Synthetic Biobots

Assembly methods

Synthetic
Biobots

Author: Cruz Francisco Osuna Aguirre

iGEM Design League

Season 2022

Introduction

This text explains the expression strategy, the assembly construction method and the type of transformation that is proposed to be used for the expression of the piperamide biosynthetic pathway genes in *Saccharomyces cerevisiae* by the Synthetic Biobots team during the 2022 season of the iGEM Design League competition, the type of expression system consists of the use of A2 peptides to allow the polycistronic expression of proteins, the construction method used is based on the one proposed in Liu et al. 2017 but with modifications in order to adapt it to the parts and standards used, finally the proposed transformation protocol is the one described in McCusker 2015.

A2 Peptides and Cloning

The proposed synthetic pathway that leads to the production of a piperamide present in *Piper nigrum* is made up of 11 enzymes. Carrying out its expression in yeast could require up to 11 genetic assemblies or perhaps less, but with complicated and/or inefficient systems. For this reason the use of A2 viral peptides was chosen, which generate relatively high levels of expression of downstream proteins compared to other multi-genetic expression strategies, in addition to their small size, which entails a lower risk of affecting the function of the expressed gene, these allow polycistronic expression in eukaryotic cells and can be used for the expression and translation of multiple proteins in metabolic engineering applications *in Saccharomyces cerevisiae*.[4;6] This expression strategy allows us to express all the enzymes of the metabolic pathway in only three assemblies, which ends up simplifying the entire piperamide production process.

The mechanism of operation of the A2 is described in Donnelly et al. 2001, consisting of ribosome skipping the formation of a glycyl-prolyl peptide bond at the C-terminal end of 2A, ending the translation of that peptide and continuing with the next one.

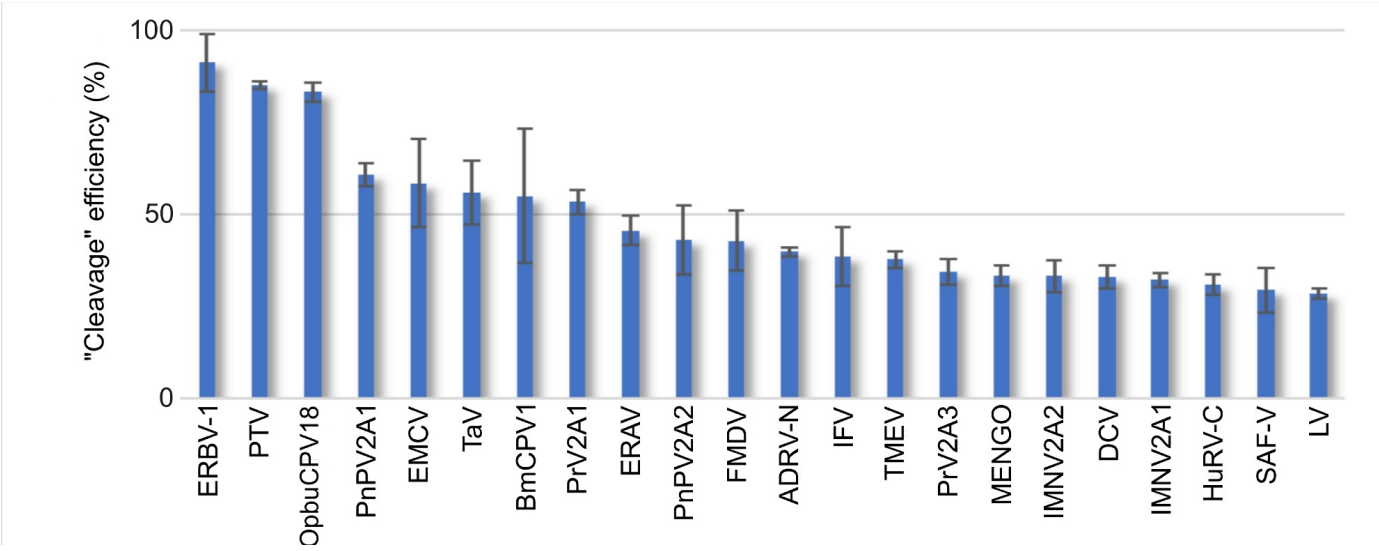

Figure 1: Cleavage efficiencies of peptides characterized in Souza-Moreira et al. 2018

There are many 2A peptides described in the scientific literature, however the functioning and efficiency of self-cleavage varies depending on the species and cellular context where it is being used. It was

found that of 22 2A peptides used for bicistronic expression in *S. cerevisiae* (Figure 1), 6 of these have a cleavage efficiency greater than 50% and three of them have high expression efficiencies (¿80%), it was decided to take the 6 previously mentioned peptides with greater efficiency as candidate peptides to be used.[6]

In addition to cleavage efficiency, how these peptides function in polycistronic arrangements of 3 and 4 proteins was taken into account, since they are also affected by factors such as how many and which other 2A peptides are present, as well as in what order they are arranged, according to Liu et al. 2017 the polycistronic arrangements with the best expression efficiency are those that carry the 2A peptide order indicated in Table 1.

Table 1: Peptide 2A order.

4-cistronic	**3-cistronic**
P2A	P2A
T2A	T2A
E2A	

Considering this and also that these peptides are among those that have a good cleavage efficiency in *S. cerevisiae*, it was decided to use the same A2 and order, their names are found in Table 2.

Table 2: Chosen A2 peptides and their abbreviation.

Name of 2A	**2A short names**
Porcine teschovirus-1 2A	P2A, PTV
Thosea asigna virus 2A	T2A, TaV
Equine rhinitis B virus 2A	E2A, ERBV-1

Assembly

In the previously mentioned research on the efficiency of polycistronic arrays (Liu et al. 2017) a standardized method for the construction of this type of expression systems is proposed, indicating which restriction sites to place between each A2 peptide and even on which vector to build it, so it was taken as a reference when making our assemblies but with some modifications to adapt it to our chassis body, to the RFC10 standard (Knight 2003) and to the pieces of our assemblies, first the restriction enzymes that flank the A2 peptides were changed, remaining as indicated in Table 3.

Table 3: Restriction sites used in 2A peptides.

2A peptide	Restriction sites
P2A	5'- NheI, 3'- AatII
T2A	5'- HindIII, 3'- XhoI
E2A	5'- BspEI, 3'- FseI

Subsequently, the list of the parts of our constructs was made and they were adapted for the assembly process and their expression in our chassis using the Benchling platform, for which the RFC10 standards and the restriction sites used in our methodology were taken into account. Therefore the restriction sites present in our sequences that were incompatible with our assembly process and RFC10 (Table 4) were eliminated, the codons of the CDS and A2 peptides were also optimized for their expression in *S. cerevisiae*.

Table 4: Restriction sites removed.

Deleted sites for RFC10	Deleted sites for assembly
EcoRI	BamHI
XbaI	NheI
SpeI	XhoI
PstI	BspEI
NotI	SalI
	HindIII
	AatII

PCR

After the optimization of the parts (their names and codes are indicated in the table 8). We proceeded to make the primers of the assembly pieces, to these we placed the restriction sites that we chose to be able to assemble our construct and to the forward of the promoters and reverse of the terminator, RFC10-like prefix and suffix sequences were added in order to create a biobrick (Table 5).

Table 5: Prefix and suffix sequences.

Prefix	Subfix
5' GAATTCGCGGCCGCTTCTAGAG '3	5' TACTAGTAGCGGCCGCTGCAG '3

For our assemblies, two different promoters were used, a constitutive promoter (pAdh) and an inducible one (pGAL1), the first was placed in constructs one and two, instead the inducible promoter was placed in construct number two. To these promoters we directly added the sequence of the RBS used at its 3' end because the primers of this piece were larger than the RBS sequence itself, so it was considered unnecessary to make a PCR product that would require different temperatures for its amplification and therefore a PCR run separated from the one that would take place with the other parts, to later unite it

again. Two versions of the terminator were also made, which differ in the restriction site present at its 5' end, the first (with SalI) was used for constructions 1 and 3, while the second (with BspEI) was used in the second assembly, this was because it differs in the restriction site to which it joins at its 5' end as there are fewer pieces.

After the creation of the primers of the pieces to be used (whose names and codes are found in Table 6), their respective PCR products were made, to which the prefix "PCR" was added to their code to identify them (Table 7). Subsequently, the assembly was carried out using the Benchling platform, the enzymes used in the digestion were those shown in table 4 whose restriction sites were previously eliminated from our sequences to avoid cuts in unwanted regions, the resulting assemblies are shown in Figures 2, 3 and 4.

Table 6: Name and code of the parts.

Part	**Code**	**Name**
Constitutive prmoter	PC	pAdh
Inducible promoter	PI	pGAL1
RBS	RBS	Designed yeast Kozak sequence
P2A	A2_2	Porcine teschovirus-1 2A
T2A	A2_6	Thosea asigna virus 2A
E2A	A2_1	Equine rhinitis B virus 2A
CDS 1	CDS_1	Pn8.2617
CDS 2	CDS_2	Pn2.84
CDS 3	CDS_3	4-coumarate 3-hydroxylase
CDS 4	CDS_4	Pn1.1317
CDS 5	CDS_5	Pn7.1626
CDS 6	CDS_6	Pn16.1198
CDS 7	CDS_7	cadA
CDS 8	CDS_8	Pn4.3222
CDS 9	CDS_9	Pn2.2377
CDS 10	CDS_10	DpkA
CDS 11	CDS_11	Pn6.2477
Terminador 1	T1	tADH1
Terminador 2	T2	tADH1
Vector	V	pRS316 yeast shuttle vector with Biobrick MCS

Table 7: List of assembly parts 1, 2 and 3 with their respective restriction sites indicated.

Assembly 1	**Restriction sites**	**Assembly 2**	**Restriction sites**	**Assembly 3**	**Restriction sites**
PCR_PC	5'- EcoRI, NotI, XbaI 3'- BamHI	PCR_PC	5'- EcoRI, NotI, XbaI 3'- BamHI	PCR_PI	5'- EcoRI, NotI, XbaI 3'- BamHI
PCR_CDS_1	5'- BamHI 3'- NheI	PCR_CDS_2	5'- BamHI 3'- NheI	PCR_CDS_6	5'- BamHI 3'- NheI
PCR_A2_2	5'- NheI 3'- AatII	PCR_A2_2	5'- NheI 3'- AatII	PCR_A2_2	5'- NheI 3'- AatII
PCR_CDS_3	5'- AatII 3'- HindIII	PCR_CDS_4	5'- AatII 3'- HindIII	PCR_CDS_9	5'- AatII 3'- HindIII
PCR_A2_6	5'- HindIII 3'- XhoI	PCR_A2_6	5'- HindIII 3'- XhoI	PCR_A2_6	5'- HindIII 3'- XhoI
PCR_CDS_5	5'- XhoI 3'- BspEI	PCR_CDS_7	5'- XhoI 3'- BspEI	PCR_CDS_10	5'- XhoI 3'- BspEI
PCR_A2_1	5'- BspEI 3'- FseI	PCR_T2	5'- BspEI 3'- PstI	PCR_A2_1	5'- BspEI 3'- FseI
PCR_CDS_8	5'- FseI 3'- SalI	V	EcoRI, XbaI, SpeI,PstI, NotI	PCR_CDS_11	5'- FseI 3'- SalI
PCR_T1	5'- SalI 3'- PstI			PCR_T1	5'- SalI 3'- PstI
V	EcoRI, XbaI, SpeI,PstI, NotI			V	EcoRI, XbaI, SpeI,PstI, NotI

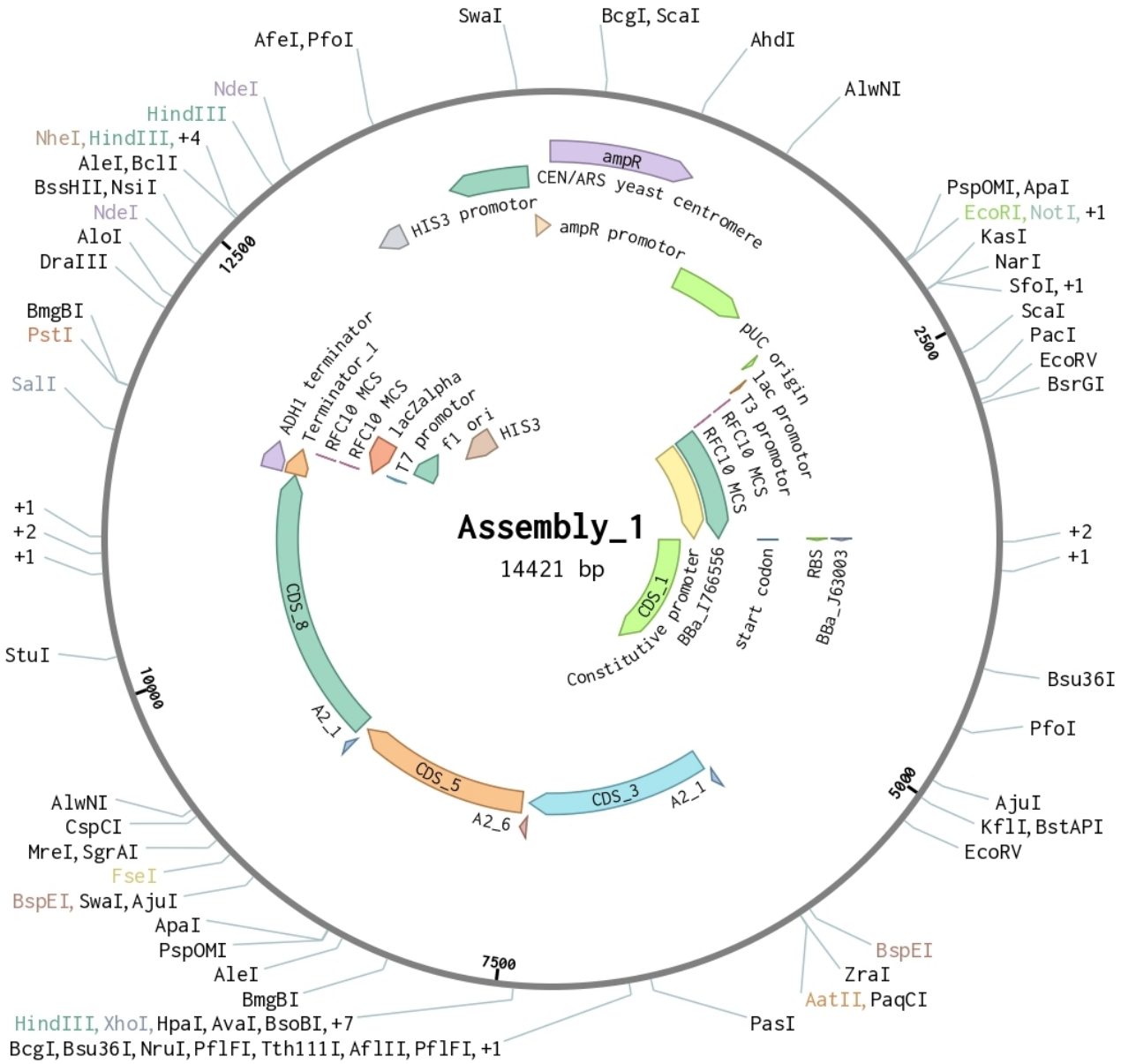

Figure 2: Assembly 1.

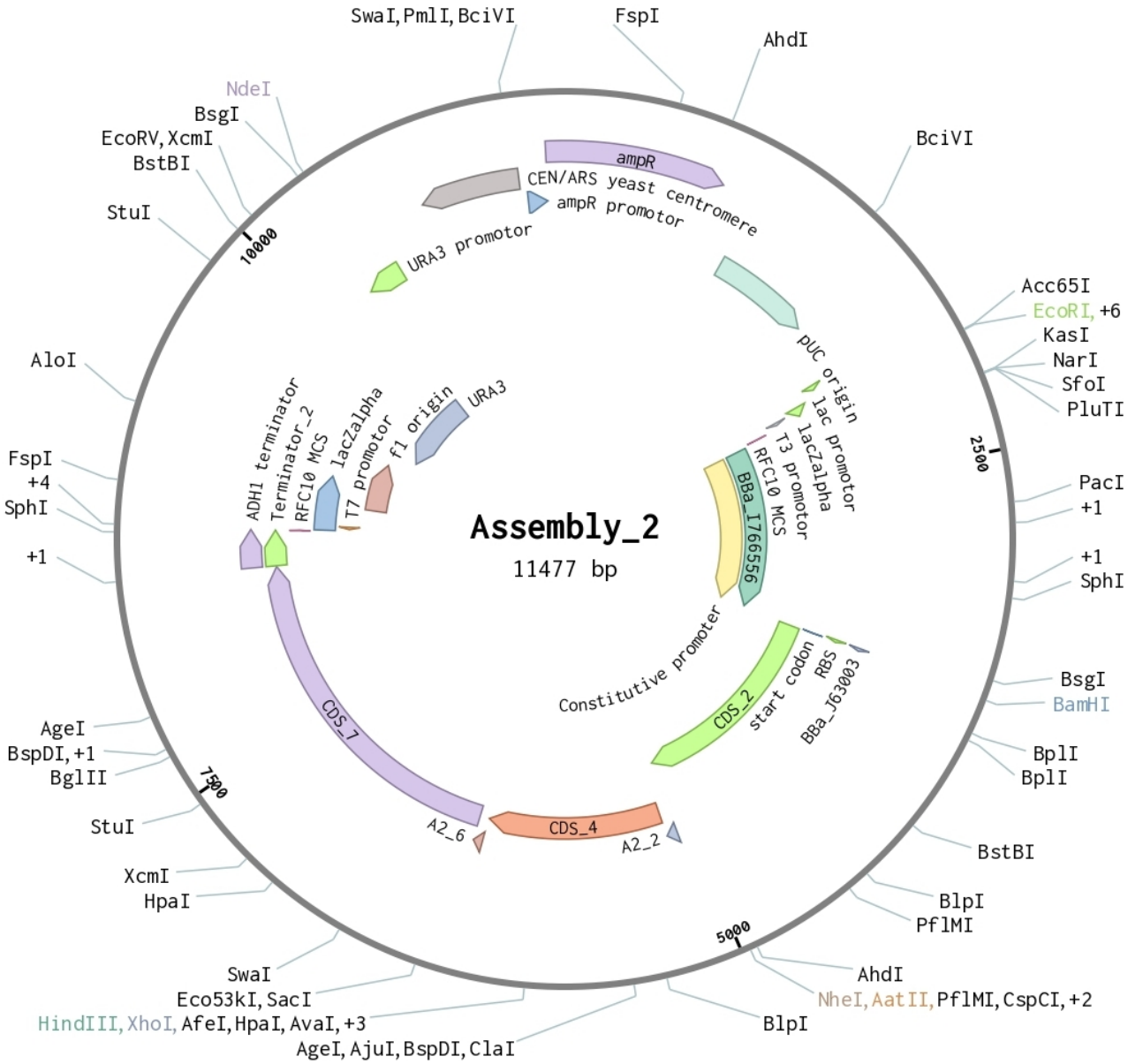

Figure 3: Assembly 2.

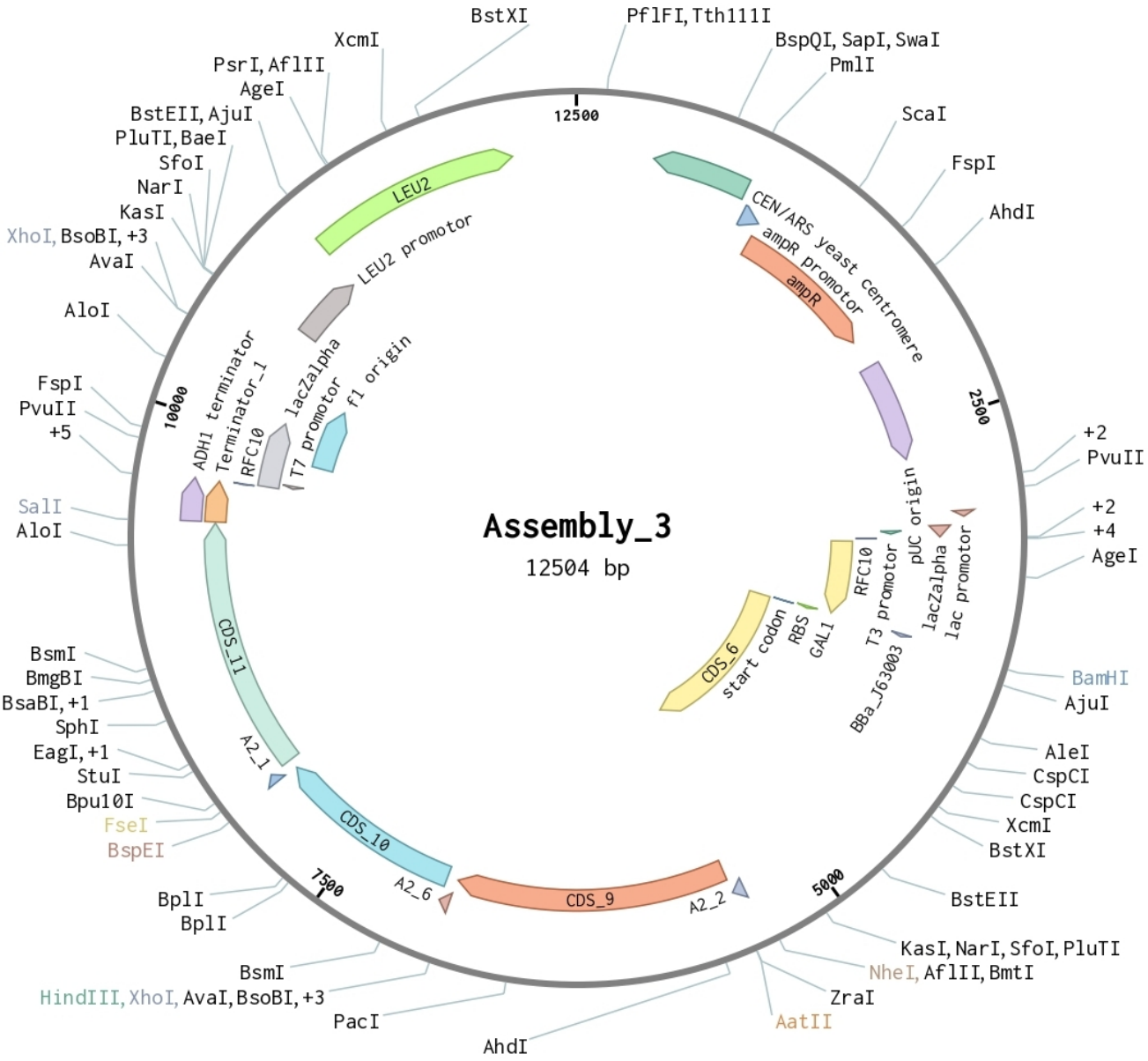

Figure 4: Assembly 3.

Transformation

It is proposed to use the *S. cerevisiae* strain BJ5464, given that it has proven to be very efficient for the expression of heterologous proteins, since its ability to form protein digestion vacuoles has been suppressed.[1] It would be used to express the 3 assemblies, and the transformation strategy that is proposed to be used is the one described in McCusker 2015, since it has been previously tested by iGEM teams and seems to give good results, the selection of the yeast cells that have been transformed will be carried out with selection media by auxotrophy of the following nutrients: leucine, uracil, uridine and

histidine, since they are the selection markers of the expression vectors used.

Appendix

Table 8: Registry name, code and length of used parts.

Description	Code	Name	Lenght (bp)
Pn8.2617	CDS_1	IDLBB_0023100	2193
Pn2.84	CDS_2	IDLBB_0023101	1524
coumarate 3-hydroxylase	CDS_3	IDLBB_0023102	1539
Pn1.1317	CDS_4	IDLBB_0023103	1083
Pn7.1626	CDS_5	IDLBB_0023104	1506
Pn16.1198	CDS_6	IDLBB_0023105	1644
cadA	CDS_7	IDLBB_0023106	2148
Pn4.3222	CDS_8	IDLBB_0023107	2337
Pn2.2377	CDS_9	IDLBB_0023108	1485
DpkA	CDS_10	IDLBB_0023109	1032
Pn6.2477	CDS_11	IDLBB_0023110	1383
GAL1	PI	BBa_J63006	549
pAdh	PC	BBa_I766556	1501
ADH terminator from S. cerevisiae	T1	BBa_J63002	225
ADH terminator from S. cerevisiae	T2	BBa_J63002	225
Designed yeast Kozak sequence	RBS	BBa_J63003	18
pRS316yeast shuttle vector with Biobrick MCS (HIS3)	V1	BBa_K1680014	4928
pRS316yeast shuttle vector with Biobrick MCS (URA3)	V2	BBa_K1680014	4848
pRS316yeast shuttle vector with Biobrick MCS (LEU2)	V3	BBa_K1680015	5979
P2A	A2_2	BBa_K1442039	66
T2A	A2_6	BBa_K1993019	54
E2A	A2_1	BBa_K4415003	69
Primer fordward CDS_1	PF_CDS_1	IDLBB_0023111	29

Primer fordward CDS_2	PF_CDS_2	IDLBB_0023112	27
Primer fordward CDS_3	PF_CDS_3	IDLBB_0023113	30
Primer fordward CDS_4	PF_CDS_4	IDLBB_0023114	31
Primer fordward CDS_5	PF_CDS_5	IDLBB_0023115	29
Primer fordward CDS_6	PF_CDS_6	IDLBB_0023116	31
Primer fordward CDS_7	PF_CDS_7	IDLBB_0023117	33
Primer fordward CDS_8	PF_CDS_8	IDLBB_0023118	30
Primer fordward CDS_9	PF_CDS_9	IDLBB_0023119	31
Primer fordward CDS_10	PF_CDS_10	IDLBB_0023120	22
Primer fordward CDS_11	PF_CDS_11	IDLBB_0023121	38
Primer reverse CDS_1	PR_CDS_1	IDLBB_0023122	26
Primer reverse CDS_2	PR_CDS_2	IDLBB_0023123	33
Primer reverse CDS_3	PR_CDS_3	IDLBB_0023124	30
Primer reverse CDS_4	PR_CDS_4	IDLBB_0023125	30
Primer reverse CDS_5	PR_CDS_5	IDLBB_0023126	26
Primer reverse CDS_6	PR_CDS_6	IDLBB_0023127	33
Primer reverse CDS_7	PR_CDS_7	IDLBB_0023128	41
Primer reverse CDS_8	PR_CDS_8	IDLBB_0023129	28
Primer reverse CDS_9	PR_CDS_9	IDLBB_0023130	28
Primer reverse CDS_10	PR_CDS_10	IDLBB_0023131	29
Primer reverse CDS_11	PR_CDS_11	IDLBB_0023132	27
Primer fordward PC	PF_PC	IDLBB_0023133	48
Primer fordward PI	PF_PI	IDLBB_0023134	47
Primer reverse PC	PR_PC	IDLBB_0023135	21
Primer reverse PI	PR_PI	IDLBB_0023136	21
Primer fordward A2_2	PF_A2_2	IDLBB_0023137	25
Primer fordward A2_6	PF_A2_6	IDLBB_0023138	32
Primer fordward A2_1	PF_A2_1	IDLBB_0023139	37
Primer reverse A2_2	PR_A2_2	IDLBB_0023140	30
Primer reverse A2_6	PR_A2_6	IDLBB_0023141	30
Primer reverse A2_1	PR_A2_1	IDLBB_0023142	31
Primer fordward T1	PF_T1	IDLBB_0023143	53
Primer fordward T2	PF_T2	IDLBB_0023144	53
Primer reverse T1	PR_T1	IDLBB_0023145	38
Primer reverse T2	PR_T2	IDLBB_0023146	38
Assembly 1	Assembly_1	IDLBB_0023147	14421
Assembly 2	Assembly_2	IDLBB_0023148	11477
Assembly 3	Assembly_3	IDLBB_0023149	12504

References

[1] Baghban, R., Farajnia, S., Rajabibazl, M., Ghasemi, Y., Mafi, A., Hoseinpoor, R., Rahbarnia, L., and Aria, M. (2019). Yeast expression systems: overview and recent advances. *Molecular biotechnology*, 61(5):365–384.

[2] Donnelly, M. L., Luke, G., Mehrotra, A., Li, X., Hughes, L. E., Gani, D., and Ryan, M. D. (2001). Analysis of the aphthovirus 2a/2b polyprotein 'cleavage'mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *Journal of General Virology*, 82(5):1013–1025.

[3] Knight, T. (2003). Idempotent vector design for standard assembly of biobricks.

[4] Liu, Z., Chen, O., Wall, J., Zheng, M., Zhou, Y., Wang, L., Ruth Vaseghi, H., Qian, L., and Liu, J. (2017). Systematic comparison of 2a peptides for cloning multi-genes in a polycistronic vector. *Scientific reports*, 7(1):1–9.

[5] McCusker, D. (2015). High efficiency yeast transformation. *iGEM Bordeaux*. Available at: https://static.igem.org/mediawiki/2015/5/5a/High$_E ficiency_Y east_T ransformation.pdf$.

[6] Souza-Moreira, T. M., Navarrete, C., Chen, X., Zanelli, C. F., Valentini, S. R., Furlan, M., Nielsen, J., and Krivoruchko, A. (2018). Screening of 2a peptides for polycistronic gene expression in yeast. *FEMS yeast research*, 18(5):foy036.

Contents

Introduction **2**

A2 Peptides and Cloning **2**

Assembly **3**

PCR 4

Transformation **9**

Appendix **10**

List of Figures

1 Cleavage efficiencies of peptides characterized in Souza-Moreira et al. 2018 2
2 Assembly 1. 7
3 Assembly 2. 8
4 Assembly 3. 9

List of Tables

1 Peptide 2A order. 3
2 Chosen A2 peptides and their abbreviation. 3
3 Restriction sites used in 2A peptides. 4
4 Restriction sites removed. 4
5 Prefix and suffix sequences. 4
6 Name and code of the parts. 5
7 List of assembly parts 1, 2 and 3 with their respective restriction sites indicated. 6
8 Registry name, code and length of used parts. 10