

The Use of Non-Canonical Amino Acids as a Novel Biocontainment Strategy

Jourdan Witt*

Abstract

Advancements in synthetic biology have led to the use of genetically modified organisms in research and industrial fields. Bacteria were one of the first organisms to be genetically engineered due to their fast growth and simple genetics and have emerged as a major scientific and commercial interest. For instance, modified commensal bacteria can be used as an oral delivery vector of therapeutics, or as probiotics to target specific pathogens in the gastrointestinal tract. The impact of the release of pathogens used in research or vaccine development could be catastrophic to the environment and public health. In addition, there is growing concern about using genetically modified organisms in open systems, as there is a possibility for unintentional proliferation into natural environments. Therefore, it is imperative that the environmental safety of genetically modified organisms is addressed, and that adequate biocontainment mechanisms are developed.

Keywords: Biocontainment, genetically modified organism, amino acid, auxotrophy

Genetically modified organisms (GMO) are a recent scientific advancement widely used in scientific research and industrial fields due to their improved characteristics. They include, but are not limited to, microorganisms containing targeted insertions of genes. For example, probiotics are live, non-pathogenic microorganisms that can be used to complement antibiotics for the treatment of foodborne pathogens and to promote gastrointestinal health. Bacteria used as probiotics or oral delivery vectors can be genetically modified to enhance specific characteristics, thereby improving their efficacy. In nature, bacteria use horizontal gene transfer, including transformation, conjugation, and transduction to increase genetic variability of the population.

To be safely used, the bacteria must be screened for pathogenic and virulence traits to help prevent horizontal gene transfer to natural flora. In addition, it is essential that the bacteria cannot uptake and incorporate virulence genes from pathogens that co-inhabit the gut (Mathipa & Thantsha, 2017). The recombinant nature of bacteria, as well as their ability to survive and propagate in natural environments, mean that effective biocontainment strategies must be employed.

Biocontainment refers to "the physical and procedural components of keeping biological organisms or agents within the confines of a barrier system" (Burnette, 2013). In the case of GMOs, there are two main strategies: active and

*Department of Biochemistry, Microbiology and Immunology, College of Arts and Science, University of Saskatchewan, Saskatoon, SK, Canada
Correspondence: jhw489@mail.usask.ca

passive biocontainment. Active biocontainment involves the genetic control of the microorganism via the activation of a killer gene or inactivation of an essential gene, which is controlled by the presence of a substance in the environment. While active biocontainment kills the host, it may result in the introduction of foreign DNA into the environment that can harm natural microorganisms if integrated into their chromosome. In contrast, passive biocontainment involves complementation of auxotrophy by providing either the intact, functional gene or an essential metabolite (one that is absent or limited in the environment). A disadvantage of this method is that it has a bacteriostatic effect, temporarily preventing the replication of bacteria without necessarily killing it (Lee, 2010). Therefore, an ideal biocontainment method would be one that bypasses these limitations.

Ross *et al.* (1990) incorporated the advantages of active and passive biocontainment to create thymine deficient strains of *Lactococcus lactis*. The researchers disrupted the essential gene *thyA*, which produces an enzyme that catalyzes the production of an intracellular de novo source of deoxythymidine monophosphate (dTMP). Thus, the bacteria would have to uptake extracellular thymidine or thymine to survive. The major downfall of this method is that it relies on the fact that thymine levels are limited in the environment. Biocontainment of these strains would be easy to achieve in a laboratory but would not transfer well to a real-world situation. Biological containment systems have emerged that exploit auxotrophy for substances completely absent in nature, such as auxotrophy for non-canonical amino acids (ncAA), which will be discussed in further detail below.

A popular example of genetic engineering comes from the book *Jurassic Park*, in which scientists genetically engineer dinosaurs to become auxotrophic for the amino acid lysine.

The lysine contingency is intended to prevent the spread of the animals in case they ever got off the island. Dr. Wu inserted a gene that creates a single faulty enzyme in protein metabolism. The animals can't manufacture the amino acid lysine. Unless they're continually supplied with lysine by us, they'll slip into a coma and die. (Crichton, 1991).

The problem with the scientist's theory is that lysine is already an essential amino acid that cannot be produced by the animal and must be provided by the environment.

Therefore, the dinosaurs could acquire lysine from plants and animals outside of the park if they escaped. Nonetheless, this piece of pop culture brings to light the importance of biocontainment for safety reasons. Similar to the modification of dinosaur genes, bacteria can be genetically modified to become auxotrophic for a particular substance, and thus would not survive in the natural environment where that substance is in low abundance.

The genetic code is the set of rules that determines how the sequence of nucleotides in deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) is translated into proteins and is conserved in all domains of life. After the genetic information contained in DNA is transcribed to messenger RNA (mRNA), it is decoded by the translational system into a functional protein. The code specifies that groups of three bases in mRNA are read as triplets called codons. To initiate translation, a preinitiation complex (containing initiation factors and the small ribosome subunit) and a methionine carrying initiator transfer RNA (tRNA) assemble on the mRNA at the AUG start codon. Once this initiation complex is formed, the large subunit of the ribosome binds, and the elongation phase begins (Snyder *et al.* 2013). An aminoacylated tRNA has an anticodon sequence in the tRNA loop that recognizes a specific three nucleotide codon found on the mRNA. There are 64 sense codons, 61 of which encode for the canonical amino acids, and 3 of which encode for the stop codons UAA (ochre), UGA (opal), and UAG (amber) (Acevedo-Rocha & Budisa, 2016). For every amino acid, there is at least one tRNA. Therefore, the genetic code can be described as redundant, because there can be multiple codons that code for the same amino acid (Lagerkvist, 1978). Aminoacyl tRNA synthetase is an enzyme that is responsible for the covalent addition of the amino acid to the tRNA (Snyder *et al.*, 2013). Most cells have an individual synthetase specific for each amino acid. The genetic code can be modified to incorporate ncAA and can be used as a tool for biocontainment. Conveniently, the limited promiscuity of synthetases means that the 20 canonical amino acids (cAA) can be replaced with closely related ncAA. Non-canonical amino acids can be incorporated either through selective pressure incorporation or stop codon suppression (figure 1). Selective pressure incorporation results in a site-specific incorporation by using the endogenous translational system to incorporate the ncAA, a close synthetic analog to the canonical amino acid. In contrast, stop codon suppression uses a novel tRNA and aminoacyl tRNA synthetase pair that are introduced into the organism to incorporate the ncAA into a tRNA that recognizes the amber stop codon (Acevedo-Rocha & Budisa, 2016).

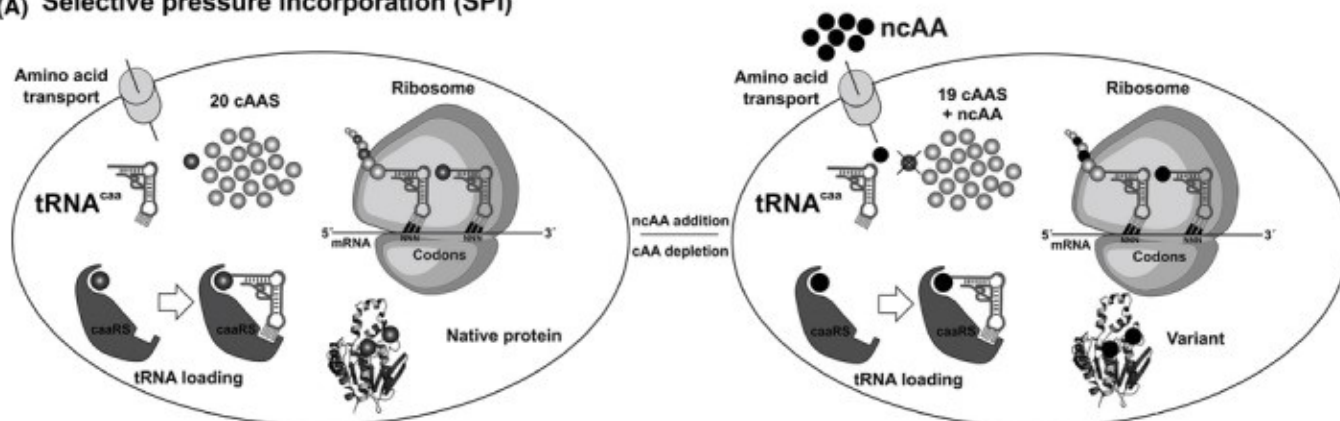
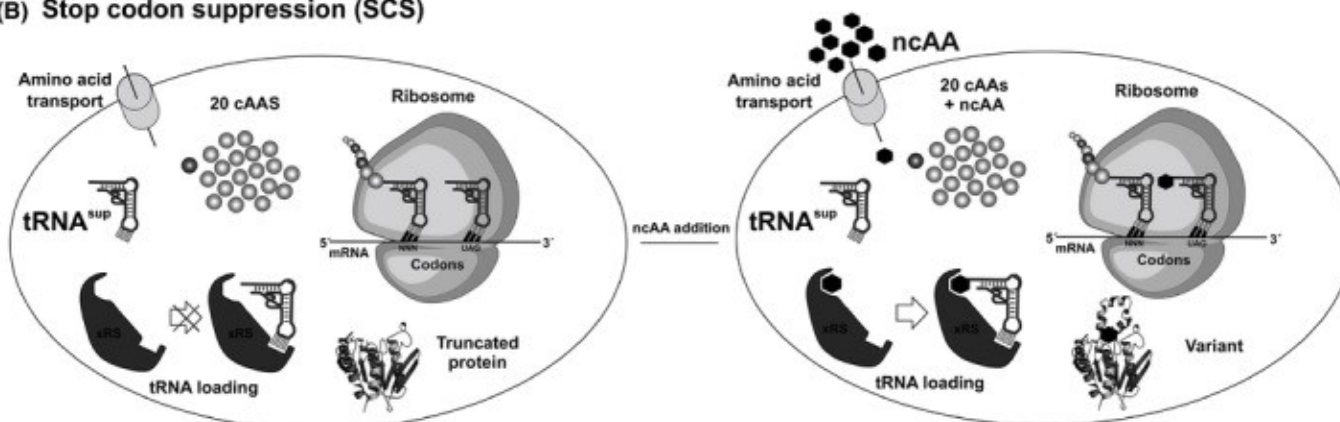
(A) Selective pressure incorporation (SPI)**(B) Stop codon suppression (SCS)**

Figure 1. Illustration depicting selective pressure incorporation and stop codon suppression methods to introduce non-canonical amino acids. Reproduced with permission from Acevedo-Rocha & Budisa (2016).

In stop codon suppression, the endogenous translation system must be altered to recognize and accept the ncAA. A novel tRNA that recognizes the codon, as well as an aminoacyl-tRNA synthetase, are required. For biocontainment purposes, it is essential that this system is orthogonal to the host's endogenous tRNAs and amino-acyl tRNA synthetases. The first condition of orthogonality is that the three-nucleotide codon assigned to the ncAA cannot already encode for a cAA. Therefore, stop codons such as UAG (amber) are often assigned. Second, the novel synthetase, but never the endogenous synthetase, can aminoacylate the novel tRNA. Finally, the novel tRNA only binds to the codon allocated to the novel amino acid (Wang *et al.* 2001).

The efficiency of stop codon suppression is limited by the release factors (RF) of the endogenous translational machinery, which are responsible for the termination of translation by recognizing stop codons on the mRNA. Prokaryotic translation is mediated by three release factors:

RF1, RF2, and RF3. RF1 recognizes UAA (ochre) and UAG (amber) stop codons, while RF2 recognizes UAA and UGA (opal) stop codons. If RF1 is present, it will interfere with the addition of the ncAA at the amber stop codon. Even though the gene for RF1 is essential in *Escherichia coli*, Johnson *et al.* (2011) showed that RF1 could be knocked out, provided that RF2 expressed constitutively in the cell. An important consideration when reassigning the amber codon is the subset of endogenous genes that use this stop codon to terminate translation in host cells. In this system, the tRNA/synthetase pair will incorporate the ncAA, extending the endogenous proteins. The researchers found that in the absence of RF1, the endogenous amber codon is in fact suppressed, since most endogenous genes ending in a UAG codon were followed by a secondary in-frame UAA or UGA codon, while others were followed by a transcriptional terminator such as a hairpin.

The endogenous proteins were extended briefly, but this had minimal effects on the bacteria, only causing a decrease in fitness and growth rates. To circumvent this problem, it has been suggested that endogenous UAG codons could be reassigned to UGA through further genetic modification, which would allow translation to be terminated at the normal position by RF2 (Johnson *et al.* 2012).

Typically, directed evolution is used to create the novel tRNA and aminoacyl tRNA synthetase by encoding them on a plasmid. Then, they can be mutated either through error-prone PCR or the use of degenerate primers. Error-prone PCR uses a Taq polymerase that does not have a proof-reading capability, and the fidelity of the incorporation of nucleotides is controlled by altering the reaction buffer. Likewise, degenerate primers can be used to increase the mutation rate; they are a population of oligonucleotide sequences in which some positions can contain several possible bases. They cover all possible nucleotide sequence combinations for the targeted protein (Iserte *et al.* 2013). Following mutation, the library of tRNA/synthetase pairs can be examined via positive and negative selection to find those that display orthogonality. In the positive selection process, the plasmid is transferred to bacterial cells that contain genes that confer antibiotic resistance only in the presence of the ncAA. If a proper and functional orthogonal pair is expressed, the ncAA will be incorporated into an antibiotic resistance gene. The gene will be functional, and confer resistance to the antibiotic, allowing the bacteria to survive. In negative selection, the plasmid is transferred to cells that contain an essential gene that contains a premature amber stop codon. In the presence of the ncAA, the novel tRNA/synthetase pair successfully incorporates the ncAA, resulting in an essential, functional gene and the survival of the bacteria (Kato *et al.* 2015).

Wang *et al.* (2001) aimed to expand the genetic code of *E. coli* by incorporating tRNA^{tyr} and tyrosyl-tRNA synthetase (TyrRS) from another organism in an attempt to avoid cross-aminoacylation while maintaining and enhancing orthogonality. A tRNA^{tyr}/TyrRS pair from the archaeobacteria *Methanococcus jannaschii* was chosen due to the ability to charge the amber suppressor tRNA with high efficiency. If the tRNA^{tyr} was to be used to deliver ncAA to growing polypeptide chains in *E. coli*, there was the possibility that the endogenous *E. coli* synthetases could improperly charge it with cAA. Therefore, it was important for the researchers to improve the orthogonality of the tRNA^{tyr}/TyrRS pair. The researchers knew the *M. jannaschii* TyrRS lacks most of its anticodon binding site and used this information to create a library of mutations in specific residues of the anticodon loop of the tRNA. By doing so, they avoided the risk that the

mutation could interfere with the recognition of tRNA^{tyr} by TyrRS or the translational machinery.

Following mutation, the researchers used a different negative selection method than described above. Rather than insert an amber nonsense codon into an essential gene, it was introduced into a toxic gene that encodes for a bacterial ribonuclease (barnase). Barnase is toxic to *E. coli* in the absence of the inhibitor barstar. If the tRNA^{tyr}/TyrRS pair was not orthogonal, *E. coli* synthetases could aminoacylate the tRNA, leading to the suppression of the amber codon. The toxic gene product would be produced, resulting in cell death. The surviving population was used in a positive screen to select for functional, orthogonal pairs, in which an amber codon was inserted into an antibiotic resistance gene. Surviving cells were resistant to ampicillin and contained tRNAs that were efficiently amino-acylated by the *M. jannaschii* TyrRS (Wang *et al.* 2001).

More recently, Kato *et al.* (2015) constructed an *E. coli* strain that was auxotrophic for the ncAA 3-iodo-L-tyrosine (IY). The strain contained two plasmids encoding for a toxin-antidote system: one with an amber codon inserted next to a synthetic antidote gene, and the other expressing a novel tRNA-synthetase pair specific for the amber codon. In the presence of the IY, the strain can express the synthetic antidote gene by incorporating the amino acid at the site of the inserted amber codon. This gene encodes for an antidote protein called ImmE3, which inhibits the toxin via the formation of a complex. The toxin gene *colE3* encodes for colicin E3 protein, a potent RNase that cleaves 16S ribosomal RNA (rRNA), effectively eliminating the sequence that is complementary to the Shine-Dalgarno sequence of mRNA. The 30S subunit of the ribosome cannot bind, which inhibits protein synthesis. Therefore, the bacteria will not be able to survive outside of the laboratory environment where IY is not present and will die due to the effects of ColE3.

To confirm the success of the toxin-antidote system in conferring auxotrophy to IY, the researchers inoculated the bacteria into either IY-containing or IYfree medium. It was found that in the absence of IY, no viable colonies were detected. In the presence of IY, however, 8.5×10^3 colony forming units were counted, indicating that IY is required for the survival of the bacteria (Kato *et al.* 2015). If this system is to be used for biocontainment, it is important that the survival of the bacteria can be controlled after it has served its purpose. To determine if the bacteria can survive for long periods of time in the absence of the auxotrophic substance, the researchers examined the rate of killing after IY is removed from the media. It was found that the survival rate of the bacteria decreased significantly after two hours, and

the half-life was determined to be approximately 50 minutes (Kato *et al.* 2015). In addition to these factors, it is important that the rate of mutation in response to IY is low to prevent the escape of bacteria and unwanted proliferation. The researchers used a fluctuation assay to determine that the frequency of mutation was 1.4 per 10^5 cell divisions, meaning that one mutation would occur every 71 thousand cell divisions (Kato *et al.* 2015). Given that the doubling time for the *E.coli* strain used in these experiments is about 60 minutes, the researchers suggest the bacteria would be killed off in the absence of IY before a successful escape mutation could occur. This genetically engineered auxotrophy for an amino acid not present in nature represents an improved method of biocontainment that could prove to be adequate if combined with other containment systems.

GMOs, including microorganisms such as bacteria, have become a significant interest in research and industrial fields due to their enhanced features that cannot be replicated to the same extent in nature. However, there is concern about using these bacteria in open systems as probiotics or vectors for oral vaccine delivery. Due to the recombinant nature of bacteria and their ability to mutate and escape the barrier system, there is the possibility they

may unintentionally proliferate into natural environments. A biocontainment method has emerged that may help avoid the mistake made by Jurassic Park researchers: it is possible to genetically modify bacteria that are auxotrophic for a substance not found outside of containment. The addition of ncAA requires the modification of biosynthetic machinery to include a novel tRNA/codon pair and an aminoacyl tRNA-synthetase. These organisms would not survive in a natural environment where the ncAA is not present, and they have been proven to have short half-lives and a low rate of escape mutations. Most importantly, this method employs orthogonality to prevent the transfer of ncAA into natural systems via horizontal gene transfer. Genetic biocontainment strategies, such as the ones described in this paper, provide a foundation for the safer use of GMOs in open systems.

References

- Acevedo-Rocha, C., Budisa, N. (2016). *Xenomicrobiology: a roadmap for genetic code engineering*. *Microb Biotechnol.* 9(5):666-76.
- Burnette, R. (2013). *Biosecurity: understanding, assessing, and preventing the threat*. (pp. 3-4). Wiley.
- Crichton, M. (1991). *Jurassic Park*. New York: Ballantine Books.
- Iserte, J., Stephan, B., Goni, S., Borio, C., Ghiringhelli, P., Lozano, M. (2013). *Family-specific degenerate primer design: a tool to design consensus degenerated oligonucleotides*. *Biotechnol Res Int.* 2013:383646.
- Johnson, D., Xu, J., Shen, Z., Takimoto, J., Schultz, M., Schmitz, R., Xiang, Z., Ecker, J., Briggs, S., Wang, L. (2011). *RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites*. *Nat Chem Biol.* 7(11):779-86.
- Johnson, D., Wang, C., Schultz, M., Schmitz, R., Ecker, J., Wang, L. (2012). *Release factor one is nonessential in Escherichia coli*. *ACS Chem Biol.* 7(8):1337-44.
- Kato, Y. (2015). *An engineered bacterium auxotrophic for an unnatural amino acid: a novel biocontainment system*. *Peer J.* 3:e1247.
- Lagerkvist, U. (1978). "Two out of three": an alternative method for codon reading. *Proc Natl Acad Sci.* 75: 1759–1762.
- Lee, P. (2010). *Biocontainment strategies for live lactic acid bacteria vaccine vectors*. *Bioeng Bugs.* 1(1):75-7.
- Mathipa, M., Thantsha, M. (2017). *Probiotic engineering: towards development of robust probiotic strains with enhanced functional properties and for targeted control of enteric pathogens*. *Gut Pathog.* 9:28.
- Munthali, M., Timmis, K., Diaz, E. (1996). *Use of colicin E3 for biological containment of microorganisms*. *Appl Environ Microbiol.* 62(5):1805-7.
- Ross, P., O'Gara, F., Condon, S. (1990). *Thymidylate synthase gene from Lactococcus lactis as a genetic marker: an alternative to antibiotic resistance genes*. *Appl Environ Microbiol.* 56(7):2164-9.
- Wang, L., Schultz, P. (2001). *A general approach for the generation of orthogonal tRNAs*, *Cell Press.* 8(9):883-90.
- Snyder, L., Peters, J., Henkin, T., Champness, W. (2013). *Bacterial gene expression: transcription, translation, and protein folding. Molecular genetics of bacteria (4th ed.)* (pp.67-123). Washington, DC: ACM