Part III - Evolutionary Studies Using Multiseq in VMD

- Aminoacyl tRNA Synthetases
- tRNA
- Aquaporins

Frankfurt, 2006, Computational Biology Workshop
Evolution of Biomolecular Structure
Class II tRNA-Synthetases and tRNA

MultiSeq Developers: Elijah Roberts, John Eargle, Dan Wright

Prof. Zan Luthey-Schulten, Patrick O'Donoghue, Anurag Sethi, Brijeeet Dhaliwal

March 2006.
Canonical Pattern & Horizontal Gene Transfer

• “The aminoacyl-tRNA synthetases, perhaps better than any other molecules in the cell, epitomize the current situation and help to understand the effects of HGT” Woese (PNAS, 2000; MMBR 2000)

• Carl Woese - Crafoord Prize 2003
Step 1: Explore active site in catalytic domain and anticodon domain in AspRS from Ecoli
Charging the tRNA

**Direct acylation**

\[
\text{Gln} + \text{tRNA}^{\text{Gln}} + \text{ATP} \xleftrightarrow{\text{synthetase}} \text{Gln-tRNA}^{\text{Gln}} + \text{AMP} + \text{PP}_i
\]

**tRNA-dependent amino acid modification**

\[
\text{Glu} + \text{tRNA}^{\text{Gln}} + \text{ATP} \xleftrightarrow{\text{synthetase}} \text{Glu-tRNA}^{\text{Gln}} + \text{AMP} + \text{PP}_i
\]

\[
\text{Gln} + \text{Glu-tRNA}^{\text{Gln}} + \text{ATP} \xleftrightarrow{\text{amidotransferase}} \text{Gln-tRNA}^{\text{Gln}} + \text{ADP} + \text{P}_i + \text{Glu}
\]

Indirect Pathway - AspRS is non-discriminating. Some organisms do not contain genes to make Asn and use Gln as source of ammonia. In that case ID pathway is the only way to obtain Asn and a direct AsnRS not found in the organism. Similar SepRS/SepCysS
Step 1: Creation of the Aminoacyl-Adenylate Complex

Amino Acid + ATP → Aminoacyl-AMP + PP\textsubscript{i}

In step 1, an O atom of the amino acid a-carboxyl attacks the P atom of the alpha phosphate of ATP. The products are Aminoacyl-AMP containing a mixed carboxy-phosphoanhydride bond and pyrophosphate.
Step 2: Creation of the Aminoacyl-tRNA

In step 2, the 2' or 3' OH of the terminal adenosine of the 3' end of the tRNA attacks the amino acid carbonyl C atom, creating a ribose ester.
Aminoacyl-tRNA Synthetase

Summary of the 2-step reaction:

1. amino acid + ATP $\rightarrow$ aminoacyl-AMP + PP$_i$

2. aminoacyl-AMP + tRNA $\rightarrow$ aminoacyl-tRNA + AMP

Overall Reaction:

amino acid + ATP + tRNA $\rightarrow$ aminoacyl-tRNA + AMP + PP$_i$

Next step: EF and Ribosome for Protein Synthesis
Evolution of Protein/RNA Interfaces:
EF-Tu/tRNA Recognition

J. Eargle, A. Sethi, L. Trabuco, A. Black, E. Roberts - in progress
Structure of the *E. coli* Ribosome

The cutaway view at right shows positions of tRNA (P, E sites) & mRNA (as orange beads).

Figure: Laboratory of Joachim Frank, Wadsworth Center cryo-EM and 3D image reconstruction
The cutaway view at right shows that the tunnel in the yeast large ribosome subunit, through which nascent polypeptides emerge from the ribosome, lines up with the lumen of the ER Sec61 channel.
VMD Movie of Ribosomal Assembly
Horizontal Gene Transfer in Protein Structure

Sequence Phylogeny
AspRS-AsnRS Group
tRNA Structure

(A) Attached amino acid (Phe)

(B) A clover leaf

(C) A helical structure

(D) 5' GCGG AUUUA GCUC A GDDGGGA GA GC GC CCAGAC UGA AYAY CUG GAG GGUC UGUG T YCG AU CCAC AGAAUU UC GC AC CA 3'

©1990 GARLAND PUBLISHING
Most RNAs have **secondary cloverleaf structure**, consisting of stem & loop domains.

Double helical **stems** arise from **base pairing** between complementary stretches of bases within the same strand.

**Loops** occur where lack of complementarity, or the presence of **modified bases**, prevents base pairing.
Hydrogen bonds link 2 complementary nucleotide bases on separate nucleic acid strands, or on complementary portions of the same strand.

Conventional base pairs: A & U (or T); C & G.
Secondary and Tertiary Interactions for tRNA$^{\text{Cys}}$ from *E. coli*

**Role of Modified Bases?**

Modified Bases in tRNA\textsuperscript{Cys} from \textit{E. coli}, \textit{S. cerevisiae}, and \textit{H. volcanii}

\begin{itemize}
  \item $\text{ms}^{2i6}\text{A}$: 2-methylthio-$N^6$-isopentenyladenosine
  \item $m^1I$: 1-methylinosine
  \item $m^5C$: 5-methylcytidine
  \item $Cm$: 2'-O-methylcytidine
  \item $m^1G$: 1-methylguanosine
  \item $G^+$: archaeosine
  \item $\Psi$ or $Y$: pseudouridine
  \item $m^1\Psi$: 1-methylpseudouridine
  \item $s^4U$: 4-thiouridine
\end{itemize}

Multiseq extension in VMD
The **genetic code** is based on the sequence of bases along a nucleic acid.

Each **codon**, a sequence of 3 **bases** in mRNA, codes for a particular amino acid, or for chain termination.

Some amino acids are specified by 2 or more codons.

**Synonyms** (multiple codons for the same amino acid) in most cases differ only in the 3rd base. Similar codons tend to code for similar amino acids. Thus effects of mutation are minimized.
tRNA Databases and Web Resources

MFOLD : Prediction of RNA secondary structure (M. Zuker)
http://bioweb.pasteur.fr/seqanal/interfaces/mfold-simple.html

Vienna RNA Package (Ivo Hofacker)
http://www.tbi.univie.ac.at/~ivo/RNA/

DOE Joint Genome Institute
http://www.jgi.doe.gov/

Compilation of tRNA sequences and sequences of tRNA genes (Mathias Sprinzl)
http://www.uni-bayreuth.de/departments/biochemie/trna/