

大腸菌由来再構成型無細胞タンパク質合成系 (PUREflex[®]) に適した鋳型DNAの配列の探索

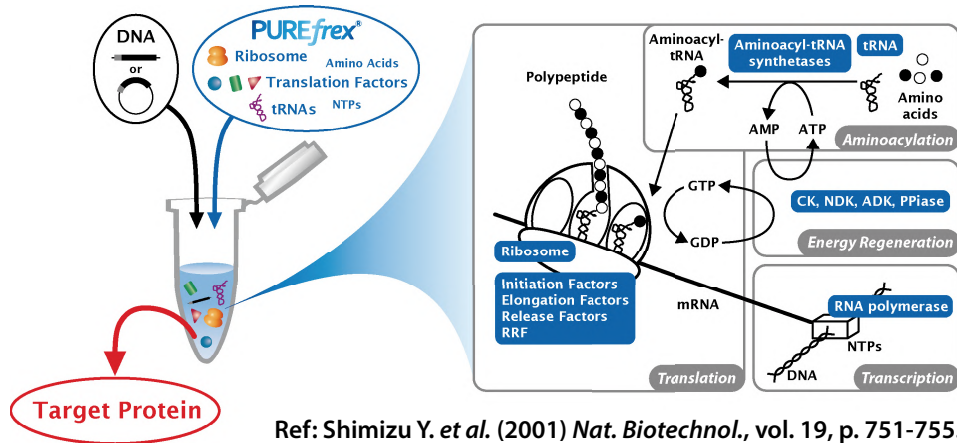
Exploration of the template DNA sequence suitable for the *E.coli*-based reconstituted cell-free protein synthesis system (PUREflex[®])

金森 崇、布施(村上) 朋重、松本 令奈 (ジーンフロンティア(株))

PUREfrefx

PUREfrefx is based on the PURE system technology.

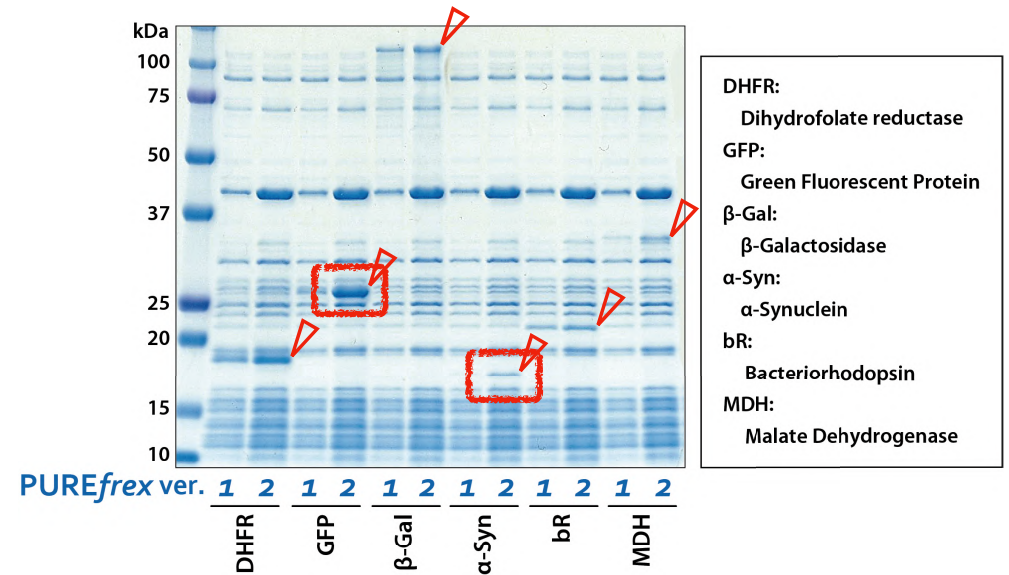
The PURE system is a reconstituted cell-free protein synthesis system, which consists only of purified factors necessary for transcription, translation and energy regeneration.



Advantage

- Low level of contamination
- Easy to adjust composition
- Usable of PCR product as a template DNA

Example of protein synthesis



Why is the synthesis efficiency different ?

- amino acid sequence
- nucleotide sequence
-
-

5' UTR of the template DNA for PURE $_{\text{frex}}$

Currently used sequence (derived from T7 *gene 10* UTR)

GAAATTAAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAAGGAGATATACCAATGNNNNNNNNNNNNNNNNNN



Stemloop region is necessary for efficient transcription reaction and stabilization of the transcript.

AT-rich region binds S1 protein of 30S ribosome subunit and increases translation efficiency.

SD (Shine-Dalgarno) sequence binds 3'-terminus of 16S rRNA and localizes mRNA to the start position of translation.

SD and Spacer



Long SD sequence

< 10 nt spacer

SD	Spacer
TAAGGAGGTG	TATAATATACCA
	TAATATACCA
	ATATACCA
	TACCA
	CCA

Currently used SD sequence

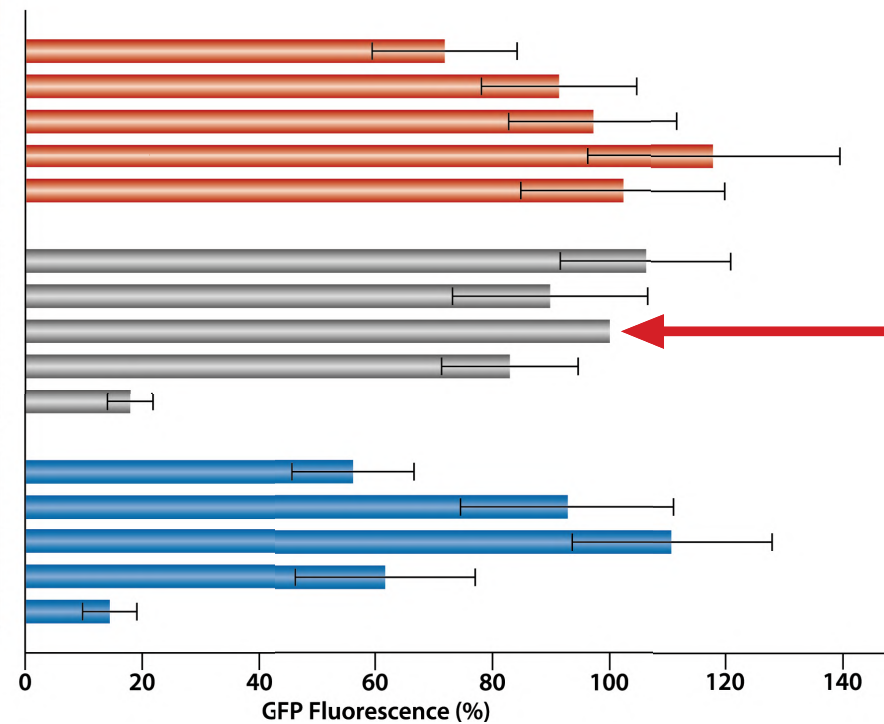
5-12 nt spacer

SD	Spacer
AAGGAG	TATAATATACCA
	TAATATACCA
	ATATACCA
	TACCA
	CCA

Short SD sequence

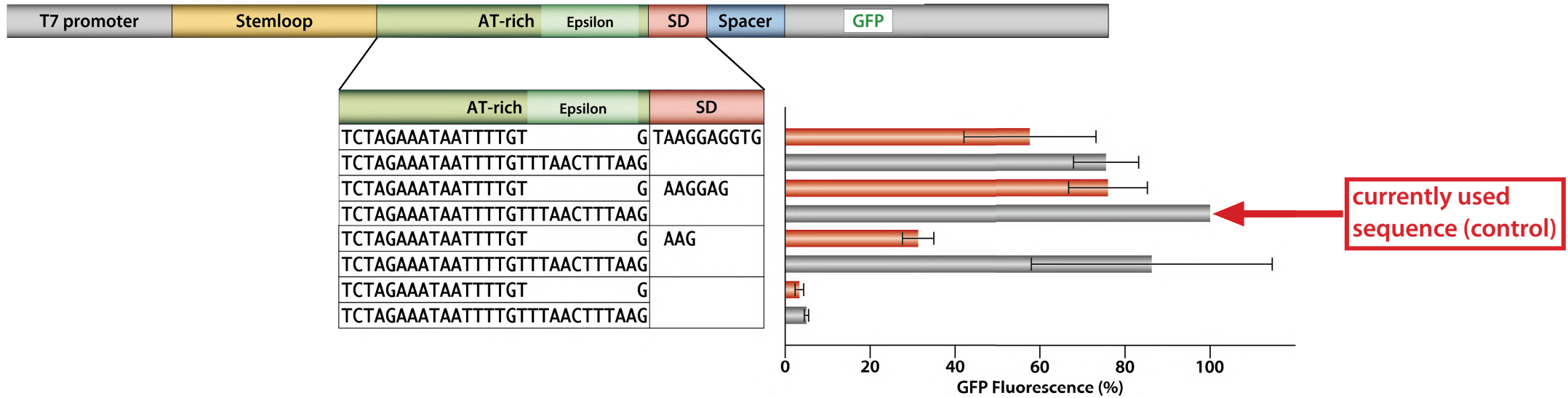
8-10 nt spacer

SD	Spacer
AAG	TATAATATACCA
	TAATATACCA
	ATATACCA
	TACCA
	CCA



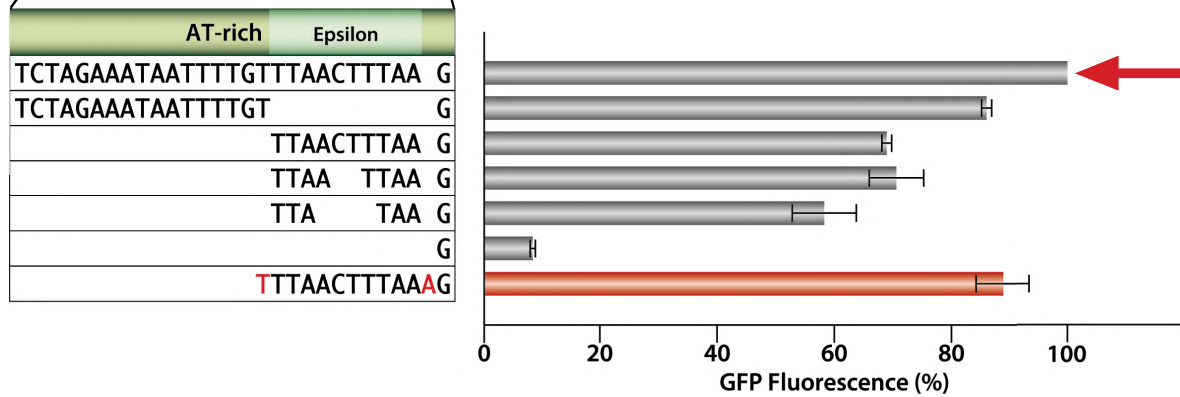
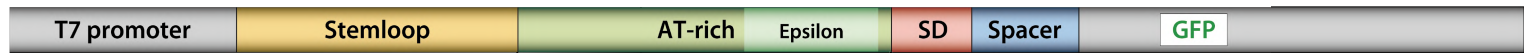
currently used sequence (control)

AT-rich (Epsilon) and SD



SD sequence	No	Short	Currently used	Long
(-) Epsilon	✗	✗	△	△
(+) Epsilon	✗	○	○	○

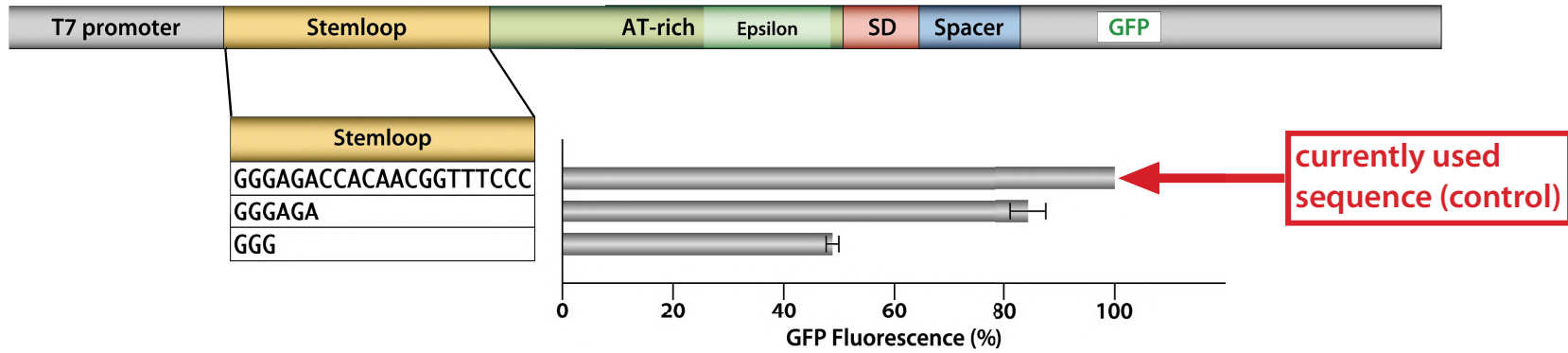
AT-rich



currently used sequence (control)

AT-rich region	No	< 10 nt	> 12 nt
	×	△	○

Stemloop



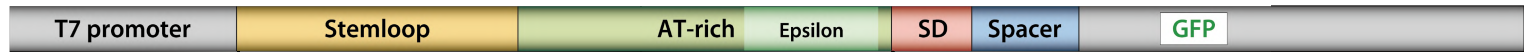
Stemloop region

3 nt

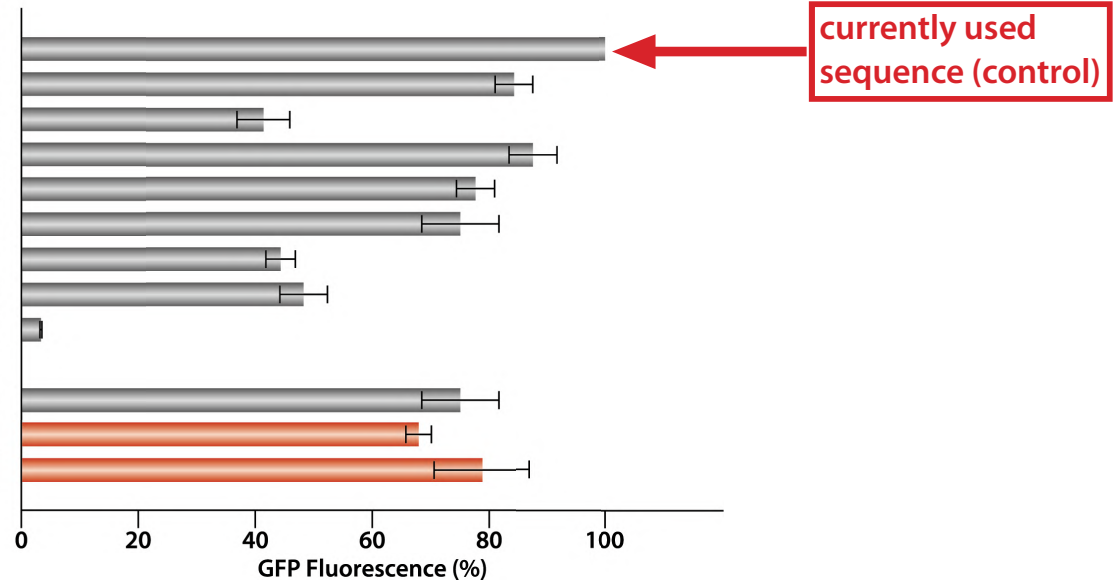
6 nt



Stemloop and AT-rich



Stemloop	AT-rich
GGGAGACCACAACGGTTTCCC	TCTAGAAATAATTTGTTTAACTTTAAG
GGGAGA	TCTAGAAATAATTTGTTTAACTTTAAG
	TCTAGAAATAATTTGT
	AATAATTTGTTTAACTTTAAG
	AATTTGTTTAACTTTAAG
	TTTGTTTAACTTTAAG
	GTTTAACTTTAAG
	TTAACTTTAAG
	G
GGGAGA	TTTGTTTAACTTTAAG
	TTTATTTAAATTTAAG
	TTTATTTAATTTAAG



6nt-SL

AT-rich region

No

< 12 nt

> 15 nt



Other proteins



GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCA

GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCC-----TTTAACTTTAAAGAAGGAGATATACCA

GAAATTAATACGACTCACTATAGGGAGA-----AATAATTTTGTTTAACTTTAAGAAGGAGATATACCA

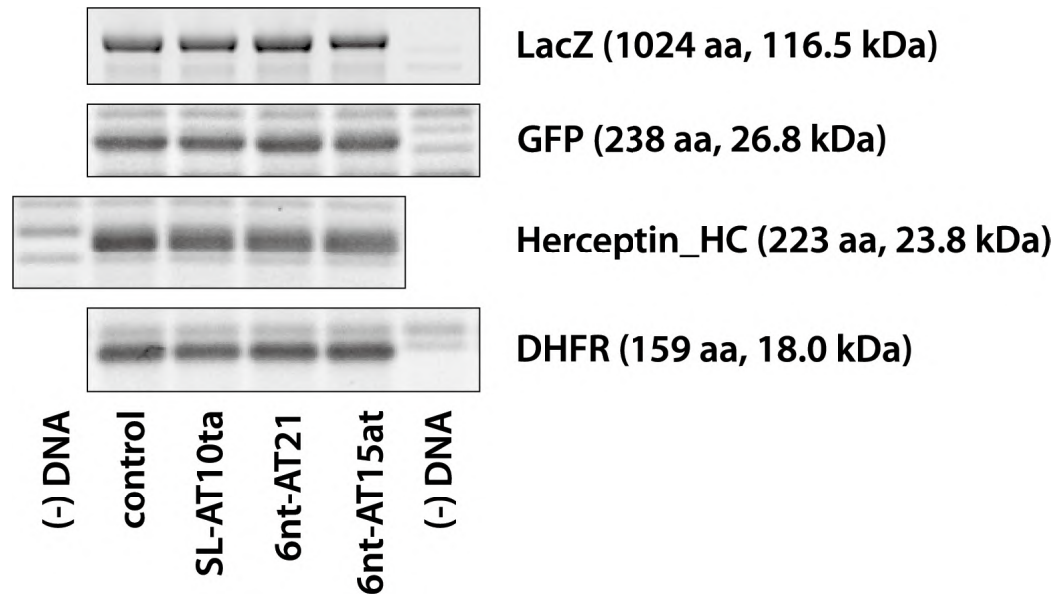
GAAATTAATACGACTCACTATAGGGAGA-----TTTATTTAATTTTAAAGAAGGAGATATACCA

control

SL-AT10ta

6nt-AT21

6nt-AT15at



Summary



Currently used sequence (control)

Efficiency

GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCA

100%

GAAATTAATACGACTCACTATAGGG-----TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCA

50%

GAAATTAATACGACTCACTATAGGGAGA-----TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCA

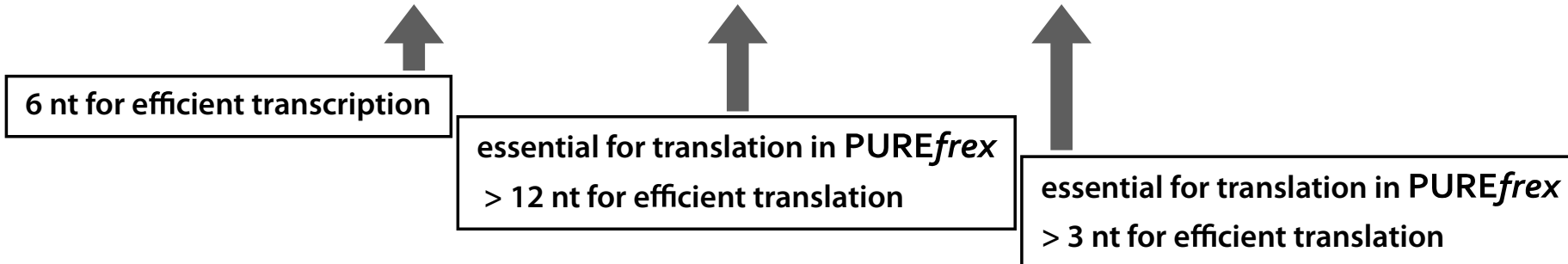
85%

GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCC-----GAAGGAGATATACCA

10%

GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAG-----ATATACCA

5%



Summary



Currently used sequence (control)

GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCA

Efficiency

100%

Length

85 bp

GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCC-----TTTAACTTTAAAGAAGGAGATATACCA

90%

70 bp (-15 bp)

GAAATTAATACGACTCACTATAGGGAGA-----AATAATTTTGTTTAACTTTAAGAAGGAGATATACCA

85%

64 bp (-21 bp)

GAAATTAATACGACTCACTATAGGGAGA-----TTTGTTTAACTTTAAGAAGGAGATATACCA

75%

58 bp (-27 bp)

GAAATTAATACGACTCACTATAGGGAGA-----TTTATTTAATTTTAAGAAGGAGATATACCA

80%

58 bp (-27 bp)

Shorter 5' UTR



Easier to prepare the template DNA by PCR