2.4 In Vitro Protein Expression In E.coli system

Usually, there are many problems in expression of eukaryotic genes in prokaryote. Due to to lack of Golgi, E.coli cannot modify the structure of protein, such as promoting the formation of the sulfur bond between two cysteine, cutting off the signal peptide. In addition, most foreign proteins were expressed as inclusion bodies in cells making it difficult to extract proteins. E.coli and eukaryotes tend to use different codons, which makes protein expression inefficient.

After careful consideration, in order to amplify the expressed protein efficiently, we decide to do five respect of work. The first one is to optimize the codons. We decide to use professional software to help our optimization to find the best codons for each of our sequences in prokaryotic expression.

	T	AA	FRQ	C	AA	FRQ	A	AA	FRQ	G	AA	FRQ
T	TTT	F	19.7	TCT	S	5.7	TAT	Y	16.8	TGT	C	5.9
	TTC	F	15	TCC	S	5.5	TAC	Y	14.6	TGC	C	8
	TTA	L	15.2	TCA	S	7.8	TAA	stop	1.8	TGA	stop	1
	TTG	L	11.9	TCG	S	8	**TAG**	**stop**	**0**	TGG	W	10.7
C	CTT	L	12	CCT	P	8.4	CAT	H	15.8	CGT	R	21.1
	CTC	L	11	CCC	P	6.4	CAC	H	13.1	CGC	R	26
	CTA	L	5.3	CCA	P	6.6	CAA	Q	12.1	CGA	R	4.3
	CTG	L	46.9	CCG	P	26.7	CAG	Q	27.7	CGG	R	4.1
A	ATT	I	30.5	**ACT**	**T**	**8**	AAT	N	21.9	AGT	S	7.2
	ATC	I	30.6	ACC	T	22.8	AAC	N	24.4	AGC	S	16.6
	ATA	I	30.7	**ACA**	**T**	**6.4**	AAA	K	33.2	**AGA**	**R**	**1.4**
	ATG	M	30.8	ACG	T	11.5	AAG	K	12.1	**AGG**	**R**	**1.6**
G	GTT	V	16.8	GCT	A	10.7	GAT	D	37.9	GGT	G	21.3
	GTC	V	16.9	GCC	A	31.6	GAC	D	20.5	GGC	G	33.4
	GTA	V	16.1	GCA	A	21.1	GAA	E	43.7	**GGA**	**G**	**9.2**
	GTG	V	16.11	GCG	A	38.5	GAG	E	18.4	GGG	**G**	**8.6**

Statistics of codon usage frequency in E.coli

The second one is to attach some tags on the protein tails, such as 6XHistidines,which is widely used in protein purification. The expression vector pET28a (+) in our experiment contains a sequence encoding poly-histidine, 6XHis-Tag. Then a recombinant target protein with His-Tag can be obtained. His-Tag can bind to Ni^{2+} ion, which is convenient for the purification of target protein. The protein added with His-Tag can be purified by Ni^{2+} affinity chromatography column under non-denaturing conditions. Besides, Glutathione S-transferase (GST) is the second most widely used recombinant protein marker after histidine. GST is fused to the N-terminal of the target protein and can be purified by glutathione-agarose affinity chromatography. Because the affinity of GST to substrate-reduced glutathione is sub-molar, the affinity chromatography resin formed by glutathione solidified in agarose has a high purification efficiency for GST and its fusion protein. GST affinity purification of fusion protein mainly includes three steps: binding, washing and elution. The whole process requires only one or two buffers, which can be purified in large quantities or in small quantities to achieve rapid purification. It is suitable for small amount preparation of recombinant E. coli protein in our experiment.

The third one is to confirm the best expression vector. According to the different promoters, vectors can be roughly divided into heat-induced promoters, such as lambda PL, cspA, and other IPTG-induced promoters, such as lac, trc, tac, T5/lac operator, etc. According to the types of expressed proteins, they can be divided into simple expression vectors and fusion expression vectors. Fusion expression vector is to add special sequences at the N or C ends of the target protein to improve the solubility of the protein, promote the correct folding of the protein, achieve rapid affinity purification of the target protein, or achieve the expression and localization of the target protein. Commonly used fusion tags for affinity purification include Poly-Arg, Poly-His, Strep-Tag II, S-tag, MBP and so on. His-Tag and GST-Tag are currently the widest used. At present, the expression

vectors commonly used are mainly pET series provided by Novagen and pQE series provided by Qiagen Company. In addition to the His tag, reduced glutathione S-transferase (GST) is another fusion tag commonly used in laboratories. It can be rapidly purified by affinity chromatography with reduced glutathione agarose. In addition, compared with His, GST can promote the correct folding of target proteins and improve the solubility of target proteins in many cases. Therefore, for those proteins which are easy to form inclusion bodies expressed by his tag, GST fusion expression can be attempted to improve. Of course, GST has a high molecular weight (26kDa), which may affect the activity of the target protein, so it is necessary to remove GST in many cases. At present, GST fusion expression system is mainly provided by GE Healthcare (formerly Amersham).

The fourth one is to detect the expression of protein (expression level, expression distribution, etc.) and to analyze the purity of the purified target protein. In order to achieve this aim, we decide to use polyacrylamide gel electrophoresis(SDS-PAGE).

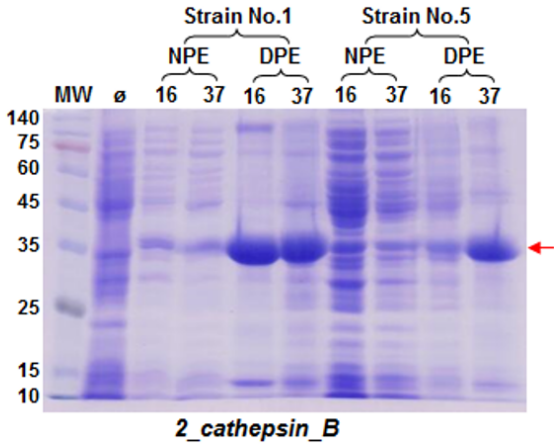

Figure 1. Expression tests of the target protein. Analysis of NPE and DPE prepared as described in §3.2.
MW. Molecular weight marker. **Ø.** Non-induced bacteria culture (negative control).
16 and 37. Incubation temperature (°C) during induction with IPTG.
Induction with IPTG 1mM during 16h at 16°C, or during 4h for other temperatures.

The last one is to establish appropriate expression conditions. There are multiple factor may influence the efficiency of our protein expression. Firstly, we need to choose the best host E.coli. As expression host bacteria, it must have several basic characteristics: stable heredity, fast growth and stable expression protein. During the specific operation, the specific expression host bacteria were selected according to the characteristics of the expression vectors used and the codon composition of the target gene. BL21 (DE3) derivative series is based on the classical T7 expression system BL21 (DE3). Novagen has developed some special expression host cells. For example: Origami (DE3), Origami B (DE3) and Rosetta-gami (DE3) strains have trxB and gor double mutations. The strains with trxB and gor mutations are more likely to promote the formation of disulfide bonds and make the protein soluble and active than those with trxB mutations alone. Secondly, to ensure the stability of the expressed proteins and avoid the degradation of proteases, we can use some protease-deficient expression hosts. Thirdly, in the process of expression, we can add some protease inhibitors to the culture system to solve the problem. Compared with stable expression of exogenous proteins, more time to ensure the soluble expression

of target proteins is the main work to establish the expression conditions. Many exogenous genes can easily form incompatible inclusion bodies when expressed in E.coli. Thirdly, there are many methods to break E. coli cells, such as repeated freeze-thaw method, osmotic impact, ultrasonic breaking, pressure breaking, etc. Ultrasound crushing and pressure crushing are common methods in our experiment because they have the least effect on the expression protein.