Influence of 5′-UTR and N-terminal region of ORF on the efficiency of protein synthesis in a reconstituted cell-free protein synthesis system (PURE*frex*®)

Tomoe Fuse-Murakami, Rena Matsumoto, Takashi Kanamori (GeneFrontier Corp.)

GeneFrontier

Abstract

PURE*frex*® is a cell-free protein synthesis system based on the PURE system, which is reconstituted only from factors involved in protein synthesis in *E. coli*. Due to the improvement of the reaction composition, the synthesis efficiency has increased up to 1 mg/mL, but it varies significantly depending on the target protein. In this presentation, we report our findings on the influence of the 5'-UTR and N-terminal region of ORF on the protein synthesis efficiency.

(1) 5'-UTR

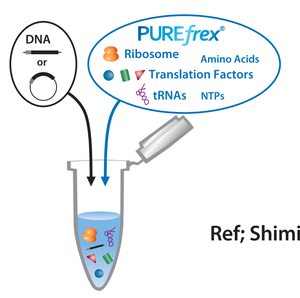
To confirm the necessity of each region in the 5'-UTR derived from T7 phage currently in use, we compared the fluorescence of sfGFP synthesized from its template DNA with different 5'-UTR. As a result, when either AT-rich region or Shine-Dalgarno (SD) sequence was removed, the fluorescence decreased to less than 10%. This result indicates that not only SD sequence but also AT-rich region in the 5'-UTR is important for the efficient translation in PURE*frex*.

(2) N-terminal region of ORF

It has been reported that the N-terminal region of ORF affects the efficiency of protein synthesis. First, for ten amino acids immediately below the start codon, the synthesis efficiency was higher when AT-rich codons were used than when codons commonly used in *E. coli* were used. For some proteins, changing a few codons in the N-terminal region increased the synthesis efficiency more than 10-fold. Second, amino acids in the N-terminal region also affected the synthesis efficiency, especially at low temperatures. For example, replacing glycine at position 4 of sfGFP with tyrosine increased the synthesis efficiency 8-fold at 23°C. These results indicate that small differences in the N-terminal sequence can cause large fluctuations in the synthesis efficiency.

1. PURE*frex*®; based on the PURE system technology

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

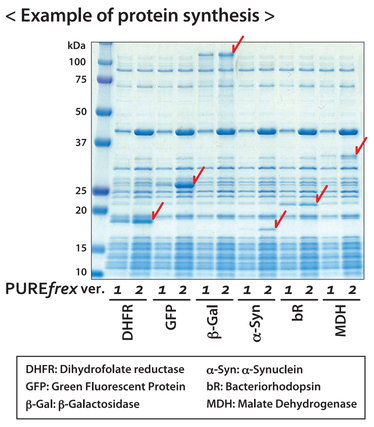


Advantage

- Low level of contamination
- Easy adjustment of the reagent composition
- PCR product usable as a temlplate DNA

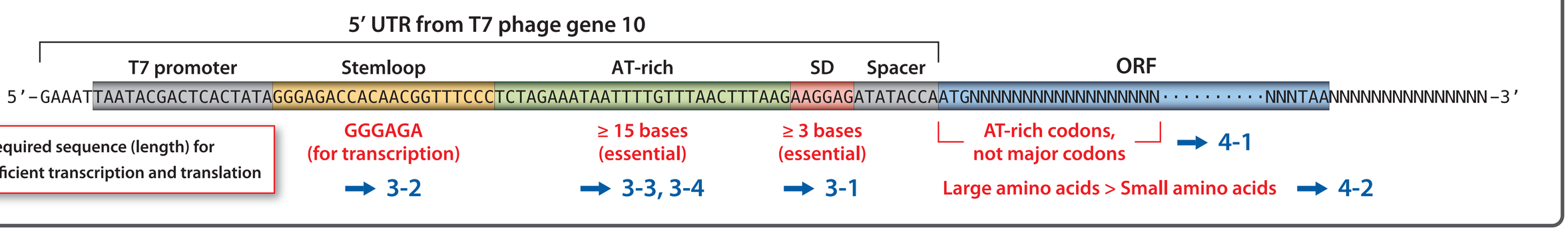
Ref; Shimizu Y. *et al.* (2001) *Nat. Biotechnol.*, vol. 19, p. 751.

< Example of protein synthesis >



DHFR: Dihydrofolate reductase
GFP: Green Fluorescent Protein
β-Gal: β-Galactosidase
α-Syn: α-Synuclein
bR: Bacteriorhodopsin
MDH: Malate Dehydrogenase

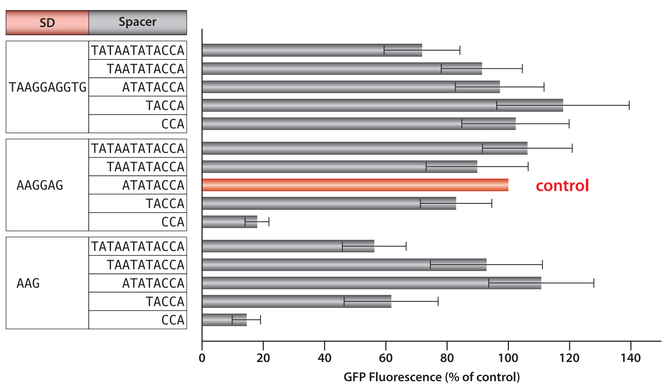
2. Construct of template DNA for PURE*frex*



3. Effect of 5′-UTR on the synthesis efficiency

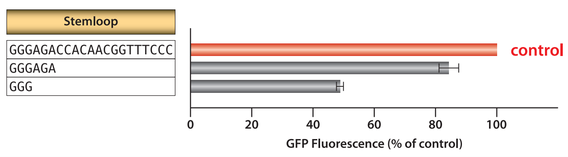
PURE*frex*®*2.1* (+GSH)
↓ + sfGFP DNA with indicated 5′ UTR
↓ incubation at 37°C for 4 h
↓ measurement of fluorescence of sfGFP

3-1. SD sequence and Spacer



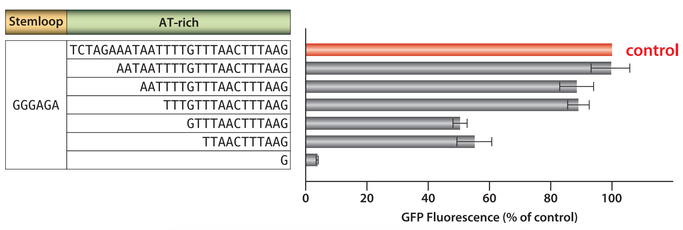
- AAG (3 bases) was enoygh for the efficient translation.
- The balance between the length of SD and spacer was important.

3-2. Stemloop



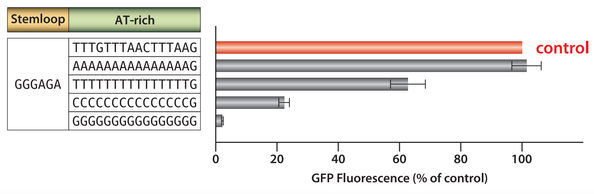
- GGGAGA, which is required for efficient transcription by T7 RNA polymerase, was sufficient.

3-3. AT-rich (Length)



- 15 bases and more was sufficient.
- AT-rich region was essential.

3-4. AT-rich (Sequence)



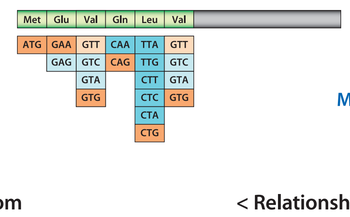
- A15 was as effective as the control, but T15 was less effective.
- In the case of C15 and G15, synthesis efficiency decreased significantly.

4. Effect of N-terminal region of ORF on the synthesis efficiency

4-1. Codons in the N-terminal region

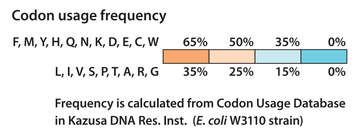
PURE*frex*®*2.1* (+GSH)
↓ + Trastuzumab HC (VH + CH1) DNA
↓ incubation at 37°C for 4 h
↓ SDS-PAGE and quantification

< N-terminal codon of Trastuzumab HC >



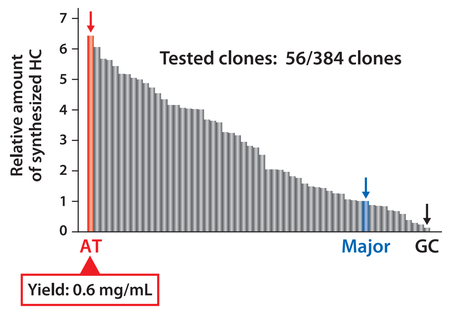
	Met	Glu	Val	Gln	Leu	Val	GC (%)	ΔG (kcal/mol)
AT	ATG	GAA	GTA	CAA	TTA	GTT	28	0.0
Major	ATG	GAA	GTG	CAG	CTG	GTG	56	-4.7
GC	ATG	GAG	GTG	CAG	CTG	GTC	61	-3.9

Codon usage frequency

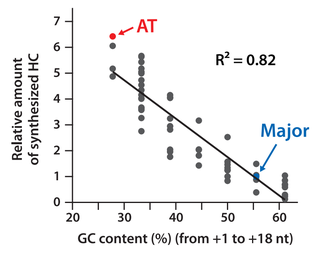


Frequency is calculated from Codon Usage Database in Kazusa DNA Res. Inst. (*E. coli* W3110 strain)

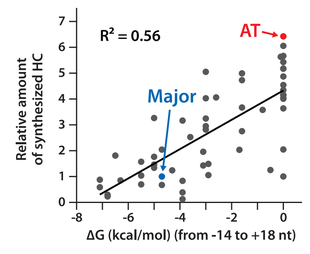
< The amount of product synthesized from the template DNA with different N-terminal codon >



< Relationship between protein synthesis and GC content >



< Relationship between protein synthesis and secondary structure of mRNA >



AT-rich codons > Major codons @ N-teminal region of ORF

4-2. Amino acids in the N-terminal region

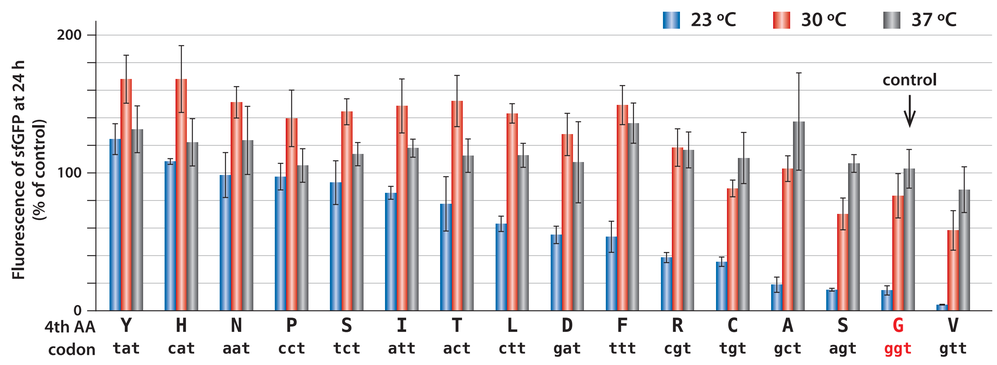
< N-terminal sequence of sfGFP >

M S K G E E
atgtctaaaggtgaagaa

other amino acids (NNT codon)

PURE*frex*®*2.1* (+ GSH)
↓ + sfGFP DNA with indicated 4th amino acid
↓ incubation at 23, 30 or 37 °C for 24 h
↓ Measurement of fluorescence of sfGFP

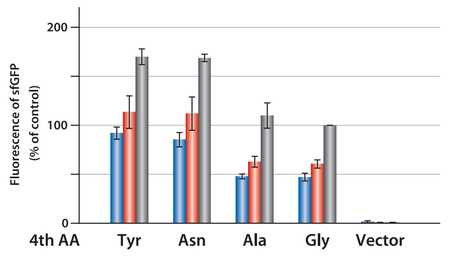
< Effects of 4th amino acid on the synthesis using PURE*frex* >



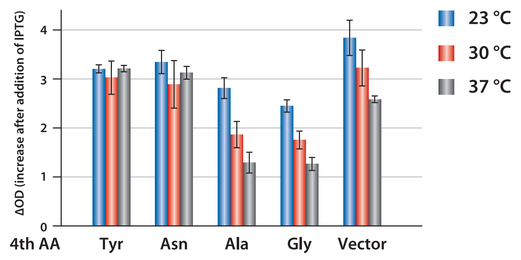
- The amount of synthesized sfGFP, especially at 23 °C, differed depending on the fourth amino acid.
- When the fourth amino acid was a large amino acid such as Tyr or His, the amount of sfGFP was increased compared to Gly.

BL21(DE3) cells harboring pET/sfGFP
↓ Cultivation at 23, 30 or 37 °C
↓ + IPTG at OD600 = 1
↓ incubation for 2 hours (30 and 37 °C) or 4 hours (23 °C)
↓ Measurement of OD600 and fluorescence

< Fluorescence of *E. coli* cells expressing sfGFP >



< Growth of *E. coli* cells expressing sfGFP >



- Not only Tyr and Asn variants but also Ala and Gly variants were expressed in *E. coli* incubated at 23 °C.
- Expression of Ala and Gly variants suppressed *E. coli* growth.

Conclusion

Ref: Fuse-Murakami *et al.* (2024) *Int. J. Mol. Sci.*, 25, 5264. https://doi.org/10.3390/ijms25105264

5′-UTR

- Not only SD sequence but also AT-rich region in 5′ -UTR is very important for efficient translation reaction.

N-terminal region of ORF

- For the N-terminal codon, AT-rich codons increase the synthesis efficiency more than commonly used codons.
- Large amino acids in the N-terminal region increase the protein synthesis efficiency, especially at low temperatures.