

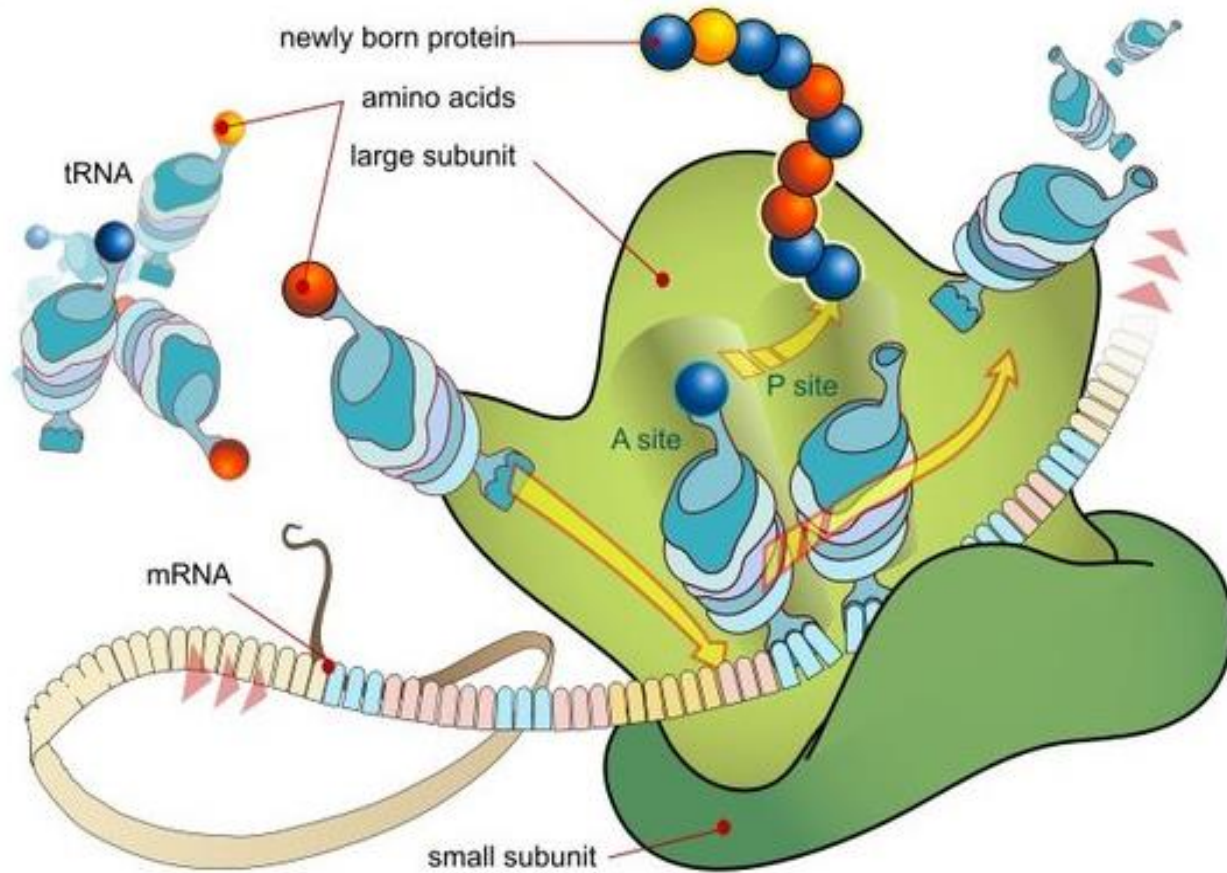
Determinants of translation efficiency and accuracy

Background

- 蛋白质的重要性：蛋白质（protein）是生命的物质基础，没有蛋白质就没有生命。因此，它是与生命及与各种形式的生命活动紧密联系在一起的物质（ Translation accuracy ）。机体中的每一个细胞和所有重要组成部分都有蛋白质参与，蛋白质占人体重量的16%~20%（ Translation efficiency ）。
- 蛋白质的翻译主要原件：mRNA，tRNA，氨基酸，核糖体，ATP，GTP，氨酰tRNA合成酶，翻译因子等。

那么哪些因素可以决定蛋白质的翻译效率和准确性呢？

蛋白质的翻译过程



定义

- 蛋白质的翻译效率：蛋白质翻译的产生量和消耗量
- 蛋白质翻译的准确性：能严格将DNA上的遗传信息转换成对应的正确蛋白质

优化蛋白质翻译效率的可能因素

- 密码子的简并性和偏爱性
- tRNA可利用量
- mRNA的结构

一 密码子的简并性和偏爱性

- 密码子的简并性

编码人的氨基酸约为**20**种，但对应的密码子却有**61**种（除去**3**个终止密码子），因此一个氨基酸可以对应多个密码子

- 密码子的偏爱性

不同物种对同义密码子的选择不是随机的，物种的进化会造成其对某些密码子具有偏爱性（**Drummond and Wilke, 2008; Cannarozzi et al, 2010; Tuller et al, 2010a**）

途径

- 1) 不同的密码子对应的细胞中tRNA含量不一样以及核糖体对tRNA的识别速度也不一样（Varenne et al, 1984; Sorensen et al, 1989）
- 2) 密码子的不同也会造成mRNA的结构的不同
- 3) 密码子的偏爱性
- 4) 密码子在mRNA上的分布

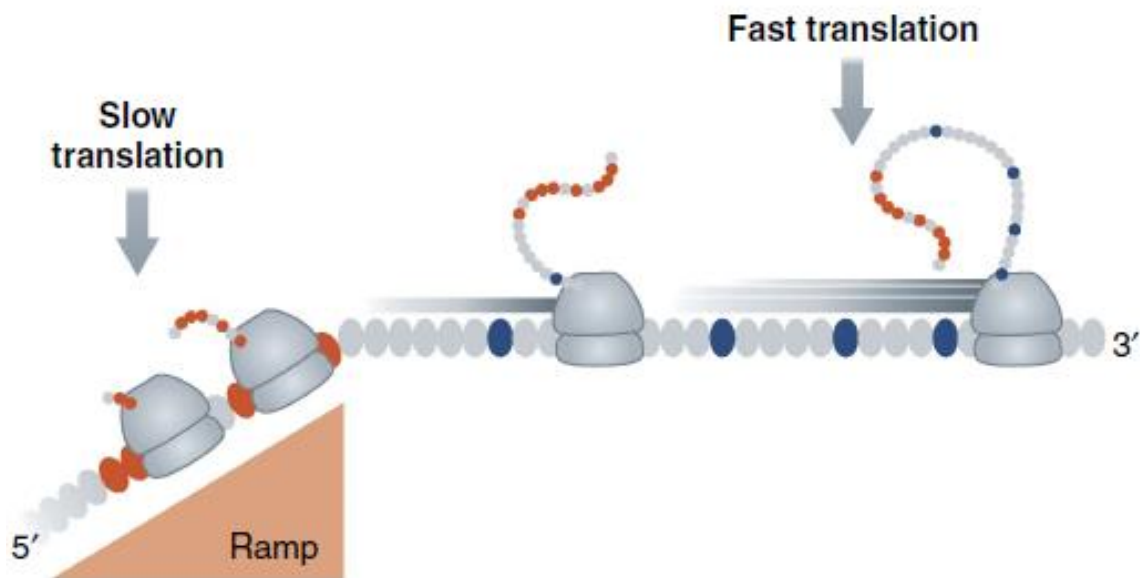
密码子的偏爱性

√蛋白质表达水平与密码子的使用存在很强的相关性
(Grantham et al, 1981; Bennetzen and Hall, 1982;
Gouy and Gautier, 1982).

√CAI (codon adaptation index)

用来表征密码子的使用情况和表达水平的相关性，编码同种氨基酸的不同密码子的使用频率不一样 (Met, Trp除外)，在一个基因对应的不同氨基酸中该物种偏爱的密码子出现的频率越高，则CAI值越大，近似的反应其翻译效率越高(Sharp and Li, 1987).

密码子的分布



研究发现，在基因的初始的 30 – 50 个密码子中，大部分的密码子倾向于选择那些对应稀有 tRNA 的密码子，尤其在高表达基因中更显著，这样会导致翻译的起始阶段变得缓慢，但是可以减少核糖体在后面的延伸过程中的脱落，提高了翻译的产量，降低了翻译的代价，进而提高了翻译效率(Tuller et al, 2010a;Huang et al, 2009.).

二 tRNA的可利用量

- tRNA是由其对应的基因转录而来，不同的tRNA对应基因的copy number是不一样的，于是将tRNA对应的基因的copy number作为tRNA丰度的指标(Dong et al, 1996; Percudani et al, 1997; Kanaya et al, 1999; Tuller et al, 2010a).
- 可利用的用于翻译的tRNA的丰度与编码该tRNA的基因拷贝数存在差异(Zaborske et al, 2009).
- Availability of 'ready-to-translate' tRNA

如何影响

√翻译的延伸速率与密码子对应的可利用的tRNA的丰度成正相关(Varenne et al, 1984).

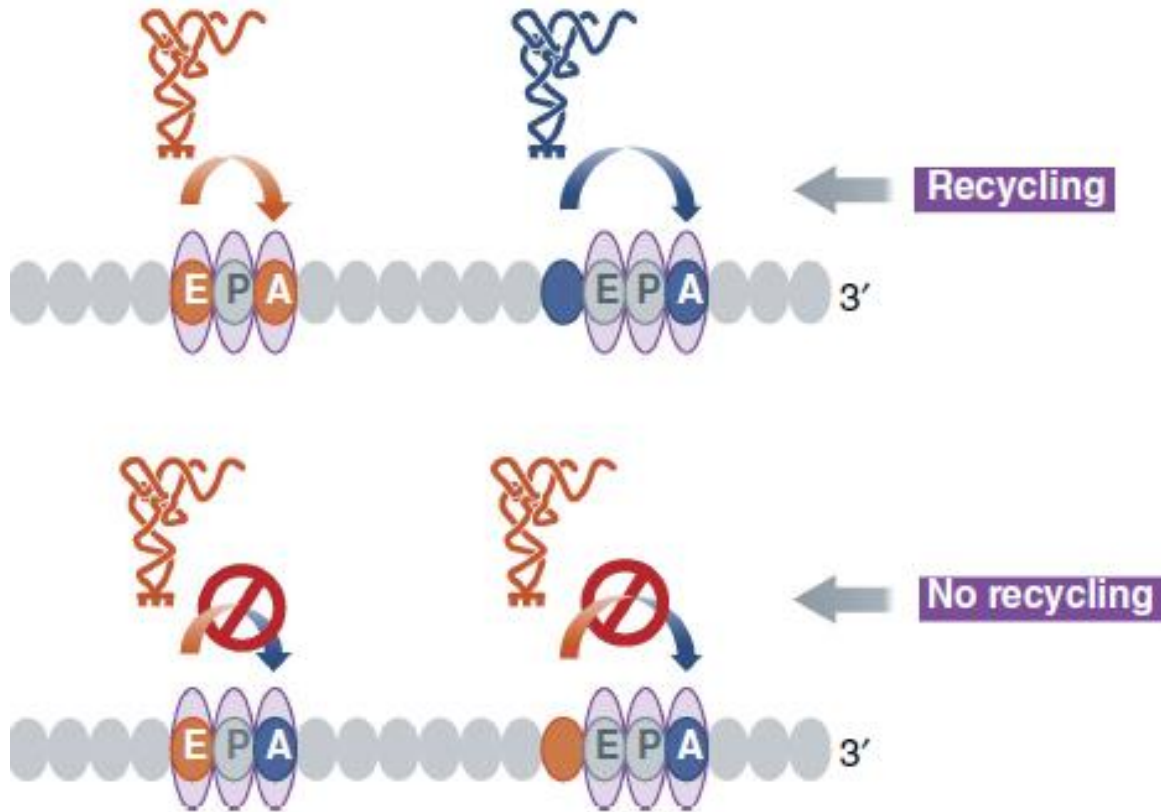
√TAI (tRNA Adaptation index) (dos Reis et al, 2004)

用可利用的tRNA含量来表示翻译效率，人们发现翻译效率与密码子对应的tRNA含量成正比。

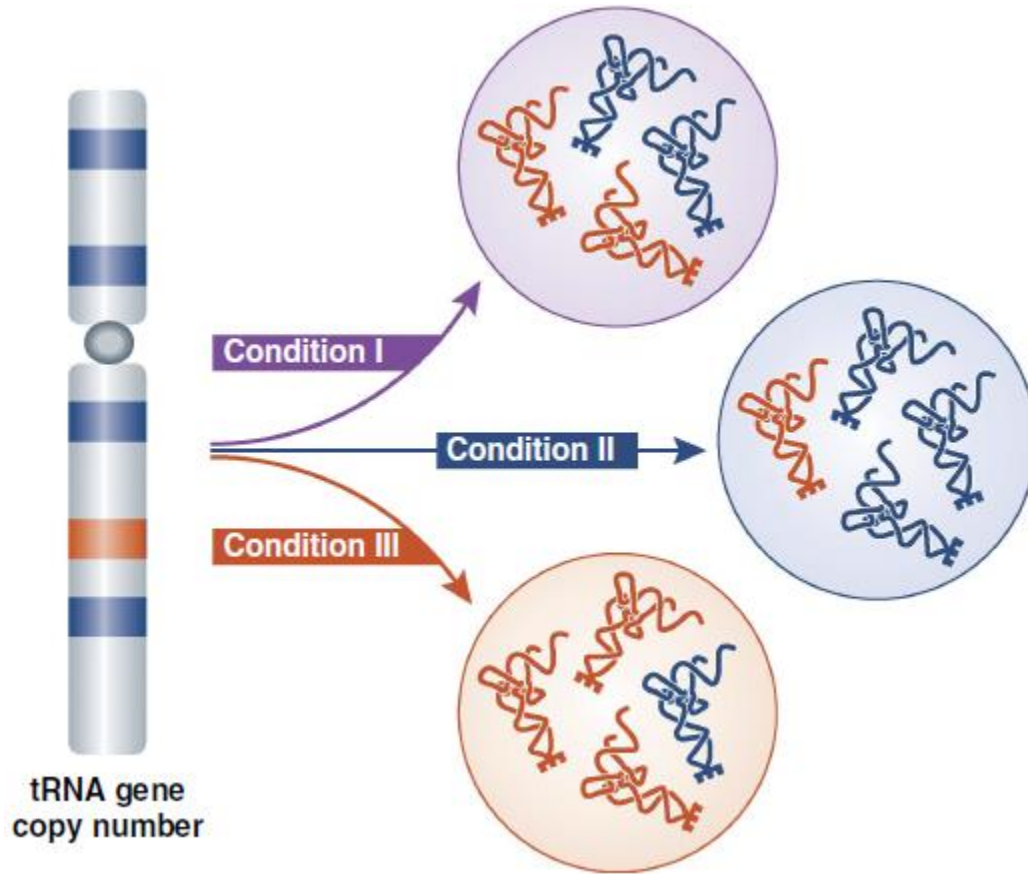
因此可利用的tRNA对翻译的效率有影响

Problem

- 翻译延长的速率与mRNA中密码子对应的可利用的tRNA含量成正相关，在一定程度上可反映蛋白质翻译的效率，但并非tRNA的含量低翻译效率都会低，也并非tRNA含量高其翻译效率也高(Cannarozzi et al, 2010).

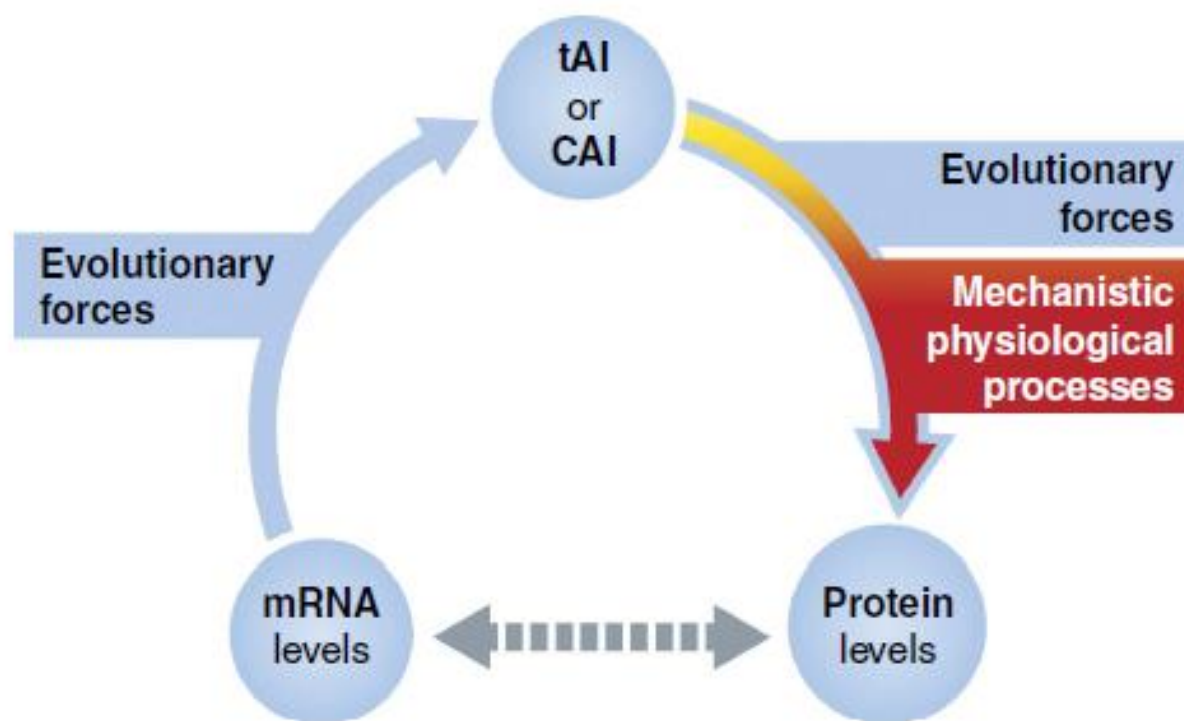


尽管在细胞整体中一些tRNA含量低，但如果上游基因翻译完后的tRNA刚好可以被下游的基因利用，则在某些低含量的tRNA的条件下基因仍可以高效表达(Cannarozzi et al, 2010).



在不同时间和空间或不同的外在条件影响下，可利用的tRNA的含量也会不同(Dittmar et al, 2006).

基因的CAI和TAI值与其对应mRNA和蛋白质的量的相关性



mRNA的水平影响翻译的水平，进而影响蛋白质的水平

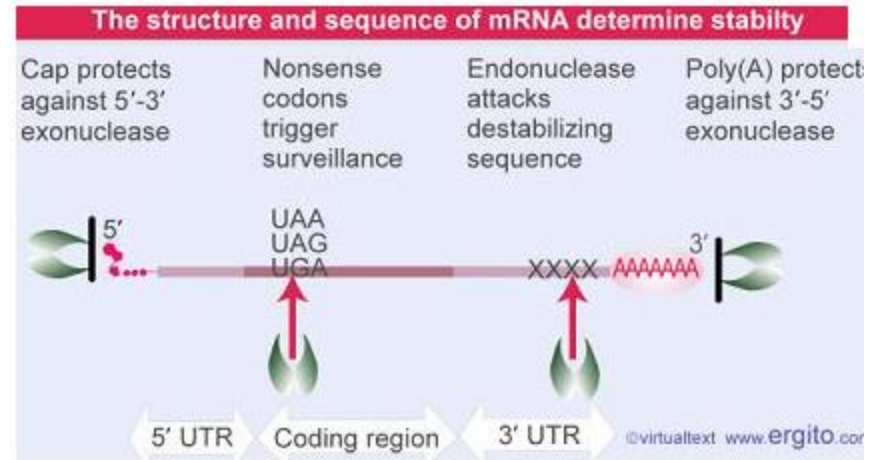
mRNA的翻译水平与用CAI或TAI测得的蛋白质的翻译效率是正相关关系，反映了进化选择压力对高表达的mRNA的密码子优化，以便不螫合太多核糖体在mRNA上，因此，在一定程度上进化选择压力对密码子的优化与mRNA的表达水平成正相关。另一方面，在一个更快的时间尺度上蛋白质的翻译效率与蛋白丰度成正相关关系，而且受进化选择压力的控制，在一定的mRNA水平上，通过对密码子优化进而提高蛋白质的翻译效率。

三 mRNA的结构

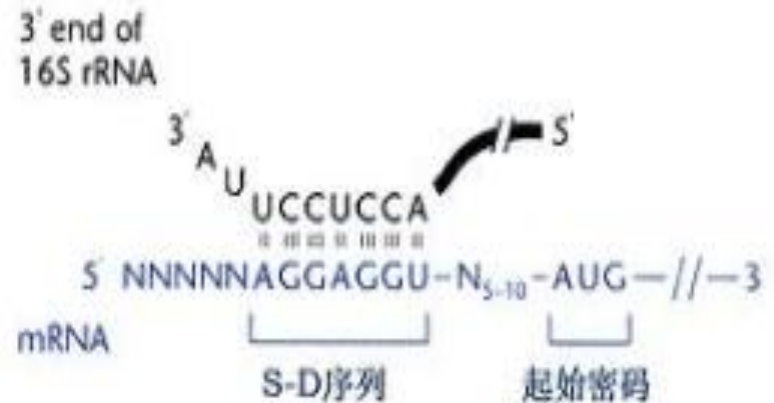
- 对翻译起始速率的影响（一级结构）
- mRNA的折叠对翻译过程中延伸的速率的影响（二级结构）

对翻译起始速率的影响

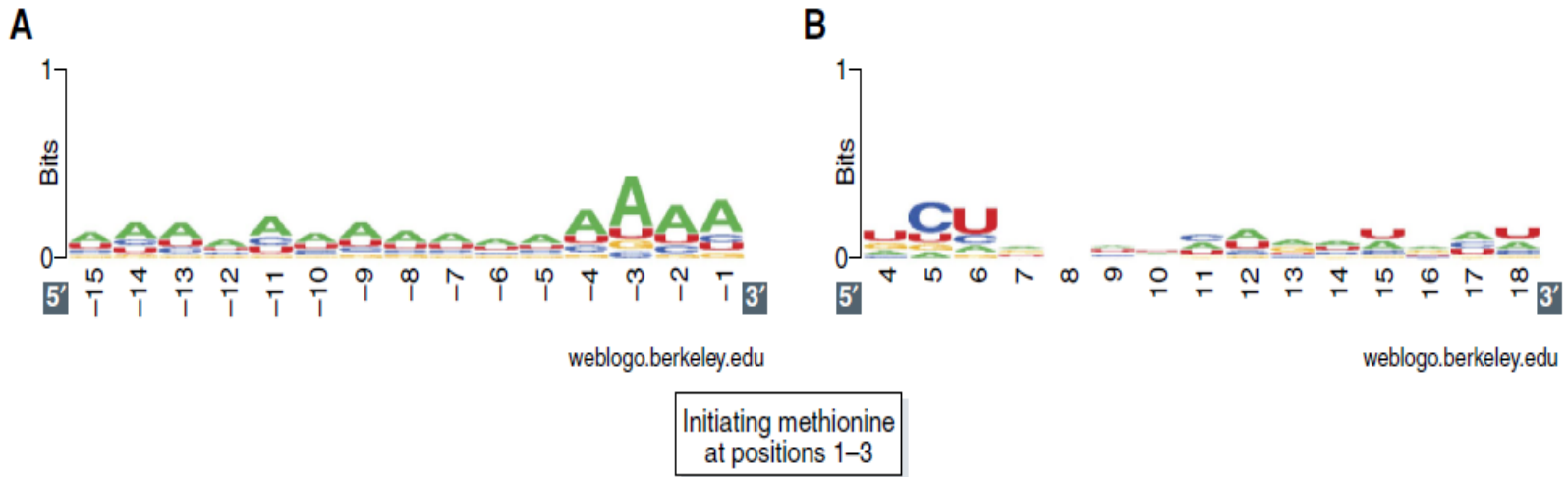
√真核生物的5'端帽子结构和3'端的poly (A)协同提高蛋白质的翻译起始效率(Gallie, 1991).



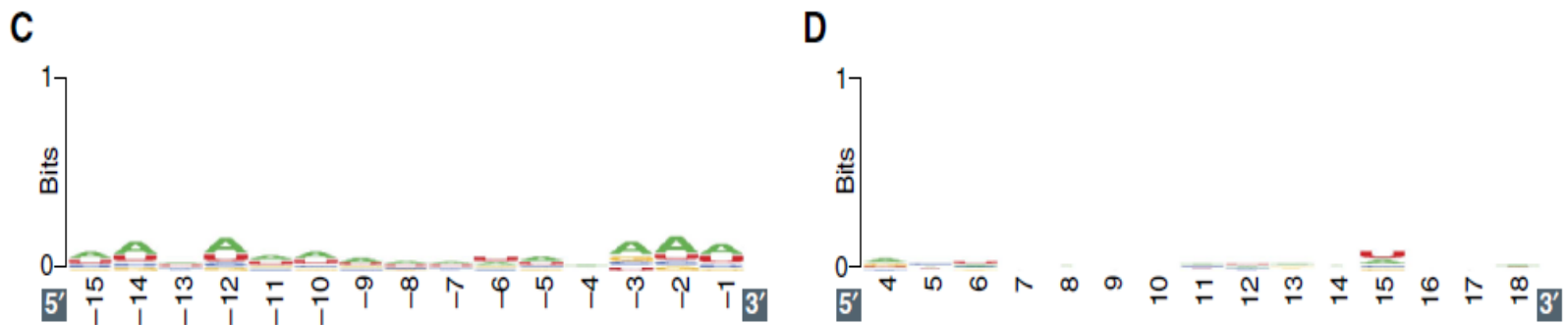
√原核生物的SD序列的RBS富含嘧啶与核糖体的16S小亚基3'端富含嘌呤正好互补(Steitz and Jakes, 1975; Jacob et al, 1987).



High ribosome-occupancy genes



Low ribosome-occupancy genes



Nagalakshmi et al (2008); Crooks et al, (2004).

mRNA的折叠对翻译速率的影响

√ mRNA分子在细胞中一般呈现二级结构或三级结构，对不同的基因结构的稳定性有的比较牢固，有的则比较松弛。

√ 研究发现mRNA5'的结构越会对翻译的效率产生影响，稳定的结构会降低翻译的效率，松弛的结构会对翻译有促进作用。

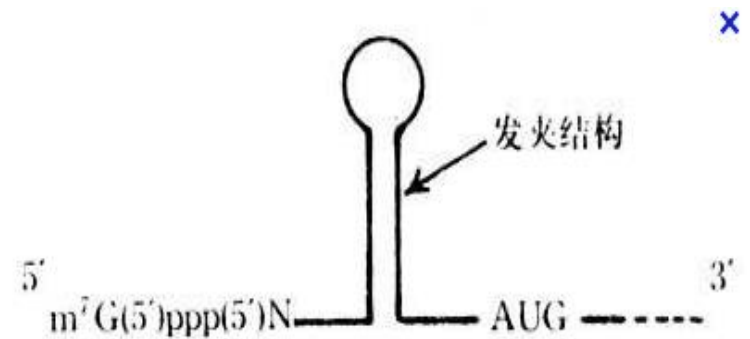


图3 真核生物 mRNA5'端帽子结构附近的发夹结构示意图

Determining the overall rate of translation:one key factor or a 'combination lock'?

mRNA的折叠在转录的起始阶段对单个基因的表达水平有显著性的影响，但是密码子偏爱性的选择对蛋白质的合成速率具有全局的增强作用，主要是通过减少核糖体在mRNA上的滞留的数目（Tuller et al , 2010b）。

人们提出以后的研究应该是分别考察单个因子对蛋白质翻译的全局速率的影响

如何对翻译的准确性进行优化

√tRNA中的反密码子与mRNA的密码子配对时允许存在一个或多个碱基的错配（ Rodnina and Wintermeyer, 2001） missense translational error 10^{-5} in yeast cells（ Stansfield et al, 1998）。

√编码错误的蛋白很可能导致蛋白质失去其应有的功能，一方面，浪费了细胞的资源，另一方面，错误的蛋白可能是有害蛋白直接导致细胞功能紊乱或致病。

In *E. coli*, 编码同种氨基酸使用的密码子对应的高丰度的tRNA与低丰度的tRNA比较其错误率降低了ninefold (Precup and Parker, 1987).

E. coli (Stoletzki and Eyre-Walker, 2007), yeast, worm, mouse and human中, 越是保守的氨基酸使用其对应的偏爱密码子的频率越高, 并且翻译的错误率也越低

Chaperonin可以提高翻译的准确率, 缓解通过密码优化途径提过翻译准确率的压力

氨基酸对应的tRNA的丰度, 对偏爱密码子的使用频率, 以及分子伴侣都会影响蛋白质翻译的准确性

Challenges

尽管我们现在对翻译的高效性和准确性认识更加的全面了，但是我们仍然面临这一个挑战，就是将现有的和以后新认识的知识整合起来建立一个完整的翻译模型，用来预测不同条件下和不同细胞类型的蛋白组。

Thank you !