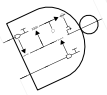
A CRISPRi NAND Gate

Benjamin Aleritsch



Big things have small beginnings.

Physik der synthetischen biologischen Systeme, Fakultät für Physik
Technische Universität München

Construction, optimization and characterization of a functionally complete logic gate based on CRISPRi in Escherichia coli

Benjamin Aleritsch

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr.rer.nat)

genehmigten Dissertation.

Vorsitzender:

Prof. Dr. Martin Zacharias

Prüfende der Dissertation:

1. Prof. Dr. Friedrich C. Simmel
2. Prof. Dr. Ulrich Gerland

Die Dissertation wurde am 07.08.2019 bei der Technischen Universität München eingereicht und durch die Fakultät der Physik am 14.01.2020 angenommen.

Contents

Abstract

In contrast to the top-down approach i. e. the attempt of understanding the functional aspects of (genetic) parts known from living organisms in synthetic biology, the bottom-up approach aims for creating man-made substitutes or alternatives for those parts based on already known principles for which one of the the long-term goals is to e. g. create artificial life. In particular, this area of research engages in generating organisational structures, both in terms of information and topology. Previous achievements are e. g. artificial pores[1] made by using the DNA origami technique or a functionally complete NOR gate[2] based on clustered regularly interspaced short palindromic repeats interference (CRISPRi). The former could be integrated into the bilipid-membrane wall of a liposome and used to organize the transport of agents into and out of that structure whereas the latter could be used to build a highly organized genetic network inside that liposome in opposition to the apparent chaos in its natural counterpart, e. g. the prokaryotic cell. This network would be at least theoretically capable of similar computation to that in a primitive electronic computer. Combining the two man-made parts mentioned beforehand, this alone already is a highly promising origin for creating artificial, programmable life before long.

This thesis describes the construction, optimization and characterization of a functionally complete logic NAND gate based on deoxyribonucleic acid (DNA) and CRISPRi which completes the category of functionally complete logic gates in that manner. The NAND gate is fully compatible with the NOR gate mentioned above and has, identical to the NOR gate, the potential to be used to build programmable computational networks in vivo in the future. The NAND gate alone can be used to build all known computational networks. In combination with the NOR or other compatible logic gates, it has the capability of decreasing circuits in complexity which benefits in reducing resource employment and achieving faster and less error-prone computation in organisms.

Introduction

It all began with Oswald Avery's discovery in 1944, that genes are made of DNA and the subsequent revelation of its structure by Watson and Crick in 1953, the famous double helix. Research in the following years yielded even more insight into the machinery of what was later called the "central dogma of molecular biology". This graceful designation incorporates the most important processes that take place between DNA, RNA and proteins in terms of the transformation of structural information or in more detail, how combinations of four distinct nucleobases i.e. DNA get transcribed into four very similar nitrogenous bases i.e. RNA and from there translated into a folded chain of amino acids i.e. proteins.

Many of those proteins have remarkable properties and a lot could already be harnessed by scientists. After the discovery of one such protein exhibiting fluorescent behavior by emitting green light after corresponding excitation in the jellyfish Aequorea victoria by Dr. Osamu Shimomura in 1962, it was possible for the first time to access an indicator of genetic behaviour i.e. by the reporter protein GFP. This enabled scientists to measure a quantitative signal from a realm that was inaccessible before. The discovery of this miniature light-beacon that could be used not only to quantitatively evaluate processes, but also give insight into the structure of e.g. a living cell and its components, bridged the gap between micro- and macrocosm.

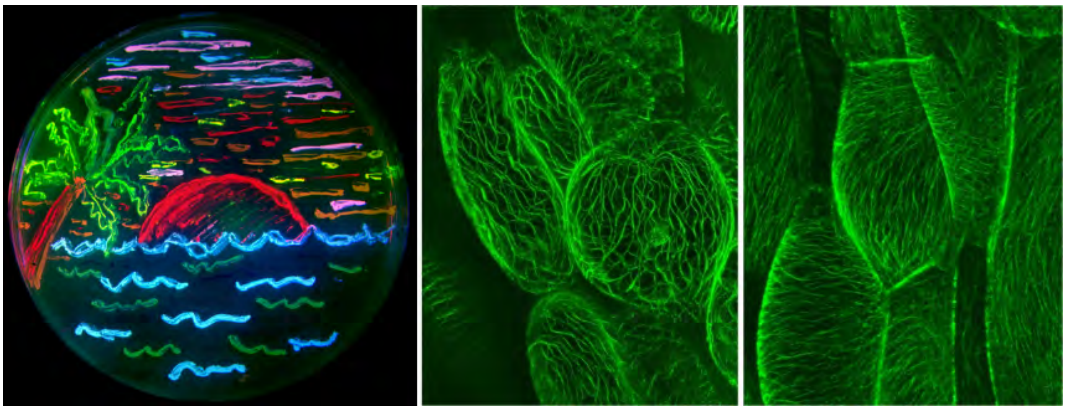


Fig. 1: Left: The famous San Diego beach scene illustrated using accordingly distributed bacteria expressing various fluorescent proteins in a petri-dish. Artwork by Nathan Shaner, photography by Paul Steinbach 2006, created in the lab of Roger Tsien, who was awarded the Nobel Prize in Chemistry for his discovery and development of the green fluorescent protein in 2008. Center and Right: Microtubules in plant cells made visible using GFP-tagging[3].

Of course, evaluating natural genetic processes was only the first step. Unlocking the entire functional capacity of this newly accessed realm shifted into focus. Findings about the application of restriction enzymes in 1970, a molecular tool to cut DNA and have it reassembled in a different combination later came in the right moment, followed by the first machines that employ the synthesis of DNA base by base in 1981.

The field continued to mature further and further as more and more inventions were made like the polymerase chain reaction by Kary Mullis in 1983, a protocol to easily automatize the amplification i.e. duplication of DNA. Around the turn of the millennium the capability in the field of molecular biology and genetics even progressed as far as concluding to decipher the entire genetic code of the human chromosome i.e. the sequencing of DNA from multiple human beings to find a composition most commonly shared.

Soon it became clear that genetic elements could not only be cut, copied and pasted as in 2003, when the gene from which GFP is created was inserted into the first genetically modified pet and a fluorescent fish (TK-1 GloFish) went on sale in Taiwan. Not only naturally occurring genes, but also entirely synthetic

components could be introduced into the genetic material of lifeforms which inaugurated the formation of the field synthetic biology.

Fig. 2: Danio rerio or Brachydanio rerio, i.e. zebrafish. The company GloFish® extended the genome of these fish by a gene from which the corresponding fluorescent protein is created.

With support of the growing capabilities of the internet, the composition and functional description of genetic components could be shared as e.g. the MIT established a Registry of Standard Biological Parts in 2003, followed by international events in this particular part of research as the International Genetically Engineered Machine (iGEM) competition.

With an exploding interest of almost the entire planets industry in the field, companies operating in molecular biology began expanding into the private sector. Concerns as Knome and 23andMe were founded around 2007, providing DNA sequencing, diagnosis as well as statistical evaluation and interpretation at low cost.

When in 2008 the TALEN technique for genome editing was invented, it paved the way for even more straight forward genetic processing when in 2013 a defense strategy of bacteria could be utilized for editing snippets of genetic code and the functional principle of CRISPR was first described. But even editing DNA almost without boundaries was not the end of the line. Based on CRISPR, genetic devices could be realized which alter the natural 50 % probability that a specific allele will be present in the offspring, completely overriding Mendelian inheritance. Currently it is evaluated by the "Target Malaria" research consortium which is funded by the Bill & Melinda Gates Foundation whether to employ this "gene drive" in order to eliminate certain mosquito

species in Africa in order to fight the spread of malaria by releasing genetically modified mosquitoes into the wild.

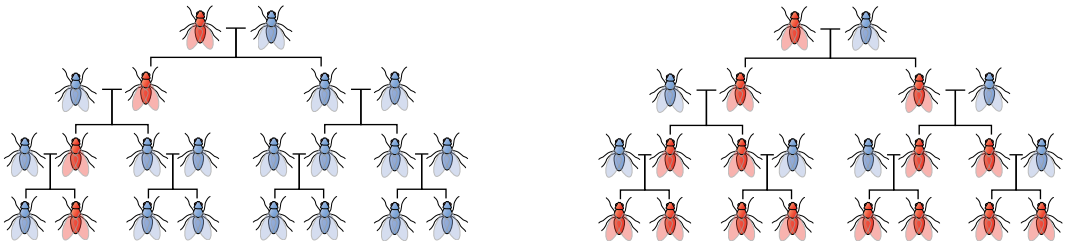


Fig. 3: Impact of the gene drive on Mendelian inheritance visualized with representative schematic fruit fly body color. Left: The natural genetic inheritance mechanism Mendelian inheritance in which recessive genes(red) are only expressed if no dominant genes(blue) are present, resulting in a population dominated by blue specimen. Right: Population evolution with the recessive gene(red) being empowered by the gene drive leads to a population dominated by red specimen. Original image taken from the Wikipedia article "Gene drive".

As time progresses the mystery of live keeps getting unraveled and its features are getting used in creative ways more often and often. Still the complexity of genetic networks is a challenge to fully grasp and an approach to recreate those networks in a familiar way was made. Synthetic genetic components could be created, not only to expand the originally four bases to up to eight by now, but also genetic components that mimic Boolean logic, analogue to how electronic circuits do in computers. Using those synthetic genetic Logic Gates in livings cells could unlock the realization of at least primitive programs, inputs and outputs. However, still not all of those Logic Gates are available in the context of synthetic biology or those available embody considerable drawbacks as a lack of the ability to connect, scale, employ or simply in terms of reliability. One construct that avoids those fallbacks is the NOR-Gate created by the group around C.Voigt. It employs two distinct strands of RNA that guide a protein to a certain predefined location of DNA, where it binds and thereby prevents transcription at that site.

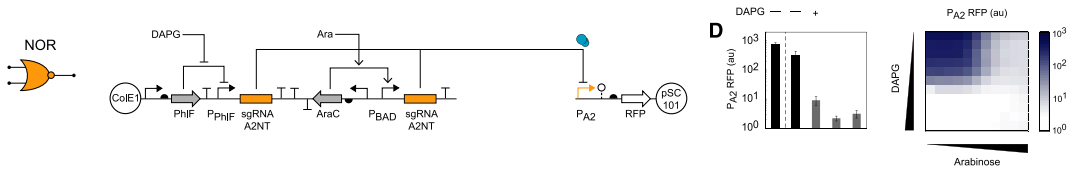


Fig. 4: Schematic and results of the NOR-Gate, published by the Voigt-Group[2]. From left to right: Electronic pendant symbol, SBOL Visual functional schematic and output characteristics as column- and density-matrix plot.

The fields of microbiology, genetics and synthetic biology keep expanding and become more important every day. Computing in the biological realm might soon prove itself a valuable and crucial tool when it comes to diagnosis, readout or predefined conditional behavior i.e. programming of biological systems. This work introduces a genetic circuit that in comparison to most other already published systems truly fulfills the conditions of a mathematically defined Logic Gate that are required for industrial employment. Its core components render it almost functionally omnipotent which make it a invaluable counterpart to the already existing also functionally complete Logic Gate.

Theory

1 Absorbance as Density Indicator for Bacteria and nucleic Acids

Absorbance is defined as the logarithm of the ratio between incident and transmitted radiant power through a sample[4] and obeys the relation of

$$A_{(\lambda)} = \int_0^l \epsilon_{(\lambda)} c_{(z)} dz \tag{1}$$

where A is the total absorbance of the sample that has been propagated by the radiant power, λ the wavelength of the radiant power, l being the total path length the radiant power has traversed in the sample, ϵ the molar attenuation coefficient of the sample and $c_{(z)}$ the molar concentration at position z along l. If $c_{(z)}$ is constant along the path, the expression can be simplified to

$$A_{(\lambda)} = \epsilon_{(\lambda)} cl \tag{2}$$

Because of its arithmetic structure, the expression of absorbance has the property of linear mapping. Thus it can be rewritten to

$$A_{(\lambda)} = \sum_{i=1}^{n} A_{i(\lambda)} = l \sum_{i=1}^{n} \epsilon_{i(\lambda)} c_i \tag{3}$$

where n is the number of substances in a mixed sample and i being the index of such. Solving for one specific k out of the accumulated absorbance in a mixed sample yields

$$A_{k(\lambda)} = A_{(\lambda)} - \sum_{i=1}^{j} A_{i(\lambda)} - \sum_{i=l}^{n} A_{i(\lambda)} \tag{4}$$

Using this expression, if the total absorbance of a mixed sample as well as the accumulated absorbance of all but that of k in the sample are known, the absorbance of k can be calculated by subtraction and also the concentration if the optical path length as well as the molar attenuation coefficient of k are known. This property enables measuring the absorbance and by that the concentration or molar attenuation coefficient of chemical agents in solution where those properties are known for the solution but not yet for the agent.

One application of this property is the measurement of the concentration of DNA or ribonucleic acid (RNA) in a solution, mostly a buffer i.e. water with the addition of specific salts. For further processing, a concentration of RNA or DNA sufficiently high enough is often required in microbiology. The concentration can be obtained by subtracting the absorbance background of the buffer alone from a sample of buffer containing the RNA or DNA of interest after measuring the absorbance by means of e. g. spectrophotometry.

Another application is the measurement of absorbance at a wavelength of 600nm as an indicator of the population density of bacteria in a medium that supports the growth of said bacteria i.e. a medium that comprises all nutrients required for bacterial cell division. Assuming that the definition of absorbance holds for the measured absorbance of bacteria as well as the absorbance of the growth medium is either unchanged or a functional relationship for its change is known throughout the entire process of the measurement, the relation is still valid.

Measuring OD600 is widespread because of the historic origin. In the middle of the 20th century, technical equipment was not as advanced as today and a choice of a particular wavelengths was not possible. Other limitations were on one hand that using light at wavelengths in the ultraviolet band has the potential of mutating DNA where on the other, most growth media absorb in the yellow band, so 600nm was most widespread in the end.

2 Indication of cellular Properties by Means of Fluorescence

Fluorescence is a type of luminescence, which describes the capability of a substance in terms of absorbing energy as electromagnetic radiation and then partial release of that energy as electromagnetic radiation at lower frequency. In the process of absorbing electromagnetic radiation i.e. excitation, the orbital configuration of the affected molecules changes to a higher energy state

$$E_0 + hf_{excitation} \rightarrow E_1 \tag{5}$$

where E_i are lower and higher state respectively, h being plank's constant and f the frequency of the incident light. In the process of releasing the energy i.e. emission,

$$E_1 \rightarrow E_0 + hf_{emission} + E_{heat} \tag{6}$$

a fraction of it dissipates to heat, leaving only a reduced portion of the energy to return to the electromagnetic manifestation. This causes a reduction in frequency which is called the stokes shift.

$$E_1 - E_0 - E_{heat} = hf_{emission} \tag{7}$$

Many different species of molecules with varying properties concerning the excitation and emission of light in terms of fluorescence exist. Among them, a subgroup are fluorescent proteins which can be produced in organisms. Fluorescent proteins are of particular interest in biophysics since they can be used to measure certain quantities as concentration or position of a particle.

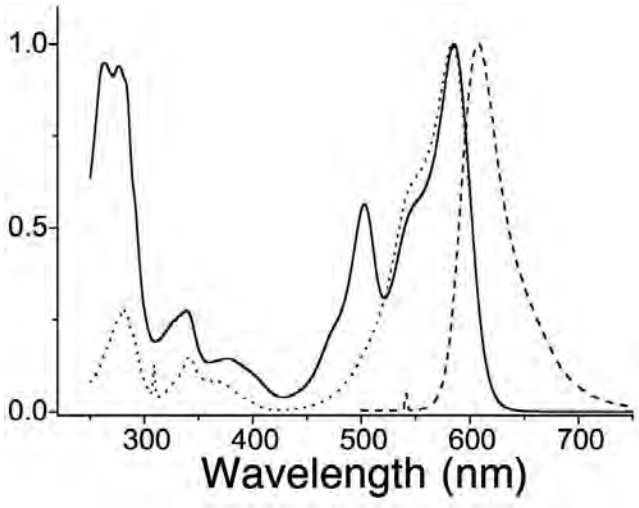


Fig. 5: Fluorescence and absorption spectra of mRFP1. The absorbance, excitation and emission spectrum are shown with a solid, dotted and dashed line respectively[5]. Maximum excitation is possible at 584nm where emission peaks at 607nm. The quantum yield of mRFP1 is 0.25. Natural time constant for maturing process is below 60 minutes where bleaching half-life is reported to be 6.2s[6].

In synthetic biology, genes from which these proteins are produced can be transferred from a natural host organism to e. g. a model organism, in order to, parallel to biophysics, also probe for certain values. Most prominent fluorescent protein supergroups are green, red, yellow and blue fluorescent protein (GFP, RFP, YFP and BFP). Fluorescence emanating from particles can be collected with photodiodes. One of the means to excite fluorescent agents is

the usage of a xenon flash lamp. After directing the light with according mirrors and reducing it to spectral bands with adequate width using optical filters it can be applied to excite fluorescent agents.

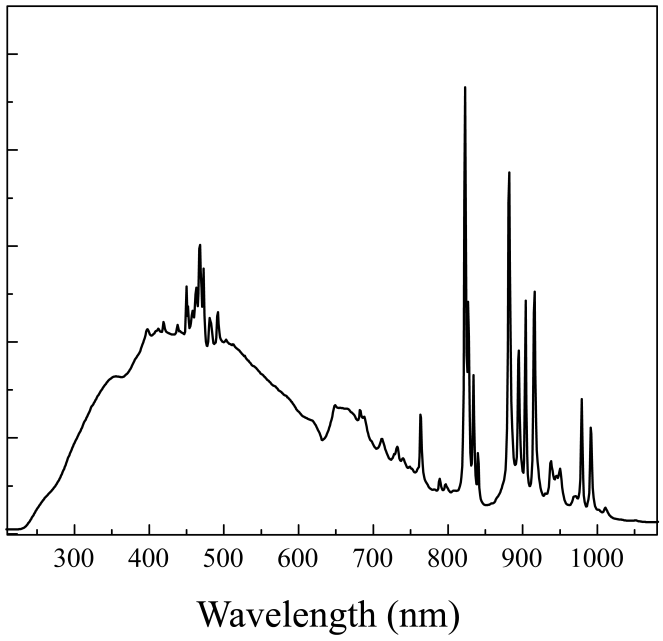


Fig. 6: Emission spectrum of a xenon arc lamp[7]. The available band is gap-less providing a free choice of wavelengths for excitation. After filtering out undesired parts of the spectrum i. e. most notably parts that overlap with the emission spectrum of the fluorophore to excite, the light of the xenon arc lamp can be used to trigger fluorescence.

A setup as such can be found in e. g. the BMG Labech FLUOstar Platereader, a device built for continuous and parallel measurement of absorbance and fluorescence while it also is able to maintain conditions as a steady temperature and oxygen supply in the growth media by shaking. These conditions allow for the growth of bacteria that are capable of the assembly of fluorescent proteins.

3 The Model Organism Escherichia coli

Model organisms are used for studying certain properties, processes or phenomena of living matter with the assumption that the discovered regularities also are inherent in other organisms. The usage of a model organism is justified by the fact that all living matter known so far can genetically be traced back to one last universal common ancestor (LUCA) and thus have common or similar patterns in terms of genotype and or phenotype and therefore e. g. metabolism and or mechanical features. The principle is further motivated by convergent evolution i. e. the fact that for certain environmental challenges, evolution often leads to the same or similar solution in terms of genotype and or phenotype for organisms with different recent phylogenetic heritage.

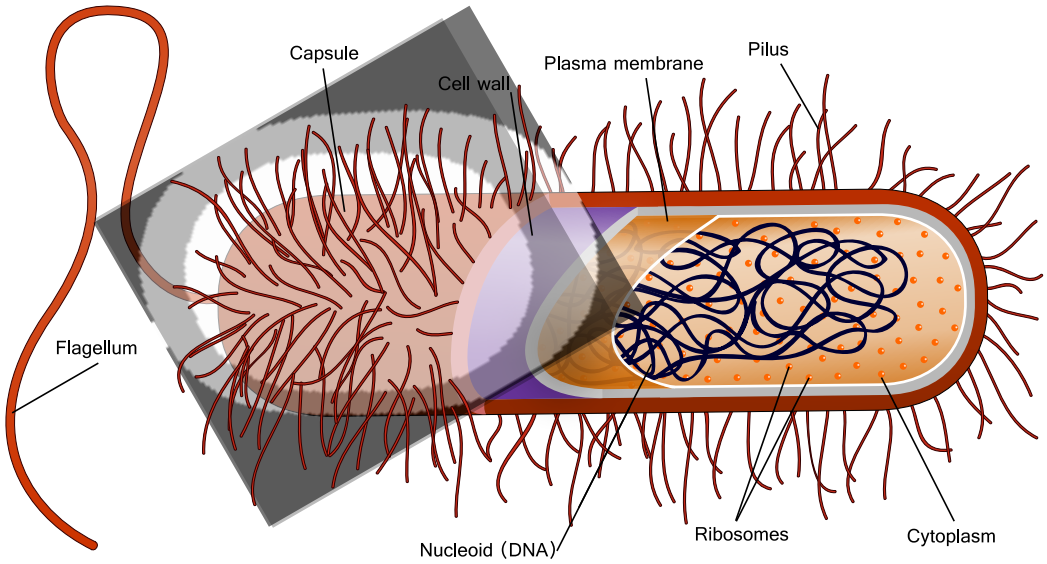


Fig. 7: *E. coli* schematic illustration[9]. The flagellum is used for propulsion and alignment where the pili can be used to copy DNA to other *E. coli* cells.

Compared to other model organisms like the house mouse, the zebrafish or the fruitfly, among others, bacterial species belong to the bottom model organisms in regards to the magnitude of genetic complexity. Bacteria are easy to cultivate and since it is not known, that bacteria feature a central nervous system, consciousness or similar, neither animal protection nor philosophical aspects apply so far. One major advantage is that a vast array of different

species within wide ranges of properties to pick from exists. *Escherichia coli* (*E. coli*) has become very popular since most of its strains are harmless for humans and thus is available in various chromosomal configurations in the wild type alone. If desired, the genetic configuration of an *E. coli* strain can artificially be altered by e. g. CRISPR by e. g. transfecting chromosome-cleaving agents like a CRISPR associated protein 9 (Cas9) and a small guiding RNA (sgRNA). A nomenclature[8] has been developed to describe what has been removed from or added to the genetic chromosomal composition compared to the wild type. Additionally, the prokaryote can be equipped with foreign DNA in loop-form called plasmids by transformation, providing even more space to implement genetic programs.

E. coli are rod-shaped organisms with a diameter of 0.25–1.0 μm and 2.0 μm length. The volume is approximately 0.6–0.7 μm^3 and dimensions vary depending on the media they have been cultivated in[10][11]. The cell wall of *E. coli* is composed of a thin peptidoglycan layer and an outer membrane. The cell wall features pores that allow for the transport of certain extracellular agents into the cell and out.

4 Environmental Conditions and Growth of E. coli

E. coli multiply by the process of binary fission i. e. the bacterial cell divides. Since this method of reproduction is asexual, if no mutations have occurred both descendants share the same genetic content. If on average more than one of these two descendants survives, the culture experiences exponential growth, thus in a contained environment i. e. no (continuous) flux of growth medium is supplied, *E. coli* cultures undergo four growth phases.

In lag phase little to none division occurs since the metabolic network of *E. coli* reacts to the environmental conditions i. e. the processable substances in the growth medium trigger the production of the corresponding RNAs and proteins that are required to break down and digest corresponding agents in the growth medium.

In logarithmic or exponential phase, *E. coli* have booted up their metabolic network sufficiently enough to initiate cell division. Also since substances in the culture medium are digested, the density of nutrients decreases whereas the density of metabolic waste-products e. g. ethanol increases. This slowly leads to a shift in the quality of the growth medium, most notably unbefitting nutrient composition and or pH. Under ideal environmental conditions, *E. coli* divide approximately every 20 minutes. In logarithmic time-resolved growth-plots, the trajectory during log-phase can be described as roughly linear. However, the slope of the trajectory must not be confused with the rate of cell duplication. It is a numeric balance of how many new cells are created over time and how many decompose.

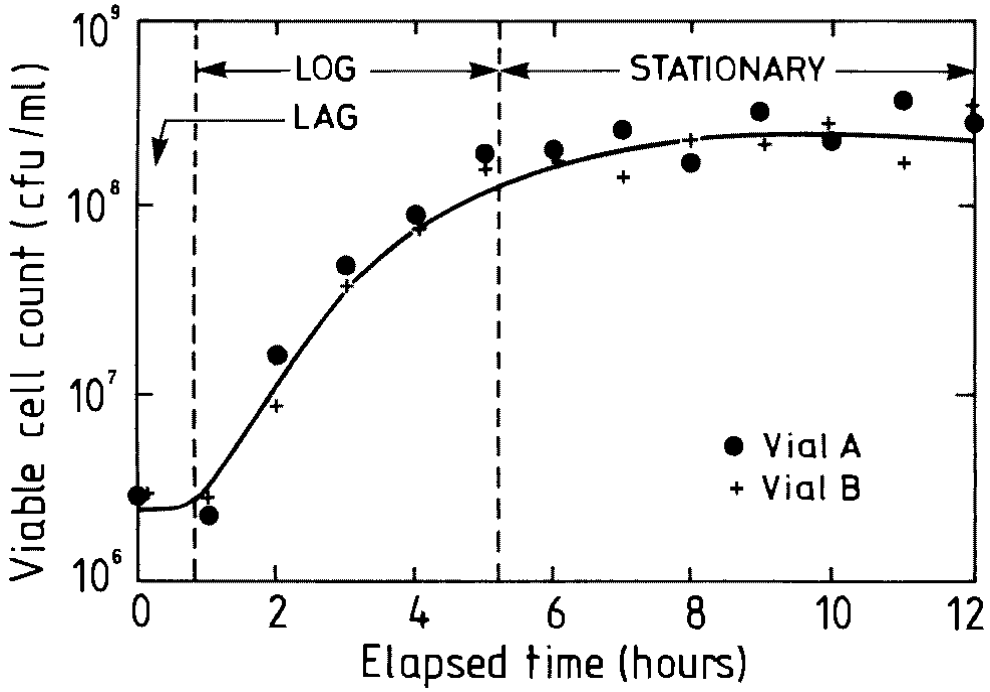


Fig. 8: Two exemplary *E. coli* growth[12] trajectories during the lag-, log-, and stationary phase.

In stationary phase, the *E. coli* culture has reached a point in time where the quality of the growth medium, i. e. substrate composition and density of

metabolic products as well as cell density has evolved to a state that creating new cells and decomposition is at equilibrium. This phase is of particular interest in biotechnology, since now secondary metabolites are produced by *E. coli*.

In death phase the rate of cell division becomes lower than the rate of decomposition, leading to a decline in cell numbers. This is caused by depleted nutrients, accumulated metabolic toxic waste-products and resulting hostile conditions in the growth medium.

Growth models of varying complexity have been developed to represent mainly the first three phases of the bacterial cell growth curve[13]. Most prominent ones are the Gompertz

$$g_{(t)} = a \cdot \exp[-\exp(b - ct)] \tag{8}$$

and Richards

$$r_{(t)} = a \cdot [1 + b \cdot \exp[c(d - t)]]^{-\frac{1}{b}} \tag{9}$$

models where the parameters do not have experimental significance and are used for fitting alone.

E. coli grow in aerobic and anaerobic conditions. Switching between the two kinds of environment requires an adaptive procedure of the organism. Thus, because of the aspect of reproducibility, a steady oxygen supply with a constant rate should be provided. Also, since *E. coli* shows varying behavior in terms of growth concerning oxygen levels[14] in the growth media, constant

conditions must be maintained in order to ensure reproducibility and comparability. This is in regards to e. g. the volume of the medium *E. coli* is grown in, the surface of the medium, the shaking velocity and pattern and the oxygen permeability of the container used for cultivation.

Many different devices for the cultivation of *E. coli* i. e. automation of nutrient and oxygen supply as well as providing optimal temperature for growth are available in biotechnology. Where chemostats support bacterial cultures with a steady flow of fresh growth medium, a turbidostat is able to obtain a representative value for cell density and only flushes through appropriate volumes of growth medium in order to keep the number of cells at a constant value. This way *E. coli* can be kept at certain phases of cell growth which is most useful for metabolite production in stationary phase, preventing the cells from changing to death- or exponential phase. However, most bioreactors are designed in a way that they supply the medium with oxygen by stirring. The stirring is mostly performed by scoops or turbines, creating shear forces. It has been shown that this can rupture even the cell walls of species that are not as fragile as those of *E. coli*[15][16]. Depending on the time that has elapsed since the start of the culture and the cell density that has been reached, a semi-state of the environment will be reached which is a mixture between the growth medium and a cell-free medium, created from the disrupted cells. This might render the experiment difficult to reproduce since the rupturing-rate contributes to the natural rate of decomposition. Where this could in theory lead to improved growth conditions, it was shown that the opposite was determined for other species than *E. coli*[17]. To overcome this, if possible, stirring in bioreactors must be avoided and replaced by shaking as is done in professional incubators.

5 Nutrient Conditions for optimal E. coli Growth

E. coli can be grown in various media as long as they contain (di)basic hydrogen, ammonium phosphate, sodium chloride, magnesium sulfate, potassium phosphate, water and a carbon source. As carbon source, mostly glucose is used but can be substituted with glycerol. Also other sugars e. g. arabinose

can be metabolized.

To optimize bacterial growth, different combinations of above ingredients i. e. growth media have been developed. One of the most popular is Lysogeny broth (LB). It contains mostly yeast extract and tryptone. Depending on the definition, glucose has to be added. Since the medium is manufactured from yeast extract, the specific composition is unclear which is why it is classified as an undefined medium. Differences in medium quality must be expected between batches, compromising reproducibility and comparability.

A medium for which ingredients are well defined is M9 minimal medium. It can easily be produced from all required agents which provide slow bacterial growth since no amino acids are included. M9 medium provides low autofluorescence and absorbance background. The medium can optionally further be enriched by the addition of specific amino acids or metabolites e. g. thiamine or casamino acids. It is suitable for cultivation of strains for production of DNA, plasmid DNA and proteins. M9 minimal medium is inexpensive and is widely used in scientific conduct and publications. This is why for this thesis all measurements have been performed in M9 minimal medium exclusively.

If medium is autoclaved i. e. sterilizing the media by exposing it to combined conditions of pressure and heat that disrupt cell walls, undesired chemical reaction may occur that render media quality between batches incomparable i. e. voids experimental reproducibility. This is carried out by e. g. the maillard reaction during which amino acids react with glucose in the process of autoclaving, adding a golden-brownish stain to the medium similar to crusts when baking bread[18]. The medium is thus not only deprived of nutrients but also agents are created that actually inhibit bacterial growth. It follows, that in order to ensure reproducibility and comparability, autoclaving growth media should be avoided. An alternative to autoclaving is to sterile-filter the medium using a filter with sufficiently small pore size.

6 The central Dogma of molecular Biology and beyond

Like in all known organisms, the interior organization of *E. coli* is governed by the central dogma of molecular biology i. e. the transformation of sequential information in terms of DNA, RNA and proteins. This sequential information is stored as a recombination of the nucleic acids adenine (A), cytosine (C), guanine (G) and thymine (T), arranged in the form of a double-helix in DNA, nitrogenous bases of adenine (a), cytosine (c), guanine (g) and uracil (u) in RNA and more than 20 aminoacids in form of a folded polypeptide-chain in proteins.

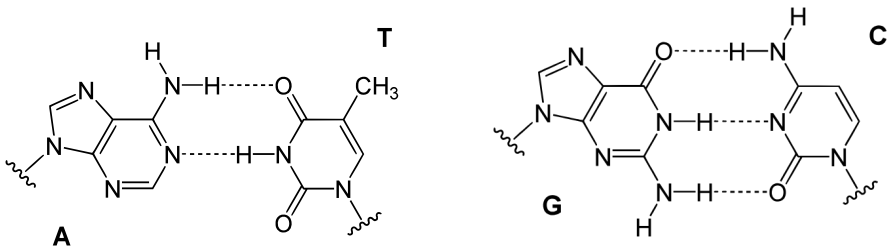


Fig. 9: Nucleobase-paring. Left: The purine adenine(A) binds to the pyrimidine thymine(T) via two hydrogenbonds. Right: The purine guanine(G) binds to the pyrimidine cytosine(C) via three hydrogenbonds. Connections to phosphate-deoxyribose backbone illustrated with curly lines.

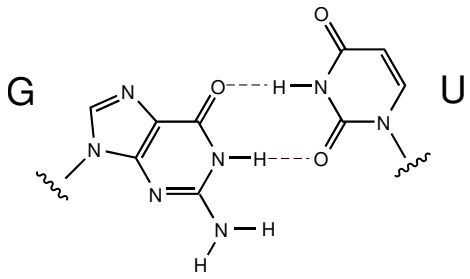


Fig. 10: Untypical ribonucleobase-paring. The binding of purine derivative guanine(G) and pyrimidine derivative uracil(u) is refered to as "wobble base pair".

The general interactions can be grouped into three major components i. e. DNA replication, translation and transcription. During DNA replication, helicase splits the double helix structure of DNA between the hydrogen bonds. Both single strands are completed each by DNA polymerase again.

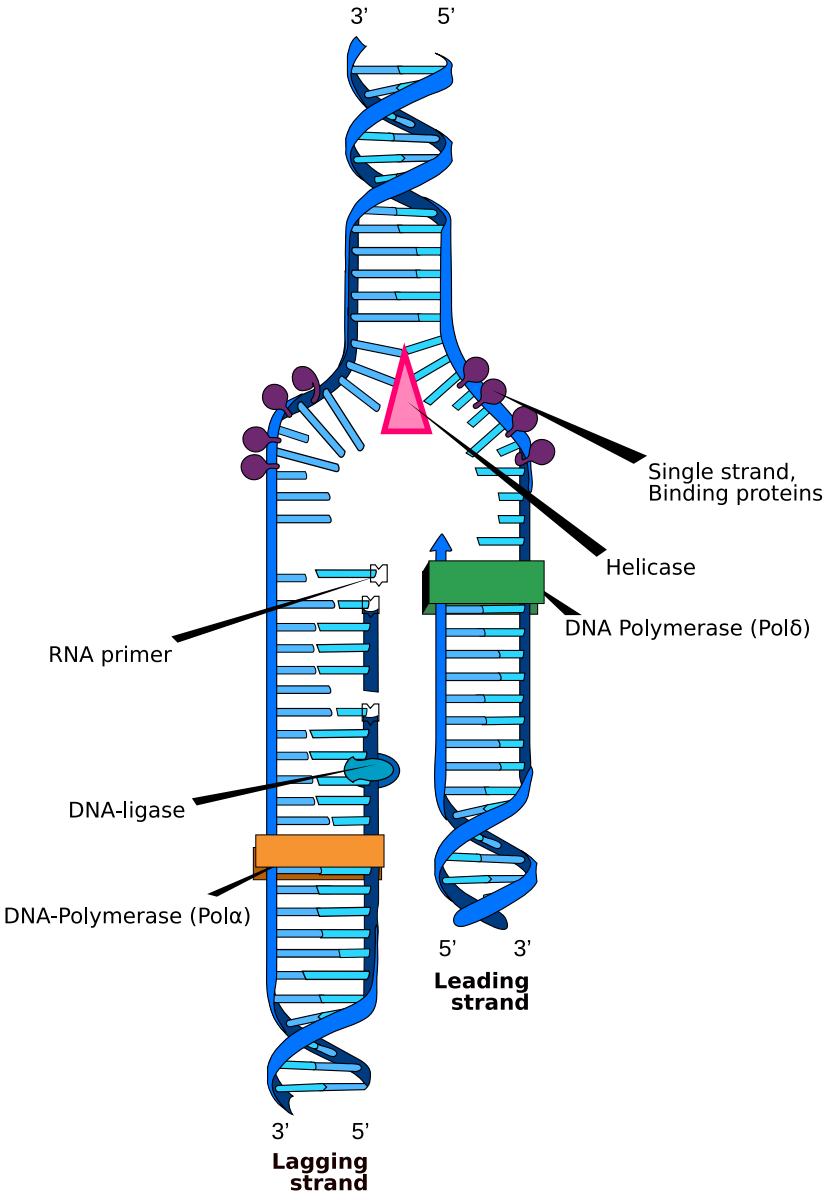


Fig. 11: The central dogma in a nutshell(1): DNA(blue) replication[19] by the helicase(pink) and DNA polymerase(green or orange).

During transcription, RNA polymerase transcribes DNA to RNA (A→a, C→c, G→g and T→u) after it has bound to a region of DNA with high RNA polymerase binding affinity i. e. a promoter, adding up free nitrogenous bases to the template strand, the part of DNA coding for RNA that has only temporarily been disconnected at the hydrogen bonds. When the RNA polymerase has reached a terminator, transcription is interrupted by e. g. the mechanical features of the RNA being produced.

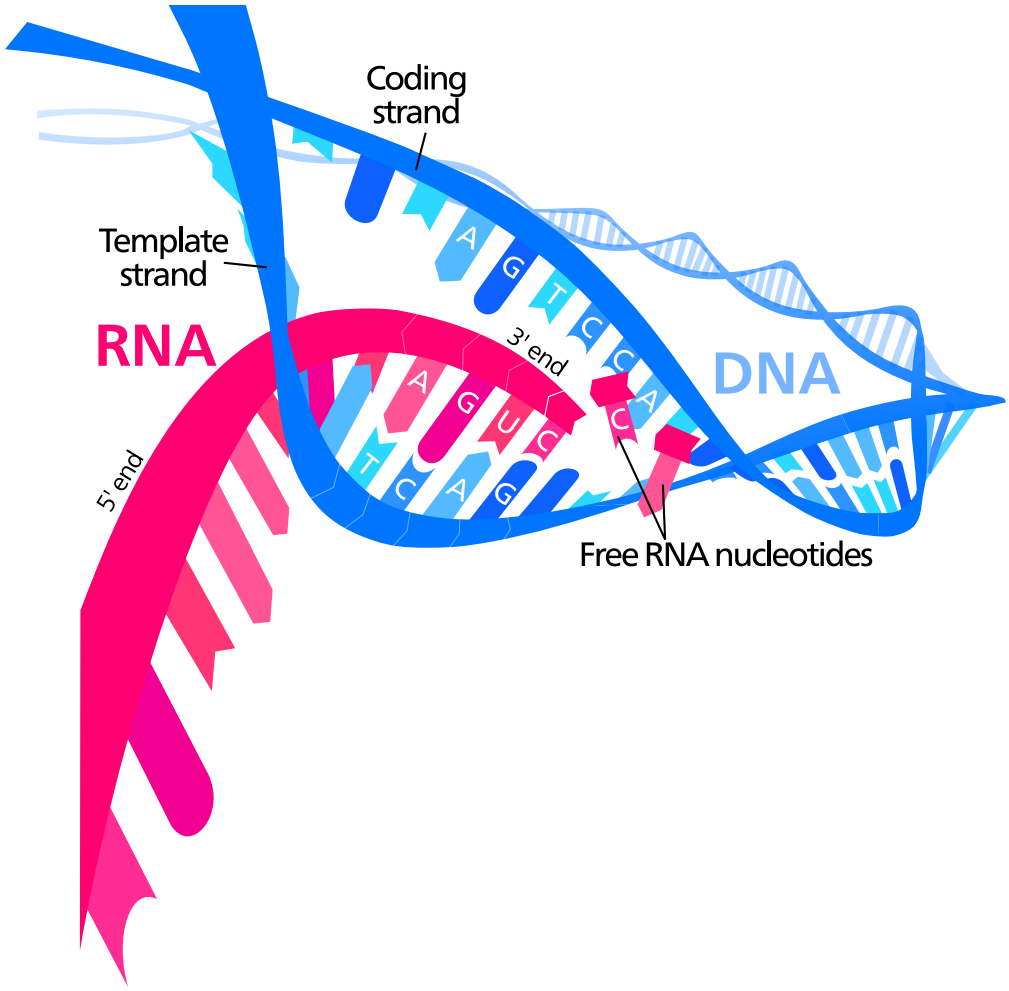


Fig. 12: The central dogma in a nutshell(2): Transcription[20] from DNA(blue) to RNA(pink) by the RNA polymerase(not shown).

During translation, after docking to the ribosome binging site (rbs) or Shine-Dalgarno sequence (sd), the small ribosomal subunit moves along the messenger RNA (mRNA) until it encounters the start codon i. e. the RNA sequence of aug, where the large ribosomal subunit and the first transfer RNA (tRNA) join it. Each tRNA is equipped with an amino acid, forming a chain linked by peptide bonds with the ribosome, a superstructure of RNA and proteins, moving along the mRNA. When reaching a stop codon i. e. the RNA sequence of uag, uaa or uga, the finished polypeptide is released.

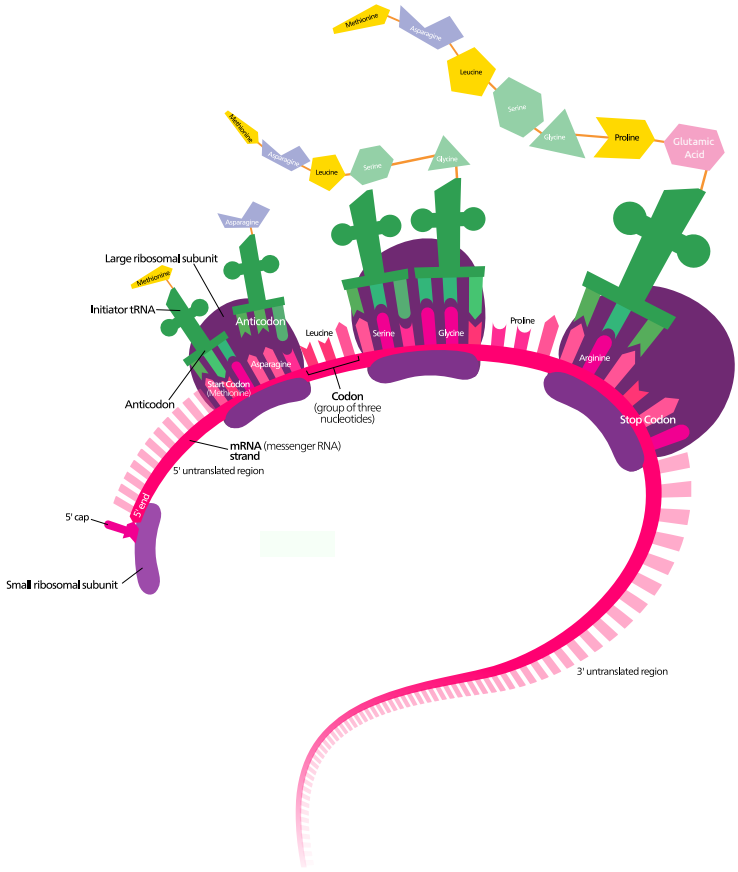


Fig. 13: The central dogma in a nutshell(3): Translation[21] from mRNA (pink) to amino acids(polygons) by the ribosome(violet) using tRNA(dark green). Three nitrogenous RNA bases (codon) code for one amino acid where designation is surjective from RNA to amino acid.

Through thermodynamic processes but also often supported by an assisting protein i. e. a chaperone, the amino acid chain then folds to a structure, reaching the state of a protein.

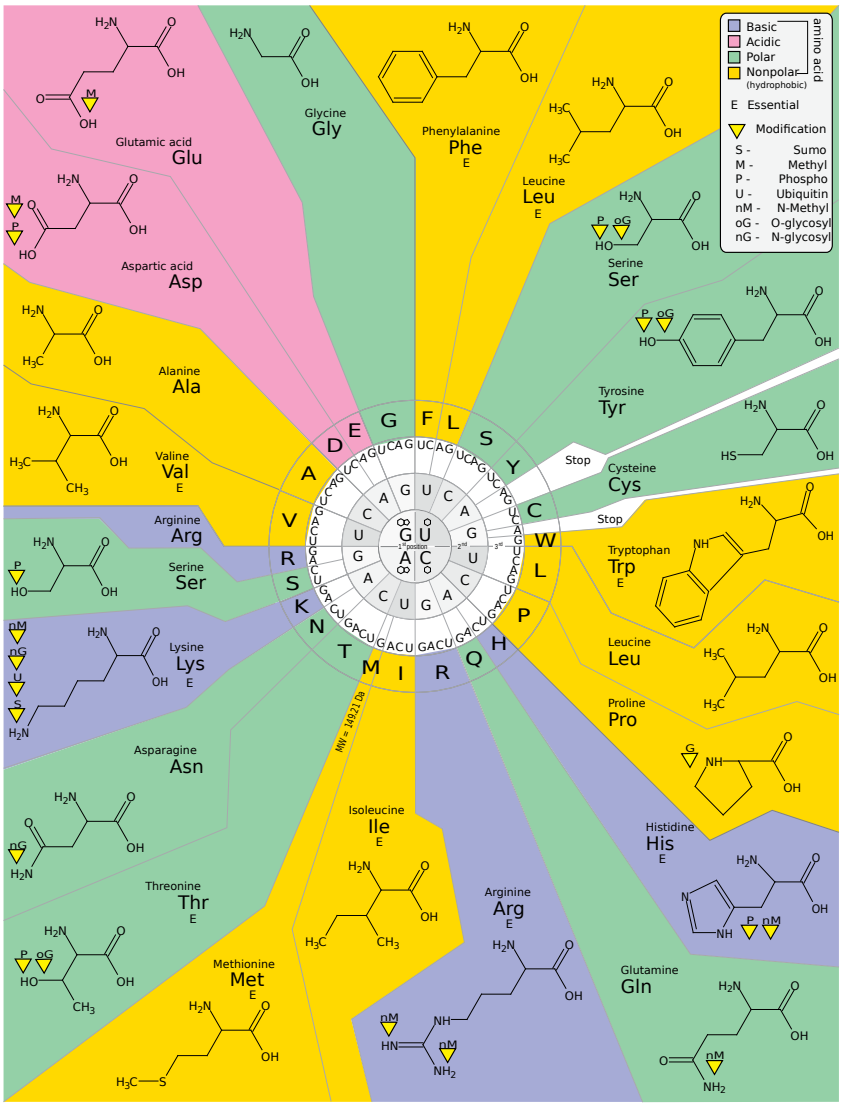


Fig. 14: The central dogma in a nutshell(4): Ribonucleic acid to amino acid conversion table[22]. Prokaryotes are not limited to only producing naturally occuring amino acids shown above. This is a major boon for the industrial employment for creation of synthetic materials.

In order to keep all participating molecules at certain levels, three members of the protein-class of enzymes namely DNAse, RNAse and protease degrade their counterpart DNA, RNA and proteins correspondingly.

For the process of transcription and translation, functional elements have been discovered on DNA and RNA. The process of transcription is initiated by a promoter, a sequence of DNA with a high binding affinity to RNA polymerase.

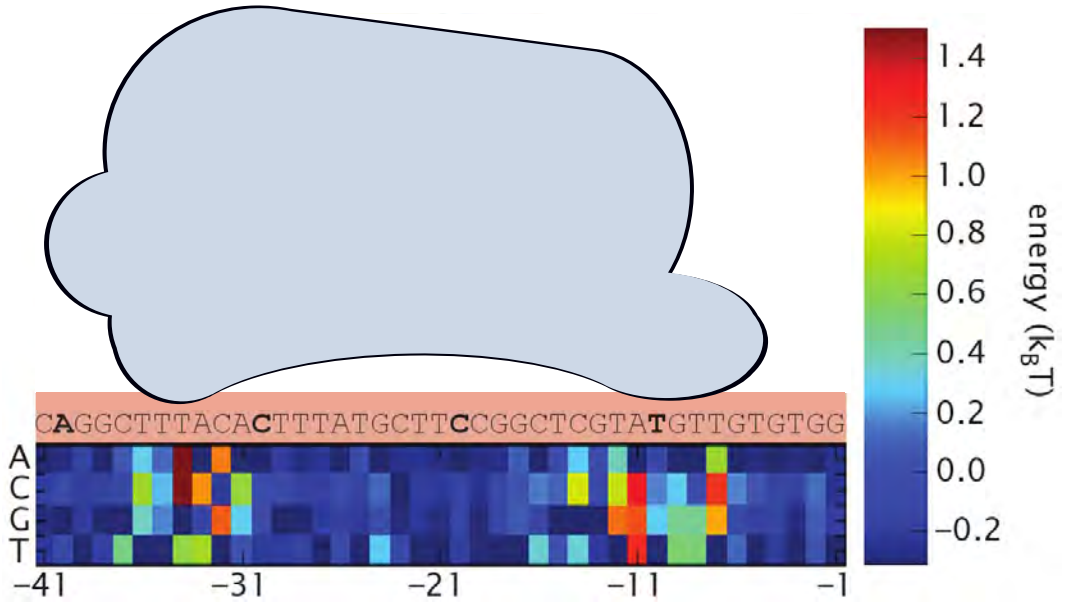


Fig. 15: Overview of how much each nucleic acid contributes to the binding affinity of the RNA polymerase complex (light blue) in the prokaryotic promoter region. Intensively contributing nucleic acid configurations i. e. promoter boxes, are shown with colors of higher value for red and lower for blue. Lesser contributing nucleic acids i. e. regions with lower binding contribution are visualized vice versa[23].

In particular, the promoter is the DNA sequence with a length of approximately 40 base pairs that is optionally followed by a ribosome binding site (RBS) and the coding sequence (CDS). As the name suggests, the RBS is a stretch of about six base pairs of DNA that codes for RNA which then will have a high binding affinity for ribosomes i. e. translation is initiated at the RBS if the DNA sequence is intended to produce a protein. In opposition to that, e. g. for RNA that is not intended to be translated into a protein, no RBS

is required. About four base pairs downstream the RBS, the CDS starts. It is the region from which after it has been transcribed into RNA, the amino acids chain is translated. Thus, the length of the CDS in basepairs must always be a multiple of three. After the CDS, a terminator concludes the transcription process.

In more detail, an E. coli promoter can be categorized into three parts. Two promoter boxes of about six base pairs which are located at an approximate position of -35 (e. g. TTGACA) and -10 (e. g. TATAAT) relative to the CDS are the main contributor for the RNA polymerase binding affinity and flank a stretch of DNA that is of lesser importance in that manner[23]. This property allows to create up to $4^{(35-10-6)} = 2^{38} \approx 275$ billion unique promoter sequences with almost identical binding affinity to RNA polymerase in reference to the sequences between the -35 and -10 promoter box alone.

The promoter region was discovered to be the element exhibiting most control for genetic expression. Given a fixed volume with a certain amount of RNA polymerase molecules in it as in an E. coli cell, the stronger the RNA polymerase binding affinity of the promoter, the higher the concentration of ribonucleic transcript and therefore optionally proteomic translation in the volume. For this relationship, a model[24] can be derived. Consider a system in two possible states where in the first state all polymerase molecules are not bound to one specific site i.e. they are localized either on non-specific sites or in the cytoplasm and where in the second state one polymerase molecule is bound to that specific site while the rest remains unbound to the specific site or in the cytoplasm as in the first state. For this system the total statistical weight can be constructed as sum over the two states

$$Z_{tot(P)} = Z_{1(P)} + Z_{2(P-1)} \tag{10}$$

where

$$Z_{1(P)} = \frac{N_{NS}!}{P!(N_{NS}-P)!} exp\left(\frac{-PE_{pd}^{NS}}{k_B T}\right) \tag{11}$$

and

$$Z_{2(P-1)} = Z_{1(P-1)} exp\left(\frac{-E_{pd}^{S}}{k_B T}\right) \tag{12}$$

where N_{NS} is the number of non-specific sites, P the amount of RNA polymerase molecules in the volume, E the binding energy for the specific(S) and average non-specific(NS) site in reference to the genomic background, T the temperature and k_B the Boltzmann constant. The first term of $Z_{1(P)}$ describes the distribution of the polymerase molecules using combinatorics which is followed by the Boltzmann weight. Assuming that P is much smaller than N_{NS}, the probability of encountering a polymerase molecule bound to a specific site can then be simplified to

$$p_{bound} = \frac{Z_{2(P-1)}}{Z_{tot(P)}} = \left[1 + \frac{N_{NS}}{P} exp\left(\frac{\Delta E}{k_b T}\right)\right]^{-1} \tag{13}$$

where $\Delta E = E_{pd}^{S} - E_{pd}^{NS}$.

7 Plasmids as Extension of the prokaryotic Chromosome

Alongside the chromosome, *E. coli* can also be equipped with additional DNA in form of a plasmid. Plasmids are DNA, mostly of circular form and a sequence-length between one and 200 kbp (kilo base pairs)[25], where there is no fixed limit from which on plasmids can be described as minichromosomes. Plasmids replicate autonomously, independent of the chromosome and the number of identical plasmids, the copy number[26], can reach up to from one to more than thousands[25]. The copy number is dependent on the length of the sequence of the plasmid, the replication mechanism, the host organism, the type of growth medium used, the amount of other plasmids present and many more parameters.

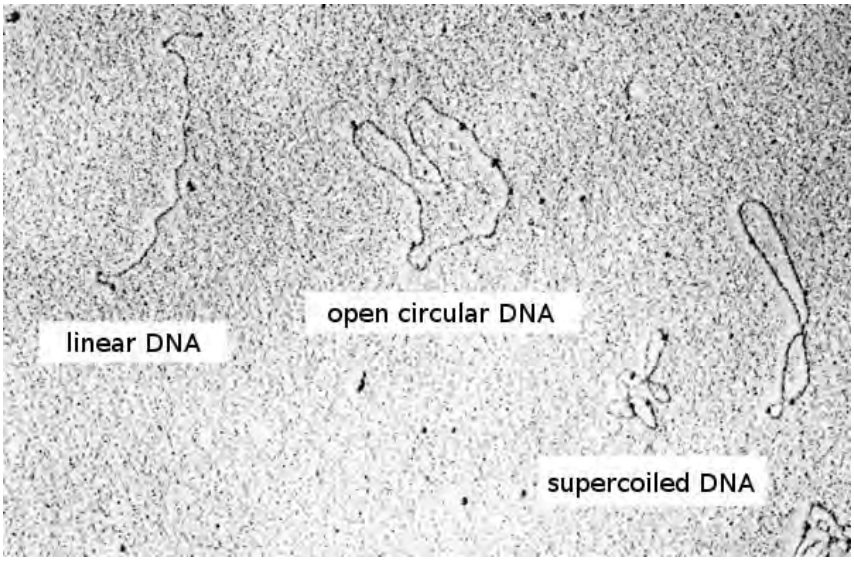


Fig. 16: Plasmids in different shapes under an electron microscope(image taken from Wikipedia Commons article "Plasmid"). Left: Additional genetic content can also be present in linear form however, it will be not as well preserved as in circular form(center). DNA can be compacted e.g. "supercoiled" (right) which is necessary for many biological processes.

For the initiation of the replication process, plasmids must possess a certain sequence of DNA i. e. the origin of replication. Roughly, the replication mechanisms can be divided into three groups. The first is able to insert itself into the chromosome of the host organism for replication. The second uses the DNA replication mechanism of the host in order to duplicate the plasmid. The third has a distinct plasmid-specific mechanism of replication where corresponding proteins (Rep) or RNA, iterons i. e. repeating sequences of DNA, DnaA boxes i. e. sequences of DNA with high binding affinity to the replication initiation protein DnaA and or an adjacent DNA sequence with high content of adenine and thymine are present[27]. Since identical replication mechanisms interfere with each other by competition for replication factors, only up to three different kinds of plasmids can be kept in *E. coli* in the first place, allocating plasmids in corresponding incompatibility groups[28].

Common Vectors	Copy Number	ORI	IG	Control
pUC	500-700	pMB1 (derivative)	A	Relaxed
pBR322	15-20	pMB1	A	Relaxed
pET	15-20	pBR322	A	Relaxed
pGEX	15-20	pBR322	A	Relaxed
pColE1	15-20	ColE1	A	Relaxed
pR6K	15-20	R6K	C	Stringent
pACYC	10	p15A	B	Relaxed
pSC101	5	pSC101	C	Stringent
pBluescript	300-500	ColE1 (derivative) and F1	A	Relaxed
pGEM	300-500	pUC and F1	A	Relaxed

Table 1: Overview of the most common plasmid types adopted from blog.addgene.org. The "Copy Number" is a rough estimate for the average number of plasmids that is present in E. Coli if in use. ORI stands for "Origin of Replication", the part of the plasmid that is responsible for the duplication process. IG stands for "Incompatibility Group" i.e. if more than one plasmid is in use, the plasmid selection must comply with the different types of plasmids being from the three different groups A, B, and C. The column "Control" indicates if the plasmid type is dependent on the host for duplication(stringent) or not(relaxed).

Plasmids can be transferred from one *E. coli* to another by horizontal gene transfer i. e. conjugation or manually by an experimenter and means of transformation.

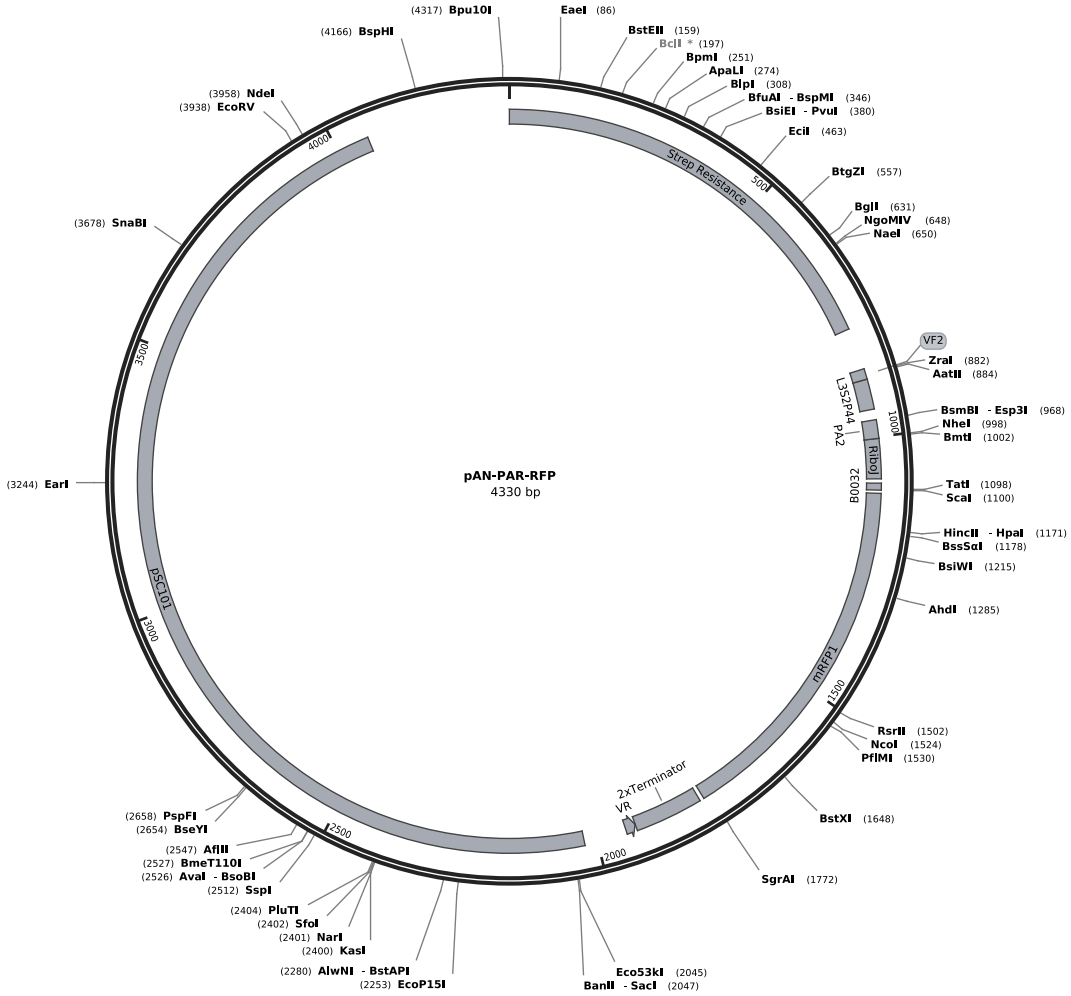


Fig. 17: Visualization of a plasmid using the software SnapGene. The name "pAN-PAR-RFP" and size "4330bp" of the plasmid is displayed in the center. Gray parts of the circle symbolize functional parts as the origin of replication "pSC101" on the left, the resistance against the antibiotic Streptomycine and the gene coding for mRFP1 on the right where non-functional parts of the plasmid DNA are left blank. The black outer rings indicate the relative position. Additional information containing possible sites for cutting or insertion is provided on the outside where the bold letters stand for the abbreviation of the restriction enzyme that could be used at that position.

8 Plasmid Incorporation into E. coli by Electroporation

One of the most efficient types of plasmid injection into *E. coli* is by electroporation. A suspension of cells and plasmids is transferred into an electroporation cuvette which has an electrode on each of two opposing sides. By applying voltage to the electrodes, the cells are exposed to a strong electric field that, because of the negative net-charge, accelerates the plasmids towards the positively charged electrode and a rapid localized rearrangement in lipid morphology takes place in the membrane of *E. coli*, leading to pore-like formations. The combination of both effects leads to plasmids penetrating into the interior of the cells. Provided the applied voltage was not too strong to disrupt the cell as well as not too weak to not allow the effects described beforehand, the electroporation was successful. After an adaptation phase in an enriched growth medium e. g. SOC medium, the transformed cells can now be further processed.

The success of the transformation depends on many factors. The electroporation cuvette should be frozen to avoid being heated too much in the process of applying voltage since by causing a denaturation i. e. re-folding of proteins, temperatures being too high would be lethal to the cells. An appropriate concentration combination for cells and plasmids must be provided. Up to three kinds of plasmids can be transformed into *E. coli* by means of electrotransformation in parallel, promising low to none mutations in the process. However, this procedure yields low efficiency so its success is strongly dependent on the skill of the experimenter.

Plasmids can get lost in the process of cell division, especially for low copy numbers. Also, they do not pose an evolutionary advantage for the host cell in the first place. Liberating the invested resources by discarding the containing plasmids and investing elsewhere might be beneficial for *E. coli* without the addition of further environmental factors. In order to force the prokaryotes to maintain a plasmid, a sequence of DNA coding for an antibiotic resistance can be added to the plasmid where the corresponding antibiotic agent is added to the growth medium.

9 Artificial antibiotic Resistance in E. coli

Mostly for low copy numbers, plasmids can get lost between generations during cell division, when all plasmids of a type only end up in one descendant. In order to maintain a homogeneous culture i. e. a bacterial culture only containing *E. coli* that are equipped with a certain plasmid, an antibiotic resistance is added to the plasmid and the corresponding antibiotic to the environment i. e. growth medium or agar plate. *E. coli* that have lost the plasmid carrying the life-saving antibiotic resistance will perish, which guarantees that the plasmid will be present in all living *E. coli*. Many different kinds of antibiotics with distinct mechanisms have been discovered in the past. Commonly used ones are kanamycin, carbenicillin and streptomycin.

Antibiotic	Stock Concentration	Working Concentration
Ampicillin	100 mg/mL	100 μ g/mL
Bleocin	5 mg/mL	5 μ g/mL
Carbenicillin*	100 mg/mL	100 μ g/mL
Chloramphenicol	25 mg/mL	25 μ g/mL
Coumermycin	25 mg/mL	25 μ g/mL
Gentamycin	10 mg/mL	10 μ g/mL
Kanamycin	50 mg/mL	50 μ g/mL
Spectinomycin	50 mg/mL	50 μ g/mL
Tetracycline	10 mg/mL	10 μ g/mL

Table 2: Overview of the common antibiotics adopted from blog.addgene.org. All antibiotics should be dissolved in H_2O where coumermycin should be stored ind DMSO and chloramphenicol in EtOH. The column "Stock Concentration" indicates the optimal quota of the agent for storage where "Working Concentration" describes the optimal portion for culture treatment. Some antibiotics are "compatible" i.e. they can be exchanged e.g. carbenicillin and ampicillin.

In *E. coli*, kanamycin interacts with parts of the bacterial ribosome i. e. the RNA strand or subunit 30S. It does not abolish the function of the ribosome but causes it to misread the RNA processed so unintentional amino acids are incorporated into the peptide chain which later forms proteins, part of the

functional basis of *E. coli*. The proteins which have inherited the translation errors do not fold as their correctly translated counterparts do nor do they have the same amino acid configuration, yielding a different set of functionality i. e. the cell loses its ability to properly organize its life-sustaining foundation, thus the errors in translation from RNA to amino acid chain caused by kanamycin are lethal for the cell[29][30]. Antibiotic resistance against kanamycin can be found in organisms that produce the antibiotic themselves. The genetic sequence that codes for the protection i. e. the production of neomycin phosphotransferase, an enzyme that modifies the molecular structure of antibiotics and by that rendering them innocuous, could be extracted from species of streptomycetes and incorporated into the DNA sequence of a plasmid[31].

Ampicillin inhibits the construction of the cell wall in *E. coli* by mimicking the D-alanyl(4)-D-alanine(5) extremity of peptidoglycan precursors. These precursors i. e. transpeptidases are responsible for the final stage of the cell wall completion i. e. the catalyzation of cross-linking peptidoglycans. The antibiotic effect of carbenicillin on *E. coli* is similar to that of ampicillin. Carbenicillin is less likely to break down in the presence of β-lactamases[32] which means slower degradation and prolonged antibiotical pressure in the presence of the penicillinase which breaks down the antibiotic by hydrolysing its β-lactam ring. This is not the only reason why carbenicillin is favored in microbiology over ampicillin even though it is less expensive. Carbenicillin can also reduce the growth of satellite colonies on plates during long-term incubation[33]. Also, its breakdown results in byproducts with a lower toxicity than ampicillin, making it the more favorable antibiotic when comparing ampicillin and carbenicillin.

Spectomycin inhibits protein synthesis by binding to the S12 protein of the 30s ribosomal subunit and blocking translation[34]. A genetic antibiotic resistance against streptomycin, a functionally similar agent, could be found in a strain of coccobacillus[35]. This gene can be used to introduce artificial antibiotic resistance in *E. coli* strains via integration into a plasmid. The aadA14 gene codes for the streptomycin 3"-adenylyltransferase that catalyzes the degradation of spectomycin and streptomycin to diphosphate and 3"-adenylylstreptomycin using adenosine triphosphate.

As for the used setup of plasmids and in the experimental framework, no degradation rate is known for the antibiotics described above, it is considered as negligible in this work.

10 E. coli as a Workhorse for DNA Amplification

Since *E. coli* replicate plasmids and are easy to cultivate, they are often used for their sole purpose of plasmid duplication i. e. amplification. Among the many strains, *E. coli DH5α* is the most commonly used because of its high DNA replication rate. After cultivating transformed *E. coli* between 12 and 20 hours, depending on the environmental conditions, plasmids with higher or lesser amounts of concentration and mutation can be extracted by plasmid preparation. For that, *E. coli* need to be washed i. e. centrifuged and separated from their medium and placed in a buffer where then the cell wall is ruptured by a specific agent i. e. the cells are lysed. Plasmids, in contrast to proteins and larger structures of nucleic acids, are not denatured, so the DNA can be filtered out by binding to specific silicates after the lyse has been stopped by appropriate counteragents. After multiple washing steps, the purified DNA can be dissolved from the silica by elusion and collected in water or a buffer from where it can be further processed as for transformations, sequencing or cloning.

11 Molecular Cloning - A cut & paste Tool for DNA

Molecular cloning is the usage of an enzymatic subgroup of proteins called restriction enzymes that cut DNA at specific sites in order to insert DNA into an existing construct or cut it out of such. These enzymes can be categorized into type-groups where specificity or usability rises with index value. The specific method of molecular cloning is stipulated by the kind of restriction enzyme used. Popular methods are blunt-end cloning, Gibson assembly and Golden Gate assembly.

In Gibson assembly within a single isothermal reaction up to five DNA fragments can be merged into e. g. a plasmid presuming a 20-40 bases overlap in DNA sequence complementarity is provided among the corresponding binding partners. Alongside the DNA sequences, also exonuclease, DNA polymerase, DNA ligase and according buffers are added. The exonuclease transforms the doublestranded DNA (dsDNA) to singlestranded DNA (ssDNA) starting at the 5' end and does not interfere with polymerase activity. The then exposed ssDNA can anneal, where missing nucleotides are inserted by DNA polymerase. The role of DNA ligase is to connect the fragments according to their complementarity so no nicks are left in the DNA. These steps result in merging all inserted pieces of DNA into one where linear and circular constructs are possible. Gibson assembly does not leave restriction enzyme scars since no digest after PCR is necessary, however occurrence of mutation is increased when DNA polymerase fills the gaps. The procedure can even be enhanced to combining 15 pieces of DNA if the exonuclease and annealing steps are done first, followed by the addition of DNA polymerase and ligase as final step.

Another method that also allows for the combination of multiple DNA sequences in one single operation is Golden Gate cloning or assembly. This is possible since digestion and ligation can be carried out simultaneously. Alongsite T4 DNA ligase, type II restriction enzymes that cut downstream their recognition site i. e. in a non-predefined region of DNA, like BsaI are used to create non-palindromic overhangs. Since for these overhangs with length of four base

pairs 256 combinations are possible, multiple DNA strands can be merged into one scarless product in a single reaction. However, also single inserts can be performed with this technique. The reaction is irreversible since the resulting DNA construct has no restriction enzyme recognition site anymore. Optimal temperatures for the operation of T4 ligases is 16°C, where 37°C is optimal for restriction enzymes, defining the boundary conditions for a Golden Gate assembly program in a thermal cycler. The quality of the product depends on the length of the overhangs, so four base pair overhangs should be favored compared to shorter ones.

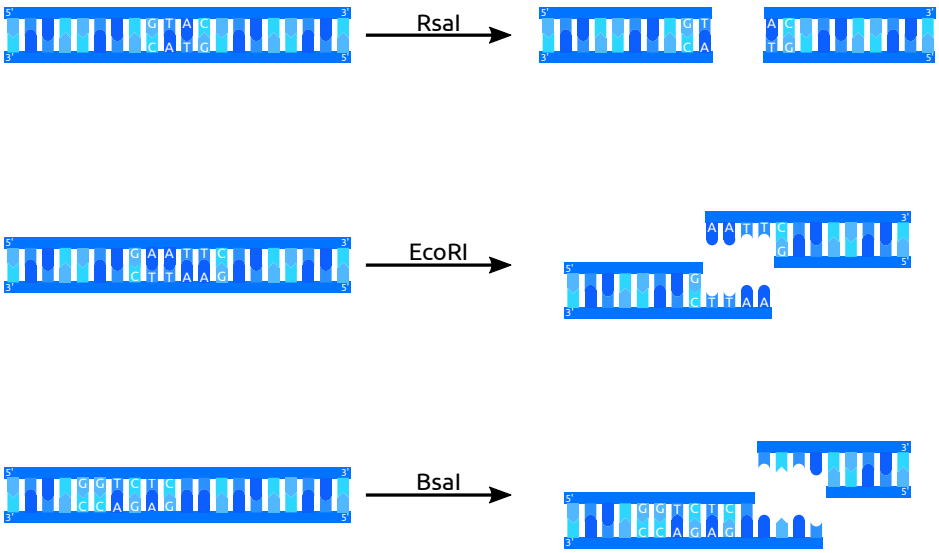


Fig. 18: Scheme of restriction enzymes used in blunt-end, Gibson- and Golden Gate assembly. Top: RsaI usage cleaves DNA to blunt ends centered at recognition site GTAC, leaving no overhangs. Center: EcoRI digests DNA leaving "sticky ends" of four bases at center of recognition site GAATTC. Bottom: BsaI also digests DNA leaving "sticky ends" but outside and downstream (towards 3' end) the recognition site GGTCTC in Golden Gate assembly. Nonspecific DNA is illustrated with absence of characters.

In contrast to Gibson- and Golden Gate assembly, blunt-end cloning does not make use of ssDNA overhangs but relies on the transient associations of the available 5' phosphate and 3' hydroxyl groups. Since this method results in less colonies and in approximately half of them, inserts have been injected with an unintended orientation or not at all, it is about ten to 100 fold less efficient. However, blunt-end cloning is one of the easiest methods for inserting dsDNA into a plasmid. Considering enzymatic digestion and subsequent purification, little to no preparation is required, making it very fast to apply. Also, an insert does not have to provide any restriction sites which makes the cloning strategy very versatile. This is why e. g. Genscript, one of the leading companies in DNA synthesis and cloning services favors this method of DNA processing.

Even more technologies of molecular cloning are available for the usage of type IV or V restriction enzymes as McrBC or the cas9-sgRNA complex from CRISPR respectively.

Many companies exist that do not only offer custom DNA synthesis but also molecular cloning service. Considering that often a time-consuming cloning strategy has to be planned beforehand, a custom designed DNA sequence has to be dissolved in special buffers after delivery, the DNA has to be amplified using additional strands of DNA called primers and specific polymerases, restriction enzymes have to be bought, stored and prepared, expensive equipment has to be used and so forth, the amount of time an experimenter has to come up with as well as the funding that has to be spent, these investments can be cut by outsourcing molecular cloning equally to DNA synthesis to a specialized provider. Genscript is the global leader in terms of providing services for DNA synthesis. Also, the company offers services for molecular cloning and DNA sequencing. An additional service of free plasmid storage renders them an ideal partner for conducting genetic experiments in microbiology.

12 Uncovering DNA Composition with Sequencing

Before further experimentally processing newly built genetic constructs, first it must be verified that the intended assembly was a success and has not failed due to errors in cloning or possible mutations to eliminate the possibility of a pointless experiment. To uncover the composition of a DNA strand, various techniques have been developed. Research in this area is still progressing rapidly and one already well established method is Sanger DNA sequencing.

During Sanger DNA sequencing, dsDNA is split up into ssDNA by heating the containing buffer. As in DNA replication, DNA polymerase can add the single normal deoxynucleosidetriphosphates (dNTPs) Adenine, Cytosine, Guanine and Thymine to the now vacated ssDNA. These single nucleotides can be replaced by altered versions, modified di-deoxynucleotidetriphosphates (ddNTPs), that terminate DNA strand elongation and are each equipped with different fluorescent proteins, replicating new DNA strands with specific length that is defined by the initiation of the DNA replication and the termination by the ddNTP.

The distribution of lengths in a sample can now be evaluated by gel electrophoresis i. e. placing the sample in a gel which is submitted to an electric field at approximately 150V. The DNA molecules with negative net-charge will be drawn towards the anode but depending on their size and complexity in structure held off by friction effects with the gel. The net-effect is a separation of DNA sequences according to length. The result can be evaluated by comparison with a ladder i. e. DNA with a predefined discrete distribution of lengths that has also been placed in the gel. The four different ddNTPs can now be separated or identified by using illumination according to the distinct excitation and emission spectra of the four fluorescent proteins bound.

Alternatively, the procedure can also be evaluated with a single stain or by means of radiation if the sample has been divided into four sub-experiments and only one of the four dNTPs was replaced by a corresponding ddNTP each.

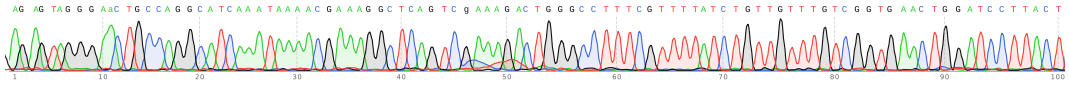


Fig. 19: Typical results for DNA sequencing. The upper letters indicate the strand elongation terminating ddNTP i. e. dNTP-replacement, the colored graphs the corresponding signal strength of the fluorescent protein used. The grey bottom letters indicate the relative position. Sequencing file received from Genscript. Image file-printed to PDF from Snapgene Viewer 4.1.9

Genscript, the leading company in DNA sequencing in North America has adapted this method of DNA analysis. The service is free of charge and if genetic constructs have been ordered at this company beforehand, it makes sense to have DNA sequencing performed directly after cloning on site.

13 Inducer Systems as Link between Macro- and Microcosm

In order to e. g. activate and repress protein synthesis or probe artificial genetic circuits in *E. coli*, a connection to the genetic content is required to enable the experimenter to toggle the setup as desired. Genetic inducer systems work by making use of the principle that externally added chemical compounds can interfere with regulative genetic products and by that alter genetic expression i. e. the transcription of RNA from DNA and also from there resulting in controllable protein synthesis. Contrary to activating systems, the three inducer systems P_{BAD}+L-arabinose, P_{phlF}+2,4-Diacetylphloroglucinol (DAPG) and P_{tet}+Tetracycline, the latter being an Anhydrotetracycline (ATc) analog or substitute, are repressor systems i. e. the corresponding chemical compound has to be added in order to release repression from the matching promoter hence enable transcription from the CDS.

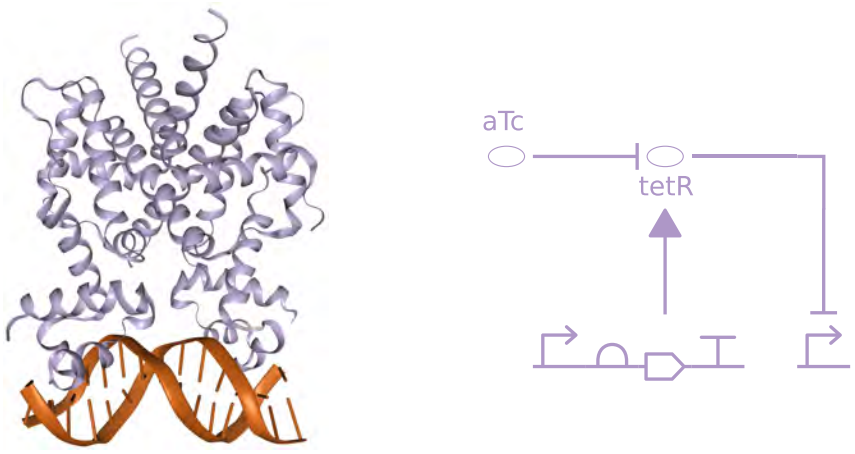


Fig. 20: TetR inducer scheme. Left: TetR protein bound to TetO boxes illustrated using NGL viewer [37]. Right: Simplified tetR inducer system illustration using SBOL Visual font.

The product of the *tetR* gene, the protein TetR has the property of binding to a specific DNA sequence, the TetO sequence. When TetR is bound to this section of DNA, the RNA polymerase is blocked from proceeding along the strand and transcription is prevented. The block can be reduced or abolished by the addition of either the chemical compound Tetracycline or Anhydrotetracycline (ATc), a functional homolog. The agent binds to TetR itself, seizing its property to be able to connect to the TetO box. The property of TetR to being able to bind to the TetO sequence can be used to further improve the protein in its repressive inducer-system function. For that, the CDS of *tetR* has been duplicated in a way, that two proteins are fused i. e. two TetR proteins are merged without losing their functional properties. This enables the TetR fusion proteins to bind to two TetO sequences simultaneously, decreasing the probability of a RNA polymerase passing the response element by approximately half through the binding of the TetR fusion protein to the operon[38][39][40]. Also, further research suggests that the activity of the fusion protein bends the DNA sequence in the close vicinity of the TetO boxes to a loop form, reducing RNA polymerase activity in this region even further[41].

It must be taken into consideration, that the concentration of the repressor-disabling agent ATc is not stable over time i. e. at a reported degradation rate

of 0.058h^{-1} in a cultured M9 medium at 37° C, only a fraction of 0.24 ATc will have remained in the system after 24h[42].

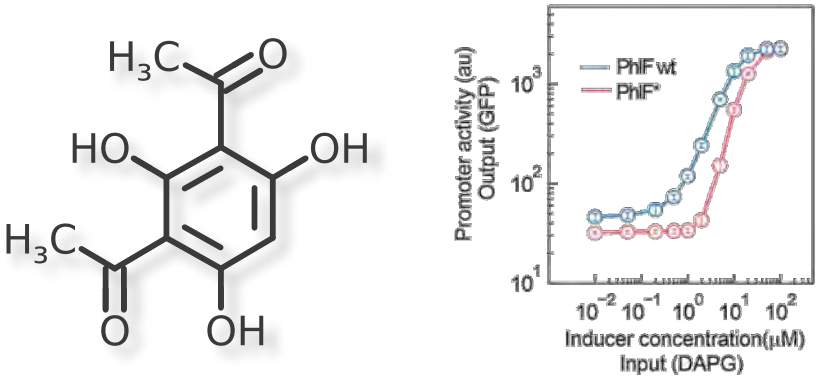


Fig. 21: Left: Chemical structure of DAPG. Right: Exemplary response curves i.e. transfer functions[36] of the regulated promoters as a function of the 2,4-diacetylphoroglucinol (DAPG) concentration, which inactivates the DNA-binding ability of the wild type PhlF (blue) and PhlF* (red) repressor.

Similar to the P_{tet}+Tetracycline inducer mechanism, the product of the *phlF* gene acts as a repressor protein on the corresponding promoter P_{phlF} i. e. it binds to this specific sequence of DNA preventing RNA polymerase to transcribe from the CDS downstream of it. If a sufficiently high amount of 2,4-Diacetylphloroglucinol (DAPG) is added, enough PhlF is bound by DAPG causing the probability for the promoter sequence being blocked by PhlF to diminish which enables transcription[43] at this site. The P_{phlF}+DAPG inducer system, which shows a very tight binding behavior i. e. when repressed, very little target gene activity is present, was also already successfully transferred into mammalian cells[44]. The inducer system is not endogenous to *E. coli*, which is why no mechanism of DAPG metabolization or breakdown in *E. coli* could be identified until this date.

The P_{BAD}[45]+L-arabinose inducer system makes use of DNA being looped[46] in the close proximity to the promoter to prevent RNA polymerase transcription initiation. This is achieved by a complex comprising two linked AraC[47]

proteins, that binds to two specific binding sites, the I_1 and O_2 region, which are located upstream the promoter. Binding can be abolished by the addition of L-arabinose to the system. L-arabinose fills the binding pocket of AraC, the N-terminal arm of AraC is released from its DNA binding domain and arabinose forces AraC to dimerise into a specific conformation, preventing it from achieving DNA-looping since instead of binding to the opposingly situated I_1 and O_2 it now binds to I_1 and its direct neighbor I_2 which unleashes P_{BAD}[48][49].

Fig. 22: P_{BAD}+L-arabinose, inducer systems illustrated using the SBOL Visual 2.3 font. Left: In a system without L-arabinose(black hexagons) the DNA containing the inducer system(light green) is looped by the binding conformation of two araC proteins resulting in a lowered RNA polymerase binding probability to the promoter due to a.o. inaccessibility, therefore achieving transcription suppression. Right: If the binding pockets(violet) of araC are filled with L-arabinose, the binding conformation of the two araC proteins undergoes a structural change which also releases the N-terminal arm(salmon) from the DNA binding region of araC. CAP-camp proteins(blue and orange) binding to the CAP-binding site(CBS) stabilize the now non-looped DNA alignment resulting in an increased RNA polymerase binding probability to the promoter and therefore transcription activation.

Another mechanism that is based on the catabolite activator protein (CAP or CRP)[50] protein contributes to the functionality of the P_{BAD}+L-arabinose inducer system. CAP binds to the CAP-binding site situated in the DNA region that can be turned into a loop (between I_1 and O_2), stabilizing the non-loop-conformation and thus contributing to the exempt configuration of P_{BAD}, promoting RNA polymerase transcription initiation. Also, CAP itself participates in transcription activation[51] by direct adoption of the RNA polymerase. CAP can only bind to the corresponding operon after a conformational change has occurred. This is mediated by the presence of cyclic

adenosine monophosphate (cAMP) where the concentration of cAMP depends on the amount of glucose in the sample i. e. cAMP concentrations are high when glucose content is low and vice versa.

The P_{BAD}+L-arabinose inducer system is endogenous to many bacteria, among them *E. coli* i. e. in unaltered strains a genetic sequence coding for the *araCBAD* operon can be found on the chromosome. This operon is responsible for the metabolization of L-arabinose, thus it must be taken into account, that L-arabinose is further processed and concentration of the agent is not stable i. e. the repressor-releasing effect of L-arabinose might diminish with time.

All inducer systems, i. e. the affected promoters and repressor-protein transcribing genes can be cloned onto plasmids which themselves can be transformed into *E. coli*, providing a configurable interface for the experimenter in vivo. Also, all promoters described earlier can be equipped with either a CDS from which a functional RNA is transcribed or a ribosome binding site followed by a CDS, which might yield the translation of a protein e. g. mRFP1.

Inducer systems can be characterized by their transfer functions i. e. for repressor systems, the relationship of the concentration of repressor-disabling agents in the system e. g. ATc, DAPG or L-arabinose and the resulting genetic expression e. g. a measurable observable e. g. a fluorophore as GFP. However, these transfer functions can not be transferred from one species to another or at least not without proper adjustment since some species lack genetic content for processing chemical compounds e. g. strains from which the araCBAD operon was removed, others don't. The transfer function does not describe the projection from repressor-disabling agent to the RNA produced, but only to the measurable property of a protein. Also, the growth medium, the distribution of different kinds of plasmids, the binding affinity of the participating promoters and much more contributes significantly to the shape of the transfer function. This is why in this work the display of a representative transfer functions for the former described inducer systems is omitted since if at all, it can only be taken into consideration as a general orientation.

14 Artificial Fluorescence in E. coli as Reporter

Analogous to how inducer systems create the access point to the input of the genetic experimental setup for the macrocosm, also the output of e. g. a genetic construct must offer some kind of obtainable form to the experimenter. The utilization of fluorophores as an experimental readout i. e. a reporter provides a quantitative performance evaluation of a genetic construct.

Many species can be found that are fluorescent by nature. Some species of the anemone genus Discosoma feature certain cells that contain a gene from which the fluorophore DsRed is produced. This gene could be isolated, transferred to a model organism and even improved[5]. DsRed is a tetramere i. e. its structure is governed by four (almost) identical amino acid chains. This tetramere could be reduced into a monomere, the monomeric red fluorescent protein (mRFP1) i. e. a fluorescent protein comprised by only one version of the former four amino acids chains. This enabled the usage of the fluorophore as a reporter[5] mostly since the edited version is produced from DNA i.e matures about ten times faster ($t_{50\%} = (21.9 \pm 1.1)$min, $t_{90\%} = (51.4 \pm 4.0)min$)[52]. As all proteins, mRFP1 is degraded by proteases. However, comparing the pace of the process with the very short maturation time, this effect can be considered as negligible.

Singlecell fluorescence measurements can be performed in a flow cytometer, however, those cells should not be used for reintegration into a time-resolved experiment, since for flow cytometry measurements mostly are performed in an aqueous solution with addition of some kind of saline i. e. a buffer is used in contrast to growth medium, possibly affecting cell behavior. Repeatedly removing cells from a culture might introduce foreign substances, change the growth conditions et cetera and thereby render parallel experiments incomparable.

In contrast to single-cell measurements, fluorescence can also be measured in a life culture by means of a platereader. Since the absorbance of cells at a wavelength of 600nm is regarded as being representative for the amount of cells in microbiology, fluorescence can be divided by absorbance yielding a fluorescence per cell volume value which offers comparability. Since fluorophores

are re-synthesized and degraded permanently, the effect of photo-bleaching is disregarded in this thesis.

15 CRISPRi - CRISPR Interference

CRISPR/Cas is a prokaryotic defense mechanism that comprises repeated stretches of palindromic DNA which are separated by spacer-DNA. The stretches are the result of viral attacks and in combination with a *Cas* gene, they pose an acquired immunity system against possible attacks by foreign genetic content present in e. g. plasmids and phages. Among the CRISPR/Cas supergroup and discovered in Streptococcus pyogenes, the CRISPR/Cas9 system which is based on the CRISPR associated protein 9 (Cas9), an endonuclease enzyme, has been intensively studied and described. The enzymatic or endonuclease activity i. e. cutting DNA at a specific site is directed by a pair of two partially bound i. e. annealed RNA molecules, the CRISPR RNA (crRNA) and trans-activating RNA (tracrRNA). For easier laboratory usage and skipping the annealing process, the system has been enhanced i. e. the DNA for the transcription of these two distinct RNA molecules has been changed in a way that transcription of only one RNA, the single guide RNA (sgRNA) which comprises the two former RNAs is sufficient.

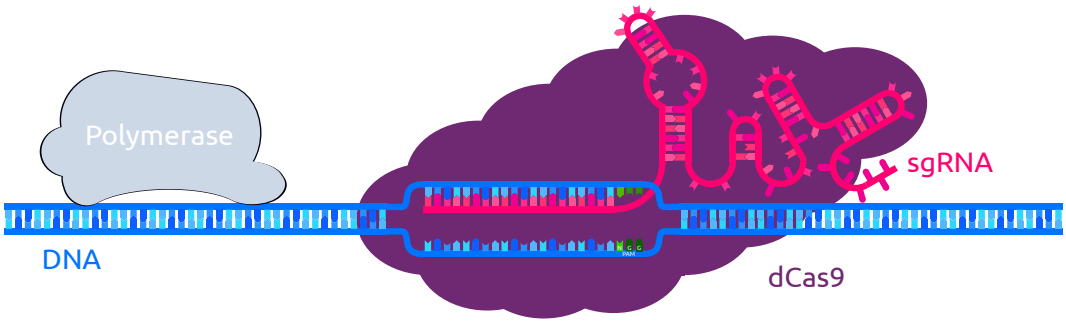


Fig. 23: CRISPRi mechanism. The dCas9-sgRNA complex binds to DNA by splitting the strand at a position that is complementary to the 20 targeting nucleotides of the sgRNA. The complex bound to the DNA prevents polymerase progression and thereby transcription of any downstream CDS. Using this mechanism, arbitrary genes can be knocked down, providing DNA of the sequence NGG, the proto adjacent motif (PAM) is present immediately downstream of the target region.

Also, a mandatory requirement for the intended DNA incision site is the presence of a protospacer adjacent motif (PAM), a DNA sequence of the structure 5'-NGG-3' where G is a placeholder for guanine and N is an arbitrary nucleic acid. The complex of Cas9 and the sgRNA can be used to cut DNA at any site providing the PAM. After the incision, the DNA repair mechanism of the host cell will rejoin the DNA. This can be used to cut out or insert DNA at certain DNA sites.

In opposition to the wild-type Cas9 enzyme, the mutated and therefore catalytically dead Cas9 (dCas9) protein lacks endonuclease activity[53]. However, it has been discovered that it is still guided by the sgRNA, but only to bind to the same specific DNA sites. If this binding region is in the area of a promoter or CDS, the dCas9 bound to the DNA will interfere with polymerase activity in the close vicinity, hence the name CRISPR Interference (CRISPRi). This method can be used to knock down genes of interest if the dCas9 protein and or the sgRNA is controlled by e. g. an inducer system. Furthermore, studies show that the sgRNA can be mutated at specific loci[54]. This renders the CRISPRi technology a potent foundation for an in vivo nucleic acid computation system.

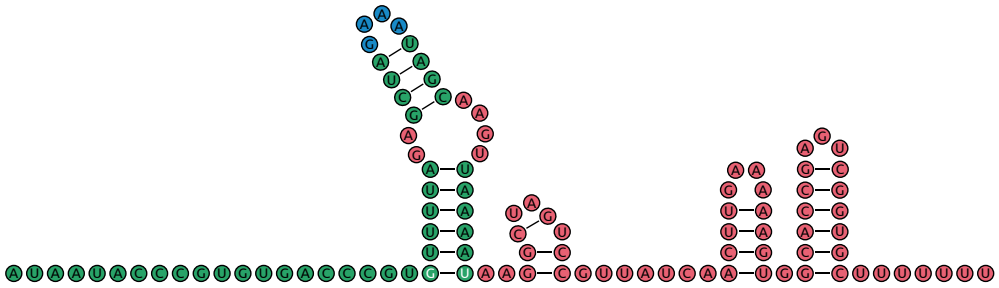


Fig. 24: Alternate illustration of sgRNA. Green background indicates RNA that can be mutated without abolishing Cas9 cleavage if baseparing complementarity is maintained, blue background marks expendable nitrogenous bases (cusp), white characters indicate wobble base pairs.

16 RNA Hybridization

In contrast to DNA, RNA is mostly abundant in its single stranded form. RNA is able to bind to other RNA molecules via hydrogen bonds and to itself. Pairing can be established by Watson-Crick base-pairing i. e. guanine binding to cytosine and adenine binding to uracil. Additionally wobble base pairs are possible where e. g. uracil binds to guanine but only two hydrogen-bonds are formed, leaving one hydrogen of the guanine vacant.

RNA hybridizes with itself i. e. folds or to other RNA molecules following the laws thermodynamics i. e. developing towards a state of RNA in composite structures in a volume where probabilistically low energy states are favored. Many properties factor into the binding energy as RNA sequence i. e. Watson-Crick base-pairing is favored over wobble base pairs, effects from neighboring base pairs have to be considered and even more, rendering an analytic calculation very difficult.

Many software packages have been developed to ease the calculation of binding energy or to evaluate the possible structures, RNA can form by (self-)hybridization. Mfold[55] is available as package or as a web application and has many options to fine-tune the calculation to the requirements of the experimenter.

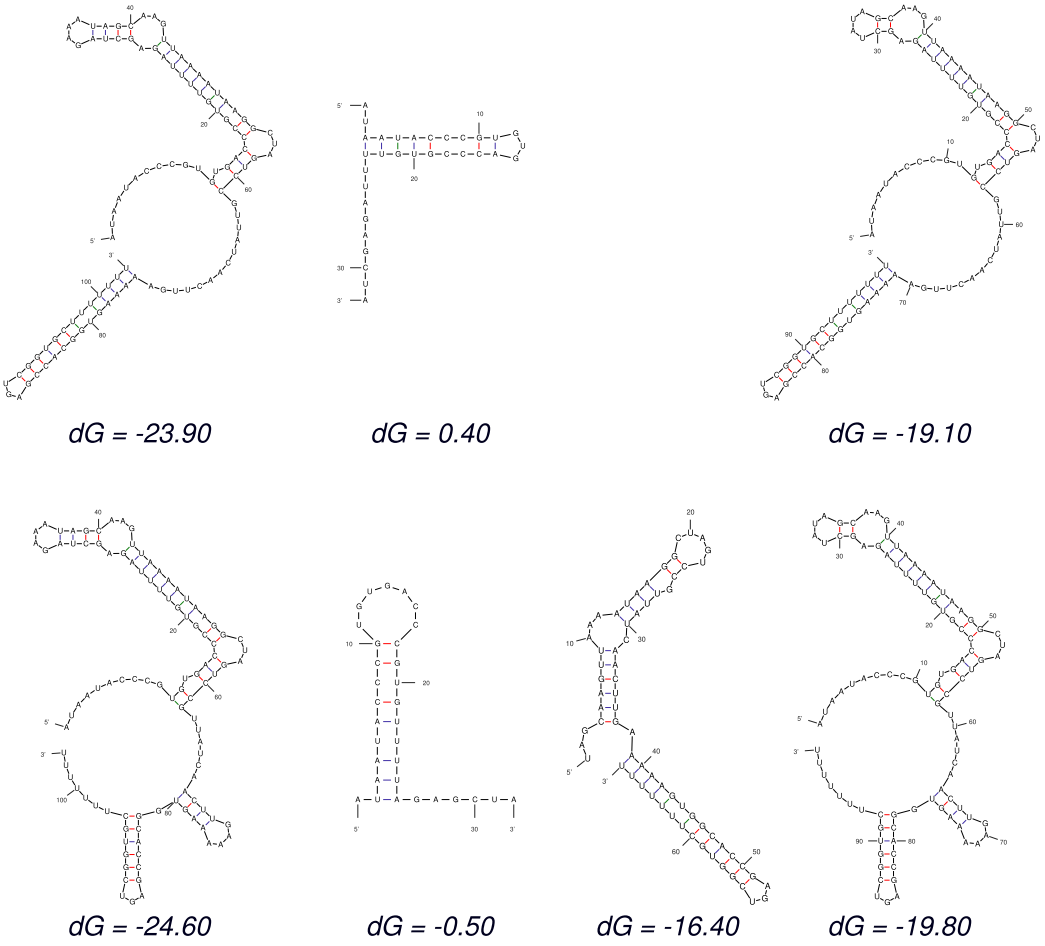


Fig. 25: Mfold structure prediction of sgRNA, crRNA, tracrRNA and crRNA-tracrRNA-complex. Left column: Two possible conformations of sgRNA. Second column: Also two possible structures for the crRNA. Third column: Only one possible conformation for the tracrRNA. Right column: Two possible structures for the hybridized crRNA and tracrRNA. Lower energies i. e. bottom structures are probabilistically favored. Free energy values are in kcal per mole. Comparing the left (sgRNA) and right (crRNA tracrRNA hybrid) column Mfold predicts that the RNA molecules will form the same structure, regardless of the separation or the removal of the cusp.

17 Revisiting Boolean Algebra

In contrast to a numeral system as the decimal, Boolean algebra describes the relationships that hold for a system only comprising two values, the binary system. These values are often denoted as true (T) and false (F) or 1 and 0. Numbers can be converted from one system into another. According to mathematical operations for the decimal, these operations also exist in the binary system. In fact, most modern computers completely perform programs on the basis of a binary system and report, mostly after conversion back in form of the decimal system.

This enables the execution of a set of logical operations in the binary system as a substitute for the corresponding mathematical operation in the decimal system. These logical operations are e. g. AND, OR, NOT, NAND, NOR or XOR.

ID	Signature	Name
00	F F F F	Contradiction
01	T F F F	Logical NOR
02	F T F F	Converse nonimplication
03	T T F F	Negation
04	F F T F	Material nonimplication
05	T F T F	Negation
06	F T T F	Exclusive disjunction (XOR)
07	T T T F	Logical NAND
08	F F F T	Logical conjunction (AND)
09	T F F T	Logical biconditional (XNOR)
10	F T F T	Projection function
11	T T F T	Material implication
12	F F T T	Projection function
13	T F T T	Converse implication
14	F T T T	Logical disjunction (OR)
15	T T T T	Tautology

Table 3: Logical Operations and Operator Names. Two inputs i. e. four possible input combinations (FF, FT, TF and TT) yield $2^4 = 16$ possible output signatures.

18 Logic Gates, functional Completeness and Networks

Two of these sixteen Boolean operations are particularly remarkable as only from each single one of them, all others can be constructed. The NOR- and NAND logic inherit this feature i. e. functional completeness and it is of vast importance to engineering and computer science.

A logic gate is a physical manifestation in approximation to the Boolean operator or the logical operation i. e. a device that attempts to mimic the behavior of a Boolean operator in the physical world. If a functionally complete logic gate can be created, by means of scalability all circuits can at least theoretically be built. In computer science and engineering, this has led to the development of the first computers since only a single element had to be designed or created, mass-produced and could be assembled into large structures afterwards. One very famous example for this principle is the first version of the Apollo Guidance Computer which was built by 4100 NOR gates in integrated circuits solely[56]. However, the larger computation networks are, the more prone to error they are. This can easily be shown:

The probability P of an entire circuit performing a correct calculation can be written as

$$P_{(n)} = p^n \tag{14}$$

with p being the probability of correct performance of a single part and n being the number of identical single parts in the network. With

$$p \in]0; 1[\tag{15}$$

and

$$n \in [1, 2, 3, ...] \tag{16}$$

It is selfevident that a circuit consisting of $n + 1$ components is larger than one consisting of n. The probability for correct calculation of the larger circuit $P_{(n+1)}$ can be deconstructed as

$$P_{(n+1)} = p^{n+1} = p^n \cdot p = P_{(n)} \cdot p \tag{17}$$

which leads to a direct comparison with the probability for correct calculation of the smaller circuit $P_{(n)}$. It can be shown, that the larger circuit has a lower probability for correct computation dividing this comparison by $P_{(n)}$,

$$P_{(n)} \cdot p < P_{(n)} \tag{18}$$

yielding

$$p < 1 \tag{19}$$

which is an initial condition, concluding the successful proof (q.e.d.). Thus computational networks should be kept as small as possible. To achieve this, it is useful to mix Logic Gates of different kinds.

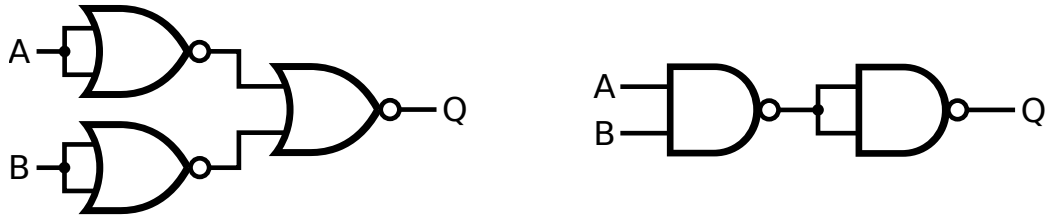


Fig. 26: Left: AND gate NOR construct. Right: AND Gate NAND construct. Given the NAND and NOR gate inherit the same probability for malperformance, a replacement of the AND gate NOR construct by NAND gates reduces the probability for error by one factor (p^3 to p^2).

Another reason to keep computation networks as small as possible is the signal propagation. The larger the network, the longer the trajectory in the carrying medium the signal propagates in, hence larger networks contribute to prolonged computation times.

19 Computation Theory

A vast amount of computation types and even more definitions for what computation exactly is exist. One of the most famous computation devices is a Turing machine i. e. a computational setup which is Turing complete. A computational network is Turing complete if it can be used to simulate a Turing machine. A Turing machine must be capable of being a pushdown automaton and combinational logic as well as being a finite-state machine. Since the NAND gate can be in one of the two finite states 0 and 1 or F and T, it fulfills this definition. Same holds by its core properties for the combinational logic as described earlier. Since from a NAND gate, memory units e. g. flip-flops can be created, it is also possible to perform pushdown automaton operations on memory arrays.

It can be shown in a more simple way that a network of NAND gates is Turing complete as follows: Since the NAND gate is functionally complete, all other logic gates can be constructed from it. From there, all feasible computational operations can be built as has been shown by the NAND gates functionally complete counterpart, the NOR gate employed in the Apollo Guidance Computer. Since the Apollo Guidance Computer is a Turing complete computer and a network of NAND gates can be used to simulate it, this network of NAND gates is Turing equivalent to the Apollo Guidance Computer and therefore also Turing complete.

20 Transfer Functions

Transfer functions are used in synthetic biology to create a functional relationship between an input to a system and the output. Prominent examples are transcription factor transfer functions i. e. the functional relationship of the inducer concentration added to the system and the result in form of e. g. the readout of a fluorescent reporter.

However, these transfer functions are very specific to the host organism they have been measured in since different hosts feature diverse metabolic systems, distributing nutritive substances in unequal ways. Even comparing transfer functions using the same host, the distribution configuration depends on many factors as cell density in the culture, temperature, nutrient availability and quality, the structure and by that time-dependent binding potential of the participating DNA and RNA, presence of disruptive factors as e. g. non-optimal pH and many more.

If additional genetic material e. g. plasmids are used in the host, the transfer functions would heavily depend on the configuration of those plasmids e. g. the amount of plasmid types in use, on their corresponding copy numbers and the number of base pairs per plasmid. Of course, plasmid reproduction is dependent on the host again, since some hosts support certain origins of replication more, some less. Subsequently, the trajectory of the transfer function is dependent on the content of the additional genetic payload. Where plasmids containing non-functional DNA would not affect the metabolism of the host much, DNA coding for heavy proteins will. In this sense, the conditional synthesis of a protein from plasmid DNA e. g. under the control of another inducer system, would also affect the transfer function.

Also, it has to be pointed out that transfer functions are measured by fluorescence readout. The main interest of this thesis is the output of crRNA and tracrRNA. Comparing the size of crRNA (32bp) or tracrRNA (67bp) with the mRNA (681bp) of the representative fluorophore mRFP1, it is evident that a factor of ten and 21 lie in between.

Additionally, to enable fluorescent readout, the translation from the mRFP1 mRNA to the protein is required. This is facilitated by ribosomes where the ri-

bosome concentration itself again is dependent on the host metabolism, hence more efficient metabolism would yield a higher rate of translation from mRNA to protein and vice versa. Also, where proteases constantly take care of degrading the mRFP1 protein, the degradation rate is not identical to this of the RNAses degrading the crRNA and tracRNA, so equilibrium between synthesis and degradation is maintained at a different level. Thus, the fluorescence readout obtained from mRFP1 can not be representative for the concentration of the two RNAs of interest.

As concluding remark, it must be considered that since environmental conditions change over time, a transfer function will not maintain its shape over the course of an experiment. This is why transfer functions should only be considered to be a rough orientation and an attempt to record or measure the transfer function of either the inducer systems in the used host and its configuration or for the output of the NAND gate in its different modes of operation can be dismissed as futile.

Experiment

1 Construction of a CRISPRi NAND Gate

Engineering functionally complete logic gates, the very foundation of computation in vivo, is a challenge in which synthetic biology research groups participate on a global scale. Prior achievements have been made e. g. creating a NOR gate by parallelizing CRISPRi i. e. targeting a promoter with dCas9 and a parallelized sgRNA under control of two distinct inducer systems.

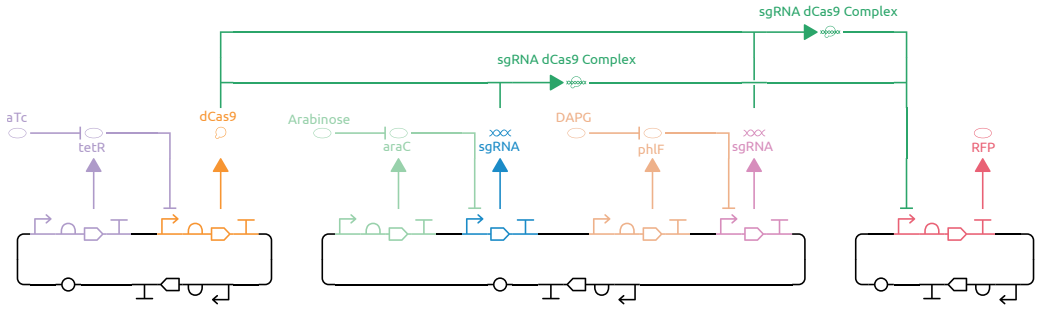


Fig. 27: NOR gate distributed on three plasmids illustrated in SBOL Visual. Left: dCas9 under control of P_{tet}+aTc with a fixed concentration of aTc which has been optimized in regards to the experimental environment (host, medium, temperature etc.) beforehand. Center: A pair of identical sgRNAs controlled by the inducer systems P_{BAD}+L-arabinose and P_{phlF}+DAPG. Right: Readout is performed through fluorescence measurement. The dCas9 molecule will interfere with the synthesis of the RFP which is representative for the next stage gate input by binding to the promoter region. It does so only if for one or both sgRNAs transcription is enabled by each of their corresponding inducer system which establishes the NOR logic. Origins of replication and antibiotic resistances shown in black on lowermost plasmid rings.

All three plasmids are available on Addgene, a company that has specialized on the storage and distribution of genetic content in the form of plasmids. Once obtained, the plasmids were amplified and purified in order to construct the NAND gate afterwards. For this they have been sent to Genscript. From the CDS under control of arabinose, starting from the blue cusp, part of the DNA and correspondlingly sgRNA has been removed (see fig. 28) by cutting at the flanking enzyme cleavage sites BsmI and SalI (not shown) and substituting the sequence in between by an accordingly reduced sequence, leaving only DNA, that codes for the first half of the sgRNA, the crRNA analogue of CRISPRi. This has also been performed for the CDS under DAPG control, making use of the flanking enzyme cleavage sites of BsiWI and SacI (also not shown). For this CDS, base pairs between and including orange and blue coloring have been removed, leaving only the DNA, that codes for the second half of the sgRNA, the tracrRNA analogue of CRISPRi.

Sections of RNA shown in dark green (see fig. 28) indicate sequences which will not affect dCas9 binding affinity if replaced by other complementary base pairs. Given that $4^{(6+4)} = 2^{20} \approx 1$ million permutations are possible for the binding region between crRNA and tracrRNA in theory and also including the vast amount of different promoter sequences, huge networks of NAND gates could be built.

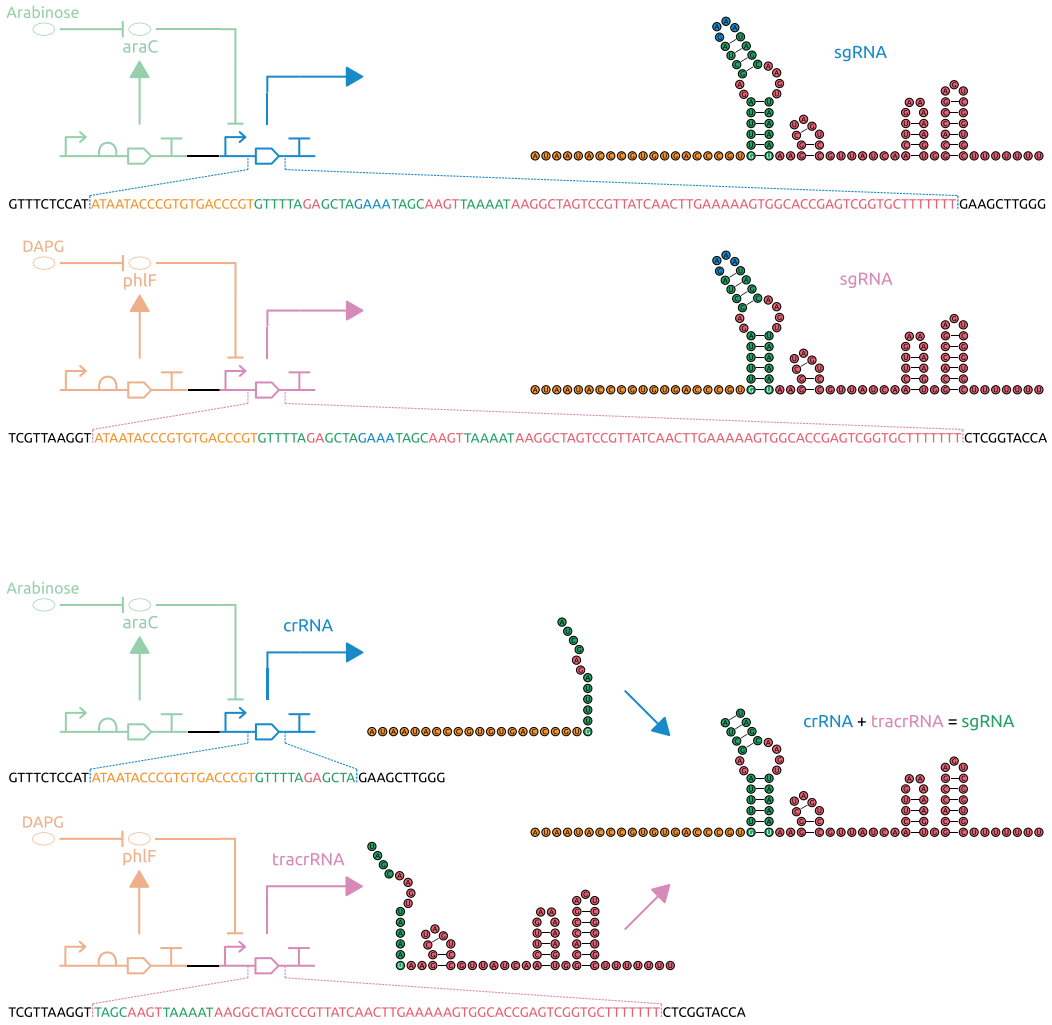


Fig. 28: NOR to NAND gate conversion. Top: NOR gate: Both the gene under control of arabinose and DAPG transcribe the same and complete sgRNA. If either is present in the host cell, dCas9 will bind and knock down a target promoter corresponding to the target sequence. Bottom: NAND gate: Both crRNA and tracrRNA are required to form the complex shown on the right which can then bind to dCas9, competent of target promoter knockdown, equivalent to the sgRNA-dCas9 complex for the NOR gate. Targeting sequence shown in orange, expendable[54] sequences for dCas9 binding in blue, mutatable[54] in dark green and base pairs mandatory[54] for dCas9 binding in salmon. Arabinose inducer system shown in light green, DAPG inducer system in light orange. SgRNA and resulting crRNA as tracrRNA genes as pathways shown in bright blue and rose respectingly. Watson-Crick base-pairing denoted in black letter, wobble base pair in white.

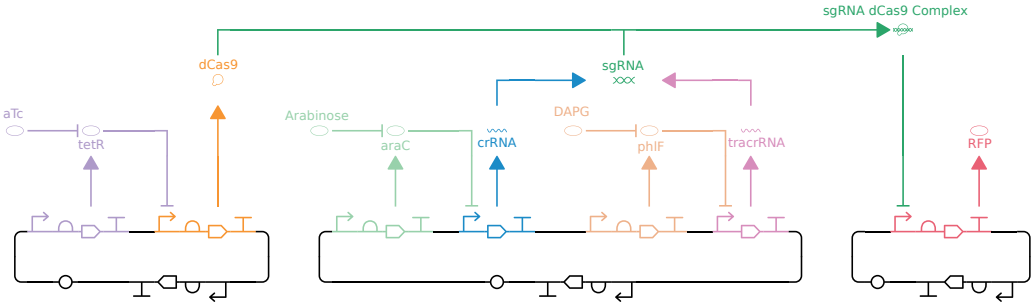


Fig. 29: NAND gate distributed on three plasmids illustrated in SBOL Visual. Left: dCas9 still under control of P_{tet}+aTc with a fixed concentration of aTc which has been optimized in regards to the experimental environment (host, medium, temperature etc.) beforehand. Center: crRNA and tracrRNA controlled by the inducer systems P_{BAD}+Ara and P_{phlF}+DAPG. Right: Readout is performed through fluorescence measurement. The dCas9 molecule will interfere with the synthesis of the RFP which is representative for the next stage gate input by binding to the promoter region. It does so only if for both crRNA and tracrRNA transcription is enabled by each of their corresponding inducer system which establishes the NAND logic.

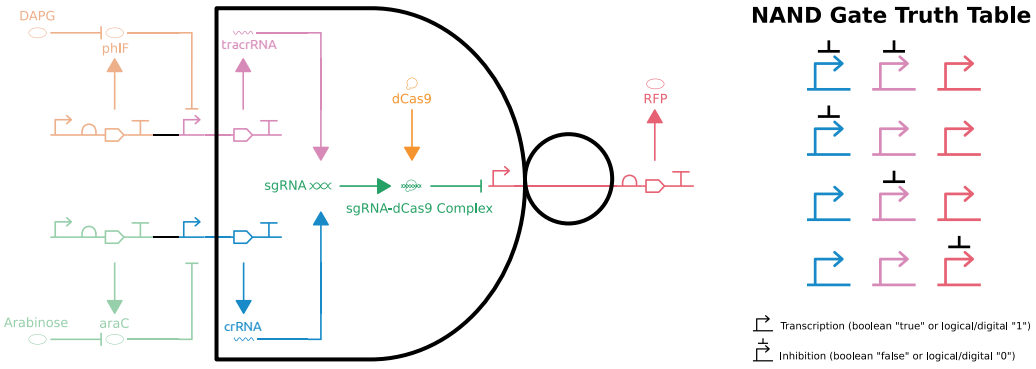


Fig. 30: NAND gate in SBOL Visual and logic/electronic illustration. Left: Promoter state (vacant or blocked) depending on presence of DAPG and arabinose fulfills boolean condition (true or false). Right: NAND gate truth table: Only if both promoters are vacant the next stage promoter is blocked.

The NAND gate described above can be accessed by the experimenter by means of inducer systems. Theses systems interfere with the metabolic network of the host and obscure the true performance of the setup as described earlier. Removing those inducer systems abolishes the control. However, where the application of fine-tuning the RNA synthesis through inducer-enabled promoter knockdown is lost, it is still possible to operate the system at digital values by removing the promoter DNA sequences according to the binary input patterns of a logic gate. The absence of a promoter would then facilitate the false-value input where the presence of a promoter would facilitate the true-value as input. These four sequence setups could be used to compare the clean binary state outputs with those of where inducer systems are still present in the plasmid, providing to baseline the fluorescence output to a level that is independent of the inducer system and derive a more accurate NAND gate transfer function. The plasmid was altered by means of molecular cloning i. e. cutting plasmid DNA at restriction enzyme sites up- and downstream of the genes responsible for the synthesis of the inducer system employed proteins and substituting with synthesized DNA where those inducer system protein genes were removed. The same was performed for the removal of the promoter sequences.

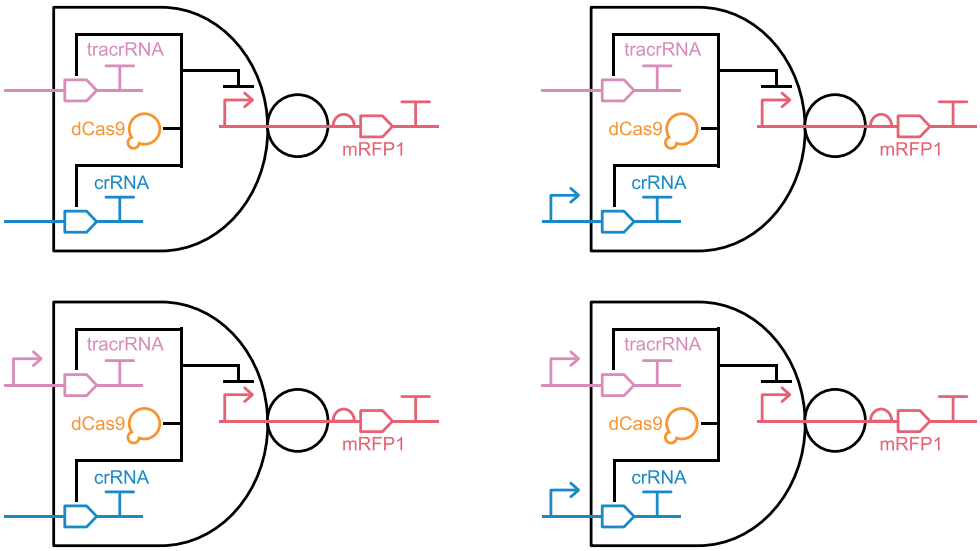


Fig. 31: Discrete NAND gate i. e. NAND gate without inducer systems. Logic input is performed by removal of promoter DNA sequences from plasmid according to input patterns. Top left: Both promoters removed representative for 00 input. Top right: Only tracrRNA promoter removed for 01 input. Bottom left: crRNA promoter removed for 10 input. Bottom right: No promoter was removed after inducer system related gene exclusion facilitating the 11 input.

To demonstrate the capability of NAND gates in a network performing more complex computation than a single gate alone, four different versions (variations in targets and RNA-RNA hybridizing sequence to facilitate orthogonality) of the NAND gate were combined into a XOR gate. This was achieved by insertion of synthesized DNA including corresponding unique and targetable promoters, NAND logic establishing combinations of crRNA and tracrRNA transcribable CDS and terminators between the described flanking enzyme cleavage sites. In this setup the output of the structure is false when the inputs are equal and vise versa (see table 3). Even though only four NAND gates are used, this small network is already capable of one-time-pad (OTP) encryption.

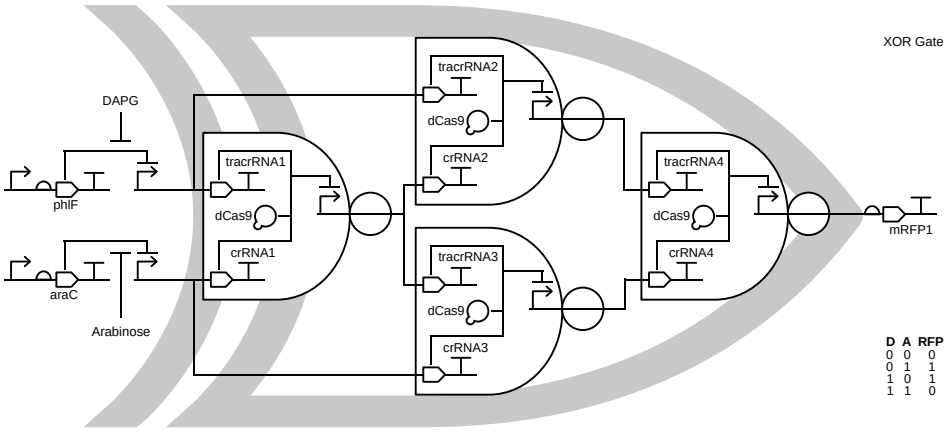


Fig. 32: XOR gate comprising four NAND gates. Target sequence affinity to promoter-sequence has been evaluated beforehand i. e. already used for the NOR gate. The circuit can be used to perform simple computations e. g. it is capable of OTP-encryption.

Turing complete machines have to be able to perform pushdown automaton operations. In order to achieve this, parallel to a computational network, an element providing a memory function is required to store information temporarily. This can be achieved by interlocking the outputs of two NAND gates into each others input to create an semi-agile system.

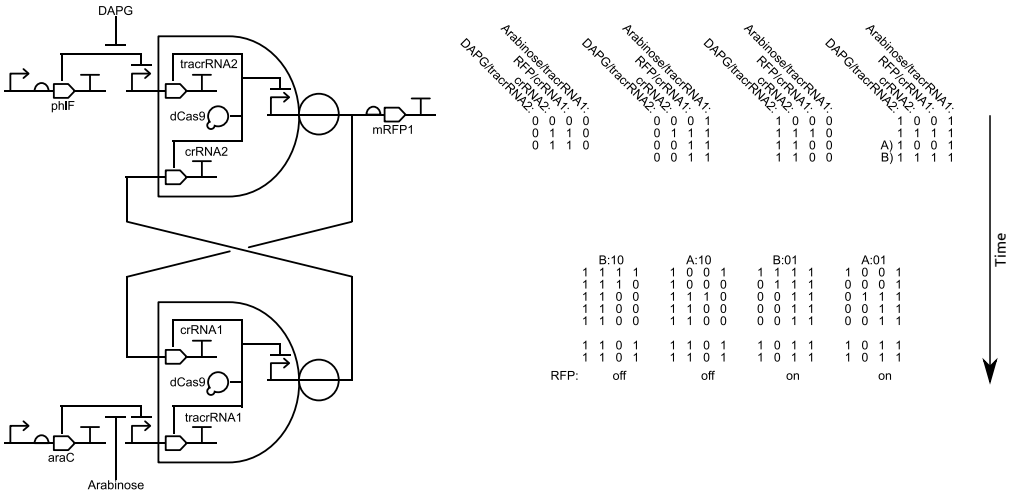


Fig. 33: Left: Flip-flop comprising two interlocked NAND gates. Right: Operation patterns: Top: Theoretical evolution (center columns) of the system-ouput for all four digital input patterns (00,01,10,11, outer columns) under control of arabinose and DAPG concentration. Where the first logic input combination leads to a stable (final and second but final state are equal) output after one turn, the second and third input pattern require two. The 11 input pattern leads to an oscillation in the output resulting in the states A) or B) (every second state is equal). Bottom: Switching input pattern from 11 to 10 or 01 for both output states B) and A). After three turns the oscillator has stabilized for all combinations where for both states A) and B) the same output is achieved depending on the input. Switching the input back to 11, the output remains, providing the capability of storing one bit.

The flip-flop can be used as memory for one information (1 bit) but also as a genetic oscillation generator. Parallelizing the up to one million possible NAND gate versions in flip-flops, theoretically 0.5 million flip-flops could be created in a single cell. This would provide a memory of $\frac{5 \cdot 10^5}{8 \cdot 10^3} = 62.5$ kilobyte (kB) per cell.

2 Inducer Growth-Impact Evaluation

Before finetuning the aTc- and corresponding dCas9-concentration in the experiment to the genetic and microbiological setup i.e. *E. coli MG1655* transformed with the three-plasmid-system, it must be ensured that the inducer-system controlling agents do not interfere with the natural growth of the cell i.e. do not interfere with its own metabolic or genetic network. This would render the results of e. g. inducing with DAPG or arabinose incomparable. For this, cell growth has been measured by a representative absorbance measurement at 600 nm wavelength over 24h in a high throughput (96 well plate) setup where 48 wells were treated with high values for concentration of DAPG and arabinose.
In literature, dCas9 was reported to be having toxic effects on cell viability[2]. To investigate this claim in the present setup but also in order to examine the impact of the induction of such a large protein on the general metabolic performace of the cell, most importantly growth, both the 48 wells which have been treated with arabinose and DAPG and the 48 which were not have been exposed to eight different values of aTc concentration.

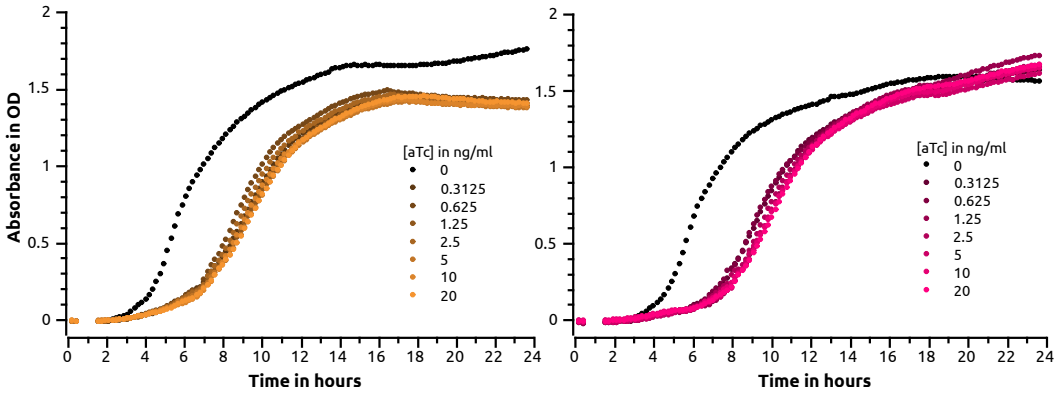


Fig. 34: Absorbance evolution in *E. coli MG1655* equipped with NAND gate plasmids for different concentrations of aTc in M9 medium incl. according antibiotics. Left: No additional inducers added. Right: 4 mM arabinose and 50 μg/ml DAPG added. Measurement performed at 37°C and 250 orbital rotations per minute. Gap is due to System Memory Error of BMG Labtech platereader FLUOstar Omega.

Figure 34 shows the absorbance evolution averaged over the signal taken from six wells. The error bars characterizing the standard deviation for each point are too close to zero to be significant and were left out in order to provide improved quality of data visualization. The variance of the fit parameters for all non-zero aTc/dCas9 concentrations is close to zero i. e. aberation between trajectories is negligible. In contrast to earlier reports, this is indicating no growth impairment nor toxicity caused by the inducer agent aTc nor the induced dCas9 protein.

Compared to the samples where dCas9 was not induced, the aTc induced growth curves are delayed in both left and right graph trajectories, but reach higher plateau level for the latter. The dCas9 protein encoded by 4107 base pairs i. e. 1369 aminoacids is a relatively large protein compared to the average protein size in bacteria[57]. In the setup the coding sequence for dCas9 was placed on a plasmid with origin of replication p15A which has an approximate copy number of ten. Where for the black trajectory dCas9 was not induced, this could have slowed down cell growth for the colored curves since the bacteria had a higher genetic payload i. e. polymerases and ribosomes were distributed to more targets and thus away from house-keeping functions to produce dCas9 by for both growth curves shown in left and right plots.

In the rightern plot the induced trajectories reach higher final values compared to the left. *E. coli MG1655* contains the ara-operon on its chromosome i. e. the prokaryote is able to metabolize the inducer as source for carbon. Additionally, part of the ara-operon, a copy of the *araC* gene is present on the NAND plasmid to ensure P_{BAD}+L-arabinose inducer system functionality. On one hand, 4.1 mg/ml being a relatively high concentration of arabinose, this could have led to improved growth conditions causing the higher absorbance levels close to the end of the measurement. Also, since RFP synthesis is decreased by CRISPRi, more nutrients are available for metabolism and growth. On the other hand where in the left graph no crRNA or tracrRNA was transcribed, in the right plot transcription of those sgRNA precursors as for dCas9 could have drawn away polymerase from binding to other genes responsible for growth. This could lead to at least a slight reduction in the evolution of absorbance. It can be hypothesized that combining the improved growth conditions by addition of arabinose and the worsened conditions by the additional transcription of the sgRNA precursors, the overall impact on growth mostly evens out, explain-

ing the only slight aberration between the non- and DAPG/arabinose-induced i. e. colored trajectories between left and right graph.

3 Initial NAND Gate Performance

To establish comparability, the first NAND gate measurement was performed with inducer concentration levels identical to those used for the precursor setup i. e. the NOR gate. Results are shown in fig. 35.

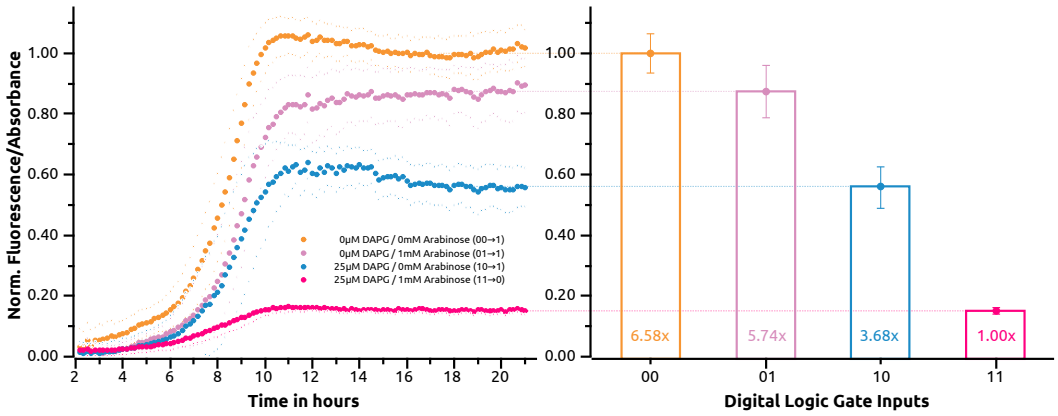


Fig. 35: Left: Normalized fluorescence per absorbance evolution in *E. coli MG1655* equipped with NAND gate plasmids for all logical inducer input combinations. Measurement was initiated at an absorbance of 0.4 OD in wells. Standard deviation illustrated with small dots. Right: Normalized final NAND gate outputs averaged over last four (17:21) hours of data. Standard deviation illustrated with error bars.

The normalized fluorescence per absorbance evolution trajectories for 00 and 11 inputs shown in fig. 19 indicate the outstanding binary quality of the outputs by both being apart more than five times to the corresponding other standard deviation. Trajectory separation occurs already almost immediately after experiment initiation. For mixed inputs however, trajectory separation is delayed and occurs about seven hours after the experiment was started. A stable equilibrium between synthesis and degradation of mRFP1 is achieved approximately 11h after the experiment started for all four outputs in parallel.

Both curves for mixed inputs are below their optimal performance ($00 \rightarrow 1$), indicating that if crRNA is induced by application of arabinose to the sample, tracrRNA is present in the system even though no induction was performed with DAPG and vice versa. The then undesired pairing of crRNA and tracrRNA would facilitate the formation of a crRNA-tracrRNA-dCas9 complex and

thus a partial reduction in mRFP1 biosynthesis which is only intended for induction with both arabinose and DAPG i.e the 11-state.

A mandatory requirement for logic gate performance is the high-grade binary quality of the NAND Gate output at a given time i.e the signals have to be in a close spread for the first three binary input combinations (00, 01, 10). Also a spacing between the outputs (0, 1) must be sufficiently large enough to ensure distinctness. The results shown in fig. 35 do not fulfill these qualities satisfactory enough, thus enhancements had to be employed.

4 Glucose Concentration Screening

Compared to P_{phlF}+DAPG, the P_{BAD}+arabinose inducer system depends on many more factors that establish proper functionality. Also, since for the initial NAND-Gate measurement the P_{BAD}+arabinose inducer system seemed to be performing worse than P_{phlF}+DAPG, it hints to performing first steps of optimization in this domain.

As described earlier, the P_{BAD}+arabinose inducer system is, among other hardly accessible factors, heavily dependent on the glucose concentration in the system. To evaluate for which glucose concentration the output of the state in which DAPG is used for induction (10) is most proximate to output of the 00 state, experiments using varying concentrations of glucose in the M9 medium have been performed to compare 10 state trajectories, hence finding an optimum glucose concentration.

Fig. 36 illustrates the results for the glucose screening. Separation of trajectories occurs approximately 9 hours after the experiment was started. For 60 mM of glucose i. e. triple the concentration in M9 medium, the 10-state was most proximate to the 00-state. This can be explained by a balanced metabolic investment of glucose into growth and synthesis of mRFP1 at this concentration configuration.

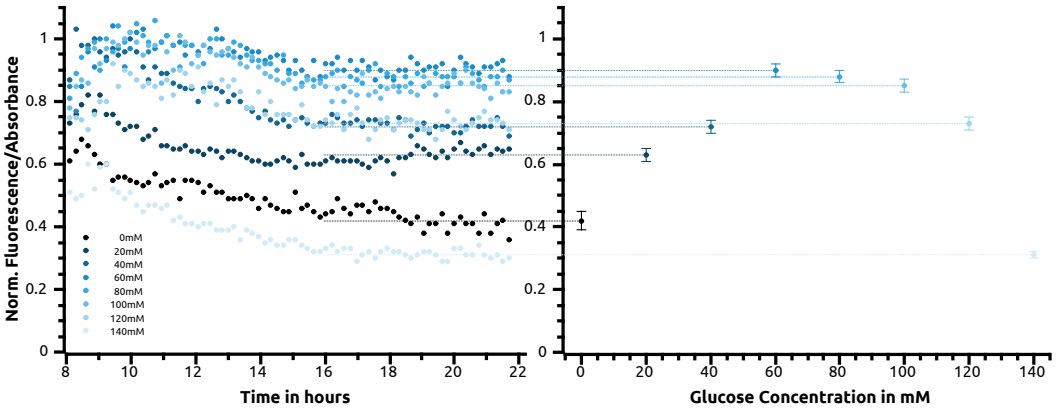


Fig. 36: Left: Normalized fluorescence per absorbance evolution of DAPG-induced state (10) with different levels of glucose in medium. Standard deviation not shown to maintain plot expressiveness. Right: Final values averaged over last six (16:22) hours of experiment. Standard deviation shown by error bars.

For a concentration of less than 60mM, lower values for fluorescence per absorbance were achieved over the course of the last six hours of the experiment. It could be reasoned that *E. coli* prioritize glucose i. e. the most prominent carbon source in the sample for means of growth, leaving less or little metabolic agents for the production of fluorophores.

Applying more than 60mM of glucose to the medium again leads to a reduction of P_{BAD}+arabinose inducer system efficiency. It could be argued that if certain levels of glucose are reached, growth of the cell or culture is favored over foreign genetic constructs, resulting in metabolic components shifted away from fluorescence to absorbance i. e. a lower value for the fraction of which.

5 Optimized NAND Gate Performance Characterization

Since for a glucose concentration of 60mM in M9 medium the 10-state of the NAND gate performed best, subsequent experiments were conducted with this concentration exclusively. Also the induction concentration for aTc was lowered to 0.3125ng/ml reducing dCas9 presence in the system in order to decrease undesired mRFP1 synthesis interference for the mixed states. Additionally the samples have been treated with inducers at an absorbance close to zero i. e. one hour after extraction from cryopreservation to ensure early reactivity to the inducers. Eventually, inducers concentrations have been increased for maximum yield.

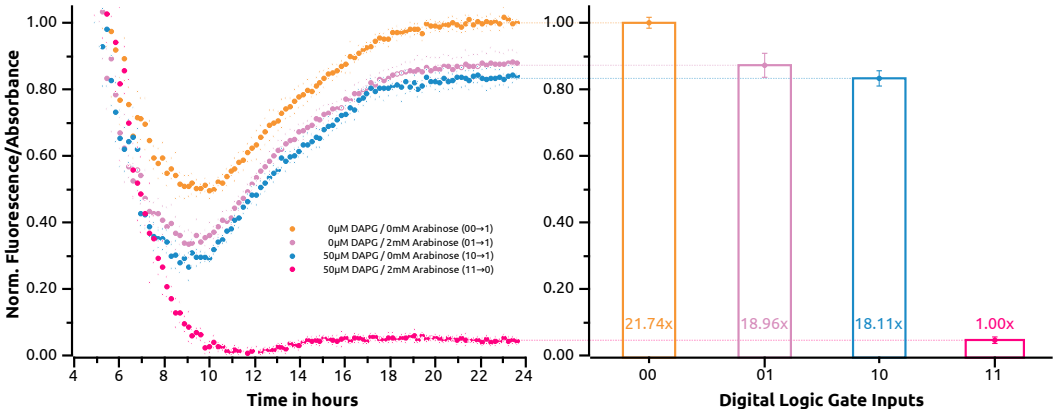


Fig. 37: Left: Normalized fluorescence per Absorbance evolution in *E. coli MG1655* equipped with NAND gate plasmids for all logical inducer input combinations. Measurement was initiated at an absorbance of 0 OD in wells. Standard deviation illustrated with small dots. Right: Normalized final NAND gate outputs averaged over last four (20:24) hours of data. Standard deviation illustrated with error bars.

Fig. 37 shows the fluorescence per absorbance evolution for the measurement of the optimized NAND gate. Compared to the initial measurement, trajectories separate later (7-8h), yet the binary quality of the final equilibrated states exceeds those of the initial NAND gate measurement by both the spread of the non-11-states states being smaller and the gap between the 11-state and others being larger.

The trajectory of both mixed states could be shifted towards the 00 state where it was also possible to improve the P_{BAD}+arabinose performance close to the level of the P_{phlF}+DAPG system. This could be credited to the reduction of aTc i. e. reduced levels of dCas9 and an increased concentration of glucose in the samples respectively. Yet still it seems that some undesired crRNA and tracrRNA remain which could be an overlapping effect from increased inducer levels.

The output of the 11-state flattens out at about 4.6% where for the initial measurement 15.2% were measured. The decrease could be explained by higher amounts of inducers used in the experiment, yielding more available crRNA and tracrRNA, resulting in a better mRFP1 synthesis interference.

The performance of the optimized NAND-Gate was evaluated to be satisfactory. To further characterize the NAND gate, measurements screening the mRFP1 synthesis interference i. e. using various combinations of arabinose and DAPG were performed. Fig. 38 shows the results of the experiment. Five density plots compiled for data extracts every five hours starting four hours after the begin of the experiment have been selected for illustration. After four hours no binary evolution can be reported yet. Cell mass is still very low so the obtained values for absorbance are close to zero, yielding in very high fluorescence per absorbance values.

A first development of a digital behavior becomes apparent after nine hours when cell mass accumulation has already contributed to a lowered fluorescence per absorbance ratio illustrated by less values for red in the squares. Compared to the three 4x4-square quadrants on the bottom and top left where only high values of one of the two inducers or low values for both have been added, the upper right quadrant content contains lower values for shades of red i. e. lowered mRFP1 presence if high values for both inducers were added. The distinction becomes more clear after 14 hours when the vast majority of the cells have entered stationary phase i. e. shifted their metabolic network from cell division to the production of e. g. mRFP1, resulting in more saturated values of red for the lesser or only partially induces quadrants, increasing in binary quality towards 19 hours.

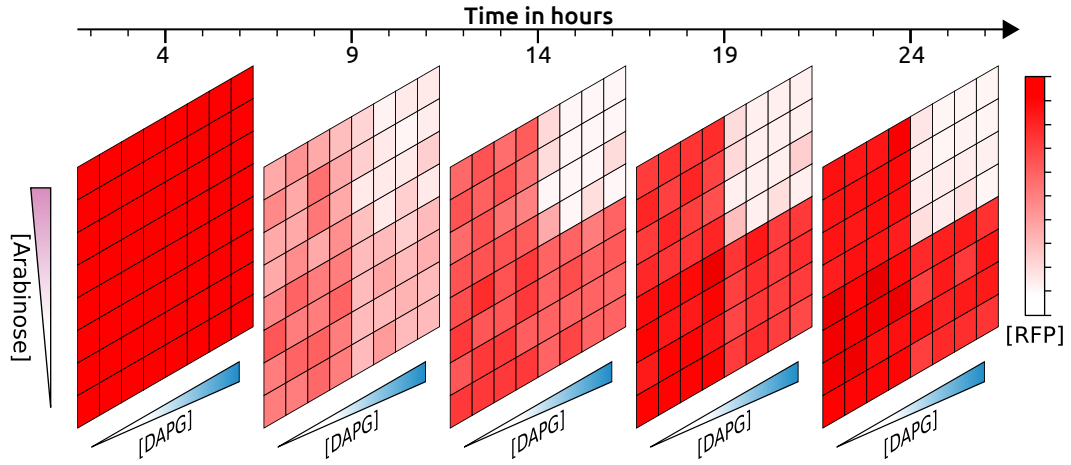


Fig. 38: Normalized NAND gate fluorescence per absorbance evolution for various inducer combinations: Bottom lines: No arabinose, top lines: 2 mM arabinose. Arabinose concentration was reduced by 1/2 to 1/2 every decreasing row number. Left columns: No DAPG, right columns: 50 μM DAPG. DAPG concentration was reduced by 2/3 to 1/3 every decreasing column number. Normalized mRFP1 fluorescence per absorbance readout in shades of red. Corner squares are equivalent to measurements performed in prior experiment.

After 24 hours close to perfect binary behavior is satisfied by the NAND gate setup in the present experiment. The corresponding concentrations of DAPG and arabinose i. e. the required dilution series to create a transition from mRFP1 expression to knockdown exactly between the four quadrants have been obtained in previous experiments.

Discussion & Outlook

1 Discussion

The presented NAND gate has shown close to perfect performance in the conducted experiments which motivates its employment in circuits and corresponding experiments. However, still many things are open for optimization.

One major setback is that no host organism without the ara-operon was available, especially the *araD* gene since the function of the protein produced from it is to further process arabinose and therefore change the arabinose concentration in the system over time. This might have had some influence on the inducer system reliability. Also, removal of the *araD* gene should further lower the 0-output state for the 11-input and thereby enhance NAND-Gate performance.

In the experiments, laboratory equipment was shared with other experimenters and also organized by such. This led to the usage of different equipment for experiments that should have been comparable later. One such piece of equipment is the adhesive cover for the microwell plates which were used. Since different vendors and products were used over time, it cannot be assumed that e.g. the exchange of gases most importantly oxygen through the barrier was sufficiently equal, compromising comparability by altered environmental parameters. Depending on the foils permeability for certain gases, *E. coli* could have switched from aerobic to anaerobic metabolism sooner or later, rendering experiments not conducted with the same foil incomparable.

In general, the growth of bacteria can only occur through processing ingredients of the growth media and by that altering the absorbance of such by changing the concentration of its ingredients and also creating new components as output caused by growth. Also, during experiments in microbiology additional agents are used in growth media in order to regulate bacterial behavior. It is a rare property of those agents, that they cannot be consumed by the bacteria, hence also the concentration of these agents changes over the time of an experiment also causing additional output products that alter the total background absorbance over time. An absorbance measurement can only be taken into consideration if the change of absorbance contributed by those effects can be neglected. This lack of scientific accuracy can at least be overcome by the usage of a chemostat i. e. a device that supports continuous flow of fresh medium through a bacterial population.

In its very definition, absorbance is clearly stated to exclude scattering effects on cell walls[4], but when measuring absorbance in a bacterial population, scattering effects do occur, not only on the cell wall but also on the interior content of bacterial cells. This property is well known and even used to compare cells concerning their size and interior complexity in e.g. flow cytometry. However, still it is very common that absorbance is used directly as value for the density of a bacterial population in microbiology i. e. the number of cells involved. It could be argued that since the average interior features, structure and size of bacteria in a culture have a derivation which is sufficiently small enough, a measurement for absorbance could still be used to at least compare the amount of cells of a certain type in a culture by relating the absorbance values as it is done in flow cytometry. Also by using correcting factors that account for the size and shape of the cell wall as well as the interior of the cells, the value for absorbance could be used to obtain the number of cells, providing these factors are known for the used type of cell (culture). Without that, absorbance at 600nm alone can not be used as a direct representation for the number of cells in a culture in general.
In fact, the physical term absorbance is misused by the scientific community in this context. Cells in a suspension mostly scatter light at a wavelength of 600nm and, since they are almost transparent, do not absorb much light at 600nm except if they are equipped with e.g. corresponding pigmentation. The term absorbance is used regardless of the true physical process which is the scattering of light and thereby a reduction of signal in the measuring device.

As a scaffold for the NAND gate, the already existing NOR gate has been used most notably because it has already been balanced in terms of promoter-strength, plasmid-copy-number, terminator-strength and so on. The performance of the NAND gate could be further improved by fine-tuning these parameters to the new NAND gate setup or even shifting the genetic payload from three plasmids to one or even into the hosts chromosome. If circuits are constructed from the NAND gate, this optimal parameter combination has to be evaluated for every kind of circuit.

Another feature of the given NOR gate scaffold are the predetermined inducer systems. Many more inducer systems exist and some of them might feature properties which match better to the now present NAND gate setup.

Even for very high concentrations of dCas9 in a cell, a perfect knockdown of the target gene can not be achieved. DNA polymerase has to process the DNA for plasmid duplication. If dCas9 would bind to the target DNA without interruption, DNA polymerase could not copy the plasmid since an ever present dCas9 would block the path, yielding in a reduction in plasmid numbers over *E. coli* generations to a subsequent loss of the plasmid. This would also mean the deprivation of the antibiotic resistance for the cell, hence lethal conditions in a medium with properly set up antibiotics.

In the performed experiments, another feature of the given NOR gate scaffold was the mRFP1 under control of an already present and targetable promoter. For better fluorescence yield, faster synthesis and even degradation or in general a better resolution in signal strength as in evolution, the usage of another fluorophore might have been more beneficial even though it would have included further steps for building the setup.

Another enhancement for not only the fluorescence synthesis but also for the inducer system participating proteins and dCas9 would be the usage of

a RBS with higher ribosome binding affinity. This, shifting away metabolic performance from growth to foreign protein synthesis could both increase the binary quality of the fluorescence signal and reduce growth, yielding an improved overall digital quality of the fluorescence per absorbance signal.

The experimental environment, the host organism and also the NAND gate itself are dependent on many factors such as concentration of glucose, arabinose, DAPG and aTc, the growth medium, shaking speed of the sample as well as its dimensions, temperature and oxygen supply, illumination, host organism, plasmid sizes, origins of replication, promoter strengths, the hosts optimization for expression of certain proteins or plasmid copy numbers. The attempt to find optimal parameters in such a vast parameter space is impossible due to the curse of dimensionality and only an attempt of improvement can only be performed by deductive measures as has been done as described above.

2 Comparison with other in vivo Logic Gates

2.1 The Ellington- & Yin-Lab AND Gate

One particular setup for the implementation of an in vivo logic gate network is the employment of multiple orthogonal types of polymerase that only transcribe after binding to corresponding promoter DNA sequences[58]. This transcription would be performed from a CDS coding for two possible distinct kinds of RNA sequences. One of the two structures of RNA would be in the form of a self-hybridized hairpin e.g. toehold switch, concealing a RBS and thus making it inaccessible for ribosome binding. The other structure would be the counterpart, being able to open up the hairpin or toehold-structure of the former RNA since a coupled and thereby opened up structure would be favored in terms of binding energy or thermodynamics. This would also excavate the RBS, unlocking translation from a transcribed downstream CDS, coding for the next-stage kind of polymerase. Since translation can only take place if both RNA strands are present in the system, this setup satisfies the AND logic.

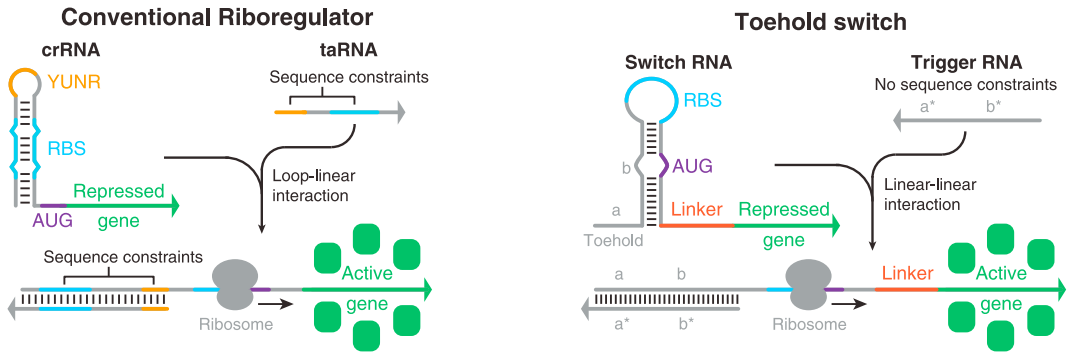


Fig. 39: Two possible AND-Gate manifestations. Left: A taRNA binds to a crRNA exposing the beforehand bound RBS hence activating gene transcription in the conventional riboregulator. Right: A Trigger RNA binds to a Switch RNA exposing the RBS and start-codon(AUG) which were bound and inaccessible due to the loop formation of the RNA correspondingly[59]. The gene coding for GFP could be replaced by coding for one specific T7-polymerase mutant, establishing next-level circuitry input allocation.

Comparing this AND gate with the NAND gate, the major drawback is that the AND Gate is not functionally complete i. e. not all circuits can be built from it. Also, the size of an AND gate network is limited by its theoretically achievable maximum size. Since only up to six orthogonal polymerase promoter pairs are available, this limits the achievable network size to six nodes. However, in vivo network sizes of six nodes will hardly be possible to implement since the employment of multiple kinds of polymerases will redirect metabolic capability away from housekeeping genes and thus the survival probability of the host cells will be increasingly compromised with growing number of parallel types of polymerases per cell.

2.2 The Buck Lab AND Gate

Similar to Ellington, the Buck laboratory also has engineered an AND Gate[60]. Instead of being based on RNA interactions or hybridization, it facilitates the AND logic by protein interaction. Identical to the former AND Gate, this also does not fulfill the property of being functionally complete. Protein interaction is required, rendering a possible network size of employed gates rather small,

i. e. if no parallelized orthogonalization is possible by e.g. protein mutation, the network size would be restricted to a single gate. Also, since protein interaction is utilized, certain amounts of metabolic payload will be redirected away from housekeeping genes.

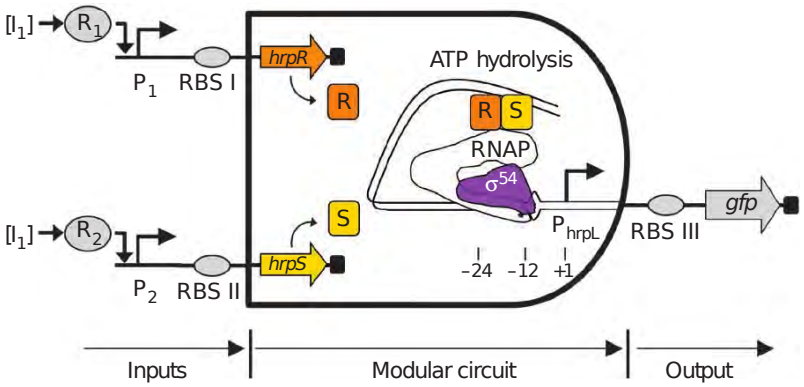


Fig. 40: Schematic display of the Buck Lab AND Gate. The inputs are carried out by the presence two different kinds of proteins which together enable RNA polymerase activity.

2.3 The Voigt Lab NOR Gate

Since the NAND gate was constructed from the NOR gate, a direct comparison is straightforward. The stated performance for the NOR gate could not be reproduced in experiment, however, an adequately prepared host organism i. e. a model organism with matching genetic layout could not be acquired. Also, the experimental setup could not be recreated because of the lack of identical instruments. Under identical experimental conditions, the performance of the NOR gate was increased in terms of the spread of the 01, 10 and 11 input i. e. in these configurations the fluorescence output was lower than the one of the NAND gate. However, the digital quality i. e. the distance of the 0 and 1 output signal of the NOR gate could not reach the one of the NAND gate, most probably since undesired translation directly leads to a knockdown for the NOR gate where for a knockdown with the NAND gate two undesired translations have to be present in parallel.

Since both the NOR and the NAND gate work using the same principle of rendering a promoter inaccessible by binding of dCas9, they can be employed

in combination. This cooperation even further offers the possibility of reducing network size by the joint operation of two functionally complete logic gates.

3 Outlook & Possible Applications

Since the CRISPR/Cas9 system has already been applied to even human cells, the NAND gate could also be integrated in such, serving as a regulatory element in the complex network of host genes which are not only primates but also less complex life forms when it comes to the genetic composition as well as model organisms which could be used in industrial processes.

The introduced structure of crRNA only includes two mandatory nucleic acids (guanine and adenine at position 27 & 28 respectively) which must not be mutated if RNA hybridization and therefore subsequent target gene knockdown is favored, but all other base pairs can. Making use of this property, almost arbitrary sequences of RNA, most notably mRNA from a host organism could be used as substitute as long as it features the two base pairs guanine and adenine. This property renders the NAND gate a potent sensor for intracellular processes and beyond. One application would be to use mRNA that is linked to quorum sensing as such an input i.e. triggering downstream genetic programs depending on bacterial density.
Using the CRISPR-Cas13a/C2c2 system, it has been shown that adopting this concept, the tracrRNA can be constructed to feature as counterpart to various bacterial or viral RNA sequences[61] that act as the crRNA. This could be used as a multiplex diagnostics principle[62] for a huge variety of bacteria and viruses in e.g. a sample of saliva as it has been demonstrated by the joint iGEM Team of LMU and TUM in 2017.

The tracrRNA or crRNA output of a NAND gate or networks of such could of course be replaced by mRNA of a specific organism and thus trigger downstream metabolic reactions as synthesis of proteins which themselves are in the focus of interest or also metabolites, where said proteins play a key-role in the production-pathway. Lycopene is such a metabolite and a NAND gate could be used to fine-tune the intermediate steps or precursor concentrations in the production of the metabolite by a sophisticated feedback-loop.

Eventually, a network constructed solely by or containing NAND gates could replace the seemingly chaotic structure of genetic networks brought up by evolution and serve the purpose of a cutting-edge tool to enable circuit-like treatment i. e. modularity and programmability of living matter and even be used to create artificial life before long.

Of course, the NAND gate could be used to create networks of its own, featuring valuable computational properties.

3.1 Flip-flop as a Memory-Unit Utilization

Using two distinct NAND gates as foundation, a flip-flop or SR NAND latch has been built which is yet to be characterized. Flip-flops are bistable multivibrators and can be used to store information. The ability to store information in an e.g. register or array is mandatory for a computational device to be Turing-complete.

The creation of addressable memory by construction of a flip-flop from NOR or NAND gates in bacteria is an essential step towards performing complex computation in living tissue before long. However, where addressing can easily be performed intracellularly using the earlier described vast variety of versions of complementary pairs of DNA and RNA in the promoter- and sgRNA-sequence respectively, for intercellular exchange addressability on a cellular level still remains to be achieved. Unlocking this, not only the complexity of DNA or RNA could be used to process and store information but these tasks could then also be distributed for parallel computation, memory function or performance in general. Even though computation on a genetic or chemical level can not compare to the speed of computation achieved in electronic or quantum computation by far, this parallelization could unlock even higher floating operations per second after all depending on the quality of addressability, culture size, error-rate and resulting error-compensating e.g. algorithms or features.

3.2 XOR-Gate and OTP Encryption

Using four distinct NAND gates as foundation, also a XOR gate has been constructed, also yet to be characterized. A XOR gate is capable of OTP encryption since its output-signal is ”1” only if the inputs differ. This property

can be used to both encrypt a secret - and then decrypt the sent message with a pre-shared key again. The OTP can not be cracked but requires a pre-shared key that can only be used once. Optionally, this key can be exchanged e.g. through Diffie–Hellman key exchange.

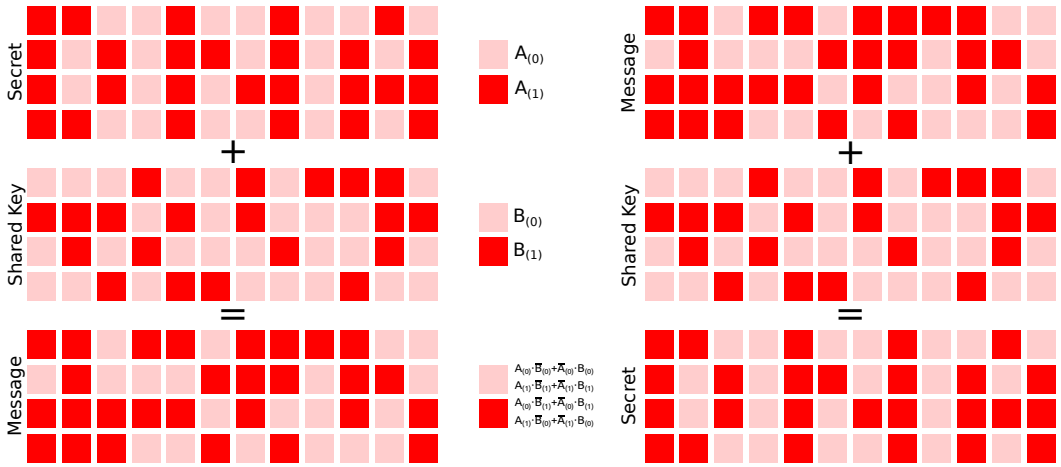


Fig. 41: Left: Encryption: Top: Suggested secret pattern of inducer arabinose ($A_{(1)}$) in e.g. a microplate filled with 4x12 *E. coli* cultures. Note the letters "DNA" in strong red with light red ($A_{(0)}$ - no arabinose added) as background. Center: Suggested secret shared key pattern of inducer DAPG ($B_{(1)}$), subsequently added to arabinose inducer pattern in top box. Bottom: Resulting encrypted message fluorescence pattern computed by XOR gate. Note that the information "DNA" can not be extracted from bottom box anymore. Right: Decryption: Encrypted message pattern of inducer arabinose ($A_{(1)}$) in e.g. a microplate filled with 4x12 *E. coli* cultures. Center: Secret shared key pattern of inducer DAPG ($B_{(1)}$), subsequently added to arabinose inducer pattern in top box. Bottom: Resulting decrypted secret fluorescence pattern computed by XOR gate. Note that the information "DNA" can now be extracted again.

The ability to process information in an algorithm that is more complex than the computational operation of a single logic gate alone is a seminal feature for a network to qualify for the property of being Turing complete.

4 Concluding Remarks

The presented NAND gate is functionally complete, making it possible to construct all desired logical operations from it and thereby enabling all feasible computation by the application of a modular core element. This might even include the realization of a Turing complete network. The system is compatible with already existing logic gates, e.g. the NOR gate. Its inputs can be configured to act as a sensor for many cellular properties where its output can lead to triggering all genetic programs thinkable. CRISPR/Cas9 is compatible with a vast variety of species, as it has been shown in the past years where it can also even be applied in cell-extract. Except to its functionally complete NOR logic counterpart, the NAND Gate by its design is superior to all in vivo logic gates known so far. These properties render it a perfect centerpiece for a universal nucleic acid computation system.

Appendix

1 Methods & Materials

1.1 Experimental Prearrangement

1.1.1 Cultivation or Incubation

Cultivation was performed in conical centrifuge tubes of varying vendors which were placed in an incubator that was set to shake cultures at 250 rotations per minute and to keep temperature at 37°C. After extraction from cryopreservation, bacteria were immersed in 5 ml LB medium, supplied with the according antibiotic(s). Optimal incubation conditions as described could not be maintained over the whole process of cultivation since outside temperatures of the lab exceeded 37°C on some occasions. Also, since the incubator was shared by multiple experimenters, opening and closing operations interrupted continuous shaking and distorted temperature equilibrium as well as the front door not being closed entirely or left widely open.

1.1.2 Cryopreservation

After transformation, a colony of *E. coli* was selected and transferred into a growth medium for twelve to 18 hours. After adding 500 μl culture liquid, 250 μl LB-medium and 250 μl glycerol to a cryogenic vial, the content was mixed by gently inverting the vial several times. The samples were frozen at -80°C by directly placing them in an appropriate freezer or by immersing them into liquid nitrogen first.
For reactivation, the frozen content of the interior of the corresponding cryogenic vial was scratched with a pipette tip which then was immersed into a growth medium to which the appropriate antibiotics had been added beforehand. The entire vial was never completely thawed and repeated freeze-thaw-cycles were avoided[63].

1.1.3 Transformation

Transformation has exclusively been performed by electroporation. For that, competent *E. coli* i.e. *E. coli* that have been thoroughly rinsed from growth medium and immersed into water were extracted from cryopreservation and thawed on ice. Samples of amplified plasmids were extracted from freezer at -20°C and also thawed on ice. At least two hours before transformation, SOC medium was produced and stored in an incubator. Transformation cuvettes were chilled to -20°C. Shortly before transformation, plasmids were introduced into the competent cell samples at varying concentrations.
For the transformation process, the mixture of competent cells and plasmids was placed into a transformation cuvette after removing it from the freezer immediately before. The cuvette was then placed into the electroporation device and a voltage of 2000V applied for 5 ms. Immediately after that, the content of the cuvette was carefully washed out with 0.3 ml SOC medium which was then placed in an electroporation tube, stored in an incubator.
After two hours, the liquids were distributed onto solid agar containing the corresponding antibiotics in petri dishes and stored in an incubator without shaking. After 14 h to 24 h, if *E. coli* had grown sufficiently to form visible colonies, pipette tips were contaminated by picking into said colonies and then suspended into LB medium containing corresponding antibiotics. If bacterial growth could be verified, from there cryostocks were created and if applicable, experiments were conducted. Petri dishes were prepared at least 4 h prior to usage. For this, solid agar containing LB-medium has been carefully heated in a microwave oven until becoming entirely fluid. As soon as the containing bottle could be touched by hand, required antibiotics were added. The fluid was vortexed and 15 ml each distributed to petri dishes after the foam diminished via an electric pipette.

1.1.4 Amplification

After extraction from cryopreservation, *E. coli* were cultivated in 5 ml LB Medium containing the corresponding antibiotics, placed in an incubator.

After 12 h to 16 h, cell cultures were further processed according to QIA Miniprep[64] yielding samples with high plasmid concentration.

1.1.5 Cloning

For cloning, samples containing amplified plasmids were sent to Genscript. Genetic synthesis and cloning i.e. establishing the genetic constructs required for experiments were discussed with employees who also performed the required working steps. After cloning and amplification, plasmids were sent to the experimenter in liphophilized form where they were transformed into *E. coli* for means of amplification and experimental usage.

1.1.6 Fluid Management

For all experiments as preparation, *E. coli MG1655* equipped with all three plasmids have been extracted from cryopreservation, cultivated in a conical centrifuge tube filled with 5 ml LB medium and corresponding antibiotics (carbenicillin, streptommycin and kanamycin). For cultivation the tube was placed inside an incubator, keeping the cultures at 37°C and shaking at 250 rotations per minute. All measurements were performed in M9 medium that had been prepared at least two hours prior to the experiment. As first step, (aTc,) kanamycin, carbenicillin and streptomycin were added to 32 ml of M9. Corresponding to the experiment, the volume was then split up into parts where the lowest number of different concentrations comparing all inducers was used. If applicable, serial dilution was then applied for the remaining inducers. As final fluid management step, LB medium containing bacteria was extracted from the initial culture and 20 μl applied to the medium in all inducer-combinations required by the experiment. The now completely prepared samples were then distributed to six or eight wells, depending on the experiment. All fluid management was performed with non-calibrated Eppendorf pipettes and in-house autoclaved Eppendorf pipette tips.

1.1.7 Absorbance and Fluorescence Measurement

During experiments, temperature was held steady at 37°C. Cultures were shaken at 250 rotations per minute. Ibidi 96 well plates were used for the experiment, which feature a transparent bottom layer. After applying 0.3 ml of *E. coli* containing sample liquids to the wells, the plate top was sealed with

varying transparent adhesive labels between experiments, assuring no liquid would contaminate the top. Light emanating from a xenon arc lamp was directed through a bandpass filter using fiberoptic cable after which it was reduced to illuminate the fluorophores in the wells at 584nm(FWHM 20nm). Using the same principle, the wells were illuminated at 600nm for absorbance measurement. Photon flux at 600nm and 620nm(FWHM 20nm) was recorded during illumination for recording absorbance and fluorescence values correspondingly in chronologically separated measurements.

1.2 Media & Agents

1.2.1 List of Media

- M9 minimal (M9) Medium
  - ddH_2O (sterile ultra-pure water)
  - 33.7 mM Na_2HPO_4
  - 22.0 mM KH_2PO_4
  - 8.55 mM NaCl
  - 9.35 mM NH_4Cl
  - 0.4 % Glucose
  - 1 mM $MgSO_4$
  - 0.3 mM $CaCl_2$
  - 1 μ g Biotin
  - 1 μg Thiamin
  - 134 μM EDTA
  - 31 μM $FeCl_3$-6H_2O
  - 6.2 μM $ZnCl_2$
  - 0.76 μM $CuCl_2$-2H_2O
  - 0.42 μM $CoCl_2$-2H_2O
  - 1.62 μM H_3BO_3
  - 81 nM $MnCl_2$-4H_2O

- Lysogeny broth (LB) Medium
  - ddH_2O
  - 10 g/l tryptone
  - 5 g/l yeast extract
  - 10 g/l NaCl
- Super Optimal Broth (SOB) Medium
  - ddH_2O
  - 0,5 % (w/v) yeast extract
  - 2 % (w/v) (20 g/l) Tryptone
  - 10 mM (0,6 g/l) sodium chloride
  - 2,5 mM (0,2 g/l) potassium chloride
  - 10 mM magnesium chloride
  - 10 mM magnesium sulfate
- Super Optimal broth with Catabolite repression (SOC) Medium.
  - SOB Medium
  - 20 mM Glucose

1.2.2 List of Agents

- L-Arabinose (Sigma-Aldrich) in ddH_2O
- 2,4-Diacetylphloroglucinol (DAPG) in ethanol
- Anhydrotetracycline (ATc) in ddH_2O
- Kanamycin (Kan) in ddH_2O
- Carbenicillin in ethanol
- Streptomycin in ddH_2O
- D-Glucose (Sigma-Aldrich) in ddH_2O

1.3 Data Evaluation

1.3.1 Digital

Data from the platereader was evaluated as such that all values of fluorescence and absorbance for wells containing bacteria were decreased by an average value which had been determined in wells containing uncontaminated growth medium, thereby "baselining" those signals. For each single value, fluorescence was divided by absorbance. Then the average of six wells was calculated. These values were normalized by relating them to the height of a linear fit parallel to the time axis, where NAND gate inputs were non present i.e. no inducers were added to the system except aTc.

1.3.2 Visual

The HSL (hue, saturation, lightness) value for the color red is 0,255,128 where this can be shaded to white (0,255,255) by increasing lightness to 255. Normalized values for fluorescence divided by absorbance $(F/A)_n$ were rescaled into shades of red by factoring their value with 127 (=255-128) and subtracting it from 255 according to:

$$R_{(F/A)_n} = 255 - (F/A)_n \cdot 127$$

The resulting values for shades of red $R_{(F/A)_n}$ representative for normalized fluorescence per absorbance were then used in fig. 38.

1.4 DNA Sequences

1.4.1 NOR Gate Plasmid - pAN-NOR

GCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG
TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAggccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg
cgtttcccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggt
atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagaca
cgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggaca
gtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaagcagcaga
ttacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatga GATTATCAA
AAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACT
TGGTCTGACAGTTA ccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacg
ggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtc
ctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtca
cgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcag
aagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattct
gagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcat CATTGGAAAACGTTCT
TCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGG
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGT
TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCC
CCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCAGTTGATAAGTCCCTAACTTTT
ACAGCTAGCTCAGTCCTAGGTATTATGCTAGCCTGAAGTACGTCTGAGCGTGATACCCGCTCACTGAAGATGGC
CCGGTAGGGCCGAAACGTACCTCTACAAATAATTTTGTTTAACTATGGACTATGTTTGAAAGGGAGAAATACTA
G ATGGCACGTACCCCGAGCCGTAGCAGCATTGGTAGCCTGCGTAGTCCGCATACCCATAAAGCAATTCTGACC
AGCACCATTGAAATCCTGAAAGAATGTGGTTATAGCGGTCTGAGCATTGAAAGCGTTGCACGTCGTGCCGGTG
CAAGCAAACCGACCATTTATCGTTGGTGGACCAATAAAGCAGCACTGATTGCCGAAGTGTATGAAAATGAAAGC
GAACAGGTGCGTAAATTTCCGGATCTGGGTAGCTTTAAAGCCGATCTGGATTTTCTGCTGCGTAATCTGTGGAA
AGTTTGGCGTGAAACCATTTGTGGTGAAGCATTTCGTTGTGTTATTGCAGAAGCACAGCTGGACCCTGCAACCC
TGACCCAGCTGAAAGATCAGTTTATGGAACGTCGTCGTGAGATGCCGAAAAAACTGGTTGAAAATGCCATTAGC
AATGGTGAACTGCCGAAAGATACCAATCGTGAACTGCTGCTGGATATGATTTTTGGTTTTTGTTGGTATCGCCT
GCTGACCGAACAGCTGACCGTTGAACAGGATATTGAAGAATTTACCTTCCTGCTGATTAATGGTGTTTGTCCGG
GTACACAGCGTTAA CTAGGGCCCATACCC CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCC
TTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTT
CTGCGTTTATA TACG cgacgtacggtggaatctgattcgttaccaattgacatgatacgaaacgtaccgtatcgttaaggt ATAATACCCGTGTGAC
CCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGT

CGGTGCTTTTTTT ctcggtaccaaagacgaacaataagacgctgaaaagcgtctttttcgttttggtcc GGAG CTCGGTACCAAATTCCAGAA
AAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTTGGTCC TACG TTCAGCCAAAAAACTTAAGACCGCCGGT
CTTGTCCACTACCTTGCAGTAATGCGGTGGACAGGATCGGCGGTTTTCTTTTCTCTTCTCAA GAATTCGCGGCC
GCTTCTAGAG TTATGACAACTTGACGGCTACATCATTCACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGG
CTGGCCCCGGTGCATTTTTTAAATACCCGCGAGAAATAGAGTTGATCGTCAAAACCAACATTGCGACCGACGGT
GGCGATAGGCATCCGGGTGGTGCTCAAAAGCAGCTTCGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAAG
ACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGACAAGCAAACATGCTGTGCGACGC
TGGCGATATCAAAATTGCTGTCTGCCAGGTGATCGCTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATC
GGTGGATGGAGCGACTCGTTAATCGCTTCCATGCGCCGCAGTAACAATTGCTCAAGCAGATTTATCGCCAGCAG
CTCCGAATAGCGCCCTTCCCCTTGCCCGGCGTTAATGATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCG
CTTCATCCGGGCGAAAGAACCCCGTATTGGCAAATATTGACGGCCAGTTAAGCCATTCATGCCAGTAGGCGCGC
GGACGAAAGTAAACCCACTGGTGATACCATTCGCGAGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGG
CGGGAACAGCAAAATATCACCCGGTCGGCAAACAAATTCTCGTCCCTGATTTTTCACCACCCCCTGACCGCGAA
TGGTGAGATTGAGAATATAACCTTTCATTCCCAGCGGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCA
ATCGGCGTTAAACCCGCCACCAGATGGGCATTAAACGAGTATCCCGGCAGCAGGGGATCATTTTGCGCTTCAGC
CAT actttt catactcccgccattcagagaagaaaccaattgtccatattgcatcagacattgccgtcactgcgtcttttactggctcttctcgctaaccaaaccggtaaccccgct
tattaaaagcattctgtaacaaagcgggaccaaagccatgacaaaaacgcgtaacaaaagtgtctataatcacggcagaaaagtccacattgattatttgcacggcgtcacactttt
gctatgccatagcatttttatccataagattagcggatcctacctgacgctttttatcgcaactctctactgtttctccat ATAATACCCGTGTGACCCGTGTTT
TAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTT
TTTTT gaagcttgggcccgaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttaaacggtctccagcttggctgttttg
gcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaacagaatttgcctggcggcagtagcgcggtggtcccacctgaccccatgccg
aactcagaagtgaaacgccgtagcgccgatggtagtgtggggtctccccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggccttt
cgttttatctgttgtttgtcggtgaactGGATCCTTACTCGAGTCTAGACTGCAGGCTTCCTC

Fig. 42: Plasmid pAN-NOR (Addgene ID: 62306, 4976 bp). Color-coding according to fig. 27: Lowercase black segments are ColE1 origin of replication and responsible for ampicillin resistance. *PhlF* CDS and subsequent terminator shown in light orange (eeb18a). *PhlF* promoter and first terminator (L3S2P55) colored in soft rose (d78dbc) lower case, enclosed sgRNA DNA sequence and second terminator (L3S2P21) in upper case letters. Sequence responsible for *araC* termination shown in capital light green (98d1ab) letters followed by *araC* CDS, also in light green. P_{BAD} promoter and terminator (TrrnB) illustrated in lower case blue (1b8bc7) letters, enclosed sgRNA DNA sequence in capital blue ones. Sequences with no relevant function or no function assigned are shown in capital black letters.

1.4.2 NAND Gate Plasmid - pAN-NAND

GCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG
TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAggccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg
cgtttcccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggt
atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagaca
cgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggaca
gtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaagcagcaga
ttacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatga GATTATCAA
AAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACT
TGGTCTGACAGTTA ccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacg
ggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtc
ctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtca
cgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcag
aagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattct
gagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcat CATTGGAAAACGTTCT
TCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGG
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGT
TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCC
CCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCAGTTGATAAGTCCCTAACTTTT
ACAGCTAGCTCAGTCCTAGGTATTATGCTAGCCTGAAGTACGTCTGAGCGTGATACCCGCTCACTGAAGATGGC
CCGGTAGGGCCGAAACGTACCTCTACAAATAATTTTGTTTAACTATGGACTATGTTTGAAAGGGAGAAATACTA
G ATGGCACGTACCCCGAGCCGTAGCAGCATTGGTAGCCTGCGTAGTCCGCATACCCATAAAGCAATTCTGACC
AGCACCATTGAAATCCTGAAAGAATGTGGTTATAGCGGTCTGAGCATTGAAAGCGTTGCACGTCGTGCCGGTG
CAAGCAAACCGACCATTTATCGTTGGTGGACCAATAAAGCAGCACTGATTGCCGAAGTGTATGAAAATGAAAGC
GAACAGGTGCGTAAATTTCCGGATCTGGGTAGCTTTAAAGCCGATCTGGATTTTCTGCTGCGTAATCTGTGGAA
AGTTTGGCGTGAAACCATTTGTGGTGAAGCATTTCGTTGTGTTATTGCAGAAGCACAGCTGGACCCTGCAACCC
TGACCCAGCTGAAAGATCAGTTTATGGAACGTCGTCGTGAGATGCCGAAAAAACTGGTTGAAAATGCCATTAGC
AATGGTGAACTGCCGAAAGATACCAATCGTGAACTGCTGCTGGATATGATTTTTGGTTTTTGTTGGTATCGCCT
GCTGACCGAACAGCTGACCGTTGAACAGGATATTGAAGAATTTACCTTCCTGCTGATTAATGGTGTTTGTCCGG
GTACACAGCGTTAA CTAGGGCCCATACCC CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCC
TTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTT
CTGCGTTTATA TACG cgacgtacggtggaatctgattcgttaccaattgacatgatacgaaacgtaccgtatcgttaaggt TAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT ctcggtaccaaagacgaacaataagacgctgaaaa
gcgtctttttttcgttttggtcc GGAG CTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTT

GGTCC TACG TTCAGCCAAAAAACTTAAGACCGCCGGTCTTGTCCACTACCTTGCAGTAATGCGGTGGACAGGA
TCGGCGGTTTTCTTTTCTCTTCTCAA GAATTCGCGGCCGCTTCTAGAG TTATGACAACTTGACGGCTACATCA
TTCACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAATACCCGCGAGAA
ATAGAGTTGATCGTCAAAACCAACATTGCGACCGACGGTGGCGATAGGCATCCGGGTGGTGCTCAAAAGCAGC
TTCGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAAGACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGA
CAGACGCGACGGCGACAAGCAAACATGCTGTGCGACGCTGGCGATATCAAAATTGCTGTCTGCCAGGTGATCG
CTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCGGTGGATGGAGCGACTCGTTAATCGCTTCCATGCG
CCGCAGTAACAATTGCTCAAGCAGATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGCCCGGCGTTAA
TGATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCGCTTCATCCGGGGCGAAAGAACCCCGTATTGGCAAAT
ATTGACGGCCAGTTAAGCCATTCATGCCAGTAGGCGCGCGGACGAAAGTAAACCCACTGGTGATACCATTCGCG
AGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGGCGGGAACAGCAAAATATCACCCGGTCGGCAAACAA
ATTCTCGTCCCTGATTTTTCACCACCCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATTCCCAGC
GGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTTAAACCCGCCACCAGATGGGCATTAAA
CGAGTATCCCGGCAGCAGGGGATCATTTTGCGCTTCAGCCAT actttcatactcccgccattcagagaagaaaccaattgtccatattgcat
cagacattgccgtcactgcgtcttttactggctcttctcgctaaccaaaccggtaaccccgcttattaaaagcattctgtaacaaagcgggaccaaagccatgacaaaaacgcgtaa
caaaagtgtctataatcacggcagaaaagtccacattgattatttgcacggcgtcacactttgctatgccatagcatttttatccataagattagcggatcctacctgacgctttttat
cgcaactctctactgtttctccat ATAATACCCGTGTGACCCGTGTTTTAGAGCTA gaagcttgggcccgaacaaaaactcatctcagaagaggatc
tgaatagcgccgtcgaccatcatcatcatcatcattgagtttaaacggtctccagcttggctgttttggcggatgagagaagattttcagcctgatacagattaaatcagaacgcaga
agcggtctgataaaacagaatttgcctggcggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtggggtctcccca
tgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaactGGATCCTTACTCGAG
TCTAGACTGCAGGCTTCCTC

Fig. 43: Plasmid pAN-NAND (4909 bp). Color-coding according to fig. 29: Lowercase black segments are ColE1 origin of replication and responsible for ampicillin resistance. PhlF CDS and subsequent terminator shown in light orange (eeb18a). *PhlF* promoter and first terminator (L3S2P55) colored in soft rose (d78dbc) lower case, enclosed tracrRNA DNA sequence and second terminator (L3S2P21) in upper case letters. Sequence responsible for *araC* termination shown in capital light green (98d1ab) letters followed by *araC* CDS, also in light green. P_{BAD} promoter and terminator (TrrnB) illustrated in lower case blue (1b8bc7) letters, enclosed crRNA DNA sequence in capital blue ones. Sequences with no relevant function or no function assigend are shown in capital black letters.

1.4.3 Discrete NAND Gate Plasmids - pAN-NANDXY

GCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG
TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAggccgcgttgctggcgtttttccacaggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg
cgtttcccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggt
atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagaca
cgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaaca
gtatttggtatctgcgctctgttgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaagcagcaga
ttacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatga GATTATCAA
AAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACT
TGGTCTGACAGTTA ccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacg
ggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggt
cctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtc
acgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtca
gaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattc
tgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcat CATTGGAAAACGTTC
TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACT
GATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAG
GGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGG
TTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTC
CCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCAGTTGATAAGTCCCTAACTTA
CG cgacgtacggtggaatctgattcgttaccaattgacatgatacgaaacgtaccgtatcgttaaggt TAGCAAGTTAAAATAAGGCTAGTCCGTTA
TCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT ctcggtaccaaagacgaacaataagacgctgaaaagcgtctttttttcgttttggtcc
GGAG CTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTTCGTTTTGGTCC TACGTTCA
GCCAAAAAACTTAAGACCGCCGGTCTTGTCCACTACCTTGCAGTAATGCGGTGGACAGGATCGGCGGTTTTCTT
TTCTCTTCTCAAGAATTCGCGGCCGCTTCTAGAG actttcatactcccgccattcagagaagaaaccaattgtccatattgcatcagacattgccgt
cactgcgtcttttactggctcttctcgctaaccaaaccggtaaccccgcttattaaaagcattctgtaacaaagcgggaccaaagccatgacaaaaacgcgtaacaaaagtgtctat
aatcacggcagaaaagtccacattgattatttgcacggcgtcacactttgctatgccatagcatttttatccataagattagcggatcctacctgacgctttttatcgcaactctctac
tgtttctccat ATAATACCCGTGTGACCCGTGTTTTAGAGCTA gaagcttgggcccgaacaaaaactcatctcagaagaggatctgaatagcgccgt
cgaccatcatcatcatcatcattgagtttaaacggtctccagcttggctgttttggcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgata
aaacagaatttgcctggcggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtggggtctccccatgcgagagtagg
gaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaact GGATCCTTACTCGAGTCTAGACT
GCAGGCTTCCTC

Fig. 44: Plasmid pAN-NAND11 (3094 bp). Color-coding according to fig. 29: Lowercase black segments are ColE1 origin of replication and responsible for ampicillin resistance. *PhlF* promoter and first terminator (L3S2P55) colored in soft rose (d78dbc) lower case, enclosed tracrRNA DNA sequence and second terminator (L3S2P21) in upper case letters. P_{BAD} promoter and terminator (TrrnB) illustrated in lower case blue (1b8bc7) letters, enclosed crRNA DNA sequence in capital blue ones. Sequences with no relevant function or no function assigend are shown in capital black letters. Where for pAN-NAND01 (3028 bp) the first lowercase soft rose sequence, for pAN-NAND10 (2783 bp) the first lowercase blue sequence was removed and for pAN-NAND00 (2717 bp) both.

1.4.4 FlipFlop Plasmid - pAN-FlipFlop

GCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG
TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAggccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg
cgtttcccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggt
atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagaca
cgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggaca
gtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaagcagcaga
ttacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatga GATTATCAA
AAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACT
TGGTCTGACAGTTA ccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacg
ggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtc
ctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtca
cgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcag
aagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattct
gagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcat CATTGGAAAACGTTCT
TCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGG
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGT
TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCC
CCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCAGTTGATAAGTCCCTAACTTTT
ACAGCTAGCTCAGTCCTAGGTATTATGCTAGCCTGAAGTACGTCTGAGCGTGATACCCGCTCACTGAAGATGGC
CCGGTAGGGCCGAAACGTACCTCTACAAATAATTTTGTTTAACTATGGACTATGTTTGAAAGGGAGAAATACTA
G ATGGCACGTACCCCGAGCCGTAGCAGCATTGGTAGCCTGCGTAGTCCGCATACCCATAAAGCAATTCTGACC
AGCACCATTGAAATCCTGAAAGAATGTGGTTATAGCGGTCTGAGCATTGAAAGCGTTGCACGTCGTGCCGGTG
CAAGCAAACCGACCATTTATCGTTGGTGGACCAATAAAGCAGCACTGATTGCCGAAGTGTATGAAAATGAAAGC
GAACAGGTGCGTAAATTTCCGGATCTGGGTAGCTTTAAAGCCGATCTGGATTTTCTGCTGCGTAATCTGTGGAA
AGTTTGGCGTGAAACCATTTGTGGTGAAGCATTTCGTTGTGTTATTGCAGAAGCACAGCTGGACCCTGCAACCC
TGACCCAGCTGAAAGATCAGTTTATGGAACGTCGTCGTGAGATGCCGAAAAAACTGGTTGAAAATGCCATTAGC
AATGGTGAACTGCCGAAAGATACCAATCGTGAACTGCTGCTGGATATGATTTTTGGTTTTTGTTGGTATCGCCT
GCTGACCGAACAGCTGACCGTTGAACAGGATATTGAAGAATTTACCTTCCTGCTGATTAATGGTGTTTGTCCGG
GTACACAGCGTTAA CTAGGGCCCATACCC CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCC
TTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTT
CTGCGTTTATA TACG cgacgtacggtggaatctgattcgttaccaattgacatgatacgaaacgtaccgtatcgttaaggt TAGCAAGTTAAAATAA
GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT ctcggtaccaaagacgaacaataagacgctgaaaa
gcgtctttttttcgttttggtcc GGAG CTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTT

GGTCC TACG TTCAGCCAAAAAACTTAAGACCGCCGGTCTTGTCCACTACCTTGCAGTAATGCGGTGGACAGGA
TCGGCGGTTTTCTTTTCTCTTCTCAA GAATTCGCGGCCGCTTCTAGAG TTATGACAACTTGACGGCTACATCA
TTCACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAATACCCGCGAGAA
ATAGAGTTGATCGTCAAAACCAACATTGCGACCGACGGTGGCGATAGGCATCCGGGTGGTGCTCAAAAGCAGC
TTCGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAAGACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGA
CAGACGCGACGGCGACAAGCAAACATGCTGTGCGACGCTGGCGATATCAAAATTGCTGTCTGCCAGGTGATCG
CTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCGGTGGATGGAGCGACTCGTTAATCGCTTCCATGCG
CCGCAGTAACAATTGCTCAAGCAGATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGCCCGGCGTTAA
TGATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCGCTTCATCCGGGGCGAAAGAACCCCGTATTGGCAAAT
ATTGACGGCCAGTTAAGCCATTCATGCCAGTAGGCGCGCGGACGAAAGTAAACCCACTGGTGATACCATTCGCG
AGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGGCGGGAACAGCAAAATATCACCCGGTCGGCAAACAA
ATTCTCGTCCCTGATTTTTCACCACCCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATTCCCAGC
GGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTTAAACCCGCCACCAGATGGGCATTAAA
CGAGTATCCCGGCAGCAGGGGATCATTTTGCGCTTCAGCCAT actttcatactcccgccattcagagaagaaaccaattgtccatattgca
tcagacattgccgtcactgcgtcttttactggctcttctcgctaaccaaaccggtaaccccgcttattaaaagcattctgtaacaaagcgggaccaaagccatgacaaaaacgcgta
acaaaagtgtctataatcacggcagaaaagtccacattgattatttgcacggcgtcacactttgctatgccatagcatttttatccataagattagcggatcctacctgacgcttttt
atcgcaactctctactgtttctccat acaaAAGTagttaTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCT
TTTTTT gaagcttgggcccgaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttaaacggtctccagcttggctgttt
tggcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaacagaatttgcctggcggcagtagcgcggtggtcccacctgaccccatgc
cgaactcagaagtgaaacgccgtagcgccgatggtagtgtggggtctccccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcc
tttcgttttatctgttgtttgtcggtgaactGGATCCTTACTCGAGACTACGCATGCGGGTAGATT TTTACACCAACGGGTCACAC
GGGTATTATGCTAGC ataatacctaggactgagctgtaactgattgt CCAATTATTGAAGGCCTCCCTAACGGGGGGCCTTTTTT
TGTTTCTGGTCTCCC GTTAGTATGTTGACCCGTGT TTTACACCTAGCTCAGTCCTAGGTATTATGCTAGC ataat
acccgtgtgacccgtgttttagagcta CTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCC
ATGTAGGTTGAGAGATCTGCAGGCTTCCTC

Fig. 45: Plasmid pAN-FlipFlop (5205 bp). Lowercase black segments are ColE1 origin of replication and responsible for ampicillin resistance. *PhlF* CDS and subsequent terminator shown in light orange (eeb18a). *PhlF* promoter and first terminator (L3S2P55) colored in soft rose (d78dbc) lower case, subsequent tracrRNA2 DNA sequence and second terminator (L3S2P21) in upper case letters. Sequence responsible for *araC* termination shown in capital light green (98d1ab) letters followed by *araC* CDS, also in light green. P_{BAD} promoter and terminator (TrrnB) illustrated in lower case blue (1b8bc7) letters, tracrRNA1 DNA sequence in capital blue ones. DNA sequence for PA2 promoter, crRNA2 and Terminator (L3S3P21) shown in violet (5c00d2), PA1 promoter, crRNA1, and Terminator (L3S2P11) in dark blue (0000ff) where lower case letters were used for both crRNAs, capital letters for both promoters and terminators. Sequences with no relevant function or no function assigend are shown in capital black letters.

1.4.5 XOR Gate Plasmid - pAN-XOR

GCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG
TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAggccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg
cgtttcccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggt
atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagaca
cgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggaca
gtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaagcagcag
attacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatga GATTATCA
AAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAAC
TTGGTCTGACAGTTA ccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgata
cgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtg
gtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtg
tcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgt
cagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcat
tctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcat CATTGGAAAACGTT
CTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAAC
TGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAA
GGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGG
GTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTT
CCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCAGTTGATAAGTCCCTAACTT
TTACAGCTAGCTCAGTCCTAGGTATTATGCTAGCCTGAAGTACGTCTGAGCGTGATACCCGCTCACTGAAGATG
GCCCGGTAGGGCCGAAACGTACCTCTACAAATAATTTTGTTTAACTATGGACTATGTTTGAAAGGGAGAAATAC
TAG ATGGCACGTACCCCGAGCCGTAGCAGCATTGGTAGCCTGCGTAGTCCGCATACCCATAAAGCAATTCTGA
CCAGCACCATTGAAATCCTGAAAGAATGTGGTTATAGCGGTCTGAGCATTGAAAGCGTTGCACGTCGTGCCGG
TGCAAGCAAACCGACCATTTATCGTTGGTGGACCAATAAAGCAGCACTGATTGCCGAAGTGTATGAAAATGAAA
GCGAACAGGTGCGTAAATTTCCGGATCTGGGTAGCTTTAAAGCCGATCTGGATTTTCTGCTGCGTAATCTGTGG
AAAGTTTGGCGTGAAACCATTTGTGGTGAAGCATTTCGTTGTGTTATTGCAGAAGCACAGCTGGACCCTGCAAC
CCTGACCCAGCTGAAAGATCAGTTTATGGAACGTCGTCGTGAGATGCCGAAAAAACTGGTTGAAAATGCCATTA
GCAATGGTGAACTGCCGAAAGATACCAATCGTGAACTGCTGCTGGATATGATTTTTGGTTTTTGTTGGTATCGC
CTGCTGACCGAACAGCTGACCGTTGAACAGGATATTGAAGAATTTACCTTCCTGCTGATTAATGGTGTTTGTCC
GGGTACACAGCGTTAA CTAGGGCCCATACCC CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGG
CCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCT
TTCTGCGTTTATA TACG cgacgtacggtggaatctgattcgttaccaattgacatgatacgaaacgtaccgtatcgttaaggt gtctAAGTgcgctTAAG
GCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT ctcggtaccaaagacgaacaataagacgctgaaaag
cgtctttttttcgttttggtcc GGAG CTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTT

GGTCC TACG TTCAGCCAAAAAACTTAAGACCGCCGGTCTTGTCCACTACCTTGCAGTAATGCGGTGGACAGGA
TCGGCGGTTTTCTTTTCTCTTCTCAA GAATTCGCGGCCGCTTCTAGAG TTATGACAACTTGACGGCTACATCA
TTCACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAATACCCGCGAGAA
ATAGAGTTGATCGTCAAAACCAACATTGCGACCGACGGTGGCGATAGGCATCCGGGTGGTGCTCAAAAGCAGC
TTCGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAAGACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGA
CAGACGCGACGGCGACAAGCAAACATGCTGTGCGACGCTGGCGATATCAAAATTGCTGTCTGCCAGGTGATCG
CTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCGGTGGATGGAGCGACTCGTTAATCGCTTCCATGCG
CCGCAGTAACAATTGCTCAAGCAGATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGCCCGGCGTTAA
TGATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCGCTTCATCCGGGCGAAAGAACCCCGTATTGGCAAAT
ATTGACGGCCAGTTAAGCCATTCATGCCAGTAGGCGCGCGGACGAAAGTAAACCCACTGGTGATACCATTCGCG
AGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGGCGGGAACAGCAAAATATCACCCGGTCGGCAAACAA
ATTCTCGTCCCTGATTTTTCACCACCCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATTCCCAGC
GGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTTAAACCCGCCACCAGATGGGCATTAAA
CGAGTATCCCGGCAGCAGGGGATCATTTTGCGCTTCAGCCAT actttcatactcccgccattcagagaagaaaccaattgtccatattgcat
cagacattgccgtcactgcgtctttttactggctcttctcgctaaccaaaccggtaaccccgcttattaaaagcattctgtaacaaagcgggaccaaagccatgacaaaaacgcgtaa
caaaagtgtctataatcacggcagaaaagtccacattgattatttgcacggcgtcacactttgctatgccatagcatttttatccataagattagcggatcctacctgacgctttttat
cgcaactctctactgtttctccat acaaAAGTagttaTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTT
TTTTT gaagcttgggcccgaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcatcattgagtttaaacggtctccagcttggctgttttg
gcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaacagaatttgcctggcggcagtagcgcggtggtcccacctgaccccatgccg
aactcagaagtgaaacgccgtagcgccgatggtagtgtggggtctccccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggccttt
cgttttatctgttgtttgtcggtgaact GGATCCTTACTCGAGTCTAGACTGCAAGTAATAATT tttacacctagctcagtcctaggtattatgcta
gc ATAATACCCGTGTGACCCGTGTTTTAGAGCTA ccaattattgaaggcctccctaacgggggggccttttttttgtttctggtctccc TAAAAGA
GAT tttacacctccacaactagctggtattatgctagc TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCA
CCGAGTCGGTGCTTTTTTTT ctcggtaccaaattccagaaaagagacgctttcgagcgtctttttttcgttttggtcc GTCGGTTTGC tttacacccgaaa
tggagcatggtattatgctagc ATAATACCtaggactgagctGtaactGAttgt ccaattattgaaggccgctaacgcggccttttttttgtttctggtctccc TGGTT
CACGT tttacacccgaaatggagcatggtattatgctagc ATAATACCAGCTAGTTGTGGGagcgcGAagac ccaattattgaaggggagcgggaaa
ccgctccccttttttttgtttctggtctccc CTTAGTACGTAGCATGGTGACACAAGCACAGTAGATCCTGCCCGCGTTTCCTATAT
ATTAAGTTAAATCTTATGGAATATAATAACATGTGGATGGCCAGTGGTCGGTTGTTACACGCCTACCGCAATGC
TGAAAGACCCGGACT FFaaagccatgacaaaaacgcgtaacaaaagtgtctataatcacggcagaaaagtccacattgattatttgcacggcgtcacactttgcta
tgccatagcatttttatccataagattagcggatcctacctgacgctttttatcgcaactctctactgtttctccat FFATAATACCATGCTCCATTTCGtcgagG
Acacg FFgacgaacaataaggcctcccaaatcgggggggccttttttattgataacaaaa CGCGCTACCTTACG cgacgtacggtggaatctgattcgttaccaatt
gacatgatacgaaacgtaccgtatcgttaaggt cgtgAAGTctcgaTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG
GTGCTTTTTTT ctcggtaccaaattccagaaaagagacgctgaaaagcgtctttttttcgttttggtcc CTAAGTAAGTGGCTTCCTC

Fig. 46: Plasmid pAN-XOR (6073 bp). Lowercase black segments are ColE1 origin of replication and responsible for ampicillin resistance. *PhlF* CDS and subsequent terminator shown in light orange (eeb18a). *PhlF* promoter and first terminator (L3S2P55) colored in soft rose (d78dbc) lower case, enclosed tracrRNA4 DNA sequence and second terminator (L3S2P21) in upper case letters. Sequence responsible for *araC* termination shown in capital light green (98d1ab) letters followed by *araC* CDS, also in light green. P_{BAD} promoter and terminator (TrrnB) illustrated in lower case blue (1b8bc7) letters, enclosed tracrRNA1 DNA sequence in capital blue ones. PA1 promoter and L3S3P21 Terminator shown in bright salmon (e96175) lower case letters, enclosed crRNA2 in capital ones. PA4 promoter and L3S2P11 Terminator also shown in bright salmon (e96175) lower case letters, enclosed tracrRNA2 in captial ones. PA3 promoter and L3S3P22 Terminator shown in violet (5c00d2) lower case letters, enclosed crRNA1 in capital ones. PA3 promoter and L3S3P00 Terminator shown in in dark blue (0000ff) lower case letters, enclosed crRNA4 in captial ones. P_{BAD} promoter and terminator (pheA-1) illustrated in bright green (2ba56aff) lower case letters, enclosed crRNA3 DNA sequence in capital ones. *PhlF* promoter and terminator (ECK120034435) colored in bright orange (f79532) lower case letters, enclosed tracrRNA3 sequence in upper case ones. RNA naming digits indicate sequences of crRNA and tracrRNA that can hybridize and then target promoter with corresponding ID. Sequences with no relevant function or no function assigend are shown in capital black letters.

1.4.6 CRISPRi/dCas9 Plasmid - pAN-PTet-dCas9

GCTTCCTCGCTCACTGACTCGCTGCACGAGGCA gacctcagcgctagcggagtgtatactggcttactatgttggcactgatgagggtgtcagtgaa
gtgcttcatgtggcaggagaaaaaaggctgcaccggtgcgtcagcagaatatgtgatacaggatatattccgcttcctcgctcactgactcgctacgctcggtcgttcgactgcggc
gagcggaaatggcttacgaacggggcggagatttcctggaagatgccaggaagatacttaacagggaagtgagagggccgcggcaaagccgtttttccataggctccgcccccct
gacaagcatcacgaaatctgacgctcaaatcagtggtggcgaaacccgacaggactataaagataccaggcgtttcccctggcggctccctcgtgcgctctcctgttcctgcctttcg
gtttaccggtgtcattccgctgttatggccgcgtttgtctcattccacgcctgacactcagttccgggtaggcagttcgctccaagctggactgtatgcacgaacccccgttcagtcc
gaccgctgcgccttatccggtaactatcgtcttgagtccaacccggaaagacatgcaaaagcaccactggcagcagccactggtaattgatttagaggagttagtcttgaagtcatg
cgccggttaaggctaaactgaaaggacaagttttggtgactgcgctcctccaagccagttacctcggttcaaagagttggtagctcagagaaccttcgaaaaaccgccctgcaaggc
ggtttttttcgttttcagagcaagagattacgcgcagaccaaaacgatctcaagaagatcatcttattaa GGGGTCTGACGCTCAGTGGAACGAAAAATC
AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC ttagaaaaactcatcgagcatcaaatgaaactgcaatttattcatatcagga
ttatcaataccatatttttgaaaaagccgtttctgtaatgaaggagaaaactcaccgaggcagttccataggatggcaagatcctggtatcggtctgcgattccgactcgtccaacat
caatacaacctattaatttcccctcgtcaaaaataaggttatcaagtgagaaatcaccatgagtgacgactgaatccggtgagaatggcaaaagcttatgcatttctttccagacttg
ttcaacaggccagccattacgctcgtcatcaaaatcactcgcatcaaccaaaccgttattcattcgtgattgcgcctgagcgagacgaaatacgcgatcgctgttaaaaggacaatt
acaaacaggaatcgaatgcaaccggcgcaggaacactgccagcgcatcaacaatattttcacctgaatcaggatattcttctaatacctggaatgctgttttcccggggatcgcagt
ggtgagtaaccatgcatcatcaggagtacggataaaatgcttgatggtcggaagaggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgct
acctttgccatgtttcagaaacaactctggcgcatcgggcttcccatacaatcgatagattgtcgcacctgattgcccgacattatcgcgagcccatttatacccatataaatcagca
tccatgttggaatttaatcgcggcctcgagcaagacgtttcccgttgaatatggctcat AACACCCCTTGTATTACTGTTTATGTAAGCAGACAGT
TTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGTACATCAGAGATTTTGAGACACAA CCAATTATTGA
AGGCCTCCCTAACGGGGGGCCTTTTTTTGTTTCTGGTCTCCC gctt TACTCCACCGTTGGCTTTTTTCCCTATCA
GTGATAGAGATTGACATCCCTATCAGTGATAGAGATAATGAGCAC TTCAAAAGATCT AAAGAGGAGAAA GGA
TCT atggataagaaatactcaataggcttagctatcggcacaaatagcgtcggatgggcggtgatcactgatgaatataaggttccgtctaaaaagttcaaggttctgggaaat
acagaccgccacagtatcaaaaaaaatcttataggggctcttttatttgacagtggagagacagcggaagcgactcgtctcaaacggacagctcgtagaaggtatacacgtcggaa
gaatcgtatttgttatctacaggagatttttttcaaatgagatggcgaaagtagatgatagtttctttcatcgacttgaagagtctttttttggtggaagaagacaagaagcatgaacgt
catcctatttttggaaatatagtagatgaagttgcttatcatgagaaatatccaactatctatcatctgcgaaaaaaaattggtagattctactgataaagcggatttgcgcttaatcta
tttggccttagcgcatatgattaagtttcgtggtcatttttttgattgagggagatttaaatcctgataatagtgatgtggacaaactatttatccagttggtacaaacctacaatcaat
tatttgaagaaaaccctattaacgcaagtggagtagatgctaaagcgattctttctgcacgattgagtaaatcaagacgattagaaaatctcattgctcagctccccggtgagaaga
aaaatggcttatttgggaatctcattgctttgtcattgggtttgacccctaattttaaatcaaattttgatttggcagaagatgctaaattacagctttcaaaagatacttacgatgat
gatttagataatttattggcgcaaattggagatcaatatgctgatttgtttttggcagctaagaatttatcagatgctattttacttttcagatatcctaagagtaaatactgaaataac
taaggctcccctatcagcttcaatgattaaacgctacgatgaacatcatcaagacttgactcttttaaaagctttagttcgacaacaacttccagaaaagtataaagaaatctttttt
gatcaatcaaaaaacggatatgcaggttatattgatgggggagctagccaagaagaattttataaatttatcaaaccaattttagaaaaaatggatggtactgaggaattattggt
gaaactaaatcgtgaagatttgctgcgcaagcaacggaccttgacaacggctctattccccatcaaattcacttgggtgagctgcatgctattttgagaagacaagaagacttttat
ccatttttaaaagacaatcgtgagaagattgaaaaaatcttgacttttcgaattccttattatgttggtccattggcgcgtggcaatagtcgttttgcatggatgactcggaagtctga
agaaacaattaccccatggaattttgaagaagttgtcgataaaggtgcttcagctcaatcatttattgaacgcatgacaaacttttgataaaaatcttccaaatgaaaaagtactacc
aaaacatagtttgctttatgagtattttacggtttataacgaattgacaaaggtcaaatatgttactgaaggaatgcgaaaaccagcatttctttcaggtgaacagaagaaagccat
tgttgatttactcttcaaaacaaatcgaaaagtaaccgttaagcaattaaaagaagattatttcaaaaaaatagaatgttttgatagtgttgaaatttcaggagttgaagatagatt
taatgcttcattaggtacctaccatgatttgctaaaaattattaaagataaagatttttttggataatgaagaaaatgaagatatcttagaggatattgttttaacattgaccttatttg

aagatagggagatgattgaggaaagacttaaaacatatgctcacctctttgatgataaggtgatgaaacagcttaaacgtcgccgttatactggttggggacgtttgtctcgaaaat
tgattaatggtattagggataagcaatctggcaaaacaatattagatttttgaaatcagatggttttgccaatcgcaattttatgcagctgatccatgatgatagtttgacatttaaa
gaagacattcaaaaagcacaagtgtctggacaaggcgatagtttacatgaacatattgcaaatttagctggtagccctgctattaaaaaaggtattttacagactgtaaaagttgtt
gatgaattggtcaaagtaatggggcggcataagccagaaaatatcgttattgaaatggcacgtgaaaatcagacaactcaaaagggccagaaaaattcgcgagagcgtatgaaac
gaatcgaagaaggtatcaaagaattaggaagtcagattcttaaagagcatcctgttgaaaatactcaattgcaaaatgaaaagctctatctctattatctccaaaatggaagagac
atgtatgtggaccaagaattagatattaatcgtttaagtgattatgatgtcgatgccattgttccacaaagtttccttaaagacgattcaatagacaataaggtcttaacgcgttctg
ataaaaatcgtggtaaatcggataacgttccaagtgaagaagtagtcaaaaagatgaaaaactattggagacaacttctaaacgccaagttaatcactcaacgtaagtttgataat
ttaacgaaagctgaacgtggaggtttgagtgaacttgataaagctggttttatcaaacgccaattggttgaaactcgccaaatcactaagcatgtggcacaaattttggatagtcgc
atgaatactaaatacgatgaaaatgataaacttattcgagaggttaaagtgattaccttaaaatctaaattagtttctgacttccgaaaagatttccaattctataaagtacgtgaga
ttaacaattaccatcatgcccatgatgcgtatctaaatgccgtcgttggaactgctttgattaagaaatatccaaaacttgaatcggagtttgtctatggtgattataaagtttatgat
gttcgtaaaatgattgctaagtctgagcaagaaataggcaaagcaaccgcaaaatatttcttttactctaatatcatgaacttcttcaaaacagaaattacacttgcaaatggagag
attcgcaaacgccctctaatcgaaactaatggggaaactggagaaattgtctgggataaagggcgagattttgccacagtgcgcaaagtattgtccatgccccaagtcaatattgtc
aagaaaacagaagtacagacaggcggattctccaaggagtcaattttaccaaaaagaaattcggacaagcttattgctcgtaaaaaagactgggatccaaaaaaatatggtggtt
ttgatagtccaacggtagcttattcagtcctagtggttgctaaggtggaaaaagggaaatcgaagaagttaaaatccgttaaagagttactagggatcacaattatggaaagaagt
tcctttgaaaaaaatccgattgacttttagaagctaaaggatataaggaagttaaaaaagacttaatcattaaactacctaaatatagtctttttgagttagaaaacggtcgtaaa
cggatgctggctagtgccggagaattacaaaaaggaaatgagctggctctgccaagcaaatatgtgaatttttatatttagctagtcattatgaaaagttgaagggtagtccaga
agataacgaacaaaaacaattgtttgtggagcagcataagcattatttagatgagattattgagcaaatcagtgaattttctaagcgtgttattttagcagatgccaatttagataa
agttcttagtgcatataacaaacatagagacaaaccaatacgtgaacaagcagaaaatattattcatttatttacgttgacgaatcttggagctcccgctgcttttaaatattttgat
acaacaattgatcgtaaacgatatacgtctacaaaagaagttttagatgccactcttatccatcaatccatcactggtctttatgaaacacgcattgatttgagtcagctaggaggtg
actaa CTCGAGTAAGGATCT CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTAT
CTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA
AATG GCGGCGCGCCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATT
CAGGGTGGTGAAT ATGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTT
TCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGC
TGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCC
AGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGT
GGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGC
GTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGT
TCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAGGACGGTACGC
GACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGT
CTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGG
AAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCG
ATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTG
CGGATATCTCGGTAGTGGGATACGACGATACCGAAGATAGCTCATGTTATATCCCGCCGTTAACCACCATCAAA
CAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGG
GCAATCAGCTGTTGCCAGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCC
CGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGA TAATCCA

GGAGGAAAAAA atgtccagattagataaaagtaaagtgattaacagcgcattagagctgcttaatgaggtcggaatcgaaggtttaacaacccgtaaactcgcccagaa
gctaggtgtagagcagcctacattgtattggcatgtaaaaaataagcgggctttgctcgacgccttagccattgagatgttagataggcaccatactcacttttgccctttagaaggg
gaaagctggcaagatttttttacgtaataacgctaaaagttttagatgtgctttactaagtcatcgcgatggagcaaaagtacatttaggtacacggcctacagaaaaacagtatga
aactctcgaaaatcaattagccttttttatgccaacaaggtttttcactagagaatgcattatatgcactcagcgctgtggggcattttactttaggttgcgtattggaagatcaagag
catcaagtcgctaaagaagaaagggaaacacctactactgatagtatgccgccattattacgacaagctatcgaattatttgatcaccaaggtgcagagccagccttcttattcggc
cttgaattgatcatatgcggattagaaaaacaacttaaatgtgaaagtgggtcctaa TAATTGGTAACGAATCAGACAATTGACGGCTCGAGGG
AGTAGCATAGGGTTTGCAGAATCCCTGCTTCGTCCATTTGACAGGCACATTATGCATCGATGATAAGCTGTCAA
ACATGAGCAGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCC
TATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCAAATATTTTATCTG

Fig. 47: CRISPRi/dCas9 Plasmid - pAN-PTet-dCas9 (Addgene ID: 62244, 8338 bp). Color-coding according to fig. 27: Lowercase black segments are p15A origin of replication and responsible for kanamycin resistance. Tetracyline inducible promoter and terminator shown in upper case bright orange (f79532) letters, *dcas9* CDS in lower case ones. *TetR* CDS is shown in lower case soft lila (ae9ac9) letters. Sequences with no relevant function or no function assigend are shown in capital black letters.

1.4.7 RFP Reporter Plasmid - pAN-PA2-RFP

atgagggaagcggtgatcgccgaagtatcgactcaactatcagaggtagttggcgtcatcgagcgccatctcgaaccgacgttgctggccgtacatttgtacggctccgcagtggat
ggcggcctgaagccacacagtgatattgatttgctggttacggtgaccgtaaggcttgatgaaacaacgcggcgagctttgatcaacgacctttttggaaacttcggcttcccctgga
gagagcgagattctccgcgctgtagaagtcaccattgttgtgcacgacgacatcattccgtggcgttatccagctaagcgcgaactgcaatttggagaatggcagcgcaatgacatt
cttgcaggtatcttcgagccagccacgatcgacattgatctggctatcttgctgacaaaagcaagagaacatagcgttgccttggtaggtccagcggcggaggaactctttgatccg
gttcctgaacaggatctatttgaggcgctaaatgaaaccttaacgctatggaactcgccgcccgactgggctggcgatgagcgaaatgtagtgcttacgttgtcccgcatttggtac
agcgcagtaaccggcaaaatcgcgccgaaggatgtcgctgccgactgggcaatggagcgcctgccggcccagtatcagcccgtcatacttgaagctagacaggcttatcttggaca
agaagaagatcgcttggcctcgcgcacagatcagttggaagaatttgtccactacgtgaaaggcgagatcaccaaggtagtcggcaaataa TACTAGCTCCGGCAA
AAAAACGGGCAAGGTGTCACCACCCTGCCCTTTTTCTTTAAAACCGAAAAGATTACTTCGCGTTTGCCACCTGA
CGTCTAAGAACTCGGTACCAAACCAATTATTGAAGACGCTGAAAAGCGTCTTTTTTTGTTTCGGTCCTCCGAAT
GACATGCGTCTCC tttacaccaacgggtcacacgggtattatgctagc AGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCG
TGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAA TACTAGAG TCACACAGGAAAG TACTAG atggcttcctccg
aagacgttatcaaagagttcatgcgtttcaaagttcgtatggaaggttccgttaacggtcacgagttcgaaatcgaaggtgaaggtgaaggtcgtccgtacgaaggtacccagacc
gctaaactgaaagttaccaaaggtggtccgctgccgttcgcttgggacatcctgtccccgcagttccagtacggttccaaagcttacgttaaacacccggctgacatcccggactacc
tgaaactgtccttcccggaaggtttcaaatgggaacgtgttatgaacttcgaagacggtggtgttgttaccgttacccaggactcctccctgcaagacggtgagttcatctacaaag
ttaaactgcgtggtaccaacttcccgtccgacggtccggttatgcagaaaaaaaccatgggttgggaagcttccaccgaacgtatgtacccggaagacggtgctctgaaaggtgaa
atcaaaatgcgtctgaaactgaaagacggtggtcactacgacgctgaagttaaaaccacctacatggctaaaaaaccggttcagctgccgggtgcttacaaaaccgacatcaaact
ggacatcacctcccacaacgaagactacaccatcgttgaacagtacgaacgtgctgaaggtcgtcactccaccggtgcttaataa TACTAGAG CCAGGCATCAA
ATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTA
GAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA GCTCACTCAAAGGCGGTAATACGGTTATC
CACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAA ct
gtcagaccaagtttacgagctcgcttggactcctgttgatagatccagtaatgacctcagaactccatctggatttgttcagaacgctcggttgccgccgggcgtttttttattggtgag
aatccaagcactagggacagtaagacgggtaagcctgttgatgataccgctgccttactgggtgcattagccagtctgaatgacctgtcacgggataatccgaagtggtcagactg
gaaaatcagagggcaggaactgctgaacagcaaaaagtcagatagcaccacatagcagacccgccataaaacgccctgagaagcccgtgacgggcttttcttgtattatgggtag
tttccttgcatgaatccataaaaggcgcctgtagtgccatttacccccattcactgccagagccgtgagcgcagcgaactgaatgtcacgaaaaagacagcgactcaggtgcctgat
ggtcggagacaaaaggaatattcagcgatttgcccgagcttgcgagggtgctacttaagcctttagggttttaaggtctgttttgtagaggagcaaacagcgtttgcgacatccttttt
gtaatactgcggaactgactaaagtagtgagttatacacagggctgggatctattcttttttatcttttttattcttctttattctataaattataaccacttgaatataaacaaaaaa
aacacacaaaggtctagcggaatttacagagggtctagcagaatttacaagttttccagcaaaggtctagcagaatttacagatacccacaactcaaaggaaaaggacatgtaatt
atcattgactagcccatctcaattggtatagtgattaaaatcacctagaccaattgagatgtatgtctgaattagttgttttcaaagcaaatgaactagcgattagtcgctatgactta
acggagcatgaaaccaagctaattttatgctgtggcactactcaaccccacgattgaaaaccctacaaggaaagaacggacggtatcgttcacttataaccaatacgctcagatg
atgaacatcagtagggaaaatgcttatggtgtattagctaaagcaaccagagagctgatgacgagaactgtggaaatcaggaatcctttggttaaaggctttgagattttccagtg
gacaaactatgccaagttctcaagcgaaaaattagaattagtttttagtgaagagatattgccttatcttttccagttaaaaaaattcataaaatataatctggaacatgttaagtct
tttgaaaacaaatactctatgaggatttatgagtggttattaaaagaactaacacaaaagaaaactcacaaggcaaatatagagattagccttgatgaatttaagttcatgttaatg
cttgaaaataactaccatgagtttaaaaggcttaaccaatgggttttgaaaccaataagtaaagatttaaacacttacagcaatatgaaattggtggttgataagcgaggccgcccg
actgatacgttgattttccaagttgaactagatagacaaatggatctcgtaaccgaacttgagaacaaccagataaaaatgaatggtgacaaaatacaacaaccattacatcaga
ttcctacctacgtaacggactaagaaaaacactacacgatgctttaactgcaaaaattcagctcaccagttttgaggcaaaatttttgagtgacatgcaaagtaagcatgatctcaa

tggttcgttctcatggctcacgcaaaaacaacgaaccacactagagaacatactggctaaatacggaaggatctgaggttcttatggctcttgtatctatcagtgaagcatcaagac
taacaaacaaaagtagaacaactgttcaccgttagatatcaaagggaaaactgtccatatgcacagatgaaaacggtgtaaaaaagatagatacatcagagcttttacgagttttt
ggtgcatttaaagctgttcaccatgaacagatcgacaatgtaac GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATG
AAGCTTCGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAAC
TGTCTGCTTACATAAACAGTAATACAAGGGGTGTTTACTAGAGGTTGATCGGGCACGTAAGAGGTTCCAACTTT
CACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAA
A

Fig. 48: mRFP1 Reporter Plasmid - pAN-PA2-RFP (Addgene ID: 62248, 4330 bp). Color-coding according to fig. 27: Lowercase black segments are streptomycine resistance and pSC101 origin of replication. PA2 promoter and terminator are shown in capital bright salmon (e96175) letters, *mrfp1* CDS in lower case. Sequences with no relevant function or no function assigend are shown in capital black letters.

2 A SBOL Visual Font for Gene Circuit Design

2.1 SBOL - The Synthetic Biology Open Language

Synthetic biology being a relatively new field of research combines disciplines as genetics, molecular biology, biophysics and a lot more. In contrast to many contributing areas, synthetic biology does not inherit a standardized principle with regards to quality and quantity which interconnects all participating branches. Yet, a foundation for the exchange of all kinds of different information is required to assure successful collaboration between laboratories and institutions. Because of this, the Synthetic Biology Open Language (SBOL)[65][66] is being developed as a standard to support the specification and exchange of biological design information.

2.1.1 SBOL Data

The SBOL standard is divided into two major parts. Where SBOL Visual is a defining framework for the graphical representation of biological systems, SBOL Data incorporates the principles that govern the digital exchange of data for these systems. The major aspects of these principles are the structural and functional elements of the corresponding genetic constructs and embody a well defined data model. SBOL Data makes use of already existing semantic web practices and resources by distinctively identifying and defining genetic design elements. The associated structures and specifications of the data model render SBOL Data a useful tool to promote global information exchange between researchers and their software.

2.1.2 SBOL Visual

SBOL Visual is a code of practice for the graphical representation of genetic constructs and their functional interaction. It accounts for all best practices and conventions that have been proven useful for graphical representation of genetic structures since the emergence of synthetic biology. SBOL Visual has been created to ensure a common repertoire of symbols which unlocks the foundation for a graphical information storage and communication, with the intention of excluding misconception. Also, one of the goals of SBOL Visual is to supply a user-friendly and highly adaptable standard. SBOL Visual is linked to the Data counterpart and since synthetic biology is a rapidly growing area

of research, SBOL has been designed to be able to easily include extensions in the future.

2.2 A Font as an Extension to the SBOL Visual Standard

2.2.1 Scope and Motivation

One of the central goals of synthetic biology is the engineering of artificial gene circuits for the implementation of synthetic functions within living organisms. These circuits are composed of gene regulatory elements e. g. promoters, operators and terminators. Synthetic Biology Open Language Visual was created to encourage the use of uniform and standardized symbols for these gene regulatory elements in synthetic biology. However, the already available software for the graphic representation of genetic designs is limited and often not completely compatible with popular software used for e. g. the preparation of publications. To overcome this, the SBOL Visual font was created, which can be installed into every operating system, then utilized in all image- and text processing programs such as LibreOffice Writer, Inkscape, Microsoft Office Word or Adobe Illustrator.

2.2.2 Font building Process

The scalable vector images in the official SBOL Visual Specifications Github Repository were extracted and prepared for the compilation into a font using Inkscape. Additionally, connecting elements for the construction of semicircular plasmid diagrams and interactive pathways were added. After scaling the images up to a width of 4096 pixels, a baseline, representing DNA was added where required. All images were edited by changing their object-like nature into paths after applicable ungrouping. The accordingly prepared graphical elements were merged into single path structures each to ensure maximum compatibility with current font standards.

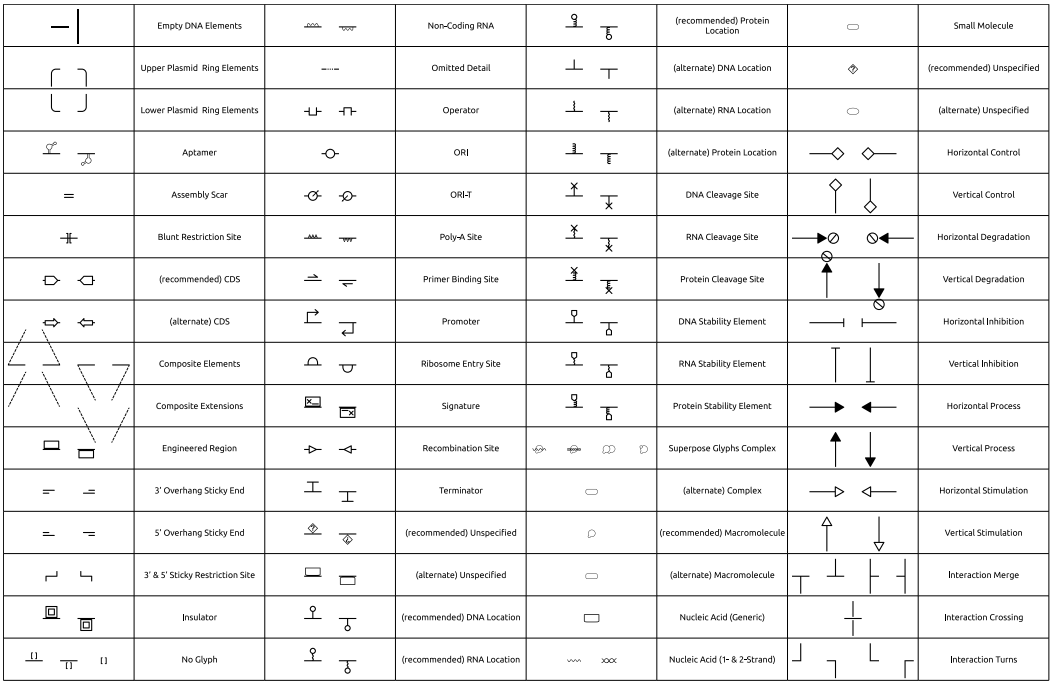


Fig. 49: SBOL Visual 2.0.0 font character set. The glyphs are placed in a private block area of the Unicode table starting at U+F0000 and can be inserted as special character. A font version in which the character allocation starts at the beginning of the Unicode table, does not exceed 256 slots and thus is compatible with older operating systems and programs is also available.

The resulting images were then ported to FontForge and placed at the private block section of the unicode table starting at index U+F0000. After further adaption of the symbols to font standard by correcting the direction of the paths, non-integral coordinates, missing extremas and intersecting paths, the SBOL Visual 2.0.0 font was generated. Since older software does not support glyph placement in the Unicode table beyond slot 256, an alternate version of the font has been created where the symbols are occupying the slots starting at the origin of the table. The font and all contributing files and information has been uploaded to Github, then, after inspection and verification approved and pulled into the master branch of the official SBOL Visual Developer Repository. Before that, a file giving detailed instructions about how to install, use, design, update and test the font has been added.

2.2.3 Installation, Workflow and Usecases

After downloading the font from the official SBOL Visual Github Repository and installing it, a typical workflow is to insert the desired special characters in a word-processing program. After the assembly of the corresponding genetic circuit via e. g. copy/cut and paste, the figure should be exported as PDF and imported in a vector-graphics editor by Poppler/Cairo import. Since the imported file is a scalable vector image, a custom size can be chosen without loss of quality. After ungrouping, all single elements can be colored independently. If required a descriptive text can be added as shown in fig. 50.

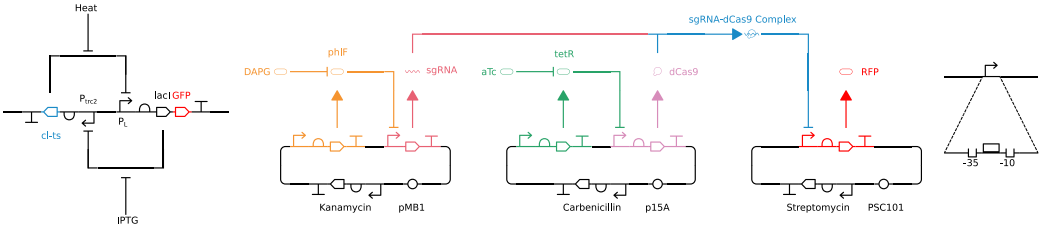


Fig. 50: SBOL Visual 2.0.0 font usage examples. Left: Toggle Switch. Center: CRISPRi with aTc/tetR and DAPG/phlF inducer systems. Right: Promoter details described using composite elements including extensions.

2.2.4 Testing Procedures

The SBOL Visual 2.0.0 font has been tested in Kubuntu 17.10 Artful Aardvark and 18.04 Bionic Beaver using LibreOffice Writer 5.4.6.2 by recreating structures shown in fig. 32 i. e. inserting via menu, copying, cutting, pasting i. e. placing symbols as demanded. The content has been exported as PDF and imported in Inkscape where further processing as coloring and minor adjustments took place. This has also been performed running Windows 7 SP1, in the Windows Version of LibreOffice Writer 5.4.6.2 as well as Microsoft Office Word 2007 Version 12.0.6668.5000.

2.2.5 Conclusion, remaining Challenges and Outlook

The Synthetic Biology Open Language (SBOL) Visual font is a collection of all symbols described in the SBOL 2.0.0 Visual specifications compiled into a font that can easily be used in most text- and image processing programs on every

graphical operating system. It enables the rapid generation of a visual design for genetic circuits according to the standard by letter-like symbols without the need for further software. Content that has been created with the SBOL font can be exported into formats as SVG and PDF and thereby processed for further use as in publications. The intention of the creation of this font is to contribute to propagating the usage of the Synthetic Biology Open Language Visual standard in the future. The development and usage of user friendly tools and or software for which no experience in computer science is required is highly encouraged to speed up and ease the publication process in the field of synthetic biology and beyond.

Software development is an ongoing circular feedback process and perfect software cannot be created, improvements will never cease to be required. This also includes the development of the SBOL Visual font, for which possible upgrades are already obvious. In word-processing programs, the width of some symbols has been observed to be changing when selecting a color other than black, causing a misalignment among the connectors. This bug was not observed when elements are colored in vector-image programs as Inkscape. Also, when constructing genetic circuits, misalignments and missing parts of the symbols are being displayed, but can be corrected by scrolling and or zooming. In contrast to the bug caused by coloring symbols withing a text-processing program, this has no implication on the image quality after being exported to a PDF.

Alongside finding the causal background and revision of these bugs with origin still unknown, the SBOL Visual font must also be maintained of course. Since synthetic biology is a rapidly growing field of research and also because of this the SBOL standard is still subject to frequent change, more symbols will have to be added into the font in the future. The framework for the generation of this font itself is also a matter of continuous development so that the current state will be obsolete before long.

The symbols compiled in the SBOL Visual font can not only be used for a representation of elements from synthetic biology but already for many genetic structures describing the basic elements and functions that constitute life at the smallest level in general. Because of this, it is reasonable that similar to how it was conducted with emoticons, the SBOL Visual font symbols should

be integrated from the private to the public part of the unicode table when the SBOL Visual standard has reached a state close enough to completion in the near future.

Bibliography

[1] Krishnan, Ziegler, Arnaut, Simmel et al
Molecular transport through large-diameter DNA nanopores
Nature Communications Volume 7, Article number: 12787 (2016)
DOI: 10.1038/ncomms12787

[2] Nielsen, Voigt
Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks
Mol Syst Biol. 2014 Nov 24;10:763
DOI: 10.15252/msb.20145735

[3] Lang, Sassmann, Schmidt, Komis
Plasmolysis: Loss of Turgor and Beyond
Plants 2014, 3, 583-593
DOI: 10.3390/plants3040583

[4] McNaugh, Wilkinson
Compendium of Chemical Terminology, 2nd ed.
Blackwell Scientific Publications, Oxford (1997)
DOI: https://doi.org/10.1351/goldbook.A00028.

[5] Campbell, Tour, Palmer, Tsien et al
A monomeric red fluorescent protein
PNAS June 11, 2002. 99 (12) 7877-7882
DOI: https://doi.org/10.1073/pnas.082243699

[6] Shaner, Campbell, Steinbach, Tsien et al
Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein
Nat Biotechnol. 2004 Dec;22(12):1567-72. Epub 2004 Nov 21.
DOI: 10.1038/nbt1037

[7] Sobarwiki
Output profile of a 150 W continuous wave ozone free xenon arc lamp[...]
Wikipedia, 19 July 2014, 15:57:13
URL: https://en.wikipedia.org/wiki/File:Xenon_arc_lamp_profile.png

[8] Demerec, Adelberg, Clark, Hartman
A proposal for a uniform nomenclature in bacterial genetics
Genetics, 1966 Jul;54(1):61-76
PMID: 5961488

[9] Ali Zifan
A diagram of a typical prokaryotic bacteria cell.
Wikipedia, 12 October 2015
URL: https://upload.wikimedia.org/wikipedia/commons/thumb/c/c5/Prokaryote_cel
Prokaryote_cell.svg.png

[10] Kubitschek
Cell Volume Increase in Escherichia coli after Shifts to Richer Media
J. Bacteriol. January 1990 vol. 172 no. 1 94-101
DOI: 10.1128/jb.172.1.94-101.1990

[11] Yu, Loo, Kong, Chan
Monitoring bacterial growth using tunable resistive pulse sensing with a pore-based technique
Appl Microbiol Biotechnol (2014) 98: 855
DOI: https://doi.org/10.1007/s00253-013-5377-9

[12] Gardner†, Craven†, Hines† et al
The prediction of bacteria type and culture growth phase by an electronic nose with a multi-layer perceptron network
J W Gardner et al 1998 Meas. Sci. Technol. 9 120
DOI: https://doi.org/10.1088/0957-0233/9/1/016

[13] Zwietering, Jongenburger, Rombouts, Riet
Modeling of the Bacterial Growth Curve
Appl Environ Microbiol. 1990 Jun; 56(6): 1875–1881
PMCID: PMC184525, PMID: 16348228

[14] McDaniel, Bailey, Zimmerli
Effect of Oxygen-Supply Rates on Growth of Escherichia coli
Appl Microbiol. 1965 Jan; 13(1): 109–114
PMCID: PMC1058202, PMID: 14264839

[15] Sowana, Williams, Dunlop, Fletcher et al
Turbulent shear stress effects on plant cell suspension cultures
Chemical Engineering Research and Design, Volume 79, Issue 8, November 2001, Pages 867-875
DOI: https://doi.org/10.1205/02638760152721370

[16] Takeda, Seki, Furusaki
Hydrodynamic Damage of Cultured Cells of Carthamus tinctorius in a Stirred Tank Reactor
Journal of chemical engineering of Japan, 27(4):466-471 · August 1994
DOI: 10.1252/jcej.27.466

[17] Rodrıguez-Monroy, Galindo
Broth rheology, growth and metabolite production of Beta vulgaris suspension culture: a comparative study between cultures grown in shake flasks and in a stirred tank
Enzyme and Microbial Technology, Volume 24, Issue 10, 1 July 1999, Pages 687-693
DOI: https://doi.org/10.1016/S0141-0229(99)00002-2

[18] Ashoor, Zent
Maillard Browning of Common Amino Acids and Sugars
Journal of Food Science, 49: 1206-1207
DOI: https://doi.org/10.1111/j.1365-2621.1984.tb10432.x

[19] Mariana Ruiz
DNA replication or DNA synthesis is the process of copying a double-stranded DNA molecule
Wikimedia Foundation
URL: https://upload.wikimedia.org/wikipedia/commons/8/8f/DNA_replication_en.sv

[20] Kelvinsong
mRNA is created in the nucleus
Wikimedia Foundation
URL: https://upload.wikimedia.org/wikipedia/commons/9/9b/MRNA.svg

[21] Kelvinsong
Diagram of RNA translation
Wikimedia Foundation
URL: https://upload.wikimedia.org/wikipedia/commons/4/44/Protein_synthesis.svg

[22] Seth Miller
The genetic code
Wikimedia Foundation
URL: https://upload.wikimedia.org/wikipedia/commons/d/d6/GeneticCode21-version-2.svg

[23] Brewster, Jones, Phillips
Tuning Promoter Strength through RNA Polymerase Binding Site Design in Escherichia coli
PLoS Comput Biol 8(12): e1002811
DOI: https://doi.org/10.1371/journal.pcbi.1002811

[24] Bintu, Buchler, Phillips et al
Transcriptional regulation by the numbers: models
Current Opinion in Genetics & Development 2005, 15:116–124
DOI: 10.1016/j.gde.2005.02.007

[25] Thomas, Summers
Bacterial Plasmids
eLS 2008, Chichester: John Wiley & Sons, Ltd
DOI: https://doi.org/10.1002/9780470015902.a0000468.pub2

[26] del Solar, Espinosa
Plasmid copy number control: an ever-growing story
Mol Microbiol. 2000 Aug;37(3):492-500.
DOI: 10.1046/j.1365-2958.2000.02005.x

[27] Nicola, Andrew
E. Coli Plasmid Vectors: Methods and Applications
Methods in Molecular Biology. 235. Humana Press. pp. 1–5
DOI: 10.1385/1592594093

[28] Kittell, Helinski
Plasmid Incompatibility and Replication Control
Clewell D.B. (eds) Bacterial Conjugation. pp 223-242
DOI: https://doi.org/10.1007/978-1-4757-9357-4_8

[29] Pestka
The use of inhibitors in studies on protein synthesis
Methods Enzymol. 1974;30:261-82
DOI: https://doi.org/10.1016/0076-6879(74)30030-4

[30] Misumi, Tanaka
Mechanism of inhibition of translocation by kanamycin and viomycin: A comparative study with fusidic acid
Biochemical and Biophysical Research Communications, Volume 92, Issue 2, 29 January 1980, Pages 647-654
DOI: https://doi.org/10.1016/0006-291X(80)90382-4

[31] Nakano, Mashiko, Ogaware
Cloning of the kanamycin resistance gene from a kanamycin-producing Streptomyces species
J Bacteriol. 1984 Jan; 157(1): 79–83.
PMCID: PMC215132, PMID: 6690428

[32] Basker, Comber, Sutherland, Valler
Carfecillin: Antibacterial Activity in vitro and in vivo
Chemotherapy 1977;23:424–435
DOI: https://doi.org/10.1159/000222012

[33] Sambrook et al
Molecular Cloning: A Laboratory Manual
2001, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
ISBN-10: 0879695773, ISBN-13: 978-0879695774

[34] Champney
Binding of spectinomycin by ribosomes fromEscherichia coli
Current Microbiology, October 1990, Volume 21, Issue 4, pp 261–265
DOI: https://doi.org/10.1007/BF02092166

[35] Kehrenberg, Catry, Haesebrouck, Schwarz et al
Novel Spectinomycin/Streptomycin Resistance Gene, aadA14, from Pasteurella multocida
Antimicrob. Agents Chemother. July 2005 vol. 49 no. 7 3046-3049
DOI: 10.1128/AAC.49.7.3046-3049.2005

[36] Hou, Zeng, Lou et al
Engineering the Ultrasensitive Transcription Factors by Fusing a Modular Oligomerization Domain
ACS Synth. Biol.2018751188-1194
DOI: https://doi.org/10.1021/acssynbio.7b00414

[37] Rose et al
NGL viewer: web-based molecular graphics for large complexes
Bioinformatics. 2018 Nov 1;34(21):3755-3758
DOI: 10.1093/bioinformatics/bty419

[38] Gossen, Bujard
Tight control of gene expression in mammalian cells by tetracycline-responsive promoters
PNAS June 15, 1992. 89 (12) 5547-5551
DOI: https://doi.org/10.1073/pnas.89.12.5547

[39] Orth, Schnappinger, Hillen, Hinrichs et al
Structural basis of gene regulation by the tetracycline inducible Tet repressor–operator system
Nat Struct Biol. 2000 Mar;7(3):215-9.
DOI: 10.1038/73324

[40] Møller, Overgaard, Nielsen, Olsen et al
Relation between tetR and tetA expression in tetracycline resistant Escherichia coli
BMC Microbiol. 2016; 16: 39.
DOI: 10.1186/s12866-016-0649-z

[41] Brunwasser-Meirom, Pollak, Goldberg, Amit et al
Using synthetic bacterial enhancers to reveal a looping-based mechanism for quenching-like repression
Nature Communications volume 7, Article number: 10407 (2016)
DOI: 10.1038/ncomms10407

[42] Politi, Pasotti, Zucca, Magni et al
Half-life measurements of chemical inducers for recombinant gene expression
J Biol Eng. 2014; 8: 5.
DOI: 10.1186/1754-1611-8-5

[43] Schnider-Keel, Seematter, Maurhofer, Keel et al
Autoinduction of 2,4-Diacetylphloroglucinol Biosynthesis in the Biocontrol Agent Pseudomonas fluorescens CHA0 and Repression by the Bacterial Metabolites Salicylate and Pyoluteorin
J Bacteriol. 2000 Mar;182(5):1215-25.
PMID: 10671440, PMCID: PMC94405

[44] Stanton, Siciliano, Ghodasara, Voigt et al
Systematic Transfer of Prokaryotic Sensors and Circuits to Mammalian Cells
ACS Synth Biol. 2014 Dec 19; 3(12): 880–891.
DOI: 10.1021/sb5002856

[45] Guzman, Belin, Carson, Beckwith
Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter
J Bacteriol. 1995 Jul;177(14):4121-30.
PMID: 7608087, PMCID: PMC177145

[46] Lobell, Schleif
DNA looping and unlooping by AraC protein
Science. 1990 Oct 26;250(4980):528-32.
PMID: 2237403

[47] Soisson, MacDougall-Shackleton, Schleif, Wolberger
Structural basis for ligand-regulated oligomerization of AraC
Science. 1997 Apr 18;276(5311):421-5
PMID: 9103202

[48] Schleif
AraC protein, regulation of the L -arabinose operon in Escherichia coli, and the light switch mechanism of AraC action
FEMS Microbiol Rev 34 (2010) 779–796
DOI: 10.1111/j.1574-6976.2010.00226.x

[49] Zhang, Schleif
Catabolite Gene Activator Protein Mutations Affecting Activity of the araBAD Promoter
J Bacteriol. 1998 Jan; 180(2): 195–200.
PMCID: PMC106871, PMID: 9440505

[50] Lawson, Swigon, Murakami, Ebright et al
Catabolite activator protein (CAP): DNA binding and transcription activation
Curr Opin Struct Biol. 2004 Feb; 14(1): 10–20
DOI: 10.1016/j.sbi.2004.01.012

[51] Busby, Ebright
Transcription activation by catabolite activator protein (CAP)
J Mol Biol. 1999 Oct 22;293(2):199-213
DOI: 10.1006/jmbi.1999.3161

[52] Balleza, Kim, Cluzel
Systematic characterization of maturation time of fluorescent proteins in living cells
Nature Methods volume 15, pages 47–51 (2018)
DOI: 10.1038/nmeth.4509

[53] Lei, Larson, Gilbert, Lim1 et al
Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression
Cell. 2013 Feb 28; 152(5): 1173–1183
DOI: 10.1016/j.cell.2013.02.022

[54] Briner, Donohoue, Gomaa, Haurwitz et al
Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality
Molecular Cell, Volume 56, Issue 2, p333–339, 23 October 2014
DOI: https://doi.org/10.1016/j.molcel.2014.09.019

[55] Zuker
Mfold web server for nucleic acid folding and hybridization prediction
Nucleic Acids Res. 31 (13), 3406-15, (2003)
PMCID: PMC169194, PMID: 12824337

[56] O'Brien
The Apollo Guidance Computer: Architecture and Operation
Springer Science & Business Media June 25, 2010
ISBN 9781441908773

[57] Brocchieri, Karlin
Protein length in eukaryotic and prokaryotic proteomes
Nucleic Acids Res. 2005 Jun 10 33(10):3390-400 table 2 pp.3392-3
DOI: 10.1093/nar/gki615

[58] Green, Silver, Collins, Yin
Directed Evolution of a Panel of Orthogonal T7 RNA Polymerase Variants for in Vivo or in Vitro Synthetic Circuitry
ACS Synthetic Biology, 4(10), 1070–1076
DOI: 10.1021/sb500299c

[59] Meyer, Ellefson, Ellington
Toehold Switches: De-Novo-Designed Regulators of Gene Expression
Cell 159, 1–15, November 6, 2014
DOI: http://dx.doi.org/10.1016/j.cell.2014.10.002

[60] Wang, Kitney, Joly, Buck
Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology
Nature Communications volume 2, Article number: 508
DOI: 10.1038/ncomms1516

[61] Gootenberg, Abudayyeh, Lee, Zhang et al
Nucleic acid detection with CRISPR-Cas13a/C2c2.
Science. 2017 Apr 28;356(6336):438-442
DOI: 10.1126/science.aam9321

[62] Gootenberg, Abudayyeh, Kellner, Zhang et al
Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6
Science. 2018 Apr 27;360(6387):439-444
DOI: 10.1126/science.aaq0179

[63] Jeffery
Frozen Permanents of E. coli Cultures
Nature Protocol Exchange, Published online 14 June 2011
DOI: 10.1038/protex.2011.228

[64] QIAGEN
QIAprep Miniprep Handbook
Fourth Edition, June 2015
Link: tinyurl.com/y8jzfbhj

[65] Cox, Madsen, Beal, Wipat et al
Synthetic Biology Open Language Visual (SBOL Visual) Version 2.0
Journal of Integrative Bioinformatics, 2018
DOI: 10.1515/jib-2017-0074

[66] Beal, Aleritsch, Voigt et al
Communicating Structure and Function in Synthetic Biology Diagrams
ACS Synth. Biol.2019
Manuscript ID: sb-2019-001397

Acknowledgment

I would like to thank the BMBF for providing funding, Daniela Ziegler for sparking the NAND gate idea by letting me know about that Briner paper, the Open Source community that maintains Wikipedia, Sci-Hub, LibGen and Leo, especially Aaron Swartz and Alexandra Elbakyan, my S.O. K.B., best friends M.S., C.M and C.R for proof-reading and their never-ending support as well as Dr. Joachim Diener, Prof. Joachim Heinzl and Prof. Angelika Görg who made this work possible in the end.

Punish your enemies,
oh lord of the night,
DESTROY THEM ALL!