COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

Volume No: 1, Issue: 1, April 2012, e201204006, http://dx.doi.org/10.5936/csbj.201204006

CSBJ

GENETIC CODE REDUNDANCY AND ITS INFLUENCE ON THE ENCODED POLYPEPTIDES

Paige S. Spencer [a], José M. Barral [a,b,c*]

Abstract: The genetic code is said to be redundant in that the same amino acid residue can be encoded by multiple, so-called synonymous, codons. If all properties of synonymous codons were entirely equivalent, one would expect that they would be equally distributed along protein coding sequences. However, many studies over the last three decades have demonstrated that their distribution is not entirely random. It has been postulated that certain codons may be translated by the ribosome faster than others and thus their non-random distribution dictates how fast the ribosome moves along particular segments of the mRNA. The reasons behind such segmental variability in the rates of protein synthesis, and thus polypeptide emergence from the ribosome, have been explored by theoretical and experimental approaches. Predictions of the relative rates at which particular codons are translated and their impact on the nascent chain have not arrived at unequivocal conclusions. This is probably due, at least in part, to variation in the basis for classification of codons as "fast" or "slow", as well as variability in the number and types of genes and proteins analyzed. Recent methodological advances have allowed nucleotide-resolution studies of ribosome residency times in entire transcriptomes, which confirm the non-uniform movement of ribosomes along mRNAs and shed light on the actual determinants of rate control. Moreover, experiments have begun to emerge that systematically examine the influence of variations in ribosomal movement and the fate of the emerging polypeptide chain.

MINI REVIEW ARTICLE

Protein synthesis and the redundancy of the genetic code

The transfer of genetic information into protein products is termed translation (Figure 1; for detailed reviews on the mechanisms of translation, please see [1-3]). Messenger RNA (mRNA), transcribed from DNA, is translated into protein by a template driven process. The template is composed of a specific combination of 61 trinucleotide codons which encode 20 amino acids. This genetic code is common to most organisms and is referred to as redundant because all amino acids, with the exception of Tryptophan and Methionine, are encoded by more than one codon (termed synonymous codons). Codons are read by adaptor molecules called transfer RNA (tRNA) that bear matching (cognate) trinucleotide sequences, or anticodons. This reading or decoding of the codon occurs by recognition through base pairing, where at least two hydrogen bonds are formed between each of the nucleotide pairs that make up the codon:anticodon minihelix. Only one position of the codon:anticodon minihelix allows pairing that can deviate from standard Watson-Crick (G:C and A:U) interactions. In the third nucleotide of the codon and the first nucleotide of the anticodon, the so-called Wobble position, nonstandard base pairing can occur and results in altered base stacking conformations that are different from that of Watson-Crick pairing yet remain within the conformational constraints of the glycosidic bonds [4]. Interestingly, there are three conserved nucleotides in the bacterial 70S ribosome which maintain decoding fidelity by monitoring the conformation of the bases in the codon:anticodon minihelix [1]. The monitoring of base conformations is much more stringent in the first two nucleotide positions of the minihelix than in the wobble position, allowing for flexibility or wobble in the decoding of this position [1]. For example, nonstandard pairing of G:U and U:G, in which one less hydrogen bond is formed compared to standard G:C and C:G pairing, is allowed only in this position. Furthermore, post-transcriptional deamination of adenosine to inosine in the first anticodon position (INN) expands the decoding capacity from strictly Watson-Crick (A:U) to other allowed "wobble" base pairing (I:U, I:C, I:A) [4]. Adenosine deamination occurs in all eukaryotic ANN anticodons; however, in bacteria, this modification is exclusive to the ACG anticodon of $tRNA^{Arg}$ [5]. There are many other base modifications throughout the tRNA molecule, but these are more variable and will not be considered here. Upon decoding, peptide bond formation is catalyzed in the peptidyl-transferase center of the ribosome and is followed by translocation of the ribosome to the next codon. While diversity exists across evolution in the complexity of the ribosome [1, 6], translation regulation factors [1, 6], and tRNA gene composition [7], the core processes of translation are remarkably conserved and consist of three general steps: initiation, elongation, and termination.

Translation rates are not uniform along an mRNA and vary with the codon composition of the message, since the individual translation rates of codons have been shown to vary by as much as 25-fold [8-10]. The non-uniformity of rates has been proposed to depend on tRNA concentration, the nature of base pairing, and/or mRNA secondary structure [10-12]. The former two will be discussed later in this review. A logical assumption is that a stable mRNA secondary structure may hinder or slow translation by either preventing the ribosome from binding or by acting as a speed bump during

[a]*Department of Biochemistry & Molecular Biology*
[b]*Department of Neuroscience & Cell Biology*
[c]*Sealy Center for Structural Biology and Molecular Biophysics*
The University of Texas Medical Branch, 301 University Blvd., Galveston, TX 77555-0620.

* Corresponding author. Tel.: +1 4097472180; fax: +1 4097472200
E-mail address: jmbarral@utmb.edu (José M. Barral)

ribosomal progression. Indeed, the presence of stable mRNA secondary structures in the ribosomal binding site have been shown to largely affect expression levels as a result of interference with translation initiation [12]. However, the role of mRNA secondary structure in determining polypeptide elongation rates has been disputed [10, 13, 14]. Once the ribosome has initiated translation, it displays powerful helicase activity capable of disrupting very stable mRNA secondary structures (T_m = 70°C) [15]. This suggests that mRNA secondary structure plays an insignificant role in the rate of translation elongation, which is the main process addressed in this review. mRNA secondary structure likely plays a much more significant role in translation initiation and termination rates, which will not be discussed here. Additionally, most of the material presented in this review pertains to the bacterial ribosome.

Polypeptide elongation rate determinants

The process of polypeptide elongation occurs by the sequential addition to the growing polypeptide chain of a single amino acid brought to the ribosome by a molecular complex with three constituents: aminoacyl tRNA (aa-tRNA), elongation factor Tu (EF-Tu), and GTP (a so-called ternary complex) bearing the correct (cognate) anticodon for the mRNA codon in the ribosomal A site (Figure 1). There are three general steps to the elongation cycle: tRNA selection, peptidyl transfer, and translocation. tRNA selection, or decoding, consists of an initial binding of the ternary complex to the ribosome followed by codon recognition. Then, the GTPase activity of EF-Tu is activated, which subsequently causes GTP hydrolysis, EF-Tu dissociation, and accommodation [16]. Accommodation is the movement of the amino acid portion of the aa-tRNA in the A site closer to the peptidyl tRNA in the P site for peptidyl transfer to occur [1]. Following peptidyl transfer, binding of elongation factor G (EF-G) and GTP hydrolysis catalyze the translocation of the ribosome one codon forward, so that the tRNAs now reside in the E and P sites, respectively [1]. The elongation cycle continues as the codon in the newly vacant ribosomal A site awaits the next tRNA arrival. Interestingly, the ribosomal A site is likely seldom vacant and is instead sampled by cognate, near-cognate, and non-cognate tRNAs [17]. The terms, near-cognate and non-cognate, have conventionally been assigned to tRNAs which have single or multiple base mismatches with a given codon, respectively. However, Plant *et al* have challenged that a functional definition, namely the ability to form a minihelix with the codon in the ribosomal A site, better distinguishes a near- from a non-cognate [18]. It is important to note, that as peptidyl transfer and translocation occur much faster, tRNA selection appears to be the rate limiting step of ribosomal progression along the mRNA during polypeptide elongation [10, 19, 20]. Independently, two groups have observed large rate differences in the steps of polypeptide elongation by performing high resolution kinetic studies of the bacterial ribosome *in vitro*. They have determined that the rate of ternary complex GTPase activation in response to codon recognition is the rate limiting step of peptidyl transfer. They found that GTP hydrolysis of the cognate ternary complex occurs 650-fold [16] or approximately 116-fold [21] faster than the near-cognate one (base mismatch in 1st codon position in these studies). The other measurable rates were similar between cognate and near-cognate tRNAs, with the exception of a faster dissociation of the near-cognate during codon recognition [16].

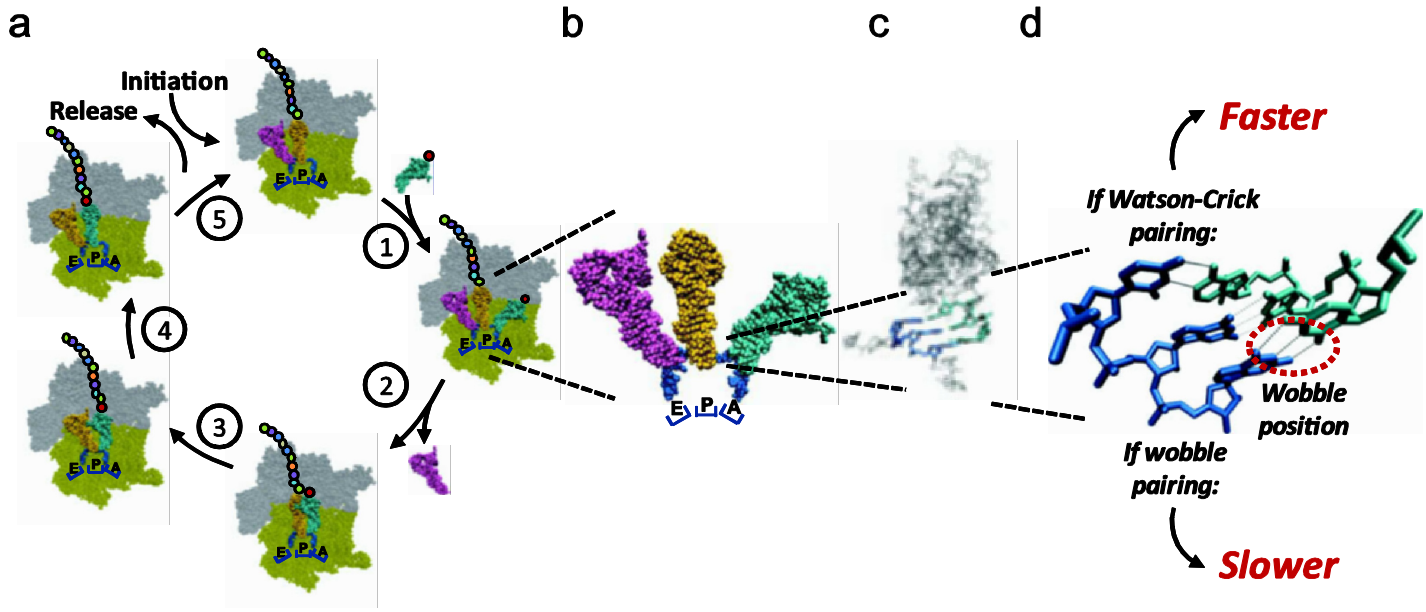

Figure 1. The nature of the codon:anticodon interaction influences translation elongation. (**a**) Summary of salient steps during bacterial translation elongation. After initiation, a ternary complex of tRNA (cyan) charged with an amino acid (red dot) and EF-Tu:GTP (not shown) binds to the A site of the 70S complex (gray/green) (1). GTP is then hydrolyzed, which results in incoming tRNA accommodation and release of EF-Tu and deacylated tRNA from the E site (2). The nascent polypeptide (chain of colored dots) is then transferred from the peptidyl tRNA in the P site to the incoming tRNA (3). EF-G binding and subsequent GTP hydrolysis (not shown) results in the critical translocation step, by which the now empty tRNA in the P site is transferred to the E site and the new peptidyl-tRNA is placed in the P site (4). EF-G release now renders the complex competent for a new round of elongation (5) or release and termination, if a stop codon is now encountered in the A site. (**b**) Space filling representation depicting an actual complex of mRNA and tRNAs in the E, P and A sites (PDB file 2Y18, from [76]. (**c**) Stick representation displaying the details of the codon (blue):anticodon (cyan) interaction in the A site shown in b (from [same as above]). (**d**) Enlarged view of actual UGG codon and $tRNA^{Trp}$ anticodon minihelix (PDB file 2Y18 [76]). Wobble position is circled to emphasize that elongation rates will be faster or slower depending on the type of interaction as indicated.