

再構成型無細胞タンパク質合成系（PUREfrex）に適した鋳型 DNA の塩基配列の探索

Exploration of nucleotide sequence of template DNA suitable for the reconstituted cell-free protein synthesis system (PUREfrex)



布施 (村上) 朋重、松本令奈、金森崇 (ジーンフロンティア (株))

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

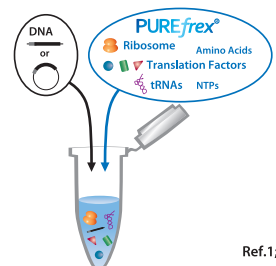
Abstract

PUREfrexは、大腸菌でタンパク質合成に関与する因子を再構成した無細胞タンパク質合成系 (PURE system) である。反応液の改良により、合成量は最大で1 mg/mL以上に増大しており、活性型のIgGを合成することも可能になっている。一方で、タンパク質の合成量が低い事例も多数報告されている。我々は、原因の一つに鋳型DNAの塩基配列があると考え、PUREfrexでのタンパク質合成に適した塩基配列を検討した。初めに、コード領域のコドンについて検討した。N末端領域について、使用頻度ではなくAT含量が高いコドンを使用した場合に合成量が高くなることを既に報告しており、今回は、コード領域全体のコドンについて検討を加えた。PUREfrexは大腸菌翻訳系に基づいた系であることから、大腸菌のコドン使用頻度を参考に塩基配列を設計し、大腸菌で使用頻度が最も高いコドンのみを使用した場合と、使用頻度に応じて数種類のコドンを使用した場合の合成量を比較した。その結果、数種類のコドンをを用いた場合に合成量が高くなることが分かった。次に、開始コドン上流の配列について検討した。その結果、リボソーム結合配列 (SD配列) は必須だが、必要なSD配列の長さは、前後のスペーサー配列や使用する反応液によって異なった。また、SD配列上流の翻訳促進配列 (イプシロン配列) により合成量が増加することも分かった。

1. PUREfrex®; based on the PURE system technology

PUREfrex® 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.

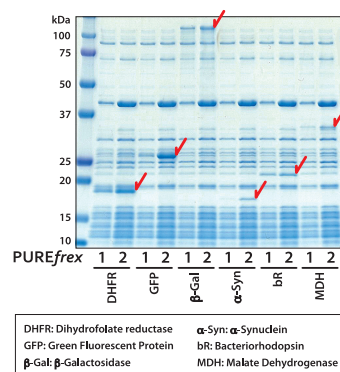


Advantage

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

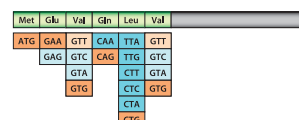
Ref.1; Shimizu Y. et al. (2001) Nat. Biotechnol., vol. 19, p. 751.

Example of Protein Synthesis

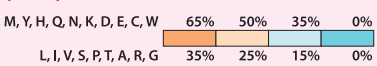


1. Effect of N-terminal codon

Trastuzumab Fab HC

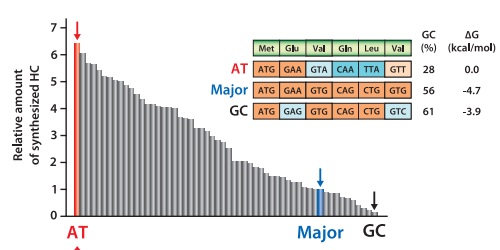


Frequency

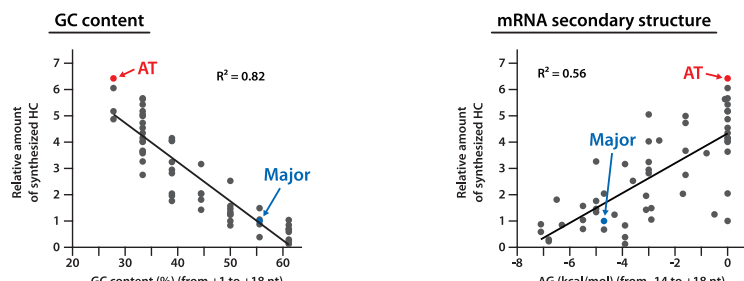


Frequency is calculated from Codon Usage Database in Kazusa DNA Res. Inst. (E. coli W3110 strain)

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



< Relationship between protein synthesis and N-terminal sequence >



AT-rich codon > Major codon @ 5'-terminus of ORF

2. Effect of codon used at entire and 5'-terminus of ORF

α-Synuclein 140 AA

	Major	W3110
Phe	TTT 0 1	TCT 0 0
Tyr	TAT 2 2	TAC 2 2
Cys	TGT 0 0	TGC 0 0
Leu	TTA 0 0	TAA 0 0
Stop	TGA 0 0	TGA 0 0
Trp	TGG 0 0	TGG 0 0
Pro	CTT 0 0	CCT 0 1
His	CTC 0 0	CCC 0 0
Arg	CGT 0 0	CGA 0 0
Ala	CTA 0 0	CCA 0 1
Gly	CTG 3 3	CCG 5 3
Met	ATT 2 1	ACT 0 1
Asp	ATC 0 1	ACC 10 5
Asn	ATA 0 0	ACA 0 1
Thr	ATG 2 2	ACG 0 3
Val	GTT 9 5	GAT 2 3
Leu	GTC 0 4	GCC 0 5
Gln	GTA 0 2	GCA 0 4
Pro	GTG 9 7	GGC 19 6
Arg		GAG 7 5
Ala		GGG 0 2

PD-L1 223 AA

	Major	W3110
Phe	TTT 3 3	TCT 1 1
Tyr	TAT 7 6	TAC 3 4
Cys	TGT 1 2	TGC 3 2
Leu	TTA 0 3	TAA 0 1
Stop	TGA 0 0	TGA 0 0
Trp	TGG 0 2	TGG 0 2
Pro	CTT 0 2	CCT 0 1
His	CTC 0 2	CCC 0 1
Arg	CGT 0 2	CGA 0 0
Ala	CTA 0 0	CCA 0 2
Gly	CTG 20 11	CCG 9 5
Met	ATT 6 7	ACT 0 3
Asp	ATC 7 6	ACC 15 8
Asn	ATA 0 0	ACA 0 2
Thr	ATG 3 3	ACG 3 5
Val	GTT 9 5	GAT 1 2
Leu	GTC 0 4	GCC 6 3
Gln	GTA 0 2	GCA 0 2
Pro	GTG 8 6	GGC 5 5
Arg		GAG 0 6
Ala		GGG 0 1

CLDN1 211 AA

	Major	W3110
Phe	TTT 8 5	TCT 0 2
Tyr	TAT 9 5	TAC 0 4
Cys	TGT 0 3	TGC 0 3
Leu	TTA 0 3	TAA 0 1
Stop	TGA 0 0	TGA 0 0
Trp	TGG 0 3	TGG 0 2
Pro	CTT 0 2	CCT 0 1
His	CTC 0 2	CCC 0 1
Arg	CGT 0 2	CGA 0 0
Ala	CTA 0 0	CCA 0 2
Gly	CTG 22 11	CCG 10 6
Met	ATT 13 8	ACT 0 2
Asp	ATC 0 5	ACC 12 6
Asn	ATA 0 0	ACA 0 1
Thr	ATG 8 8	ACG 0 3
Val	GTT 9 5	GAT 0 4
Leu	GTC 0 3	GCC 0 6
Gln	GTA 0 2	GCA 0 5
Pro	GTG 15 6	GGC 23 8
Arg		GAG 0 1
Ala		GGG 0 2

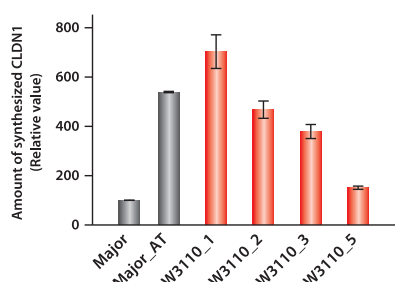
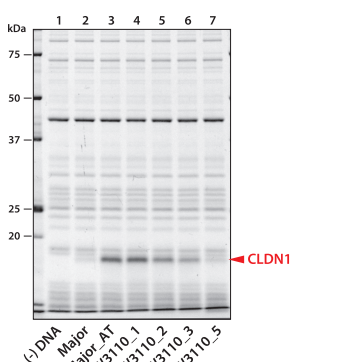
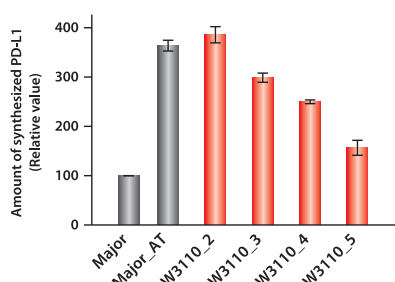
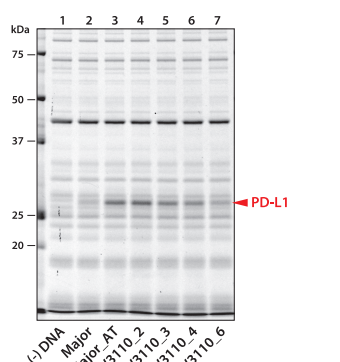
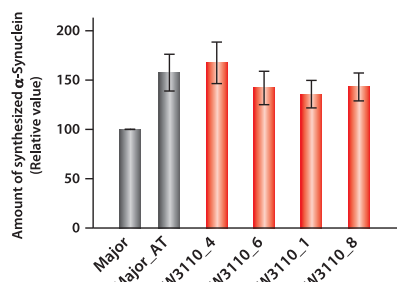
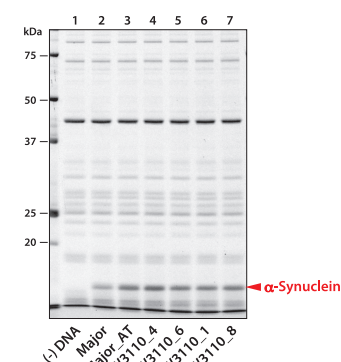
	Major	Major_AT	W3110_4	W3110_6	W3110_1	W3110_8
GC (%)	40.0	33.3	26.7	26.7	23.3	23.3
ΔG (kcal/mol)	-5.4	-3.6	-3.6	-5.7	-6.5	-7.2

	Major	Major_AT	W3110_2	W3110_3	W3110_4	W3110_5
GC (%)	56.7	43.3	33.3	40.0	40.0	53.3
ΔG (kcal/mol)	-4.0	-5.1	-2.0	-4.2	-3.3	-4.0

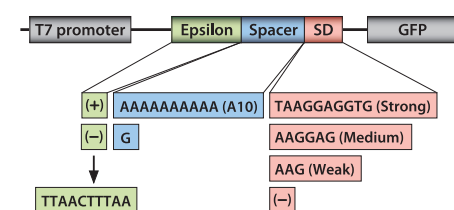
	Major	Major_AT	W3110_1	W3110_2	W3110_3	W3110_5
GC (%)	73.3	53.3	-1.6	-7.5	-7.3	73.3
ΔG (kcal/mol)	-6.6	-6.4	-1.6	-7.5	-7.3	-6.6

Major: preferentially use the most frequently used codon (designed by Supplier A)
W3110: use the codon according to the codon usage in E.coli W3110 strain (designed by CodHonEditor (Ref.2))
Ref.2; Takai K. (2016) Nucleosides Nucleotides Nucleic Acids, vol. 35, p. 223.

GC (%): GC content calculated from +1 to +30 nt.
ΔG (kcal/mol): Minimum free energy calculated from -20 to +30 nt.



3. Effect of 5'-UTR

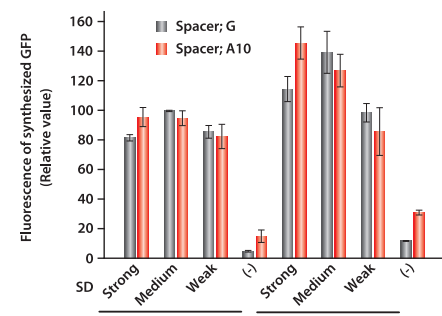


Ref.3; O'Connor M. et al. (2001) Nucleic Acids Res., vol. 29, p. 1420.
Ref.4; Takahashi S. et al. (2013) J. Am. Chem. Soc., vol. 135, p. 13096.

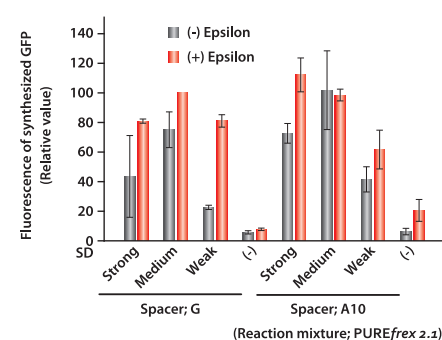
Currently used construct; (+) - G - Medium

< Effect of Epsilon, Spacer and SD sequence >

Comparison of spacer region



Comparison of presence or absence of Epsilon sequence



SD sequence;
Medium > Strong > Weak > (-) in case of G
Strong > Medium > Weak > (-) in case of A10
Epsilon sequence;
(+) > (-) in all case

Conclusion

コドン:

- 一つのアミノ酸に対し、大腸菌の使用頻度に従って複数のコドンを使用したほうが良い。
- ORF全体で使用するコドンに関係なく、N末端のコドンは、ATリッチのコドンを使用したほうが良い。

5'UTR:

- SD配列は必須であり、3塩基 (AAG) でも機能するが、6残基以上にしたほうが良い。
- Epsilon配列は、翻訳効率を増加させる。特に、強いSD配列や弱いSD配列の場合に増加効果が大きい。

Acknowledgement

CodHonEditorについて情報を提供していただいた愛媛大学の高井和幸教授に感謝いたします。

@ entire ORF

Codon according to codon usage ≥ Major codon

@ 5'-terminus of ORF

AT-rich codon > Major codon