

原核表达常见问题解析

一、不知道蛋白的特性及结构？

1. 首先，要确定一件事，那就是这几个蛋白质有人研究过没有？还是最新发现的蛋白质？如果没有人研究过，那就得先用测部分氨基酸，然后设计引物克隆了。
2. 如果有人研究过，可以根据软件来预测。如用 swiss—pdb 软件，但这个是要有氨基酸序列，只知道基因序列，可以在 NCBI 上进行 blastx，得到蛋白。

二、如何选择蛋白表达的宿主菌？

原核系统和真核系统偏爱的密码子有不同，因此，在用原核系统表达真核蛋白的时候，真核基因中的一些密码子对于原核细胞来说可能是稀有密码子，从而导致表达效率和表达水平很低。一般选择菌主可看下表：

| 原核表达现象 | 建议尝试菌株 |
|----------|---|
| 蛋白为毒蛋白 | BL21(DE3)pLysS T7pLysY OverExpressC43(DE3) |
| 蛋白不表达 | 序列含有稀有密码子 Rosetta(DE3) |
| 蛋白明显降解 | ArcticExpress(DE3) |
| 蛋白表达为包涵体 | OrigamiB(DE3) ArcticExpress(DE3) |
| 蛋白表达不理想 | 二硫键错误折叠 OrigamiB (DE3) Shuffle T7-B Shuffle T7-K12 |
| 过高的本底表达 | BL21(DE3)pLysS |

三、质粒测序正确，蛋白无法表达

1. 分析一下稀有密码子，如果比较多，可以尝试 Rosetta (DE3)
2. 可能是基因本身的问题。RNA3' 的特殊结构可能导致转录出现问题，这种情况可以尝试融合表达，譬如 pET-32a。

3. 也许是表达量太低，也可以试一下 western blot，定性的检测一下。

四、如果 IPTG 诱导后细胞停止了生长，是不是表示细胞死了？

T7 RNA 聚合酶非常活跃，T7 转录和翻译信号也极强，因此一旦诱导，细胞的主要生理活动都向着表达目的蛋白的方面转化。在通常情况下，细胞将停止生长，形成克隆的能力大大降低，但并未死亡。也有一些例外情况，例如一些特别的基因或者极为严紧的载体/宿主菌组合（比如含有 pLysE 的宿主菌）等，这时在诱导后菌落还是会继续生长。

五、如何提高重组蛋白在原核细胞里的表达水平，特别是可溶性表达？

这个问题是最困扰做原核蛋白表达纯化的人的。比如大肠杆菌表达蛋白本身表达量就大，但是表达的大都是包涵体，想要获得可溶性蛋白，就需要做复性，或是再设计实验时就想办法让其在上清中表达。一般就要通过基因优化，载体宿主优化筛选，表达条件优化，诱导条件优化等等。

1. 降低重组蛋白合成的速率

可溶性蛋白的产率取决于蛋白的合成速率，蛋白的折叠速率，以及聚集的速率。高水平表达时，肽链聚集的速率一旦超过折叠速率，就会形成包涵体。因此，降低重组蛋白合成的速率有利于提高重组蛋白的可溶性表达。

2. 密码子优化

密码子优化就是根据表达系统对密码子的偏好性进行优化筛选。经过优化的基因序列往往能提高 mRNA 二级结构的稳定性，有利于新生肽段的正确折叠，提高外源活性蛋白的表达。

3. 表达温度的选择

大肠杆菌的最适生长温度在 37 ~ 39°C 之间，但此温度下极易生成包涵体蛋白，降低可溶性蛋白的表达，而低温培养条件下表达外源蛋白能有效地增加可溶蛋白的比例。

4. 诱导条件优化

摇瓶培养时，应选用低菌体浓度诱导，因为在低菌浓度下菌体处于对数生长期，生长活跃，有利于表达可溶性蛋白。如果能保证合理的补料与充分的通气，在较高菌浓度下诱导也同样可能获得可溶蛋白的高效表达。在某些情况下，诱导剂的流加能显著提高可溶蛋白的表达水平。

六、目的蛋白总是以不可溶的形式出现

在原核蛋白表达纯化中目的蛋白经常发生错误的折叠，并聚集成为包涵体。经过诱导，目的蛋白通常可达细胞总蛋白的 50% 以上。虽然有一定比例的蛋白以可溶的单体形式存在，而多达 95%（甚至更多）的蛋白则在包涵体中。

实践中，有很多实验室采取降低诱导温度，例如 25–30°C，或降低 IPTG 浓度(0.01–0.1mM)并延长诱导时间，还有采用特别的培养基等方法获得更多的可溶蛋白。然而，让目的蛋白以包涵体形式聚集也并非总是坏事。不溶态在某些情况下非常有利：

1. 形成包涵体是目的蛋白表达量很高的表现。
2. 作为初步分离，将目的蛋白的包涵体纯化出来非常简便。用核酸酶处理，并经简单洗涤，通常可以获得纯度达 75–95%的目的蛋白。
3. 存在于包涵体中的目的蛋白通常可以免于蛋白酶的水解破坏作用。研究者采用下述方法加强蛋白重新折叠的效果，并在很多蛋白上取得了好结果：当蛋白还结合在树脂上时，使用 6M–8M 梯度盐酸胍、1mM 还原型及 0.2mM 氧化型谷胱甘肽处理，继后用咪唑正常洗脱。有些实验室在透析去除变性剂的过程中加入底物或类似物，也有帮助酶折叠的效果。

七、表达纯化全长蛋白时，为什么有截短的产物？

1. 最直接原因就是蛋白水解了。
2. 可能是翻译起始问题。起始密码子 AUG (Met) 的上游在 RNA 编码区有类似于核糖体结合位点序列 (AAGGAGG)的间隔序列(通常是 5–13 个核苷酸)时，容易出现这种问题。截短蛋白在纯化全长蛋白的操作时往往造成麻烦。解决途径之一即是选用两端都带有融合标签的 pET 载体，例如 pET-28 和 pET-30 系列载体在 N-端和 C-端都有 His-Tag。这样，全长蛋白就可以通过提高咪唑浓度，在洗脱时将截短蛋白与全长蛋白区分开。pET-29 和 pET-30 系列载体在 N-端融合有 S-Tag 序列而 C-有 His-Tag，因此可以先后用固定化 S- 蛋白和 His- Bind 树脂亲和纯化以获取全长蛋白。

八、His 标签包在蛋白质结构内部还能用蛋白酶切除吗？

如果重组蛋白纯化标签被包在内部，一般可以增加所提蛋白质样品的溶解度来解决，比如适量增加变性试剂浓度或类型使得蛋白质的肽链松散开暴露出组氨酸标签。另外，还可以加大离子型、非离子型或两性离子表面活性剂的浓度或类型，加大 DTT 浓度（DTT 浓度 < 1mmol/L，否则 DTT 会与 Ni²⁺ 结合），适当调整调高盐浓度（氯化钠溶液低于 2mol/L，从低浓度到高浓度逐渐分梯度上升），总之，作用就是增加蛋白质的可溶性，因为 6His 的亲水性较强，增加可溶性之后 6His 容易暴露在水相。加大 DTT 浓度也可以切断二硫键使得蛋白质的肽链松散，进而使 6His 容易暴露在水相，从而方便纯化和酶切。再者重新构建，不加 His-tag。

九、目的蛋白大于 100KD 或含有多个亚基，是否可以用原核表达系统？

在细菌中已成功表达了很多大蛋白。多亚基复合物则可以通过分别表达各亚基，然后在有尿素的溶液中，将各组分以适当比例混和，再透析除去变性剂。

十、是否所有的位点特异性蛋白酶都有一样的酶切特性，只是识别位点不同？

位点特异性蛋白酶(例如凝血酶、肠激酶和 Xa 因子)通常被用来切割融合蛋白。这些酶的活性和第二点酶切倾向性很不相同。凝血酶在这三种中是特异性活性最高的，能够有效切质量比仅为 1:2000 的蛋白。Xa 因子似乎对于切点周围的序列很敏感，经常会出现特异性位点切割不理想却发生别的位点被切割的情况。肠激酶的专一性是上述三种中最好的，但由于切割效率低（通常要求质量比达到 1:10）而显得比较昂贵。另一个需要考虑的问题是切割完成之后，是否需要去除蛋白酶。在质量比较高的酶切反应进行完后，通常要通过色谱法处理。比较方便可行的方法是采用生物素化凝血酶结合链亲和素琼脂糖一起使用。尽管没有一种蛋白酶完美无缺，但凝血酶还算是活性高、专一性好的典型。

十一、如果去除信号肽，会不会影响蛋白本身的表达？

蛋白表达强度不受信号肽调控，所以去除信号肽不会影响蛋白质表达。但是信号肽一方面参与蛋白质折叠成特定空间结构，另一方面它还决定蛋白最终被定位到特定的亚细胞区，一般蛋白只有在特定亚细胞区才能发挥其功能，所以信号肽对蛋白最终活性还是有很大影响的。

十二、E. coli 会去除成熟蛋白 N 端的甲硫氨酸吗？这种加工对目的蛋白的稳定性有何影响？

N 端 fMet 是否被去除受倒数第二个氨基酸影响。这个加工过程由甲硫氨酸氨基肽酶催化，并受以下关系支配：去除的困难程度随倒数第二位氨基酸的支链大小的增加而降低。研究者在实验中发现以下氨基酸种类出现在倒数第二的位置上时，此加工过程极少发生或根本没有发生：His、Gln、Glu、Phe、Met、Lys、Tyr、Trp 和 Arg。而当倒数第二位上是其它种类的氨基酸时，加工过程发生的可能从 16%到97%，确定了在大肠杆菌中蛋白氨基末端氨基酸与其稳定性之间的关系，也即“N-端原则”。具报导：如果下列氨基酸位于氨基端时蛋白的半衰期仅为 2 分钟：Arg、Lys、Phe、Leu、Trp 和 Tyr。相反，其它氨基酸位于氨基端时可以使受检蛋白的半衰期长达 10 小时以上。综上两条研究可以大致得出结论：Leu 在氨基末端时很容易因为 fMet 被去除而暴露出来，从而导致蛋白的迅速水解。显然，当采用 pET 载体上的 Nco I 或 Nde I 位点表达非融合蛋白时，完全不必担心 Leu 的密码子出现在倒数第二的位置上。