Biosafety Implications in the Context of Semisynthetic Bacterial Genomes

iGEM CeBiTec-Bielefeld 2017

CONTENTS

INTRODUCTION

Many teams need to deal with biosafety issues during the competition. *Is our organism hazardous? What will happen if it escapes from lab?* For this reason, the iGEM HQ offers various opportunities to obtain information like the safety policies, safe lab work guidelines or the white list of allowed organisms. Further, many teams are working on the development of multilateral biosafety systems to prevent the release or at least the survival of genetically modified organisms (GMOs) in the environment. With the intention to summarize all relevant levels of biosafety systems, we worked on our biosafety review article: Auxotrophy to Xeno-DNA: An Exploration of Combinatorial Mechanisms for a High-Fidelity Biosafety System for Synthetic Biology Applications. Biosafety can begin on different levels, starting from a bioreactor to a complete synthetic DNA. In our review, we describe different containment approaches previously attempted by former iGEM teams.

PHYSICAL CONTAINMENT

We define physical containment as the separation of cells from the environment due to physical materials like the wall of a bioreactor (Bakanidze, Imnadze, & Perkins, 2010; Kimman, Smit, & Klein, 2008). Besides this “full containment”, cell retainment systems like micro encapsulation offer a broad range of applications (Sewell, 1995; Gaudioso, 2004; Cohen, 2013; Curtiss III, Alexander, Alexander, Clark-Curtiss, Goldschmidt, & Hsu, 1978; Kailasapathy, 2002; Tatiya, Hedaoo, Mahulikar, & Gite, 2013; Puddu, Paunescu, Stark, & Grass, 2014; Hedaoo & Gite, 2014). Often, silica gels are used for encapsulation. They build a physical barrier between cells, preventing cell aggregation and physical interaction between cells (Nassif, Bouvet, Noelle Rager, Roux, Coradin, & Livage, 2002; Miller & Selgelid, 2008).

A Physical containment strategy was developed by the iGEM team Paris Saclay 2015, to embed bacteria in silica beds by developing a novel protocol based on two publishes papers (Cenciarelli, Gabbarini, Pietropaoli, Malizia, Tamburrini, Lucovici, et al., 2013-2015; Van Aken & Hammond, 2003).

The iGEM team Paris Bettencourt 2012, developed alginate beads for application with *E. coli*. They stabilized the cells with polyethyleneimine. To demonstrate the viability of encapsulated

E. coli cells, blue white staining was applied. Beads which were stabilized showed a 1,000-fold reduced cell-release rate.

Sensors

By defining physical containment in a broader sense, it can be understood as the containment by physical parameters like temperature or light. These factors can be regulated in a bioreactor; by sensing these environmental conditions, release of bacterial cells from a controlled environment can be linked to other biosafety systems like kill switches.

Temperature

If released from the physical containment, a heat or cold shock might occur if the temperature controlled environment is not given anymore. The majority of heat shock proteins function as molecular chaperons or proteases (Singer & Soll, 1973; Berg, Baltimore, Nathans, Yober, Roblin, & Cohen, 1974; Daegelen, Studier, Lenski, Cure, & Kim, 2009; Zhao, Liu, & Burgess, 2005; Nonaka, 2006; Wade, Roa, Grainger, Hurd, Busby, Struhl, et al., 2006), while cold shock proteins allow for low temperatures by maintaining efficient RNA translation and membrane fluidity (Molin, Klemm, Poulsen, Biehl, Gerdes, & Andersson, 1987; Blattner, Plunkett, Bloch, Perna, Burland, Riley, et al., 1997; Ermolenko & Makhatadze, 2002; Horn, Hofweber, Kremer, & Kalbitzer, 2007). Methyl-accepting chemotaxis proteins are involved in temperature-dependent changes in the movement of *E. coli* (Mizuno & Imae, 1984; Maeda & Imae, 1979; Salman & Libchaber, 2007; Nara, Lee, & Imae, 1991; Guyer, Reed, Steitz, & Low, 1983; Hanahan, 1983; Smith & Johnson, 1988; William Studier, Rosenberg, Dunn, & Dubendorff, 1990).

The DNA and RNA levels offer a more direct level of temperature based control. The DNA topology, especially supercoiling, affects the transcription efficiency (Tenover, Arbeit, Goering, Mickelsen, Murray, Persing, et al., 1995; Datsenko & Wanner, 2000; Anderson, 1975; Pruss & Drlicat, 1988; Dorman & Corcoran, 2009; Dorman, 1996). Riboswitches offer a control mechanism on RNA level. They control biosynthesis through temperature mediated structural changes, like the temperature unstable hairpin loop (Altuvia, Kornitzer, Teff & Oppenheim, 1989; Schweder K Hofmann & Hecker, 1995).

Integrating those temperature sensing RNA molecules has been accomplished by several iGEM teams (Schweder, Schmidt, Herrman, Neubauer, Hecker, & Hofmann, 1992; Kimman, Smit, & Klein, 2008; Guan, Schmidt, Pei, Wei, & Ma, 2013; Vilanova & Porcar, 2014). Team

Paris Saclay, 2015, used an RNA based temperature sensor technology with upper (42°C) and lower (32°C) boundaries for growth of *E.coli*. The RNA molecule ROSE (Andrup, Hielsen & Kølvraa, 1990; Chowdhury, Maris, Allain, & Narberhaus, 2006) from *Bradyrhizobium japonicum* (Narberhaus, Weiglhofer, Fischer, & Hennecke, 1996; Walton, 1981) was proposed to control the expression of three essential genes in *E. coli*. The iGEM teams TU Delft, 2008, and NCTU Formosa, 2011, used the RNA molecule FourU from *Salmonella entericaserovar* Typhimurium M556. This temperature sensing RNA molecule controls the expression of essential genes, permitting *E. coli* to grow only at 37 °C.

Light

Due to the usually dark environment inside a bioreactor, light sensors (Dröge, Pühler, & Selbitschka, 1998; Snow, Andow, Gepts, Hallerman, Power, & Tiedje, 2005; Schwarz, Bohne, Wang, Cejudo, & Nickelsen, 2012; Taylor & Zhulin, 1999) might be applicable in microbial biosafety systems when coupled with kill switches.

One approach is the exchange of the oxygen-sensing module of the histidine kinase FixL (Chan, Lee, Cameron, Bashor, & Collins, 2015; Gilles-Gonzales, Ditta, & Helinski, 1991) of *B.japonicum* (iGEM, 2017; Fischer, 1994) with the LOV photo sensor module of YtvA from *B. subtilis* (Guan, Schmidt, Pei, Wei, & Ma, 2013; Losi, Polverini, Quest, & Gärtner, 2002). This results in a histidine kinase YF1, which can be inhibited by blue light. Another approach is the use of phytochromes as a membrane-bound extracellular sensor for light detection regulating transcription (Vilanova & Porcar, 2014; Yeh, 1997). A synthetic biology approach involves the fusion of this photoreceptor and an intracellular *E. coli* histidine kinase domain (Wright, Delman, Stan, & Ellis, 2015; Levskaya, Chevalier, Tabor, Simpson, Lavery, Levi, et al., 2005).

The iGEM teams of Uppsala, 2011, and Cornell, 2011, conceptualized light-dependent biosafety systems. Cornell had a lysis cassette be expressed by green light. Upsalla, 2011, planned the combination of *ccaS* and *ccaR* on one plasmid in combination with the phycocyaobilin biosynthesis genes *ho1* and *pcyA* (BBa_K597105).

AUXOTROPHIES

An organism is regarded as auxotrophic, if it cannot synthesize every essential component, making it unable to survive or replicate without receiving the missing components from an external source (Moe-Behrens, Davis, & Haynes, 2013). Auxotrophic organisms are cost-efficient and easily applicable biocontainment systems. However, instances such as metabolic cross-feeding and overexpression pose possible risks to their reliability (Wright, Stan, & Ellis, 2013). The first auxotrophy was characterized in 1941 (Beadle & Tatum, 1941), while the first application for reasons of biosafety was developed in 1977 (Curtiss, Inoue, Pereira, Hsu, Alexander, & Rock, 1977). For practical reasons, auxotrophy systems can be categorized into three groups; strains auxotrophic for natural amino acids, strains auxotrophic for other natural components, and strains auxotrophic for xenobiotic compounds.

Strains auxotrophic for natural amino acids

As antibiotic resistance genes as selection markers are potentially dangerous for the environment, (Vidal, Finsach, Striedner, Caminal, & Ferrer, 2008; Rajasekaran, Seleem, Contreras, Purwantini, Schurig, Sriranganathan, et al., 2008; Dong, Xiang, & Shao, 2010), auxotrophies have been prominently used as substitutes for antibiotic resistance markers on plasmids. For example, the modified *E. coli* strain M15 is auxotrophic for glycine and used as part of an antibiotic-free expression system with a plasmid providing the necessary glyA gene (Vidal, Finsach, Striedner, Caminal, & Ferrer, 2008).

Strains auxotrophic for other natural components

An alternative to the auxotrophy system described above is to prevent the organism from building a cell wall. For example, *Corynebacterium glutamicum* (Tauch, Götker, Pühler, Kalinowski, & Thierbach, 2002) and *Mycobacterium smegmatis* (Chacon, Feng, Harris, Cáceres, Adams, & Barletta, 2002; Milligan, Tran, Strych, Cook, & Crause, 2007) lack the alr gene, which encodes for an alanine racemase responsible for converting L-alanine to D-alanine, an essential component of the peptidoglycan layer. Yet, such natural auxotrophies are still potentially hazardous, as the auxotrophic strain may come in contact with the required compound in the natural environment.

Strains auxotrophic for xenobiotic compounds

As an alternative for natural auxotrophies, it is possible to construct organisms auxotrophic for synthetic amino acids. Most prominently, codons of organisms have been redesigned to express an alternative genetic code and thus, enhance biosafety. For example, the TAG codon has been recoded to a sense codon for non-canonical amino acids through an orthogonal translation system (Rovner, Haimovich, Katz, Li, Grome, Gassaway, et al., 2015). This way, the resulting organism is dependent on the corresponding non-canonical amino acid.

Auxotrophy systems in the iGEM competition

As the iGEM Headquarter neither collects nor provides auxotrophic systems, their application in the iGEM competition is relatively rare. In 2013, team Bielefeld Germany successfully developed a multi-layered biosafety system containing a D-alanine auxotrophy, by means of deleting alr and dadX. In the following year, team Bielefeld-CeBiTec further enhanced the system for antibiotic-free selection purposes.

KILL SWITCH

Another approach is to prevent the survival of a cell by five different types of bacteriolytic toxin-antitoxin systems (TA) (Unterholzner et al., 2014).

The inhibition of Type I systems is mediated by antisenseRNA (Hayes and Van Melderen, 2011; Gerdes et al., 1997), whereas the toxin gene on the sense strand as well as the antitoxin gene on the antisense strand are transcribed. The antisense transcript prevents the translation of the toxin transcript by blocking off the ribosome (Gerdes and Wagner, 2007; Van Melderen and De Bast, 2009).

Type II systems like *ccdA-ccdB* are usually organized as two genes that form an operon. The resulting toxin and antitoxin form a complex, so that the toxin has no effect to the host cell. Since the antitoxin is less stable than the toxin, a ‘point of no return’ can be reached under stress conditions, resulting in growth inhibition or cell death (Christensen et al., 2004; Aizenman et al., 1996; Amitai et al., 2004; Grady and Hayes, 2003). Not just one iGEM team made use of this system. For example, team Paris Bettencourt 2012 (Paris Bettencourt 2012) used a Col E2 toxin and Col E2 antitoxin, encoded on two different plasmids (BBa_K914001, BBa_K914002), while another plasmid lead to the degradation of the antitoxin plasmid to switch on the kill switch.

The third type of TA system is based on the interaction of antitoxin RNA with the toxin protein (Fineran et al., 2009; Samson et al., 2013), but no iGEM team made use of it so far. Recently, type IV and V TA systems have been identified like *yeeU-yeeV* (Masuda et al., 2012) or *ghoS-ghoT*, encoding a sequence-specific endoribonuclease activity which cleaves the GhoT mRNA and inhibits the translation of the toxin protein (Wang et al., 2012).

In 2013, Team Bielefeld Germany (Bielefeld Germany 2013) improved a two-part biosafety system based on barnase and a D-alanine auxotrophic strain. Both parts control each other. In case of an inactive auxotrophy part, the barnase is expressed and will lead to cell death.

More complex kill switches like *Deadman* and *Passcode* were recently developed (Chan et al., 2015). Both systems are based (Bielefeld Germany 2013) on one or multiple synthetic molecules necessary for the cell´s survival. The *Deadman* switch uses anhydrotetracycline (aTc) as the synthetic molecule, which prevents inhibition of *lacI* transcription by TetR. LacI prevents cell death by inhibiting the transcription of the toxin. The *Passcode* system follows another approach by using three different 'input signals' (galactose, cellobiose, IPTG). Input A and B together lead to the inhibition of LacI-ScrR expression, which represses the toxin expression. Input C could inhibit LacI-ScrR, so the cell survives only in presence of galactose and cellobiose and in absence of IPTG.

SELF DESTRUCTION

Although Kill switch systems are widely used as biosafety systems, genetic information can potentially be released and spread through HGT depending in the TA type (I to V) (Dröge, Pühler, & Selbitschka, 1998; Velkov, 2001; Thomas & Nielsen, 2005; Wright, Delmans, Stan, & Ellis, 2015). The uptake and retention of genetic information usually correlates with an evolutionary advantage. Adapting a combination of toxin-antitoxin systems and self-destruction systems prevents the spread of recombinant DNA by the degradation of nucleic acids, while the targeted cells are killed. Systems which are based on the destruction of nucleic acids can be subdivided into three systems depending in the sequence specificity of used system.

Specific systems are based on nucleases which hydrolyze nucleic acids at sequence specific sites know as restriction sites. These are toxin-antitoxin systems based in nucleases applied to kill cells with minor risk of DNA leakage. The nuclease (toxin), encoded in the plasmid, is

constitutively expressed, while the expression of nuclease- methyltransferase (antitoxin), encoded on the chromosome, is induced (Gallagher, Patel, Interiano, Rovner, & Isaacs, 2015). A known toxin-antitoxin system is EcoRI/EcoRI methyltransferase (Greenes, Gupta, Boyer, Francisco, Francisco, Brown, et al., 1980; Willliams, 2003; Roberts, Belfort, Bestor, Bhagwat, Bickle, Bitinaite, et al., 2003; Roberts, 2005; Roberts, Vincze, Posfai, & Macelis, 2007), which is also used by the iGEM team University College London 2012 (University College London, 2012). They developed a threefold active biological containment system in combination with holin/anti-holin endolysin and colicin-E3/colicin Immunity E3 (BBa_K729009, BBa_K729010) to minimize HGT via bacterial conjugation.

Unspecific systems utilize nucleases which hydrolyze DNA and are not sequence specific. They are combined with an inducible promoter, a ribosome binding-site and a start codon, introducing single strand breaks (Molin, Boe, Jensen, Kristensen, Givskov, Ramos, et al., 1993; Ahrenholtz, Lorenz, & Wackernagel, 1994). One example is nucA. The team TU Munich 2013 (TU Munich, 2013) used the thermonuclease NucA of Staphylococcus aureus (BBa_K1159105) to degrade DNA of a genetically modified moss. Team NTU-LIHPAO-Taiwan 2015 (NTU LIHPAO Taiwan, 2015) also used the thermonuclease NucA of S. aureus (BBa_K1159105) under control of a lambda cI-regulated promoter (BBa_R0051) to degrade DNA, thus killing Lactobacillus casei if the cl protein (BBa_C0051) concentration is low. The system is designed to inhibit HGT from L.casei to bacteria in the human gastrointestinal tract while controlling the proliferation of the cells.

The third system is based in a combination of sequence specific and unspecific parts like the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) system (Bhaya, Davison, & Barrangou, 2011; Terns & Terns, 2011; Wiedenheft, Sternberg, & Doudna, 2012; Gasiunas, Barrangou, Horvath, & Siksnys, 2912; Jinek, Chylinski, Fonfara, Hauer, Doudna, & Charpentier, 2012). The sequence-specific CRISPR is used to guide the sequence-unspecific Cas nuclease to its target, thus making this system highly regulated without the risk of uncontrollable cleavage (Marraffini & Sontheimer, 2008; Gomaa, Klumpe, Luo, Selle, Barrangou, & Beisel, 2014; Caliando & Voigs, 2015). The team HKU Hongkong 2015 (HKU Hongkong, 2015) designed a CRISPR/Cas9 system with a specific sgRNA (BBa_K1774000) to target the DNA polymerase III alpha subunit (dnaE) thus inhibiting replication of the bacteria.

SEMANTIC CONTAINMENT

One important approach of semantic containment to prevent the spread of recombinant DNA fragments is xenobiology. It describes new biological systems with variations in their genetic information like DNA and RNA or its expression products through changes of amino acids, while those systems are included in orthogonal systems (Schmidt, 2010). It restrains horizontal gene transfer (HGT) (Plos et al., 1989) through transduction (de la Cruz and Davies, 2000), conjugation, and transformation (Davison, 1999). So Xeno-DNA (XNA) could be an important biosafety tool by preventing HGT as it is inaccessible for wild-type DNA processing enzymes (Schmidt and de Lorenzo, 2012).

The most interesting approach of XNA is to replace or expand the standard genetic code (Martinot and Benner, 2004; Jiang and Seela, 2010; Kwok, 2012; Zhang et al., 2017; Yamashige et al., 2012; Seela et al., 2005; Switzer et al., 1989; Yang et al., 2011). First conceptions extend the four nucleotide genetic code by replacing thymidine with 5-chlorouracil (Dunn and Smith, 1957; Marlière et al., 2011). While other approaches deal with the insertion of artificial bases like dP (2-amino-8-(1‘-β-D-2‘-deoxyribofuranosyl)-imidazol[1,2-α]-1,3,5-trizan-4(H)-one), and dZ (6-amino-5-nitro-3-(1‘-β-D-2‘-deoxyribofuranosyl)-2(1H)-pyridone) (Yang et al., 2010), which pair with three hydrogen bonds. The concept of DNA pairing systems can be extended by pairings dependent on metal ion coordination (Takezawa and Shionoya, 2012; Atwell et al., 2001; Weizman and Tor, 2001) or hydrophobic interactions (Seo et al., 2009). One example of base pairing through hydrophobic interactions are d5SICS – dMMO2 and d5SICS-dNaM (Seo et al., 2009). The Romesberg group worked even on the nucleoside triphosphate transporter for dNAM and d5SICS and a CRISPR-Cas retention system (Malyshev et al., 2014) (Zhang et al., 2017).

Equally important are approaches concerning the change of DNA topology. The benzo expansion of pyrimidines creates dxT and dxC, therefore results in expanded DNA (xDNA), with a more stable helix to heat and an increased size about 2.4 Å (Liu et al., 2003, 2004),

Alternative methods are a modified backbone by integrating substitutes for deoxyribose and ribose, while the altered backbone needs to meet the requirement to build a functioning helix that does not interact with natural replication enzymes, instead requiring adjusted or even synthetic enzymes (Herdewijn and Marlière, 2009).

Changing a stop codon into a sense codon has been done by various groups (Heider et al., 1992; Blight et al., 2004; Mukai et al., 2010; Ohtake et al., 2012; Lajoie et al., 2013b; Miller

et al., 2015; Crnković et al., 2016; Miller et al., 1979). Especially the amber codon (UAG) is one interesting candidate. The Church lab presented an organism in which all 314 amber stop codons were replaced by ochre stop codons. The deletion of release factor 1 (encoded by *prfA*), which recognizes UAG and UAA, then allows for recoding of the amber stop codon (Isaacs et al., 2011; Lajoie et al., 2013a). Through an engineered aminoacyl-tRNA synthetase (aaRS) and corresponding tRNA, a new orthogonal translation system was created (Liu and Schultz, 2010). The HGT of sequences encoding an aaRS and the corresponding tRNA are either lethal or very detrimental in natural, non-recoded organisms as they will lead to mistranslation of amber containing genes. Also, shuffle codes are a thinkable way. They are used to encode the identical polypeptide as expected in the wild type, but are based on tetranucleotides (Magliery et al., 2001).

Two iGEM teams worked on such projects quite successfully. In 2012, team Paris Bettencourt (Paris Bettencourt 2012) worked on a far-reaching biosafety project. The semantic containment based in an amber mutation in the gene conferring a kanamycin resistance (BBa_P1003). This system aimed at the prevention of the expressed antibiotic resistance gene in wild type cells after an HGT event. In 2016, team TU Darmstadt (TU Darmstadt 2016) combined auxotrophic incorporation of a non-canonical amino acid and a reporter for low levels of the ncAA (BBa_K1416000, BBa_K1976025) (Zhang et al., 2005) designed by the team Austin Texas 2014 (Austin Texas 2014).

MINIMALGENOME

The minimal genome is the minimal set of genes which are essential for growth and survival of the cell (Mushegian, 1999). Most approaches rely on transposon mutagenesis (Smalley et al., 2003) or antisense RNA (Forsyth et al., 2002; Ji et al., 2001) to identify essential genes (Juhas et al., 2011). It can be constructed via ’top down‘, while redundant genes are deleted systematically (Maniloff, 1996) or ’bottom up‘ approaches, which describes the combination of single genes until the minimal required set is reached (Maniloff, 1996).

The most popular work on the minimal genome is the *Mycoplasma mycoides* JCVI-syn1.0 (Gibson et al., 2010) by the J.C. Venter lab. An synthetic 1.08Mbp genome sequence of *M. mycoides* was assembled and transplanted into an *M. capricolum* recipient cell (Hutchison et al., 2016). Based on JCVI-syn1.0, the genome reduction was achieved in 2016 in the new synthetic 532 kbp genome of JCVI-syn3.0 (Hutchison et al., 2016).

As biosafety tool, a minimal genome offers various applications, for example cells would depend on complex media as well as on stable conditions.

CELL FREE SYSTEMS

From a biosafety perspective, the possible release of GMOs into the environment and HGT between engineered and wildtype organisms are of major concern. Most of these risks can be eliminated by the application of cell-free protein synthesis (CFPS). By now the feasibility of CFPS has been shown for a variety of applications, such as synthetic biology, vaccine production and protein engineering (Hodgman and Jewett, 2012; Yang et al., 2005; Smith et al., 2013).

Two main strategies exist for CFPS. The first one is based on crude cell extracts from the desired cells. The necessary crude extracts themselves are easy to prepare, but fast energy depletion and degradation by proteases and nucleases pose two big problems (Kitaoka et al., 1996; Matveev et al., 1996; Shimizu et al., 2001). These limitations can be circumvented using the PURE ("protein synthesis using recombinant elements") system developed by Shimizu et al. 2001. This cell-free system is based on purified (His)-tagged translation factors and can be programmed by natural mRNA (Shimizu et al., 2001). To effectively remove all living cells before deploying a CFPS system outside of the lab, efficient cell lysis is important. Standard methods for the preparation of cell-extracts are already highly effective, but they still do not provide a completely cell-free extract (Shrestha et al., 2012; Smith et al., 2015). The aforementioned PURE system can be deployed to fully circumvent the risk of an unsterile cell-extract. This system is not only safer, but it was also shown that the energy consumption is much lower when compared to S30 cell extract systems (cell extracts cleared from heavier components by centrifugation at 30,000 xg), and that higher productivities can be reached.

Cell-free systems have been deployed by several iGEM teams for their projects. Bielefeld-CeBiTec 2015 used cell extracts of *E. coli* KRX and ER2566 strains to produce sfGFP for their paper-based biosensor (Bielefeld-CeBiTec 2015). Teams Edinburgh 2015 and Exeter 2015 also worked on the development of a biosensor using cell-free protein synthesis. The desired enzymes were expressed in *E. coli* BL21 by Edinburgh 2015 and fused to cellulose-binding domains (Edinburgh 2015). To detect bovine tuberculosis, Exeter 2015 developed a biosensor using cell-free systems. A commercially available S30 cell-free kit was used to

express GreenFET1J as a response to a trigger RNA (Exeter 2015). Furthermore, Team Freiburg 2015 tried to build a microchip for simultaneous detection of several infectious diseases (Freiburg 2015).

As shown above, our review provides an easier overview about possible applications of biosafety for the first time and we hope future teams can adapt different systems or a combination of systems to their project by our simplified overview.