquid handling stations (LHS) and usage of smaller MTP formats (384 or 1536 MTP's) the throughput can be increased up to 10^5 variants per day[39]. Over the course of this thesis a microfluidic screening system was established that is capable of a throughput of 10^5-10^6 samples per day[78] and therefore exceeds the capacities of these traditional screening approaches by two orders of magnitude. In contrast to most reported microfluidic screening systems, it relies on an absorbance-based assay, instead of a fluorescence readout. For the last 14 years, droplet-based microfluidic sorters were predominantly fluorescent activated droplet sorters (FADS), since they can reach even higher throughputs of 10^8 or even 10^9 droplets per day[75,42]. The difference in sample throughput is mainly derived by the possibility to encapsulate the enzyme holding cells in smaller droplets. While the absorbance-based system works with 80µm droplets, FADS systems usually work with droplet sizes around 20µm. This is possible since the fluorescent signal intensity of the readout is depending on the

emitted light, which does not directly corelate to the pathlength of the droplet. The emitted light can be amplified by photo multiplier tubes (PMT's), while the absorbance signal measures the reduction of the transmitted light and cannot be amplified. With smaller droplets, higher flowrates can be applied in the microfluidic chip, without the risk to disrupt the droplet by shear forces on the channel walls or by dielectrophoretic sorting. Disrupting the droplets would potentially lose the genotype-phenotype linkage of the enzyme derived assay signal and the corresponding DNA of the cell. This would ultimately mean that the sequence for the sorted hit cannot be retrieved. Indeed, FADS sorters have a two orders of magnitude higher sample throughput[82] and three orders of magnitude higher sensitivity[152], compared to the absorbance activated droplet sorter (AADS) established in this work. However, there are also several advantages of the AADS system compared to the FADS:

First, most fluorescent assays for microfluidic screenings are based on a surrogate substrate. These surrogate substrates consist of the actual substrate and a fluorescent moiety, which is directly linked to the substrate. This moiety is only excited, once the screened enzyme performs either a transfer of an additional group or cleaves the bond between the substrate and the fluorescent moiety. However, by changing the properties of the native substrate the enzyme variants found in the screening campaign might have adapted to the fluorescent moiety of the surrogate substrate and not the native substrate. For example, Aharoni *et al.*[44] screened a directed evolution library of a glycosyltransferase with several different fluorescently labeled sugars. They found a variant with a 153-fold increase transfer activity towards a BIODIPY-tagged substrate. Since the new variant just showed improved transfer activity for this fluorescent moiety, it indicates an adaptation of the variant selectivity towards the fluorescent dye instead of the acceptor sugar moiety. Consequently, changing the substrate properties by adding (fluorescent) moieties might lead to mutations, which are only beneficial for the surrogate substrate, but not necessarily for the native substrate. It can be assumed that this problem appears more occasionally, but due to the publication bias (only successful results are published) this problem is widely neglected[57]. The AADS presented in this work, was directly established with a cofactor coupled absorbance assay (chapter 1.2.4) always using native substrates and avoiding the mutation problems towards surrogate substrates mentioned above. In the meantime, several papers for FADS have been published, exploiting cofactor coupled reactions[82,153,154].

Second, the established absorbance-based system is less complex than a FADS system. It does not require lasers or PMT's, which reduces necessary precautions (e.g. laser safety) and decreases the price for the sorting system[113]. Also, instead of an FPGA (Field programable gate array) card, the absorbance-based system works with a simple A/D-converter (analog to digital). The advantage of this converter is that the corresponding sorting program can be directly accessed in LabVIEW and it is not mandatory to compile and transfer the program to

a FPGA. This simplifies the debugging of the program and reduces the obstacle to establish a functional program, which allows online observation of the screening campaign and transfer of the acquired data to the computer for later analysis (chapter 3.1.4). Additionally, the required space for the sorting system is reduced since the AADS does not require a (separate) dark room, but can be placed on a bench in an ordinary laboratory.

Third, the absorbance-based assay with WST-1 is extremely robust and allows investigations over several days. Fluorescent dyes can show leakage[155], degradation either by light[156,157] or thermal degradation[158] and tend to generate a high background over time[159]. This makes it difficult to investigate the long-term performance of an enzyme. In contrast to that, WST-1 showed great potential in terms of leakage and stability. Initial tests in microtiter plates revealed that the product WST-1 formazan is not only stable at room temperature but also at 45 °C for 48 and 24 hours, respectively (chapter 3.1.3). The leakage of the dye between microfluidic droplets was tested for room temperature and 60 °C and for both experiments, no leakage could be observed over the course of 27 hours (chapter 3.1.5). Another advantage of WST-1, is the low background reaction. The first enzyme engineering experiments (chapter 3.2.1) revealed that the background reaction is strongly coupled to the concentration of the substrate. By reducing the xylose concentration from 400 mM to 100 mM the background reaction was decreased by a factor of ~ 6, which resulted in an improved resolution for hit identification. The robustness of WST-1 formazan even at elevated temperatures enabled the screening not only for specific activity of the enzyme, but also for thermostability over longer periods of time (chapter 3.4).

4.2. Enabling microfluidic ultrahigh-throughput screenings for long term incubations at elevated temperatures

Many industrial processes in food processing, fine chemical synthesis or waste water treatment cannot be performed at room temperature, but require elevated temperatures[160,161]. Therefore, the enzymes need to withstand elevated temperatures for the whole time of the enzymatic conversion, which can be several hours or even days[162]. Naturally, enzymes with high thermostability can be found in places, which are constantly warm e.g. hot springs[163], deserts[164] or hydrothermal vents[165]. Over the course of billion years these micro organisms adapted to the prevailing conditions and only microorganisms, which expressed enzymes with a high thermostability (thermozymes) were able to survive. However, screening these thermophiles can be very challenging. Especially the low sample throughput of conventional screening systems (agar and microtiter plates) makes it difficult to find improved enzyme variants in reasonable time. Therefore, high-throughput screening methods for thermozymes

are strongly desired[166]. Previous microfluidic studies for engineering thermostability of enzymes used tubing to apply heat to all enzyme candidates after droplet encapsulation. The tubing was guided either through a hot water bath for 30 seconds[167] or bound around a heated metal coil for 10 minutes[103]. The incubation time and temperature were adapted in a way to denature thermolabile enzymes. After incubation, the screening of the surviving enzymes was then performed at ambient temperatures. This led to the problem that enzymes, which can withstand high temperatures and are highly active at elevated temperatures, but less active at ambient temperatures, may not be found. Additionally, variants with improved specific activity, but slightly reduced thermostability, might be selected since the short heat shock was not sufficient to completely denature the enzyme. In chapter 3.4 it was shown, that it is crucial to screen the enzyme library for the complete time at the desired temperature (Figure 35 and Figure 36). Only then it is possible to monitor the heat-related enzyme denaturation over time and finally choose the variant with the highest activity and thermostability for the given industrial process parameters. However, for the successful screening of thermozymes with a microfluidic screening platform the two main difficulties for screenings at elevated temperatures had to be solved:

I. **Finding optimal screening conditions:** One main advantage of droplet-based microfluidics over conventional screening methods like agar plates or microtiter plates is the higher sample throughput[57]. However, for finding optimal screening conditions, it is more important to parallelize the experiments and test the influence of different reaction parameters e.g. substrate concentration (I), cofactor concentration (II), enzyme concentration (III) and temperature (IV) at different time points (V). The influence of different concentrations for assay components (I - III) can be tested by creating several droplet populations and capturing them in different compartments, measuring at different time points (V). However, testing the optimal conditions to engineer thermostability can be a very time-consuming task for droplet-based microfluidic screenings. Especially if epPCR was used for library preparation, it is hard to estimate how the mutations influence the functionality of the enzyme variants at different temperatures (IV). Encapsulation of the whole library can only be performed for a limited combination of assay component concentrations and the incubation at elevated temperatures is critical, because it can only be performed with an incubator. This tremendously reduces the options for parallelization and can extend the investigation time for the optimal screening conditions to several days. The result might be batch-to-batch variations within the expression levels of the enzyme variants. To keep this error as small as possible, enzyme-holding cells were lysed and a gradient incubation was performed in a thermocycler to mimic different incubation temperatures. After defined timepoints,

the residual activity for each incubation temperature was measured (chapter 2.6.1). As a result, up to 12 different assay component concentrations (I-III) can be incubated at 8 different temperatures (IV) at up to 8 different timepoints (V). This means that from one experiment 768 measurement points can be obtained and help to find optimal screening conditions with one enzyme batch. More importantly, it could be shown that with small adaptions (incubation temperature needs to be reduced by 5 °C) the results from the PCR-tubes are transferable to the microfluidic droplet experiment (chapter 3.4). The reduction in incubation temperature can be explained by the lower enzyme concentration in the droplets, which makes the enzymes more prone to denaturation.

II. **High temperatures for prolonged time**: Constant high temperatures are one of the biggest threats for enzyme engineering campaigns. If the process is not conducted in a closed system, the volume is reduced over time since the liquid phases tend to evaporate. Moreover, the concentration of the non-evaporating particles (e.g. salts) increases and therefore change the assay conditions within the reaction chamber, which eventually influences the intensity of the assay signal. For droplet-based microfluidic systems, this problem is even more significant. The continuous phase consists of fluorinated oil, which easily evaporates at room temperature within several minutes. With higher temperatures, the evaporation is even accelerated. Evaporation of the continuous phase will finally lead to the merging of the microfluidic compartments and the loss of the unique phenotype-genotype connection. To avoid evaporation, it is necessary to keep the droplet compartments within a closed carrier system. As described above, several groups have tried to mimic constant high temperatures, by applying heat challenge to all enzyme variants within a tubing. Since the assay incubation afterwards is performed at room temperature, this might lead to the loss of variants, which are only active at elevated temperatures. In order to avoid the loss of these highly active variants, elevated temperatures in this study were applied over several hours to replicate industrial process conditions. To realize this objective, all droplets were collected in a droplet collection chamber[104] and in- and outlets of the chamber were sealed after droplet production. The chamber was then placed in an incubator with a predefined temperature (45 °C - 65 °C). With this approach, evaporation of the fluorinated oil could be reduced to a minimum. However, after reinjection the droplet population showed an increased number of merged droplets. This can be explained by the expansion of the droplets at higher temperatures[168]. The expansion of the droplets increased the pressure within the droplet collection chamber and raised the chances for merging of droplets. By releasing some of the fluorinated oil within

the chamber and replacing the space by air, the droplet quality could be improved. Apparently, the droplet chamber allows gaseous exchange and therefore reduced the overall pressure on the stored droplets. This step decreased the merging of the droplets to a neglectable level and allowed investigations of enzyme variants after incubation at 45 °C to 65 °C. The incubation time was chosen in a way to get the highest difference between template variant and library variants.

4.3. Application of AADS for enzyme engineering, possible improvements and future perspectives

Over the course of this thesis, an absorbance-based microfluidic screening system was established and the functionality was shown by engineering three different enzymes towards higher activity and/or thermostability. However, the two main bottlenecks of the absorbance activated droplet sorting remain the low sensitivity and (compared to FADS) the lower throughput. The high concentrations of assay product that are needed to differentiate improved variants from the template variant (at least 100 µM[78]) resulted in long incubation times. The incubation time was shortened by applying higher temperatures to the encapsulated enzymes and thereby increasing the enzymatic reaction velocity. Still, it was necessary to incubate the enzyme variants for several hours to reach reasonable separation of improved variants. This bottleneck was used as an advantage to observe long-term stability of the designated enzymes at elevated temperatures. However, enzymes with a low conversion rate (~k_{cat} = 1 s^{-1}) were not screenable with the AADS, since the background reaction did not allow distincitions even after several hours/days of incubation. This excludes the screening of many interesting wildtype enzymes. With the current capacities, the AADS system can be applied for the niche purpose of engineering temperature stability of enzymes with a decent conversion rate. However, with the low sensitivity and a throughput of 10^5 – 10^6 samples per day, the AADS established in this work is still clearly inferior to already established FADS systems. Since the intial publication of AADS in 2016[78], several papers have been published, which have potential to improve the performance of the AADS:

4.3.1 Increasing the sensitivity of the AADS system

The intensity of the absorbance-based signal is defined by the Beer-Lambert law (chapter 1.2.4). In order to increase the signal intensity either the (I) product concentration, (II) pathlength or the (III) molar extinction coefficient of the assay product need to be increased.

(I) **Product concentration:** Zurek *et al.*[169] improved the product concentration by increasing the number of enzymes within each microfluidic compartment. After cell encapsulation, the

droplets were captured in a droplet chamber and incubated at 37 °C to allow cell proliferation. They showed that differences in cell growth could be kept at a standard deviation below 10% by applying continuous oxygenation of the encapsulated cells with fresh fluorinated oil. After sufficient cell growth, the droplets were reinjected into a second chip, where lysis agent and substrates were added to the droplet via picoinjection to start the assay reaction. After another 2 hours incubation the droplets were sorted. The higher cell density led to higher enzyme concentrations within the droplets and resulted in a 12-fold higher product conversion. The increased enzyme concentration allowed the sorting of an amine dehydrogenase library with a spec. activity of $k_{cat}/K_M = 0.5\ s^{-1}\ mM^{-1}$ for the template variant, which could not be identified by single cell encapsulation.

(II) **Pathlength:** Yang *et al*.[170] developed a microfluidic screening system with an absorbance-based readout. To increase the signal intensity of their fluorescein droplets, the measurement chip was designed with a very narrow analysis channel. The dimensions were 35 x 26 x 800 µm (height x width x length). They used droplets with a diameter of 96 µm and squeezed them into the narrow analysis channel. This results in a 16-fold increase in pathlength compared to the microfluidic screening chip used in this work. With this pathlength they could reach a theoretical sensitivity limit of 276 nM fluorescein ($\varepsilon_{Fluorescein}$=59,125 M^{-1} cm^{-1}). However, since the droplets are rather big and need to completely leave the analysis channel before the next measurement can be performed, a decent spacing of droplets is necessary. Therefore, a throughput of only 15 droplets per second could be demonstrated. A similar approach was presented by Duncombe *et al*. [171]. They measured UV-vis spectra in a z-shaped analysis channel. In order to reach higher sensitivity but retain high-throughputs, 7 different dimensions for the analysis channel were tested. Finally, the group decided on the following dimensions 105 x 50 x 100 µm (height x width x length) for the analysis channel. For their sortings, they used a throughput of 30 Hz, but they mentioned that up to 100 Hz could be possible for a screening. They evaluated the sensitivity of their system by sorting a mixture of droplets with different BSA-concentrations. By increasing the integration time of the measurement (and reducing the throughput) they were able to sort droplets with a BSA concentration as low as 10 µM, which would be a 10-fold improvement compared to the initial AADS publication. Another interesting way to extend the optical pathlength was shown by Rushworth *et al.*[172]. They used cavity mirrors to enhance the optical pathlength. For the measurement, a white diode is focused on the encapsulated sample. The light is guided through the sample and onto a cavity mirror. This mirror reflects 98% of the light back into the sample and to the cavity mirror on the opposite side of the channel. By reflecting the light several times and guiding it through the sample the pathlength for the measurement can be increased by a factor of 28 and the resulting sensitivity is enhanced by a factor of 5. The main bottleneck of this

approach remains the long integration time of 50 ms, which dictates the later sample throughput of the system.

(III) **Molar extinction coefficient:** Another option to increase the sensitivity of an absorbance-based screening system is the application of different assay dyes with higher molar extinction coefficients. In many microtiter plate assays the conversion of NAD^+ to $NADH^+ + H^+$ is monitored at 340 nm to obtain the specific activity of an enzyme. For most microfluidic measurements, this approach is not suitable since the pathlength for the measurement is reduced from several millimeters to a few micrometers. In order to retain the sensitivity of the screening system, Gielen *et al.*[78] approached this problem by application of WST-1 dye ($\varepsilon_{WST\text{-}1}$=37,000 M^{-1} cm^{-1}). Compared to NADH this dye has a 6-fold higher molar extinction coefficient, which increases the overall sensitivity of the system. Reviewing literature a brought spectrum of dyes can be found (Table 10).

Table 10: List of absorbance dyes for potencial application in microfluidic assays.

Dye	Absorption max [nm]	Ext. coefficient [M^{-1} cm^{-1}]	Reaction
INT Formazan	503	19.300 [173]	NAD-Cofactor
MTT	570	13.000 [174]	NAD-Cofactor
MTS	490	26.900 [175]	NAD-Cofactor
XTT	470	21.600 [175]	NAD-Cofactor
NBT	530	25.400 [176]	NAD-Cofactor
WST-1	435	37.000 [78,177]	NAD-Cofactor
WST-3	436	30.100 [177]	NAD-Cofactor
WST-4	537	14.000 [178]	NAD-Cofactor
WST-8	460	30.700 [177]	NAD-Cofactor
Guaiacol	470	26.600 [179]	Peroxidase
Pyrogallol	430	2.470 [179]	Peroxidase
ABTS	436	29.300 [179]	Peroxidase
DMP	469	49.600 [180]	Mn Peroxidase
DA-64	727	6.500 [181]	Peroxidase
2ACy-2	537	81.700 [182]	NAD-Cofactor

However, most dyes have a lower extinction coefficient than WST-1 and would not improve the sensitivity of the system. The most promising candidate 2ACy-2, was tested but found not suitable for microfluidic assays since it showed leakage towards the oil phase. By changing the chemical structure to a more polar molecule, leakage could be minimized and allow product detection with a higher extinction coefficient.

As a conclusion, the sensitivity of the AADS could be improved by more than two orders of magnitude by applying the cell growth in droplets described by Zurek *et al.* [169] (12-fold improvement) and combining it with the chip design published by Yang *et al*.[170] (up to 16-fold improvement). To avoid the loss of sample throughput the dimensions for the measurement channel could be adapted by accepting a certain loss in throughput.

4.3.2 Increasing the sample throughput of the established AADS - system

The throughput of a sorting system mainly defines how many samples per second can be analyzed. But additionally, the sample throughput also decides about the expected standard deviation, when measuring the enzyme activity of each droplet. In conventional screenings, the longer a sorting takes, the more enzyme turnover is conducted and factors like cell viability and nutrition availability (medium and oxygen) can play an important role for the overall enzyme turnover. Moreover, droplet production is approximately ten times faster than droplet sorting and the order of droplets is not kept. This also leads to higher standard deviations depending on the total amount of droplets being analyzed. With higher standard deviations, more conservative sorting thresholds need to be applied to keep the number of false positive hits low. For the absorbance-based assay used in this work, these errors can be assumed insignificant. Since all cells are immediately lysed after encapsulation, the amount of produced enzymes should be comparable and factors like cell viability and nutrition availability have minor influence. Furthermore, the incubation time for the thermophilic enzymes studied in this work was always chosen in that way that the conversion at elevated temperatures led to the majority of the measured product compared to the sorting process at room temperature. However, the throughput of a sorting system ultimately decides how many samples can be measured overall in a reasonable time and it therefore defines the library size, which can be investigated. The AADS-system established in this work is capable of a sorting rate between 100 – 150 Hz. Compared to state of the art FADS systems this is one order of magnitude less throughput, which can be directly translated into the possible library size that can be investigated. In order to increase the throughput of the AADS system, several advances have been published. For example, Medcalf *et al*.[113], improved the AADS previously developed by their group[78]. They changed the chip design and added a side oil channel behind the fiber measurement. The channel is guided in a 45° angle, which allows pushing the droplets faster out of the sorting junction and reduces the number of false positive hits by passive flow. The addition of the oil channel after the narrow analysis channel decreases the initial velocity of droplets in the sorting junction and thereby reduces the risk to disrupt the droplets. However, to achieve their goal of a 1 kHz sorting rate, they reduced the droplet size to 50 µm diameter since bigger droplets would most likely be disrupted by the dielectrophoretic sorting or the sorting divider. In order to keep the sensitivity of their sorting system in the low micromolar range, they removed the masking side peaks of the absorbance signal by matching the refractive index of oil and droplets (water). They used a 35% 1,3-bis(trifluoromethyl)-5-bromobenzene solution mixed with the HFE 7500 fluorinated oil and obtained sharp, unmasked peaks. Another approach was shown by Richter *et al.*[126]. They established an absorbance-based sorting system with an acoustic sorting (AcAADS). The integrated tapered inter digitated transducer generates a traveling surface acoustic wave (TSAW), which can

precisely deflect the droplet into the sorting channel. The group showed that with this approach, sorting rates of 1.3 kHz are possible for 70 µm droplets. Additionally, they use the same fiber-based measurement system, but added two bi-concave lenses in front of the fibers. The lenses help to avoid masking of the droplets signal peaks, which are derived by the different refractive indexes of water and fluorinated oil. By incorporation of the lenses, they removed the masking signal peaks in their detection and reduced the minimal detection limit down to 30 µM of tartrazine. A more general approach was published by Sciambi *et al.*[42]. They changed the geometry of the sorting chip by replacing the conventional hard wall divider with a gapped divider (Figure 43).

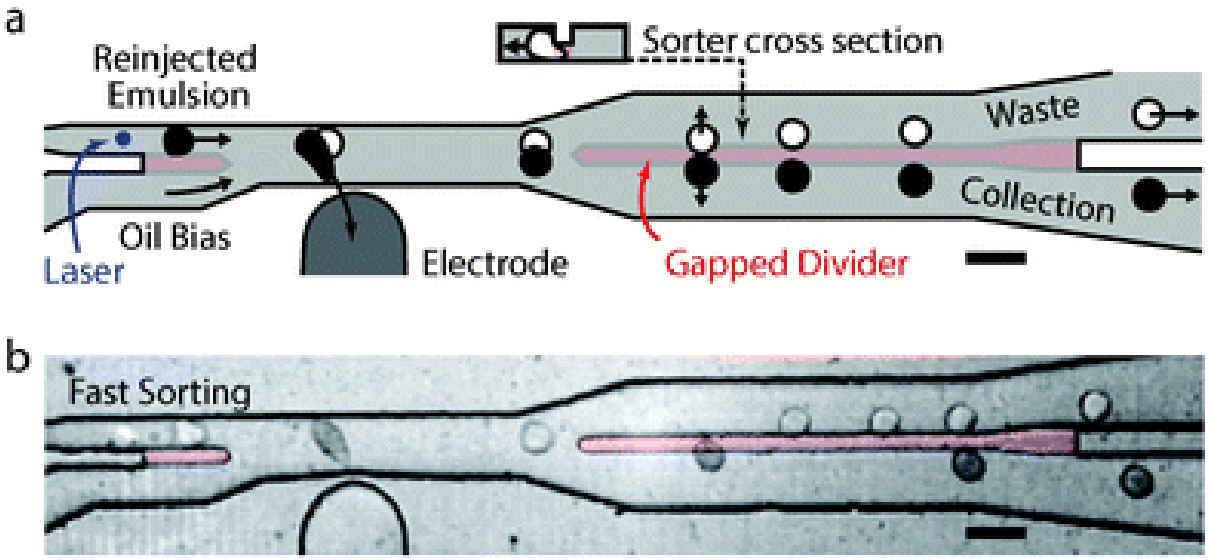


Figure 43: Principle of function gapped divider. Only dark droplets enable sorting electrode. (A) Schematic of gapped divider. White walls are formed with solid PDMS reaching from top to the bottom of the chip. Light red walls only have a certain height within the channel structure and allow droplets to go below (see sorter cross section sketch) and therefore do not disrupt the droplets. (B) High speed camera showing the sorting junction of the chip. Colorless droplets are not triggered and therefore go to the waste channel (top), while dark droplets are shortly dragged by the electrode towards the collection channel (bottom). Figure adapted from Sciambi et al.[42]

This gapped divider reaches only part of the way from the channel ceiling to the floor and therefore extends the way and the time a droplet can travel before moving in the sorting channel. Droplets are squeezed under the gapped divider, but due to the droplet Laplace pressure they eventually move in the direction, where the majority of the droplet is located once entering the squeezed zone under the gapped divider. Droplet disruption at the divider or by applying too high dielectrophoretic sorting forces can be avoided. With this change, the group could increase the sorting speed up to 30 kHz for 25µm droplets. Implementing a similar gapped divider into an absorbance-based sorting chip could increase the sorting speed by at least one order of magnitude.

4.4 Outlook

Over the course of this thesis, the AADS was used for enzyme engineering campaign of three different enzymes. The established microfluidic screening system proved to be a valuable tool for enzyme engineering and especially for the investigation of long-term thermophilic activity. It was shown that AADS could be applied for different screening goals: substrate specificity (chapter 3.2), cofactor specificity (chapter 3.3) or thermostability (chapter 3.5). However, until now the application of AADS is very limited. The following paragraph gives an outlook with possible options how to extend the application possibilities of AADS:

In this work, all engineered enzymes belong to the group of dehydrogenases. As already pointed out in chapter 4.3.1, there are only a few assay dyes with high extinction coefficients, that could be applied in absorbance-based microfluidics. Hence, there is a high demand to develop novel, stable assay dyes to extend the applications for droplet-based AADS sortings. For example, Kračun *et al*.[183] describe how to synthesize a variety of chromogenic substrates based on four different chlorotriazine dyes. They combine polysaccharides and proteins with the dyes and show their application for hydrolases, lytic polysaccharide monooxygenases and proteases. This would enable the sorting for another three enzyme groups and expand the applicability of AADS. A different way to observe enzymatic reactions, where no assay substrate is available was shown by Debon *et al.*[82]. The group used an enzyme cascade to engineer a cyclohexylamine oxidase. The reaction from the substrate amine to the product imine reduces the cofactor FAD to $FADH_2$. By oxygen-dependent recycling, the cofactor is recycled and equimolar amounts of hydrogen peroxide are produced. The hydrogen peroxide is then converted from a horseradish peroxidase, yielding a fluorescent product that can be detected by the FADS system. Another cascade reaction was shown by Kolaitis *et al*.[153]. The group engineered an aldehyde dehydrogenase towards the oxidation of L-guluronic acid. In this reaction, the cofactor NAD is converted to NADH. By implementation of a NADH oxidase (NOX) hydrogen peroxide is produced and can then be detected by a horseradish peroxidase, which reduces Amplex UltraRed, resulting in a strong fluorescent signal. Learning from these two approaches, an enzyme cascade could also be developed for the AADS system and therefore expand the application to different enzyme classes. Another aspect is that the published AADS system relies on only one color read-out. Multi-color measurements would allow studying protein-protein interactions or normalization of enzyme activity. Following the example from Zurek *et al.*[169], where cells are proliferated within the droplet, it would be a great addition to compare the number of enzymes that contributed to the result to the total enzyme turnover. This could be achieved by coexpressing another protein with a high extinction coefficient together with the enzyme. Like this, the measured signal intensity could be normalized and false positive hits by higher cell density or expression level could be avoided.

This dual or multichannel analysis could be realized by measuring UV-Vis spectra as shown by Duncombe *et al.*[171]. Another option would be the implementation of a second fiber-based measurement as demonstrated by Cole *et al*.[184] for a FADS system. Overall, the AADS system has great potential to broaden its applications and be, at least for some applications, a serious competitor to FADS systems.

5 References

1. McGovern, P. E. *et al.* Fermented beverages of pre- and proto-historic China. *Proc. Natl. Acad. Sci.* **101**, 17593–17598 (2004).
2. Olsen, H. S. & Falholt, P. The role of enzymes in modern detergency. *J. Surfactants Deterg.* **1**, 555–567 (1998).
3. Khan, U. & Selamoglu, Z. Use of Enzymes in Dairy Industry: A Review of Current Progress. *Arch. Razi Inst.* **75**, 131–136 (2020).
4. Barman, S., Sit, N. & Badwaik, L. S. Pectinase production by Aspergillus niger using banana (Musa balbisiana) peel as substrate and its effect on clarification of banana juice. *J. Food Sci. Technol.* **52**, 3579–3589 (2015).
5. Sudhakaran, V. K. & Shewale, J. G. Enzymatic splitting of penicillin V for the production of 6-APA using immobilized penicillin V acylase. *World J. Microbiol. Biotechnol.* **9**, 630–634 (1993).
6. Ranganathan, S. V., Narasimhan, S. L. & Muthukumar, K. An overview of enzymatic production of biodiesel. *Bioresour. Technol.* **99**, 3975–3981 (2008).
7. Chien, A., Edgar, D. B. & Trela, J. M. Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *J. Bacteriol.* **127**, 1550–1557 (1976).
8. Li, R. *et al.* Computational redesign of enzymes for regio- and enantioselective hydroamination article. *Nat. Chem. Biol.* **14**, 664–670 (2018).
9. Liu, L. & Corma, A. Metal Catalysts for Heterogeneous Catalysis: From Single Atoms to Nanoclusters and Nanoparticles. *Chem. Rev.* **118**, 4981–5079 (2018).
10. Fritz, B., Aichele, C. & Schmidt, M. Environmental impact of high-value gold scrap recycling. *Int. J. Life Cycle Assess.* **25**, 1930–1941 (2020).
11. Zdarta, J., Pinelo, M., Jesionowski, T. & Meyer, A. S. Upgrading of Biomass Monosaccharides by Immobilized Glucose Dehydrogenase and Xylose Dehydrogenase. *ChemCatChem* **10**, 5164–5173 (2018).
12. Ma, Y., Li, B., Zhang, X., Wang, C. & Chen, W. Production of Gluconic Acid and Its Derivatives by Microbial Fermentation: Process Improvement Based on Integrated Routes. *Front. Bioeng. Biotechnol.* **10**, 1–12 (2022).
13. Ramachandran, S., Fontanille, P., Pandey, A. & Larroche, C. Gluconic Acid: Properties, Applications and Microbial Production. *Food Technol. Biotechnol.* **44**, (2006).
14. Niu, W., Molefe, M. N. & Frost, J. W. Microbial Synthesis of the Energetic Material Precursor 1,2,4-Butanetriol. *J. Am. Chem. Soc.* **125**, 12998–12999 (2003).
15. Werpy, T. & Petersen, G. Top Value Added Chemicals from Biomass Volume I. *Us Nrel* Medium: ED; Size: 76 pp. pages (2004). doi:10.2172/15008859
16. Wong, C. H., Drueckhammer, D. G. & Sweers, H. M. Enzymatic vs. Fermentative

Synthesis: Thermostable Glucose Dehydrogenase Catalyzed Regeneration of NAD(P)H for use in Enzymatic Synthesis1. *J. Am. Chem. Soc.* **107**, 4028–4031 (1985).

17. Rollin, J. A., Tam, T. K. & Zhang, Y. H. P. New biotechnology paradigm: Cell-free biosystems for biomanufacturing. *Green Chem.* **15**, 1708–1719 (2013).
18. Zhou, X. *et al.* Coupled (R)-carbonyl reductase and glucose dehydrogenase catalyzes (R)-1-phenyl-1,2-ethanediol biosynthesis with excellent stereochemical selectivity. *Process Biochem.* **50**, 1807–1813 (2015).
19. Longwell, C. K., Labanieh, L. & Cochran, J. R. ScienceDirect High-throughput screening technologies for enzyme engineering. *Curr. Opin. Biotechnol.* **48**, 196–202 (2017).
20. Kuo, Y.-M., Henry, R. A. & Andrews, A. J. Measuring specificity in multi-substrate/product systems as a tool to investigate selectivity in vivo. *Biochim. Biophys. Acta* **1864**, 70–76 (2016).
21. Scherer, M., Fleishman, S. J., Jones, P. R., Dandekar, T. & Bencurova, E. Computational Enzyme Engineering Pipelines for Optimized Production of Renewable Chemicals. *Front. Bioeng. Biotechnol.* **9**, 673005 (2021).
22. Schmidt-Dannert, C. & Arnold, F. H. Directed evolution of industrial enzymes. *Trends Biotechnol.* **17**, 135–136 (1999).
23. Crow, J. F. & Abrahamson, S. Seventy years ago: Mutation becomes experimental. *Genetics* **147**, 1491–1496 (1997).
24. Sega, G. A. A review of the genetic effects of ethyl methanesulfonate. *Mutat. Res.* **134**, 113–142 (1984).
25. Kimball, R. F. The mutagenicity of hydrazine and some of its derivatives. *Mutat. Res. Genet. Toxicol.* **39**, 111–126 (1977).
26. Leung, D. W., Chen, E. & Goeddel, D. V. A Method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* (1989).
27. Cadwell, R. C. & Joyce, G. F. Randomization of genes by PCR mutagenesis. *Genome Res.* **2**, 28–33 (1992).
28. Potapov, V. & Ong, J. L. Examining sources of error in PCR by single-molecule sequencing. *PLoS One* **12**, 1–19 (2017).
29. McCullum, E. O., Williams, B. A. R., Zhang, J. & Chaput, J. C. Random mutagenesis by error-prone PCR. *Methods Mol. Biol.* **634**, 103–109 (2010).
30. Patrick, W. M., Firth, A. E. & Blackburn, J. M. User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries. *Protein Eng. Des. Sel.* **16**, 451–457 (2003).
31. Juili L. Lin-Goerke, D. J. R. and J. D. B. PCR-Based Random Mutagenesis Using Manganese and Reduced dNTP Concentration. *Biotechniques* **23**, 409–412 (1997).
32. Copp, J. N., Hanson-Manful, P., Ackerley, D. F. & Patrick, W. M. Error-prone PCR and

effective generation of gene variant libraries for directed evolution. *Methods Mol. Biol.* **1179**, 3–22 (2014).

33. Hermes, J. D., Blacklow, S. C. & Knowlest, J. R. Searching sequence space by definably random mutagenesis : Improving the catalytic potency of an enzyme. *Proc. Natl. Acad. Sci.* **87**, 696–700 (1990).
34. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).
35. Lacks, S. & Greenberg, B. A deoxyribonuclease of Diplococcus pneumoniae specific for methylated DNA. *J. Biol. Chem.* **250**, 4060–4066 (1975).
36. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51–59 (1989).
37. Xiao, H., Bao, Z. & Zhao, H. High throughput screening and selection methods for directed enzyme evolution. *Ind. Eng. Chem. Res.* **54**, 4011–4020 (2015).
38. Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **16**, 379–394 (2015).
39. Wunder, F., Kalthof, B., Muller, T. & Huser, J. Functional Cell-Based Assays in Microliter Volumes for Ultra-High Throughput Screening. *Comb. Chem. High Throughput Screen.* (2008). doi:10.2174/138620708785204054
40. Kintses, B. *et al.* Picoliter cell lysate assays in microfluidic droplet compartments for directed enzyme evolution. *Chem. Biol.* **19**, 1001–1009 (2012).
41. Cossarizza, A. *et al.* Guidelines for the use of flow cytometry and cell sorting in immunological studies. *Eur. J. Immunol.* **47**, 1584–1797 (2017).
42. Sciambi, A. & Abate, A. R. Accurate microfluidic sorting of droplets at 30 kHz. *Lab Chip* (2015). doi:10.1039/c4lc01194e
43. Herzenberg, L. A., Sweet, R. G. & Herzenberg, L. A. Fluorescence-activated cell sorting. *Sci. Am.* (1976). doi:10.1038/scientificamerican0376-108
44. Aharoni, A. *et al.* High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat. Methods* **3**, 609–614 (2006).
45. Ruff, A. J., Dennig, A., Wirtz, G., Blanusa, M. & Schwaneberg, U. Flow cytometer-based high-throughput screening system for accelerated directed evolution of P450 monooxygenases. *ACS Catal.* **2**, 2724–2728 (2012).
46. Thorsen, T., Roberts, R. W., Arnold, F. H. & Quake, S. R. Dynamic pattern formation in a vesicle-generating microfluidic device. *Phys. Rev. Lett.* **86**, 4163–4166 (2001).
47. Holtze, C. *et al.* Biocompatible surfactants for water-in-fluorocarbon emulsions. *Lab Chip* **8**, 1632 (2008).
48. Baret, J. C. Surfactants in droplet-based microfluidics. *Lab Chip* **12**, 422–433 (2012).

49. Ostafe, R., Prodanovic, R., Ung, W. L., Weitz, D. A. & Fischer, R. A high-throughput cellulase screening system based on droplet microfluidics. *Biomicrofluidics* **8**, 6–10 (2014).
50. Beneyton, T. *et al.* Droplet-based microfluidic high-throughput screening of heterologous enzymes secreted by the yeast Yarrowia lipolytica. *Microb. Cell Fact.* **16**, (2017).
51. Zinchenko, A. *et al.* One in a million: Flow cytometric sorting of single cell-lysate assays in monodisperse picolitre double emulsion droplets for directed evolution. *Anal. Chem.* **86**, 2526–2533 (2014).
52. Bauer, W.-A. C., Fischlechner, M., Abell, C. & Huck, W. T. S. Hydrophilic PDMS microchannels for high-throughput formation of oil-in-water microdroplets and water-in-oil-in-water double emulsions. *Lab Chip* **10**, 1814–1819 (2010).
53. Romanowsky, M. B., Abate, A. R., Rotem, A., Holtze, C. & Weitz, D. A. High throughput production of single core double emulsions in a parallelized microfluidic device. *Lab Chip* **12**, 802–807 (2012).
54. Mastrobattista, E. *et al.* High-Throughput Screening of Enzyme Libraries : In Vitro Evolution of a b -Galactosidase by Fluorescence-Activated Sorting of Double Emulsions. *Chem. Biol.* **12**, 1291–1300 (2005).
55. Adams, L. L. A., Kodger, T. E., Kim, S., Shum, C. & Weitz, D. A. Single step emulsification for the generation of multi-component double emulsions. *Soft Matter* 10719–10724 (2012). doi:10.1039/c2sm25953b
56. Fischlechner, M. *et al.* Evolution of enzyme catalysts caged in biomimetic gel-shell beads. *Nat. Chem.* **6**, 791–796 (2014).
57. Markel, U. *et al.* Advances in ultrahigh-throughput screening for directed enzyme evolution. *Chem. Soc. Rev.* **49**, 233–262 (2020).
58. Alallam, B., Oo, M. K., Ibrahim, W. N. & Doolaanea, A. A. 3D culture of cancer cells in alginate hydrogel beads as an effective technique for emergency cell storage and transportation in the pandemic era. *J. Cell. Mol. Med.* **26**, 235–238 (2022).
59. Yu, L. *et al.* Core-shell hydrogel beads with extracellular matrix for tumor spheroid formation. *Biomicrofluidics* **9**, 24118 (2015).
60. Kleine-Brüggeney, H. *et al.* Long-Term Perfusion Culture of Monoclonal Embryonic Stem Cells in 3D Hydrogel Beads for Continuous Optical Analysis of Differentiation. *Small* **15**, 1–11 (2019).
61. Nossal, G. J. & Lederberg, J. Antibody production by single cells. *Nature* **181**, 1419–1420 (1958).
62. Talmage, D. W. Diversity of antibodies. *J. Cell. Physiol. Suppl.* **50**, 229–246 (1957).
63. Burnet, F. M. A Modification of Jerne's Theory of Antibody Production using the Concept

of Clonal Selection. *CA. Cancer J. Clin.* **26**, 119–121 (1976).

64. Rotman, B. Measurement of activity of single molecules of beta-D-galactosidase. *Proc. Natl. Acad. Sci. United States* **47**, 1981–1991 (1961).
65. Duffy, D. C., McDonald, J. C., Schueller, O. J. A. & Whitesides, G. M. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). *Anal. Chem.* **70**, 4974–4984 (1998).
66. McDonald, J. C. *et al.* Fabrication of microfluidic systems in poly(dimethylsiloxane). *Electrophoresis* **21**, 27–40 (2000).
67. Stankova, N. E. *et al.* Optical properties of polydimethylsiloxane (PDMS) during nanosecond laser processing. *Appl. Surf. Sci.* **374**, 96–103 (2016).
68. Kim, M. C., Lam, R. H. W., Thorsen, T. & Asada, H. H. Mathematical analysis of oxygen transfer through polydimethylsiloxane membrane between double layers of cell culture channel and gas chamber in microfluidic oxygenator. *Microfluid. Nanofluidics* **15**, 285–296 (2013).
69. Thorsen, T., Maerkl, S. J. & Quake, S. R. Microfluidic large-scale integration. *Science (80-.).* **298**, 580–584 (2002).
70. Ahn, K., Agresti, J., Chong, H., Marquez, M. & Weitz, D. A. Electrocoalescence of drops synchronized by size-dependent flow in microfluidic channels. *Appl. Phys. Lett.* **88**, 5–8 (2006).
71. Ahn, K. *et al.* Dielectrophoretic manipulation of drops for high-speed microfluidic sorting devices. *Appl. Phys. Lett.* (2006). doi:10.1063/1.2164911
72. Link, D. R. *et al.* Electric Control of Droplets in Microfluidic Devices. *Angew. Chemie* **118**, 2618–2622 (2006).
73. Baret, J. C. *et al.* Fluorescence-activated droplet sorting (FADS): Efficient microfluidic cell sorting based on enzymatic activity. *Lab Chip* **9**, 1850–1858 (2009).
74. Agresti, J. J. *et al.* Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc. Natl. Acad. Sci. U. S. A.* (2010). doi:10.1073/pnas.0910781107
75. Colin, P. Y. *et al.* Ultrahigh-throughput discovery of promiscuous enzymes by picodroplet functional metagenomics. *Nat. Commun.* **6**, (2015).
76. Fischlechner, M. *et al.* Evolution of enzyme catalysts caged in biomimetic gel-shell beads. *Nat. Chem.* **6**, 791–796 (2014).
77. Sjostrom, S. L. *et al.* High-throughput screening for industrial enzyme production hosts by droplet microfluidics. *Lab Chip* (2014). doi:10.1039/c3lc51202a
78. Gielen, F. *et al.* Ultrahigh-throughput – directed enzyme evolution by absorbance-activated droplet sorting (AADS). *Proc. Natl. Acad. Sci.* **113**, E7383–E7389 (2016).
79. Beneyton, T., Coldren, F., Baret, J. C., Griffiths, A. D. & Taly, V. CotA laccase: High-throughput manipulation and analysis of recombinant enzyme libraries expressed in E. coli using droplet-based microfluidics. *Analyst* **139**, 3314–3323 (2014).

80. Sun, A. C. *et al.* A droplet microfluidic platform for high-throughput photochemical reaction discovery. *Nat. Commun.* **11**, 1–6 (2020).
81. Weinreich, D. M., Delaney, N. F., DePristo, M. A. & Hartl, D. L. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science (80-.).* **312**, 111–114 (2006).
82. Debon, A. *et al.* Ultrahigh-throughput screening enables efficient single-round oxidase remodelling. *Nat. Catal.* **2**, 740–747 (2019).
83. Miller, E., Rotea, M. & Rothstein, J. P. Microfluidic device incorporating closed loop feedback control for uniform and tunable production of micro-droplets. *Lab Chip* (2010). doi:10.1039/b925497h
84. Glawdel, T. & Ren, C. L. Global network design for robust operation of microfluidic droplet generators with pressure-driven flow. *Microfluid. Nanofluidics* (2012). doi:10.1007/s10404-012-0982-y
85. Xu, J. H., Li, S. W., Tan, J., Wang, Y. J. & Luo, G. S. Preparation of highly monodisperse droplet in a T-junction microfluidic device. *AIChE J.* **52**, 3005–3010 (2006).
86. Umbanhowar, P. B., Prasad, V. & Weitz, D. A. Monodisperse emulsion generation via drop break off in a coflowing stream. *Langmuir* **16**, 347–351 (2000).
87. Marín, A. G., Campo-Cortés, F. & Gordillo, J. M. Generation of micron-sized drops and bubbles through viscous coflows. *Colloids Surfaces A Physicochem. Eng. Asp.* **344**, 2–7 (2009).
88. Kang, D., Ali, M. M., Zhang, K., Pone, E. J. & Zhao, W. Trends in Analytical Chemistry Droplet microfluidics for single-molecule and single-cell analysis in cancer research , diagnosis and therapy. *TRENDS Anal. Chem.* (2014). doi:10.1016/j.trac.2014.03.006
89. Gañán-Calvo, A. M. Generation of Steady Liquid Microthreads and Micron-Sized Monodisperse Sprays in Gas Streams. *Phys. Rev. Lett.* (1998). doi:10.1103/PhysRevLett.80.285
90. Anna, S. L., Bontoux, N. & Stone, H. A. Formation of dispersions using 'flow focusing' in microchannels. *Appl. Phys. Lett.* **82**, 364–366 (2003).
91. Yobas, L., Martens, S., Ong, W. L. & Ranganathan, N. High-performance flow-focusing geometry for spontaneous generation of monodispersed droplets. *Lab Chip* (2006). doi:10.1039/b602240e
92. Takeuchi, S., Garstecki, P., Weibel, D. B. & Whitesides, G. M. An axisymmetric flow-focusing microfluidic device. *Adv. Mater.* (2005). doi:10.1002/adma.200401738
93. Kalantarifard, A., Alizadeh-haghighi, E. & Elbuken, C. A microfluidic droplet system for ultra-monodisperse droplet generation : A universal approach. *Chem. Eng. Sci.* **261**, 117947 (2022).
94. Malloggi, F. *et al.* Monodisperse colloids synthesized with nanofluidic technology.

Langmuir (2010). doi:10.1021/la9028047

95. Xu, X. *et al.* High aspect ratio induced spontaneous generation of monodisperse picolitre droplets for digital PCR. *Biomicrofluidics* **12**, 14103 (2018).
96. Seemann, R., Brinkmann, M., Pfohl, T. & Herminghaus, S. Droplet based microfluidics. *Reports Prog. Phys.* (2012). doi:10.1088/0034-4885/75/1/016601
97. Garstecki, P. *et al.* Formation of monodisperse bubbles in a microfluidic flow-focusing device. *Appl. Phys. Lett.* (2004). doi:10.1063/1.1796526
98. Cramer, C., Fischer, P. & Windhab, E. J. Drop formation in a co-flowing ambient fluid. *Chem. Eng. Sci.* (2004). doi:10.1016/j.ces.2004.04.006
99. Lee, W., Walker, L. M. & Anna, S. L. Role of geometry and fluid properties in droplet and thread formation processes in planar flow focusing. *Phys. Fluids* **21**, 32103 (2009).
100. Edd, J. F. *et al.* Controlled encapsulation of single-cells into monodisperse picolitre drops. *Lab Chip* **8**, 1262–1264 (2008).
101. Frenz, L., Blank, K., Brouzes, E. & Griffiths, A. D. Reliable microfluidic on-chip incubation of droplets in delay-lines. *Lab Chip* **9**, 1344–1348 (2009).
102. Al., S. et. Reactions in Droplets in Microfluidic Channels Angewandte. *Angew. Chemie* 7336–7356 (2006). doi:10.1002/anie.200601554
103. Romero, P. A., Tran, T. M. & Abate, A. R. Dissecting enzyme function with microfluidic-based deep mutational scanning. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7159–7164 (2015).
104. Neun, S., Kaminski, T. S. & Hollfelder, F. *Single-cell activity screening in microfluidic droplets*. *Methods in Enzymology* **628**, (Elsevier Inc., 2019).
105. Niu, X., Gulati, S., Edel, J. B. & Demello, A. J. Pillar-induced droplet merging in microfluidic circuits. *Lab Chip* (2008). doi:10.1039/b813325e
106. Mazutis, L. *et al.* Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* (2013). doi:10.1038/nprot.2013.046
107. Priest, C., Herminghaus, S. & Seemann, R. Controlled electrocoalescence in microfluidics: Targeting a single lamella. *Appl. Phys. Lett.* **89**, 134101 (2006).
108. Brouzes, E. *et al.* Droplet microfluidic technology for single-cell high-throughput screening. *Proc. Natl. Acad. Sci. U. S. A.* (2009). doi:10.1073/pnas.0903542106
109. Abate, A. R., Hung, T., Marya, P., Agresti, J. J. & Weitz, D. A. High-throughput injection with microfluidics using picoinjectors using picoinjectors. *Proc. Natl. Acad. Sci. U. S. A.* (2010). doi:10.1073/pnas.1006888107
110. Van Loo, B. *et al.* High-Throughput, Lysis-Free Screening for Sulfatase Activity Using Escherichia coli Autodisplay in Microdroplets. *ACS Synth. Biol.* **8**, 2690–2700 (2019).
111. Hosokawa, M. *et al.* Droplet-based microfluidics for high-throughput screening of a metagenomic library for isolation of microbial enzymes. *Biosens. Bioelectron.* **67**, 379–385 (2015).

112. Maceiczyk, R. M., Hess, D., Chiu, F. W. Y., Stavrakis, S. & deMello, A. J. Differential detection photothermal spectroscopy: towards ultra-fast and sensitive label-free detection in picoliter & femtoliter droplets. *Lab Chip* (2017). doi:10.1039/C7LC00946A
113. Medcalf, E. J., Gantz, M., Kaminski, T. S. & Hollfelder, F. Ultra-High-Throughput Absorbance-Activated Droplet Sorting for Enzyme Screening at Kilohertz Frequencies. *Anal. Chem.* (2023). doi:10.1021/acs.analchem.2c04144
114. Chabert, M. & Viovy, J. L. Microfluidic high-throughput encapsulation and hydrodynamic self-sorting of single cells. *Proc. Natl. Acad. Sci. U. S. A.* (2008). doi:10.1073/pnas.0708321105
115. Joensson, H. N., Uhlén, M. & Svahn, H. A. Droplet size based separation by deterministic lateral displacement- separating droplets by cell-induced shrinking. *Lab Chip* (2011). doi:10.1039/c0lc00688b
116. Surenjav, E., Priest, C., Herminghaus, S. & Seemann, R. Manipulation of gel emulsions by variable microchannel geometry. *Lab Chip* (2009). doi:10.1039/b808160c
117. Tan, S. H., Nguyen, N. T., Yobas, L. & Kang, T. G. Formation and manipulation of ferrofluid droplets at a microfluidic T-junction. *J. Micromechanics Microengineering* (2010). doi:10.1088/0960-1317/20/4/045004
118. Wu, L., Chen, P., Dong, Y., Feng, X. & Liu, B. F. Encapsulation of single cells on a microfluidic device integrating droplet generation with fluorescence-activated droplet sorting. *Biomed. Microdevices* (2013). doi:10.1007/s10544-013-9754-z
119. Abate, A. R. & Weitz, D. A. Single-layer membrane valves for elastomeric microfluidic devices. *Appl. Phys. Lett.* (2008). doi:10.1063/1.2945797
120. Yap, Y. F. *et al.* Thermally mediated control of liquid microdroplets at a bifurcation. *J. Phys. D. Appl. Phys.* (2009). doi:10.1088/0022-3727/42/6/065503
121. Franke, T., Abate, A. R., Weitz, D. A. & Wixforth, A. Surface acoustic wave (SAW) directed droplet flow in microfluidics for PDMS devices. *Lab Chip* (2009). doi:10.1039/b906819h
122. Bi, F. Z. & Barber, B. P. Frank Z. Bi and Bradley P. Barber. *IEEE Microw. Mag.* 65–80 (2008).
123. Schmid, L., Weitz, D. A. & Franke, T. Sorting drops and cells with acoustics: Acoustic microfluidic fluorescence-activated cell sorter. *Lab Chip* (2014). doi:10.1039/c4lc00588k
124. Siegel, A. C., Bruzewicz, D. A., Weibel, D. B. & Whitesides, G. M. Microsolidics: Fabrication of three-dimensional metallic microstructures in poly(dimethylsiloxane). *Adv. Mater.* (2007). doi:10.1002/adma.200601787
125. Sciambi, A. & Abate, A. R. Generating electric fields in PDMS microfluidic devices with salt water electrodes. *Lab Chip* **14**, 2605–2609 (2014).

126. Richter, E. S. *et al.* Acoustic sorting of microfluidic droplets at kHz rates using optical absorbance. *Lab Chip* (2022). doi:10.1039/d2lc00871h
127. Crawford, D. F., Smith, C. A. & Whyte, G. Image-based closed-loop feedback for highly mono-dispersed microdroplet production. *Sci. Rep.* **7**, 1–9 (2017).
128. Weber, T. *et al.* Sensors and Actuators : B . Chemical Recovery and isolation of individual microfluidic picoliter droplets by triggered deposition. *Sensors Actuators B. Chem.* **369**, 132289 (2022).
129. Cole, R. H. *et al.* Printed droplet microfluidics for on demand dispensing of picoliter droplets and cells. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 8728–8733 (2017).
130. Haidas, D. *et al.* Microfluidic Platform for Multimodal Analysis of Enzyme Secretion in Nanoliter Droplet Arrays. *Anal. Chem.* **91**, 2066–2073 (2019).
131. Perry, J. M., Soffer, G., Jain, R. & Shih, S. C. C. Expanding the limits towards 'one-pot' DNA assembly and transformation on a rapid-prototype microfluidic device. *Lab Chip* **21**, 3730–3741 (2021).
132. Shih, S. C. C. *et al.* A droplet-to-digital (D2D) microfluidic device for single cell assays. *Lab Chip* **15**, 225–236 (2015).
133. Sinha, H., Quach, A. B. V, Vo, P. Q. N. & Shih, S. C. C. An automated microfluidic gene-editing platform for deciphering cancer genes. *Lab Chip* **18**, 2300–2312 (2018).
134. Karbaschi, M., Shahi, P. & Abate, A. R. Rapid, chemical-free breaking of microfluidic emulsions with a hand-held antistatic gun. *Biomicrofluidics* (2017). doi:10.1063/1.4995479
135. Zachos, I. *et al.* Hot Flows: Evolving an Archaeal Glucose Dehydrogenase for Ultrastable Carba-NADP+Using Microfluidics at Elevated Temperatures. *ACS Catal.* (2022). doi:10.1021/acscatal.1c04320
136. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* (1970). doi:10.1038/227680a0
137. C., P. N., Felix, V., Lanette, F., Gerald, G. & Theronica, G. Applications de la Micro Extraction en Phase Solide (SPME) pour la maîtrise de la qualité et de la sécurité des vins. *Prorein Sci.* 2411–2423 (1995).
138. Vázquez-Figueroa, E., Chapparro-Riggers, J. & Bommarius, A. S. Development of a thermostable glucose dehydrogenase by a structure-guided consensus concept. *ChemBioChem* **8**, 2295–2301 (2007).
139. Li, J. *et al.* Ala258Phe substitution in Bacillus sp. YX-1 glucose dehydrogenase improves its substrate preference for xylose. *Process Biochem.* **56**, 124–131 (2017).
140. Martínez, A. T. *et al.* Oxidoreductases on their way to industrial biotransformations. *Biotechnol. Adv.* **35**, 815–831 (2017).
141. Zheng, Y., Yin, H., Yu, D., Chen, X. & Tang, X. Recent advances in biotechnological

applications of alcohol dehydrogenases. *Appl. Microbiol. Biotechnol.* 987–1001 (2017). doi:10.1007/s00253-016-8083-6

142. Zachos, I., Döring, M., Tafertshofer, G., Simon, R. C. & Sieber, V. carba-Nicotinamid-Adenin-Dinukleotid-Phosphat: Robuster Cofaktor für die Redox-Biokatalyse. *Angew. Chemie* **133**, 14822–14828 (2021).
143. Acevedo-Rocha, C. G., Hollmann, F., Sanchis, J. & Sun, Z. A Pioneering Career in Catalysis: Manfred T. Reetz. *ACS Catal.* 15123–15139 (2020). doi:10.1021/acscatal.0c04108
144. Guterl, J.-K. *et al.* Cell-free metabolic engineering: production of chemicals by minimized reaction cascades. *ChemSusChem* **5**, 2165–2172 (2012).
145. Gmelch, T. J., Sperl, J. M. & Sieber, V. Molecular Dynamics Analysis of a Rationally Designed Aldehyde Dehydrogenase Gives Insights into Improved Activity for the Non-Native Cofactor NAD+. *ACS Synth. Biol.* (2020). doi:10.1021/acssynbio.9b00527
146. Teshima, M. *et al.* Development of a Highly Selective NAD+-Dependent Glyceraldehyde Dehydrogenase and its Application in Minimal Cell-Free Enzyme Cascades. *ChemSusChem* (2024). doi:10.1002/cssc.202301132
147. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
148. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science (80-.).* (2021). doi:10.1126/science.abj8754
149. Kunka, A. *et al.* Advancing Enzyme's Stability and Catalytic Efficiency through Synergy of Force-Field Calculations, Evolutionary Analysis, and Machine Learning. *ACS Catal.* (2023). doi:10.1021/acscatal.3c02575
150. Van Der Flier, F. *et al.* What makes the effect of protein mutations difficult to predict? *bioRxiv* (2023).
151. Shembekar, N. *et al.* Single-Cell Droplet Microfluidic Screening for Antibodies Specifically Binding to Target Cells Resource Single-Cell Droplet Microfluidic Screening for Antibodies Specifically Binding to Target Cells. *CellReports* **22**, 2206–2215 (2018).
152. Mair, P., Gielen, F. & Hollfelder, F. Exploring sequence space in search of functional enzymes using microfluidic droplets. *Curr. Opin. Chem. Biol.* **37**, 137–144 (2017).
153. Kolaitis, G. *et al.* Development of a Universal NADH Detection Assay for High Throughput Enzyme Evolution Using Fluorescence Activated Droplet Sorting. *ChemRxiv* (2022).
154. Penner, M. *et al.* Sub-single-turnover quantification of enzyme catalysis at ultrahigh throughput via a versatile NAD(P)H coupled assay in microdroplets. *bioRxiv* 2023.11.22.568356 (2023). doi:10.1101/2023.11.22.568356
155. Scheler, O., Kaminski, T. S., Ruszczak, A. & Garstecki, P. Dodecylresorufin (C12R)

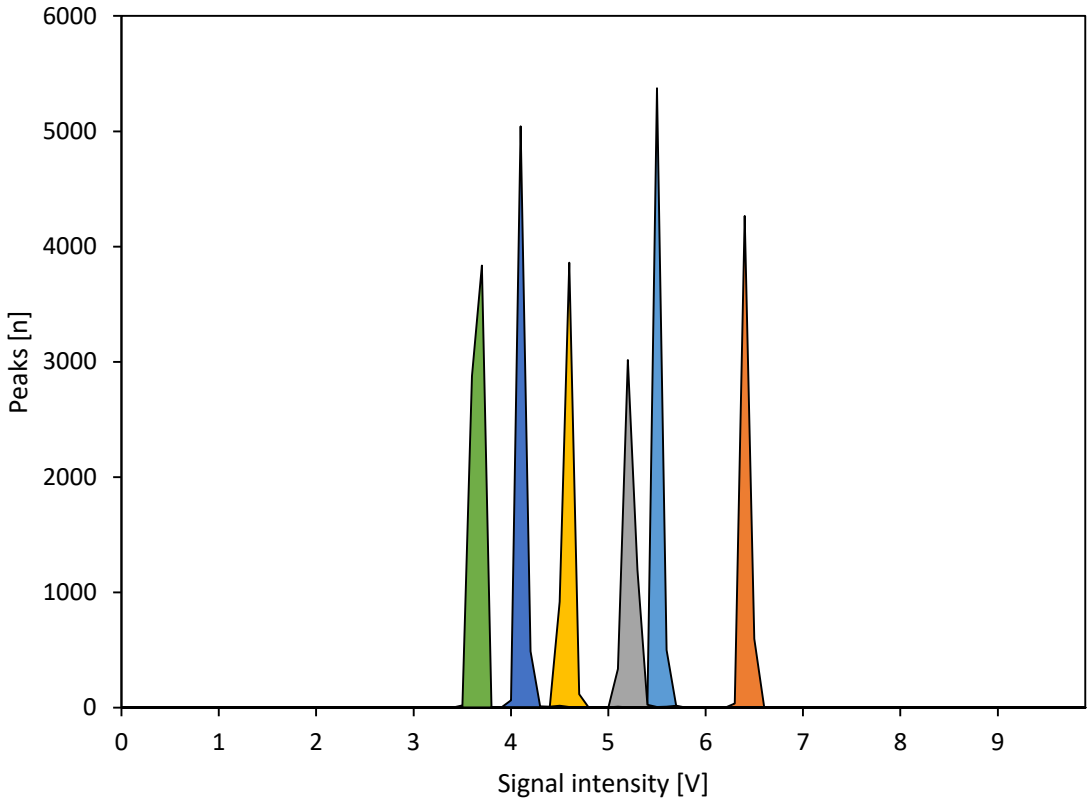
Outperforms Resorufin in Microdroplet Bacterial Assays. *ACS Appl. Mater. Interfaces* **8**, 11318–11325 (2016).

156. Eggeling, C., Volkmer, A. & Seidel, C. A. M. Molecular photobleaching kinetics of Rhodamine 6G by one- and two-photon induced confocal fluorescence microscopy. *ChemPhysChem* **6**, 791–804 (2005).
157. Morawski, P. A., Motley, S. J. & Campbell, D. J. Rapid Light-Dependent Degradation of Fluorescent Dyes in Formulated Serum-Free Media. *ImmunoHorizons* **3**, 585–592 (2019).
158. Lee, E. M., Gwon, S. Y., Son, Y. A. & Kim, S. H. Switching properties of fluorescent photochromic poly(methyl methacrylate) with spironaphthoxazine and D-π-A type pyran-based fluorescent dye. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **86**, 600–604 (2012).
159. Mary, P., Chen, A., Chen, I., Abate, A. R. & Weitz, D. A. On-chip background noise reduction for cell-based assays in droplets. *Lab Chip* **11**, 2066–2070 (2011).
160. López-López, O., Knapik, K., Cerdán, M. E. & González-Siso, M. I. Metagenomics of an alkaline hot spring in Galicia (Spain): Microbial diversity analysis and screening for novel lipolytic enzymes. *Front. Microbiol.* **6**, 1–18 (2015).
161. Miguel-Ruano, V. *et al.* Biochemical and Structural Characterization of a novel thermophilic esterase EstD11 provide catalytic insights for the HSL family. *Comput. Struct. Biotechnol. J.* **19**, 1214–1232 (2021).
162. Samoylova, Y. V., Sorokina, K. N., Piligaev, A. V. & Parmon, V. N. Application of Bacterial Thermostable Lipolytic Enzymes in the Modern Biotechnological Processes: A Review. *Catal. Ind.* (2019). doi:10.1134/S2070050419020107
163. Gupta, R., Govil, T., Capalash, N. & Sharma, P. Characterization of a glycoside hydrolase family 1 β-galactosidase from hot spring metagenome with transglycosylation activity. *Appl. Biochem. Biotechnol.* **168**, 1681–1693 (2012).
164. Adriaenssens, E. M. *et al.* Metagenomic analysis of the viral community in Namib Desert hypoliths. *Environ. Microbiol.* **17**, 480–495 (2015).
165. Anderson, R. E., Sogin, M. L. & Baross, J. A. Evolutionary strategies of viruses, bacteria and archaea in hydrothermal vent ecosystems revealed through metagenomics. *PLoS One* **9**, (2014).
166. DeCastro, M. E., Rodríguez-Belmonte, E. & González-Siso, M. I. Metagenomics of thermophiles with a focus on discovery of novel thermozymes. *Front. Microbiol.* **7**, 1–21 (2016).
167. Beneyton, T. *et al.* Droplet-based microfluidic high-throughput screening of heterologous enzymes secreted by the yeast Yarrowia lipolytica. *Microb. Cell Fact.* **16**, 18 (2017).

168. Khater, A., Mohammadi, M., Mohamad, A. & Nezhad, A. S. Dynamics of temperature-actuated droplets within microfluidics. *Sci. Rep.* (2019). doi:10.1038/s41598-019-40069-9
169. Zurek, P. J., Hours, R., Schell, U., Pushpanath, A. & Hollfelder, F. Growth amplification in ultrahigh-throughput microdroplet screening increases sensitivity of clonal enzyme assays and minimizes phenotypic variation. *Lab Chip* (2021). doi:10.1039/d0lc00830c
170. Yang, T., Stavrakis, S. & DeMello, A. A High-Sensitivity, Integrated Absorbance and Fluorescence Detection Scheme for Probing Picoliter-Volume Droplets in Segmented Flows. *Anal. Chem.* **89**, 12880–12887 (2017).
171. Duncombe, T. A., Ponti, A., Seebeck, F. P. & Dittrich, P. S. UV-Vis Spectra-Activated Droplet Sorting for Label-Free Chemical Identification and Collection of Droplets. *Anal. Chem.* (2021). doi:10.1021/acs.analchem.1c02822
172. Rushworth, C. M., Jones, G., Fischlechner, M., Walton, E. & Morgan, H. On-chip cavity-enhanced absorption spectroscopy using a white light-emitting diode and polymer mirrors. *Lab Chip* **15**, 711–717 (2015).
173. Babson, A. L. & Babson, S. R. Kinetic colorimetric measurement of serum lactate dehydrogenase activity. *Clin. Chem.* **19**, 766–769 (1973).
174. Nakamura, A., Urabe, I. & Okada, H. Anchimeric assistance in the intramolecular reaction of glucose-dehydrogenase-polyethylene glycol NAD conjugate. *J. Biol. Chem.* **261**, 16792–16794 (1986).
175. Sutherland, M. W. & Learmonth, B. A. The tetrazolium dyes MTS and XTT provide new quantitative assays for superoxide and superoxide dismutase. *Free Radic. Res.* **27**, 283–289 (1997).
176. Che, Y. *et al.* Water-induced disproportionation of superoxide ion in aprotic solvents. *J. Phys. Chem.* **100**, 20134–20137 (1996).
177. Ishiyama, M., Miyazono, Y., Sasamoto, K., Ohkura, Y. & Ueno, K. A highly water-soluble disulfonated tetrazolium salt as a chromogenic indicator for NADH as well as cell viability. *Talanta* **44**, 1299–1305 (1997).
178. Oritani, T., Fukuhara, N., Okajima, T., Kitamura, F. & Ohsaka, T. Electrochemical and spectroscopic studies on electron-transfer reaction between novel water-soluble tetrazolium salts and a superoxide ion. *Inorganica Chim. Acta* **357**, 436–442 (2004).
179. Ogola, H. J. O. *et al.* Molecular characterization of a novel peroxidase from the cyanobacterium Anabaena sp. strain PCC 7120. *Appl. Environ. Microbiol.* **75**, 7509–7518 (2009).
180. Wariishi, H., Valli, K. & Gold, M. H. Manganese (II) Oxidation by Manganese Peroxidase from the Basidiomycete Phanerochaete chrysosporium. *J. Biol. Chem.* **267**, 23688–23695 (1992).

181. Takagi, K. *et al.* A sensitive colorimetric assay for polyamines in erythrocytes using oat seedling polyamine oxidase. *Clin. Chim. Acta* **340**, 219–227 (2004).
182. Fomin, M. A. *et al.* Two-Acceptor Cyanine-Based Fluorescent Indicator for NAD(P)H in Tumor Cell Models. *ACS Sensors* **1**, 702–709 (2016).
183. Kračun, S. K. *et al.* A new generation of versatile chromogenic substrates for high-throughput analysis of biomass-degrading enzymes. *Biotechnol. Biofuels* (2015). doi:10.1186/s13068-015-0250-y
184. Cole, R. H., De Lange, N., Gartner, Z. J. & Abate, A. R. Compact and modular multicolour fluorescence detector for droplet microfluidics. *Lab Chip* **15**, 2754–2758 (2015).

6 Appendix



Appendix 3.1.2.1: Standard curve for measurement program. Six populations with different concentrations of tartrazine were produced [0 mM (orange), 1 mM (light blue), 2 mM (gray), 3 mM (yellow), 4 mM (dark blue) and 5 mM (green)] and measured with the self customized LabView program. The average signal intensity of each population shown was calculated and plotted in a standard curve (Figure 17)

Appendix 3.1.4.1: Description LabVIEW peak detector

For determination of the product concentration within microfluidic droplets, a self-made LabVIEW program was composed. The program receives a voltage signal from a photo detector, which converts the light intensity into a digital value, making it accessible for the LabVIEW program. The concentration of each droplet can be identified by finding local minima (peaks). The higher the product concentration of a droplet, the lower is the transmitted light and the resulting peak value. Droplets with very high product formation are especially interesting as they hold potentially improved variants. These droplets have to be separated from the majority of other droplets. This process step is called “sorting”.
The next pages explain how the program works and give a brief overview about the algorithm. In general, a LabVIEW program consist of a user interface (front panel) and a programming layer (block diagram). The block diagram defines the abilities of the program, but is later not required for program operation. Both, front panel and block diagram should be shown and explained. Since the block diagram is quite complex, it is presented in five different parts.

Part 1: Data acquisition
Part 2: Peak finding
Part 3: Front panel output: peak detector
Part 4: Front panel output: heat map
Part 5: Output of data files and calculation of parameters

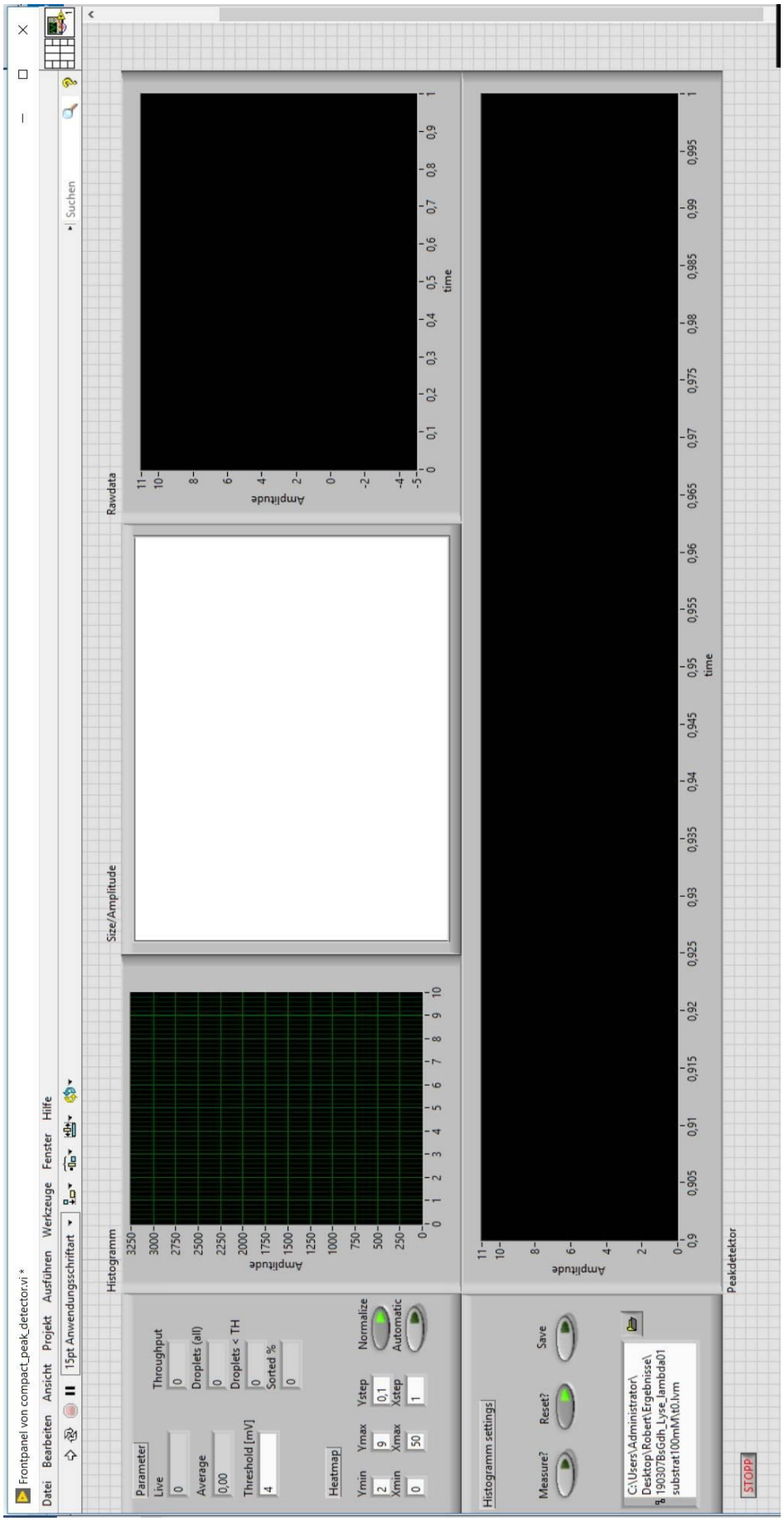


Figure 1: Frontpanel peak detector

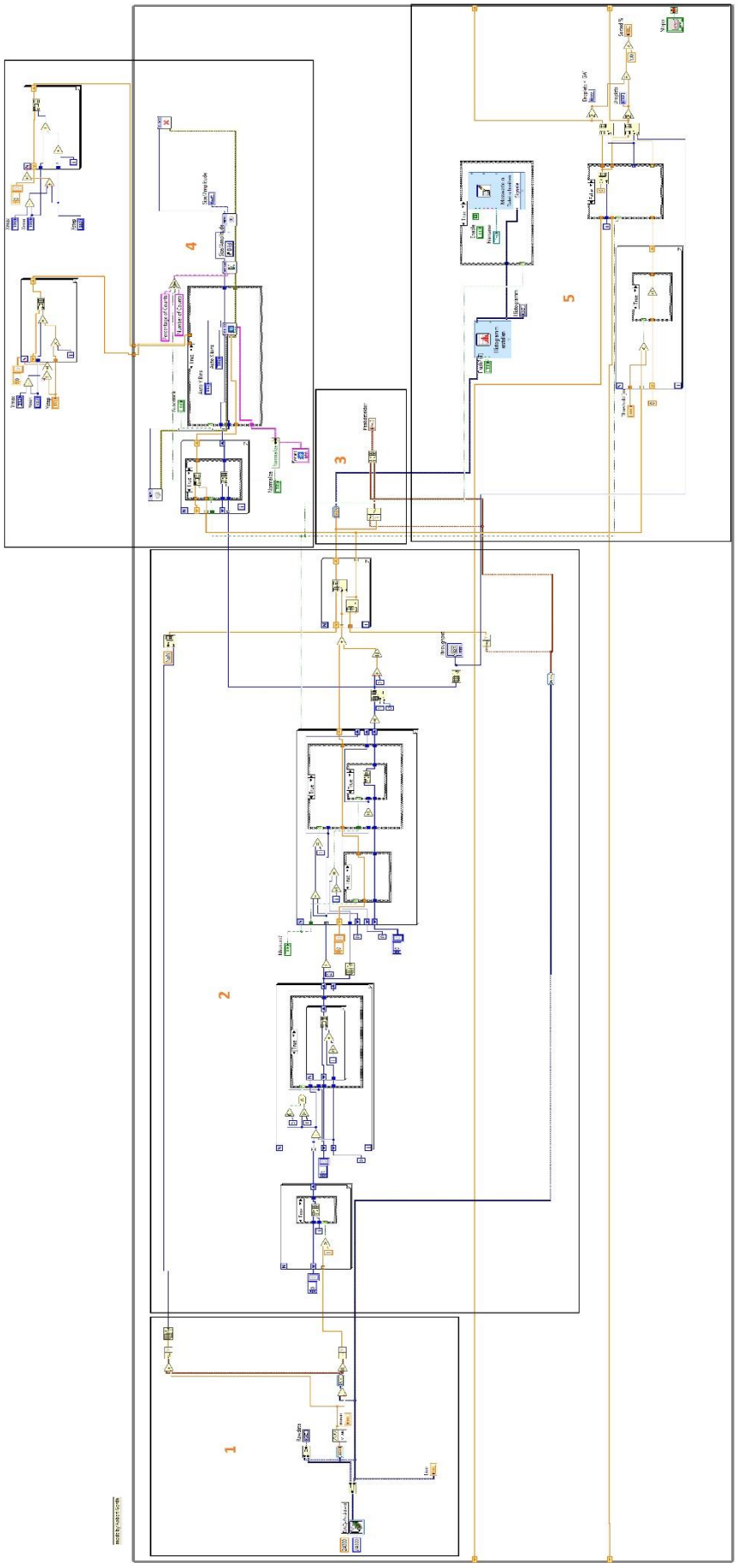


Figure. 2: Block diagramm peak detector

Part 1: Data Acquisition

Within the first part of the program, the sample rate and sample number are assigned via the DAQ-assistant (e.g. 24.000). The assistant also defines the inputs for the analog signals (A0/A3). A0 provides the voltage signal from the photo detector, while A3 provides the signal from the voltage amplifier, which is equal to the later sorting event and allows the user to identify sorted droplets. The joint signal is separated into two individual channels for analysis. A0 is used to show the live signal of the peak detector and present it on the GUI (front panel). Additionally, the dynamic data type (DDT) of the original signal is converted into an array format to allow application of the averaging function of LabVIEW. The obtained mean-value gives the average value of all 24.000 voltage values measured within one second. Most of these values represent the baseline value, which is later subtracted from all measured values and allows an automatic baseline detection (e.g. 8 V). The DDT signal is converted into a signal curve. The modulus of all values is determined to avoid negative values in the next steps. The signal curve is divided into corresponding x/y-values and further analyzed in part 2 (Figure 2). In parallel a vector with the size of the complete measured signal within one second (here 24.000 values) is created and filled with NaN (Not a number) – entries which are later filled up.

Part 2: Determination of peak values

The second part of the program determines the peak values from the signal curve. This peak value directly correlates with the concentration of assay product, but is not necessarily the highest value within a droplet signal. In fact, the concentration is rather reflected by the middle value of all signal points belonging to one droplet signal. This means, if a droplet consists of 23 measurement values, the product concentration correlates to the 12^{th} point within this signal series. The array created in part 1 consisting of normalized signal values. It is forwarded to part 2 of the program. In the first for-loop the distance to the baseline is compared for every value. Only values that exceed a distance of 1 are considered as peak values. This procedure ensures that baseline noise is not considered as peak values. If one of the measured values fulfills the distance criterium, the index of this value is noted. This procedure is performed for all measurement points within one second (24.000).

The array with all collected indizes is forwarded to the second for-loop. This for-loop has the task to fill up indizes, which are close to each other. Between each droplet signal there is a minimum of 20 values that are around the baseline. Depending on the product concentration, some values within the droplet signal can go below the threshold set in for-loop 1. In this case, the program would identify two (small) droplets, which would appear right after each other. In order to avoid double counting and wrong size distribution of droplets, the second loop identifies these small gaps. If there is less than 5 and more than 1 index gap between the

consecutive droplet signals, the true case within the loop is activated and missing indizes are filled up.
Example: Array derived from loop 1: 1, 2, 3, 5, 37, 38, 39…
Loop 2 recognizes that the gap between indize 3 and 5 is smaller than 5. Therefore, this gap is closed by adding the indize 4 in between.
New Array: 1, 2, 3, **4**, 5, 37, 38, 39….
Simultaneously, the loop structure recognizes that the gap between indize 5 and indize 37 is bigger than 5. This activates the false-case within the loop and no indizes are filled up.

The obtained array with the filled up indices is passed to loop 3. Loop 3 consists of two true/false cases. The outer true/false case is used to activate the measurement or reset the values obtained up to that point to zero (false case). A query takes place before the right true/false case. It asks whether the consecutive indices have a distance of 1 (True). If the outer case structure is true, the inner case structure always counts up by one until the outer case structure changes to false. This gives you the total number of measured values that form a peak signal. If the outer case structure goes into the false case, the index of the first measured value belonging to the peak signal is exported (orange data line). The lower blue data line that comes from the inner True/False structure contains the list of peak lengths of all peaks within one second. The first value is removed from the array because the False criterion must be met the first time the loop is run through (0-0=0 and not 1) to mark the beginning of a new droplet. The number of measured values of a droplet is then halved and rounded down (for odd lengths) in order to reach the center of the droplet. The index of the first measured value of the droplet is added to half the length of the droplet and the actual mean index of each droplet is thus calculated. The array with the length of all droplet signals is forwarded to part 4 (top) and the size of the array - i.e. the number of entries - is forwarded downwards, which corresponds to the throughput of droplets per second.
In the fourth loop, the array of the size of all measured values (NaN array) from part 1 is now filled with the previously found peak values. For this purpose, the previously calculated droplet peaks are used to “index the array”. Only the calculated indices are selected in the original raw data signal and using the “Replace partial array” function, the NaN array is overwritten at the corresponding indices and the calculated peak values are inserted into the raw signal. The array with the values of the peaks is passed to part 3.

Part 3: Front panel output: Peak detector

The NaN array obtained in part 2, which was filled with the peak values, is now converted into a signal curve (brown line). However, only the data at the location of the peaks is forwarded - all entries with NaN remain empty. The signal curve obtained in this way is coupled with the

complete signal curve of all raw data and displayed in the peak detector diagram. The calculated peaks in the graph are displayed in different colors via the front panel, so that the calculated values are clearly distinguishable from the raw signal. The array is also converted into dynamic data (blue line) and used to create a histogram, which is further explained in part 5.

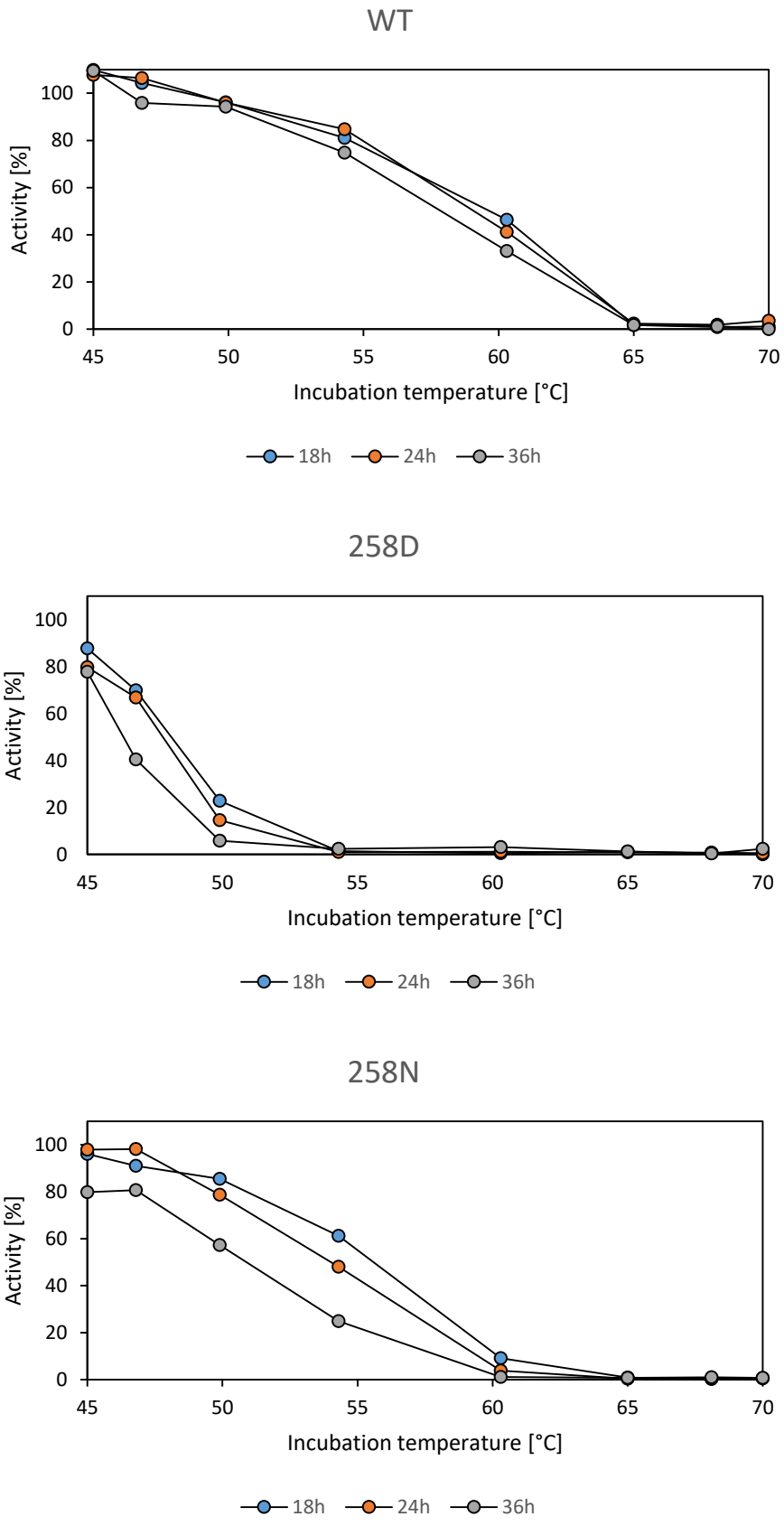
Part 4: Front panel output: Heat map

In addition to the optical assignment of the peaks to the raw signal, it is also interesting to see how the distribution of the signal intensity corresponds the droplet size. In order to make these relationships visible, a 2D histogram application from the “Advanced Plotting toolkit” toolbox was used and the data of peak size (X values) and signal intensity (Y values) were forwarded to the toolbox as a 1D array. Outside the while loop you can see two more for-loops. These define both the maximum and minimum values of the X- and Y-axes, as well as their step sizes. The obtained vector determines the division of the X- and Y-axes in the front panel.

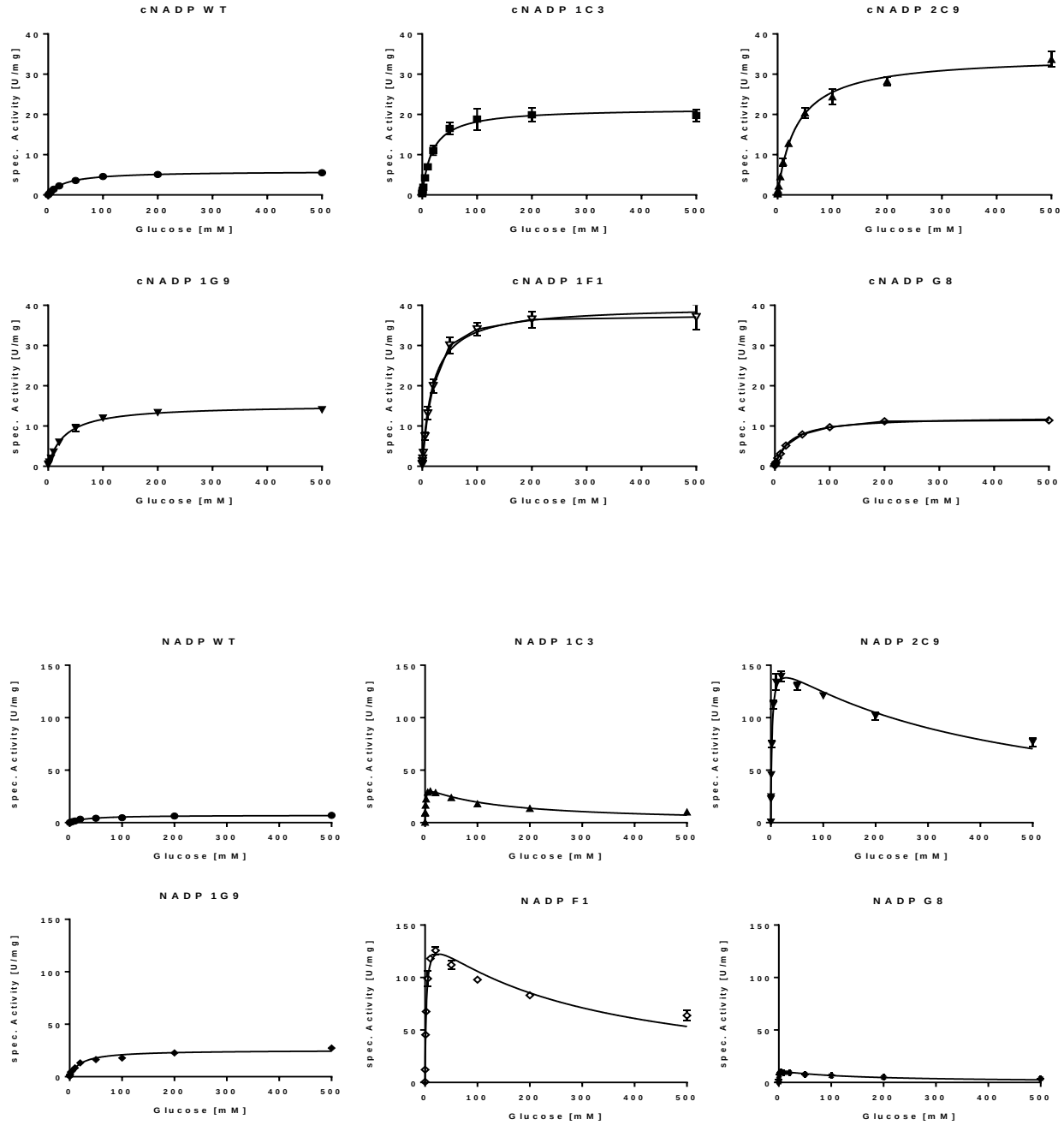
Part 5: Output of data files and calculation of parameters

The signal intensities obtained for the calculated peak positions from part 2 are transferred from the array form into the dynamic data type to meet the requirements for the predefined “Create histogram” VI. Here, the individual values are arranged for a 1D histogram and displayed graphically on the front panel. In addition, the data obtained in the upper case structure is written to an .lvm file using the predefined “Write Measured Values to File” VI. The resulting file can later be opened with an editor and the results can be used for further analysis. The structure is activated using the “Measure?” button located at the front panel. The “Reset?” button allows you to reset the collected data. The storage location of the .lvm file can be determined using the “Filename” field anchored in the front panel. The for-loop in the lower part only counts droplets that are below a certain limit set by the user. The limit can be set via the front panel.

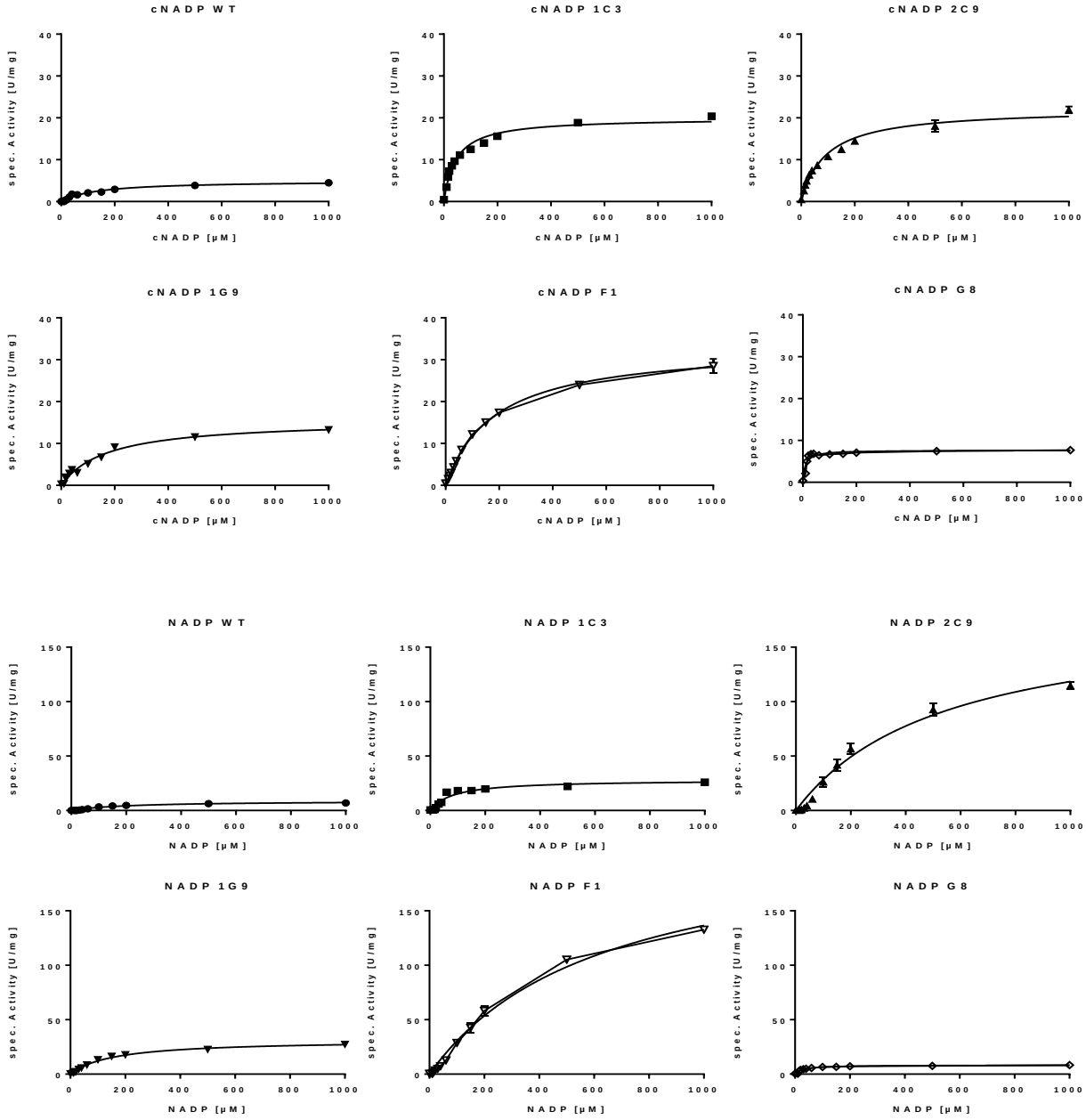
The case structure next to the for-loop is used to reset the added values and start new measurements. To the right of the case structure, the total number of droplets, the number of droplets that are smaller than the given limit value and the quotient of both values (sorted droplets) are then calculated and displayed on the front panel.



Appendix 3.1.6: Residual activity of BsGDH template variant Q252L, E170K (top), template variant with additional mutation 258D (middle) and template variant with additional mutation 258N (bottom) after 18 hours (blue), 24 hours (orange) and 36 hours (grey)



Appendix 3.3.1: Substrate kinetics for *Ss*GDH hits found in overlap extension screening round. Both measurements were performed with 1 mM of cNADP/NADP in 50 mM Tris HCl pH 8 buffer at 65 °C. Absorption of cNADP (360 nm) and NADP (340 nm) were monitored over time and slopes were taken to calculate specific activity.



Appendix 3.3.2: Cofactor kinetics for *Ss*GDH hits found in overlap extension screening round. Both measurements were performed with 50 mM glucose in 50 mM Tris HCl pH 8 buffer at 65 °C. Absorption of cNADP (360 nm) and NADP (340 nm) were monitored over time and slopes were taken to calculate specific activity.