

## 基因优化实验案例

1. 案例概述	2
2. 实验设计	2
3. 密码子使用偏性的优化	2
4. GC 含量的优化	3
5. mRNA 二级结构的优化	4
6. 核糖体结合位点的优化	5
7. 表达检测结果	5
8. 蛋白纯化与二聚体检测	6
9. 服务网址	6

## 1. 案例概述

基因转录翻译是一个复杂的过程，每一步都关系到蛋白质的最终表达，其中密码子使用频率、mRNA 二级结构和 GC 含量等都对蛋白表达有着重要影响，因此需要对其进行优化。德泰生物综合考虑蛋白质表达过程中的影响因素，自主研发了 MaxCodon™ 软件，可以优化这些关键因素，从根本上提高蛋白表达的效率。

某客户在表达一个二聚体蛋白时，遇到了蛋白表达量低、无法检测到二聚体等问题。公司接手后，用自主研发的密码子优化软件对基因序列进行优化，显著提高了蛋白表达的效率，成功纯化得到了二聚体蛋白。

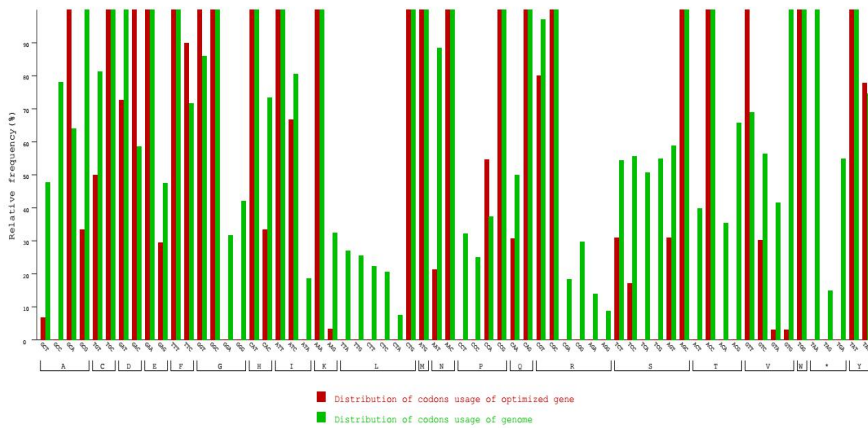
## 2. 实验设计

将客户提供的目的基因序列，输入德泰自主研发的 MaxCodon™ 软件中进行优化。用宿主细胞使用频率高的密码子替换稀有密码子，提高目的基因在宿主细胞中的 CAI（密码子适应指数），预测基因序列的 mRNA 二级结构，减少复杂二级结构的数量，并调节将序列的 GC 含量调整到合适的水平。MaxCodon™ 软件还可以通过优化算法，减少核糖体起始位点的类似序列，有效避免翻译起始的紊乱。

完成基因优化后，在序列 C 端添加 His 标签，方便纯化。之后合成基因，插入载体，转到大肠杆菌中表达。培养后，破碎菌体，收集上清，电泳检测。然后纯化上清，在还原与非还原条件下电泳，检测二聚体是否存在。

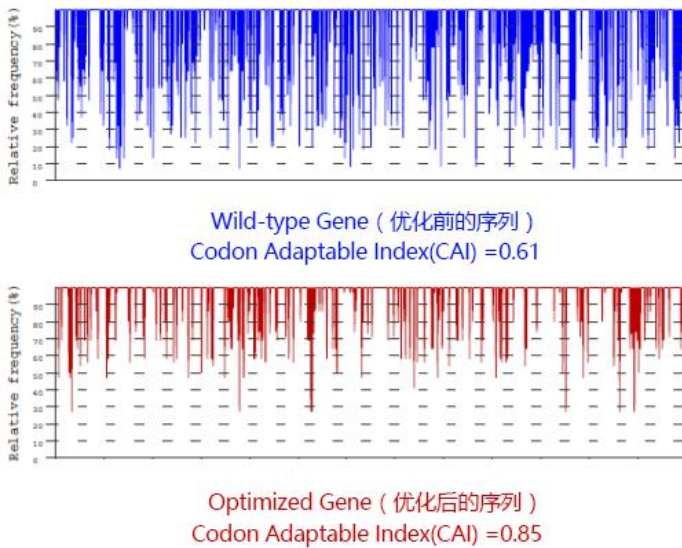
## 3. 密码子使用偏性的优化

不同物种的同义密码子使用频率是不同的，这种密码子使用偏性与 tRNA 的丰度有关。由于 tRNA 是 mRNA 翻译成蛋白质的执行者，tRNA 的丰度与蛋白质表达的效率直接相关。因此德泰生物在优化密码子时将目的基因中的稀有密码子替换成宿主细胞常用的密码子，在蛋白质序列不变的情况下，提高密码子适应指数（CAI），从而提高目的基因表达水平。



红色条带：优化后基因密码子使用分布  
绿色条带：基因原始序列密码子使用分布

图 3-1 密码子相对使用频率分析

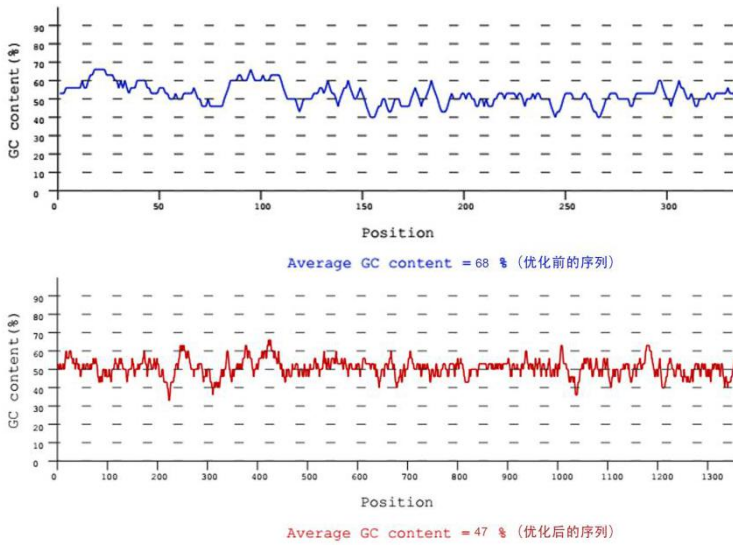


密码子适应指数 (CAI)：反映编码区同义密码子与密码子最佳使用相符合的程度 取值范围在 0-1 之间。CAI 的值越高，该基因与宿主细胞密码子使用偏性的符合程度就越高。

图 3-2 优化前后序列的密码子适用指数

## 4. GC 含量的优化

在碱基配对过程中，形成的氢键数量越多，配对就越稳定。由于 GC 配对能形成 3 个氢键，比 AU 或者 AT 配对都来得更加稳定。因此 GC 含量不仅直接影响 PCR 退火温度，也影响着 mRNA 二级结构的稳定性，间接关系到 mRNA 翻译成蛋白质的效率。过低 GC 含量可能会导致转录终止，过高 GC 含量则容易使 mRNA 形成稳定二级结构，阻碍蛋白表达。我们研发的 MaxCodon™ 软件可以通过基因优化，调节 GC 含量，提高蛋白表达效率。



目的基因的 GC 含量为 68%，该基因原始序列 GC 含量过高，不利于蛋白表达。对其进行优化后，GC 含量降低到了 47%。

图 4-1 GC 含量分析

## 5. mRNA 二级结构的优化

mRNA 形成的二级结构是影响翻译过程的重要因素，mRNA 二级结构（特别是核糖体结合位点（RBS）附近）越复杂越稳定，肽链的延长速率就越低，翻译的效率和蛋白表达水平也就越低。MaxCodon™ 在用同义密码子替换稀有密码子的同时，能够预测 mRNA 二级结构，减少序列中的碱基配对，从而避免产生复杂稳定的 mRNA 二级结构，提高蛋白表达的效率。

二级结构种类	序列优化前	序列优化后
小型发夹结构	12	4
中型发夹结构	2	0
大型发夹结构	1	0

表 5-1 优化前后二级结构数量对比

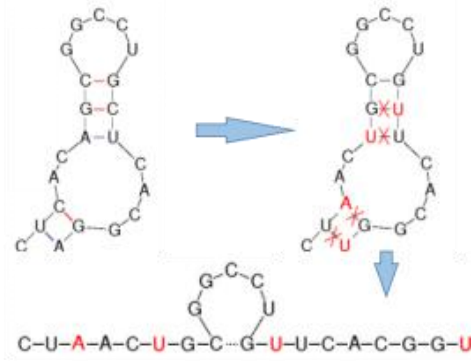


图 5-1 基因优化前后 mRNA 二级结构

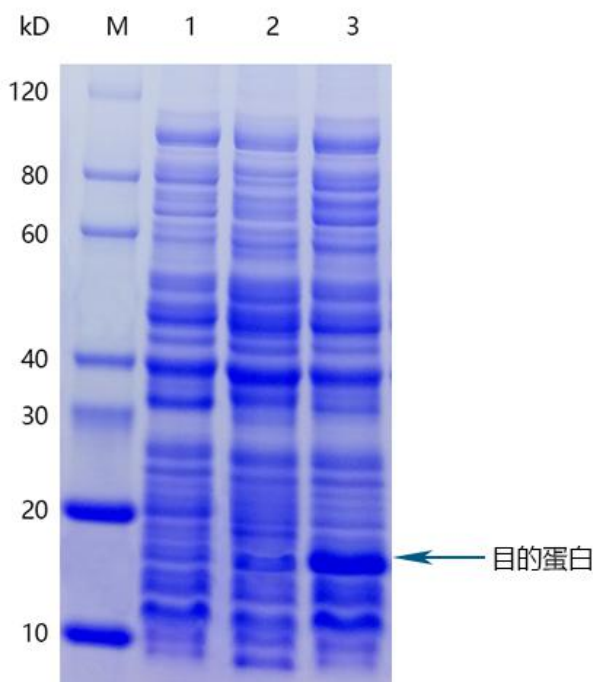
通过图 5-1 所示的同义密码子替换，用软件分析预测，达到减少二级结构数量的目的。

## 6. 核糖体结合位点的优化

核糖体结合位点是从转录起始位点开始延伸的一段碱基序列，核糖体结合位点对形成翻译起始复合物是必需的。类似的序列会造成翻译起始的紊乱。MaxCodon™ 软件通过优化算法，可以有效避免这一情况发生。

## 7. 表达检测结果

目的基因序列经过优化以后，蛋白表达效率明显提高，目的蛋白所在的条带比优化前更为清晰，蛋白表达量显著增加。



Lane M : Protein marker  
Lane 1 : 阴性对照  
Lane 2 : 未经过优化的蛋白表达  
Lane 3 : MaxCodon™ 优化后的蛋白表达

图 7-1 基因优化前后的蛋白表达量

## 8. 蛋白纯化与二聚体检测

重组表达的蛋白添加了 His 标签，使用 Ni-IDA 亲和层析，层析过柱之后，在还原和非还原状态下分别进行电泳，检测是否有二聚体存在。最终结果显示，在非还原条件下电泳，在单体蛋白两倍左右分子量的位置上确实出现了清晰的条带，证实上清中确实存在二聚体，和客户的描述一致。

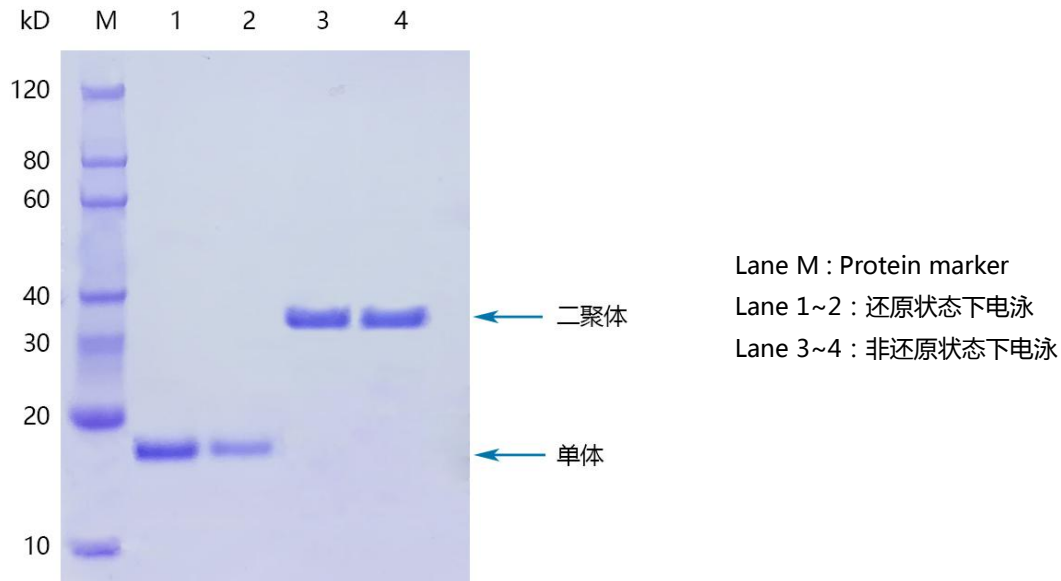


图 8-1 还原和非还原状态下电泳

## 9. 服务网址

[www.detaibio.com/gene-to-protein-service.html](http://www.detaibio.com/gene-to-protein-service.html)