Effect of N-terminal amino acid sequence on the protein synthesis using a reconstituted cell-free protein synthesis system

GeneFrontier

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

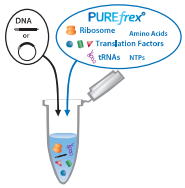
Abstract

The *Escherichia coli* (*E. coli*)-based protein synthesis using recombinant elements (PURE) system is a cell-free protein synthesis system reconstituted from purified factors essential for *E. coli* translation. The PURE system is widely used for basic and synthetic biology applications. One of the major challenges associated with the PURE system is that the protein yield of the system varies depending on the protein. Studies have reported that the efficiency of ribosomal translation is significantly affected by nucleotide and amino acid sequences, especially in the N-terminal region. Here, we investigated the inherent effect of various N-terminal sequences on protein synthesis using the PURE system. We found that a single amino acid substitution in the N-terminal region significantly altered translation efficiency in the PURE system, especially at low temperatures. This result gives us useful suggestions for the expression of the protein of interest *in vitro* and *in vivo*.

1. Introduction

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.

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



- One of the challenges;
  Differences in the synthesis efficiency depending on the proteins
- Background;
  - The 3rd to 5th amino acid residues in the N-terminal region are important for the efficient translation.
  - Large amino acids, which are encoded by AT-rich codon, facilitate early elongation reaction.

Ref; Verma M. *et al.* (2019) *Nat. Commun.*, Jia B. *et al.* (2021) *Sci. Rep.*

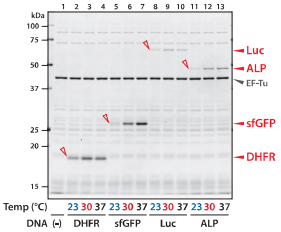
- Aim of this study;
  Evaluation of the effect of amino acids in the N-terminal region on the synthesis efficiency at 23, 30 and 37 °C.

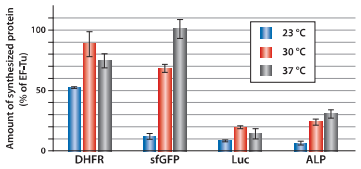
2. Protein synthesis at 23, 30, or 37 °C

< Information of proteins used in this study >

Protein		M.W.	N-terminal AA Seq
DHFR	*E. coli* dihydrofolate reductase	18.0 kDa	MISLIAALAV...
sfGFP	Superfolder GFP	26.8 kDa	MSKGEELFTG...
Luc	Firefly luciferase	60.7 kDa	MEDAKNIKKG...
ALP	*E. coli* alkaline phosphatase	47.3 kDa	MRTPEMPVLE...

PURE*frex*®2.1 (+ 4 mM GSH)
↓ + DNA (PCR product)
↓ Incubation at 23, 30 or 37 °C for 24 h
↓ SDS-PAGE and quantification





The decrease in the amount of synthesized protein at 23 and 30 °C relative to 37 °C varied depending on the protein.

3. Influences of the second and fourth amino acid on sfGFP synthesis

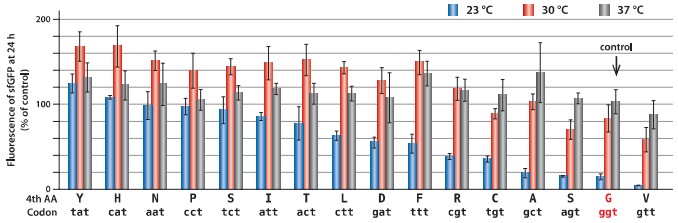
PURE*frex*®2.1 (+ 4 mM GSH)
↓ + sfGFP variant DNA (PCR product)
↓ Incubation at 23, 30 or 37 °C for 24 h
↓ Measurement of fluorescence

< N-terminal sequence of sfGFP >

M S K G E E
atgtctaaaggtgaagaa....

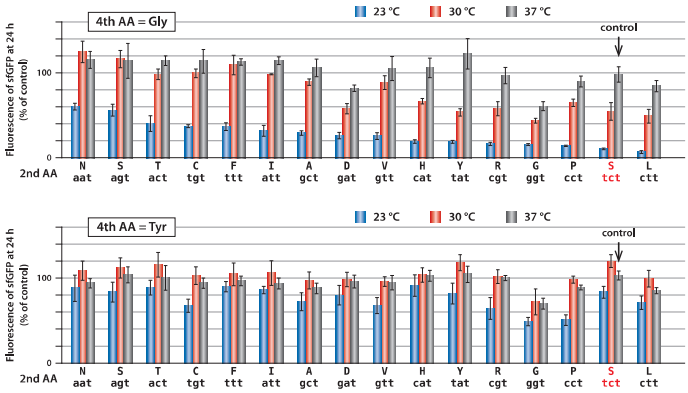
other amino acids

3-1. Synthesis of sfGFP with a different amino acid at the fourth position



- 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.

3-2. Synthesis of sfGFP with a different amino acid at the second position

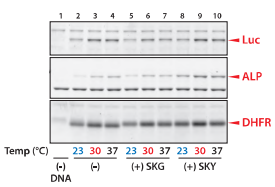


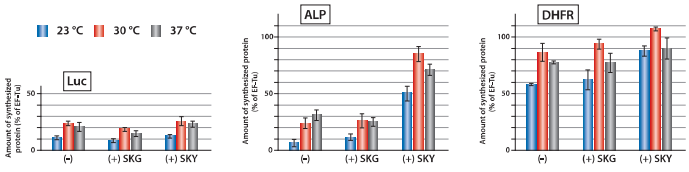
- When the fourth amino acid was Gly, the amount of sfGFP synthesized at 23 °C differed depending on the second amino acid.
- When the fourth amino acid was Tyr, there was little effect of the second amino acid except for Gly.

4. Effect of the N-terminal additional sequence

PURE*frex*®2.1 (+ 4 mM GSH)
↓ + DNA (PCR product)
↓ Incubation at 23, 30 or 37 °C for 24 h
↓ SDS-PAGE and quantification

Luc MEDAKNIKKG...
ALP MRTPEMPVLE...
DHFR MISLIAALAV...
△
SKG or SKY



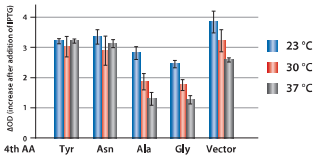


Addiotion of Ser-Lys-Tyr at N-terminal region significantly increased the amount of synthesized ALP.

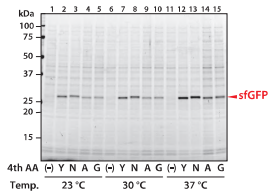
5. Expression of sfGFP with a different fourth amino acid in *E. coli*

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

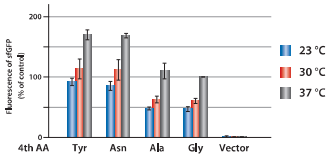
< Growth of *E. coli* cells expressing sfGFP >



< SDS-PAGE of *E. coli* cells with sfGFP >



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



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

Conclusion

- The efficiency of synthesis especially at low temperature varied from protein to protein.
- The fourth amino acid of sfGFP significantly affected the synthesis efficiency *in vitro* at 23 °C.
- Addition of Ser-Lys-Tyr to N-terminal region significantly increased the amount of ALP synthesized at low temperatures.
- Ala and Gly variants could be expressed in *E. coli* even at 23 °C, but suppressed the growth.

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