Part III - Evolutionary Studies Using Multiseq in VMD

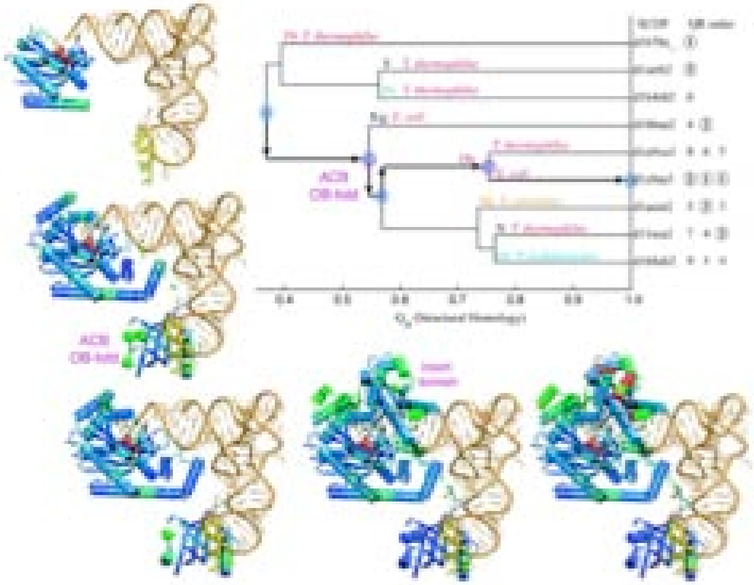
- Aminoacyl tRNA Synthetases
- tRNA
- Aquaporins

Frankfurt, 2006, Computational Biology Workshop

University of Illinois at Urbana-Champaign
Luthey-Schulten Group
Theoretical and Computational Biophysics Group

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
Brijeet Dhaliwal

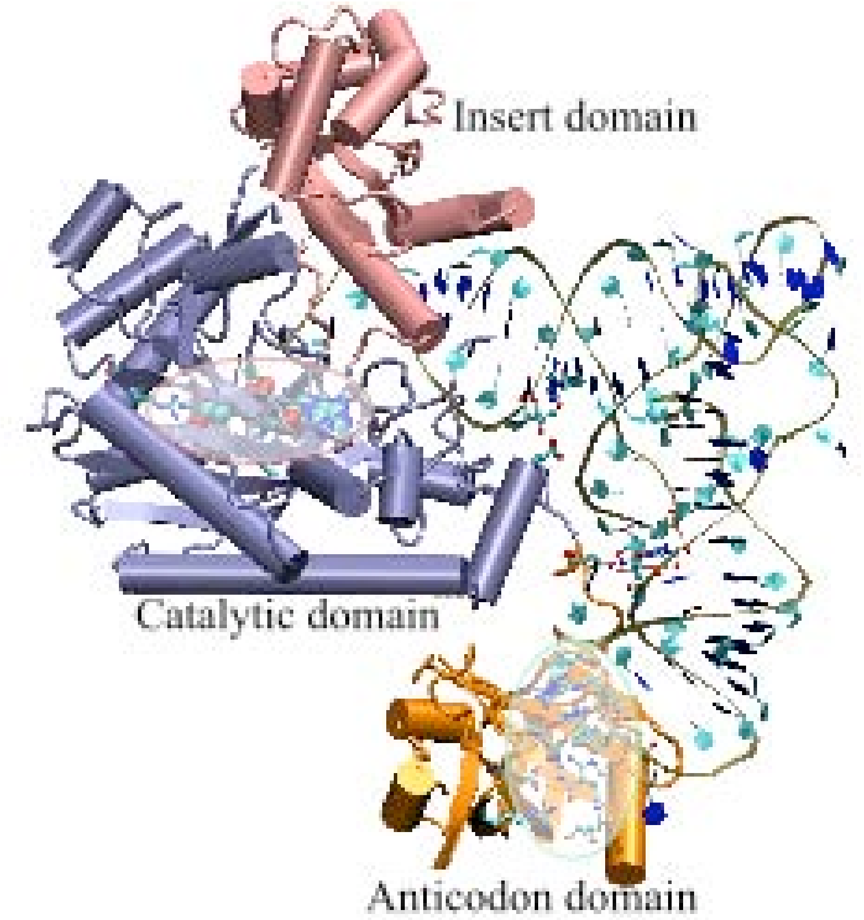
March 2006.

Canonical Pattern & Horizontal Gene Transfer

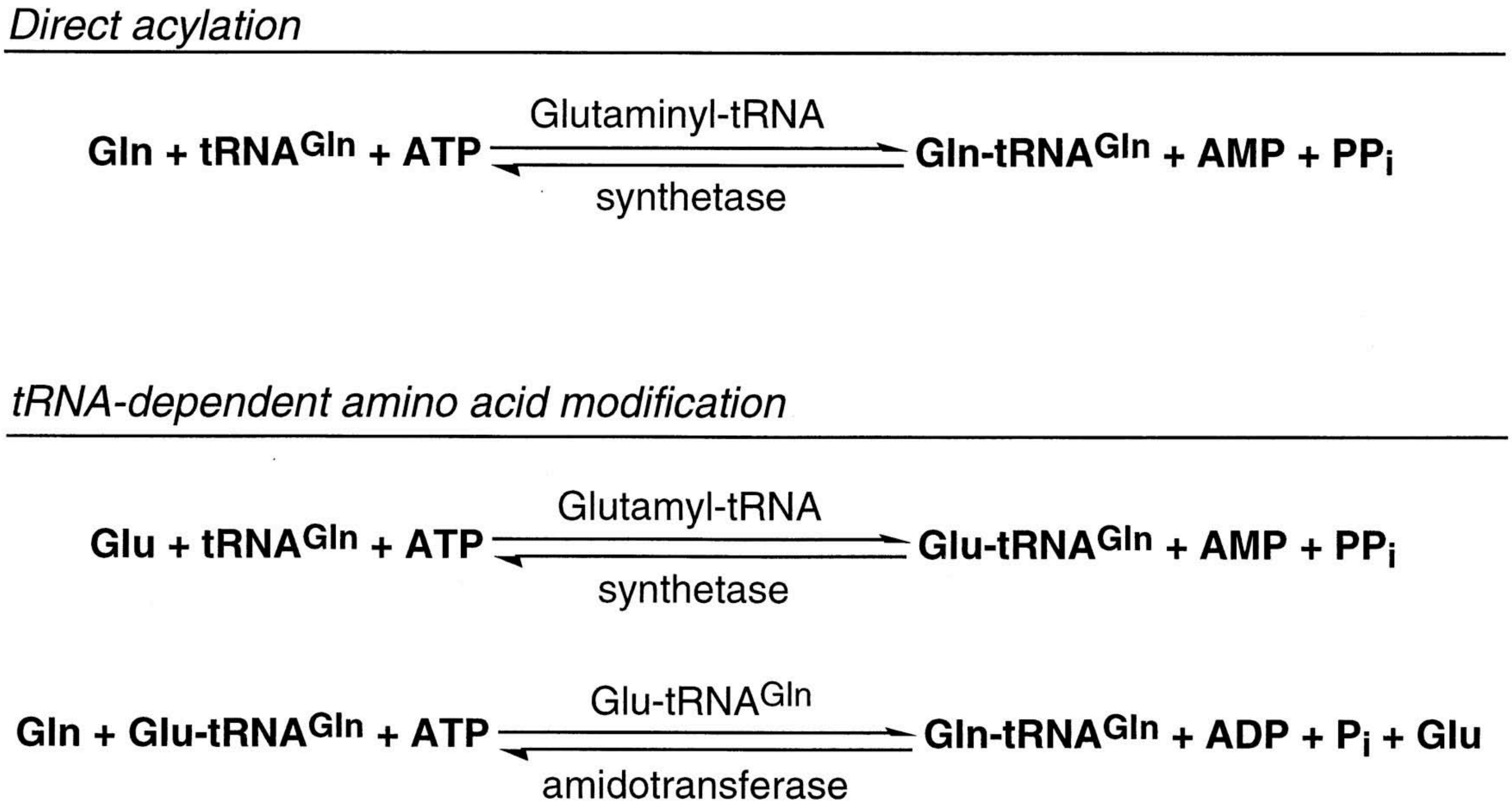
- “The aminoacyl-tRNA synthetases, perhaps better than any other molecules in the cell, eptiomize the current situation and help to understand the effects of HGT” Woese (PNAS, 2000; MMBR 2000)

- 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

$$\mathrm{Gln} + \mathrm{tRNA}^{\mathrm{Gln}} + \mathrm{ATP} \xrightleftharpoons[\text{synthetase}]{\text{Glutaminyl-tRNA}} \mathrm{Gln\text{-}tRNA}^{\mathrm{Gln}} + \mathrm{AMP} + \mathrm{PP_i}$$

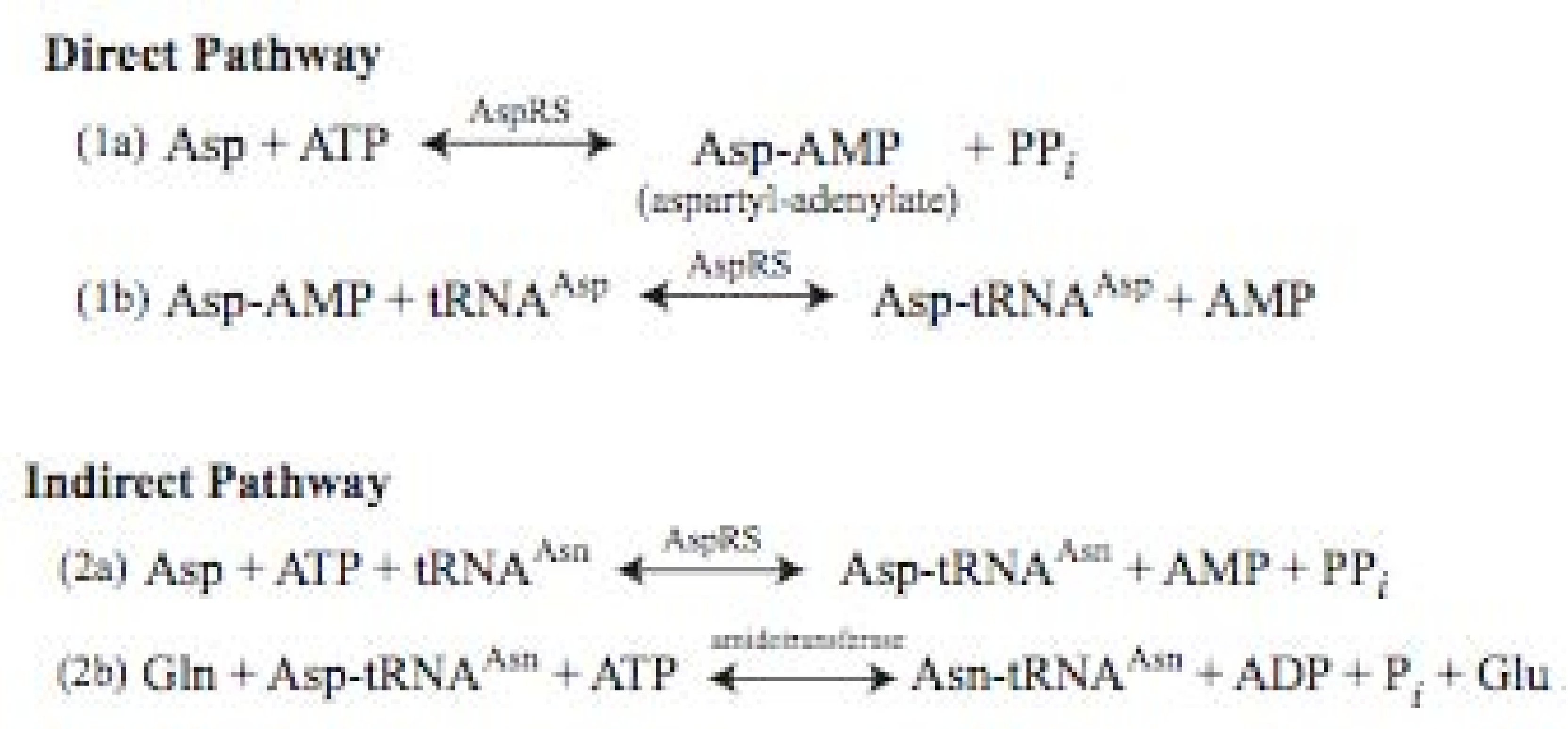
tRNA-dependent amino acid modification

$$\mathrm{Glu} + \mathrm{tRNA}^{\mathrm{Gln}} + \mathrm{ATP} \xrightleftharpoons[\text{synthetase}]{\text{Glutamyl-tRNA}} \mathrm{Glu\text{-}tRNA}^{\mathrm{Gln}} + \mathrm{AMP} + \mathrm{PP_i}$$

$$\mathrm{Gln} + \mathrm{Glu\text{-}tRNA}^{\mathrm{Gln}} + \mathrm{ATP} \xrightleftharpoons[\text{amidotransferase}]{\text{Glu-tRNA}^{\text{Gln}}} \mathrm{Gln\text{-}tRNA}^{\mathrm{Gln}} + \mathrm{ADP} + \mathrm{P_i} + \mathrm{Glu}$$

Woese, Olsen (UIUC), Ibba (Panum Inst.), Soll (Yale) *Micro. Mol. Biol. Rev.* March 2000.

Amino Acid Biosynthesis and tRNA Charging



Direct Pathway

(1a) $\mathrm{Asp} + \mathrm{ATP} \xleftrightarrow{\mathrm{AspRS}} \underset{\text{(aspartyl-adenylate)}}{\mathrm{Asp\text{-}AMP}} + \mathrm{PP}_i$

(1b) $\mathrm{Asp\text{-}AMP} + \mathrm{tRNA}^{\mathrm{Asp}} \xleftrightarrow{\mathrm{AspRS}} \mathrm{Asp\text{-}tRNA}^{\mathrm{Asp}} + \mathrm{AMP}$

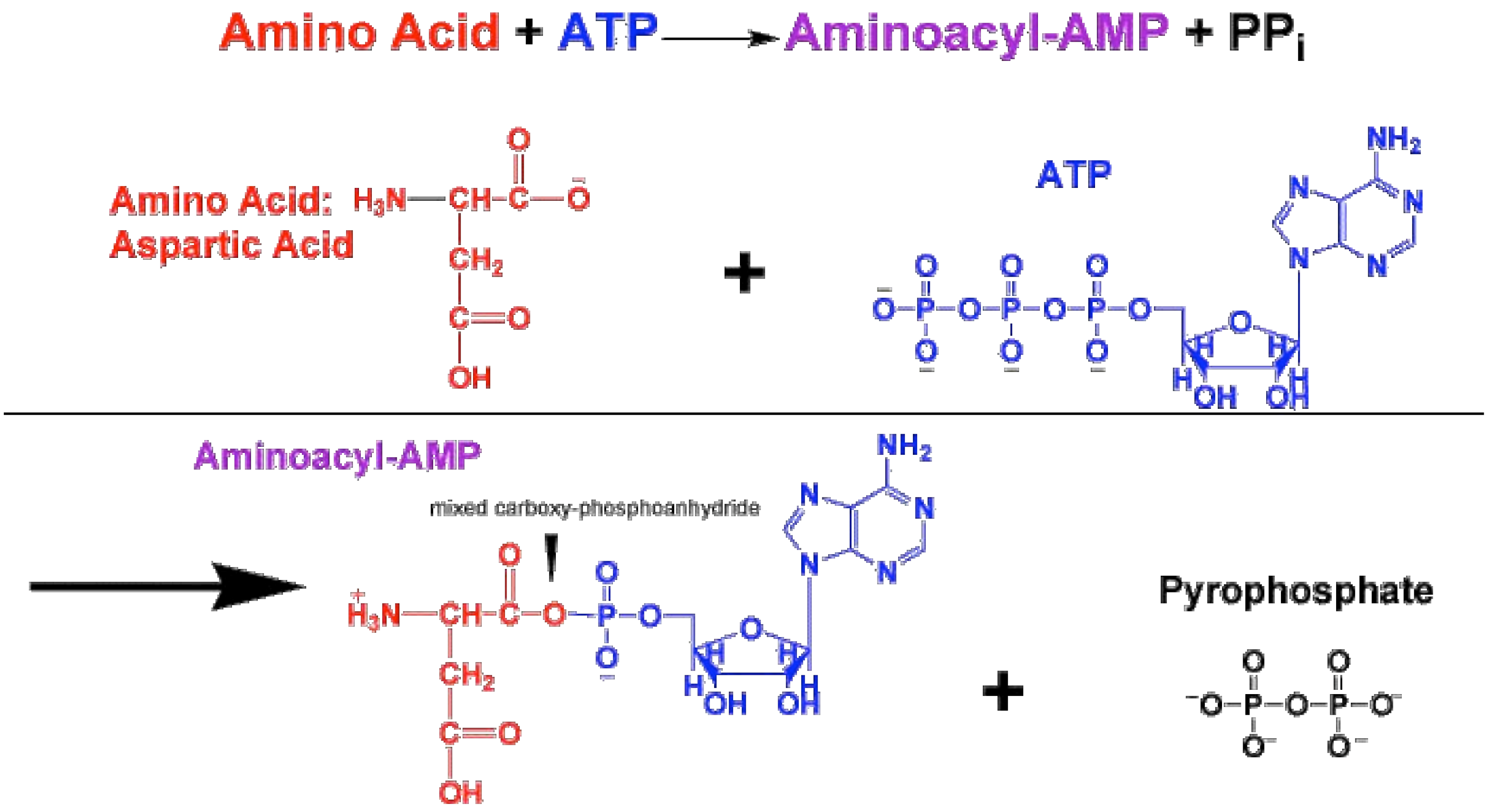
Indirect Pathway

(2a) $\mathrm{Asp} + \mathrm{ATP} + \mathrm{tRNA}^{\mathrm{Asn}} \xleftrightarrow{\mathrm{AspRS}} \mathrm{Asp\text{-}tRNA}^{\mathrm{Asn}} + \mathrm{AMP} + \mathrm{PP}_i$

(2b) $\mathrm{Gln} + \mathrm{Asp\text{-}tRNA}^{\mathrm{Asn}} + \mathrm{ATP} \xleftrightarrow{\text{amidotransferase}} \mathrm{Asn\text{-}tRNA}^{\mathrm{Asn}} + \mathrm{ADP} + \mathrm{P}_i + \mathrm{Glu}$

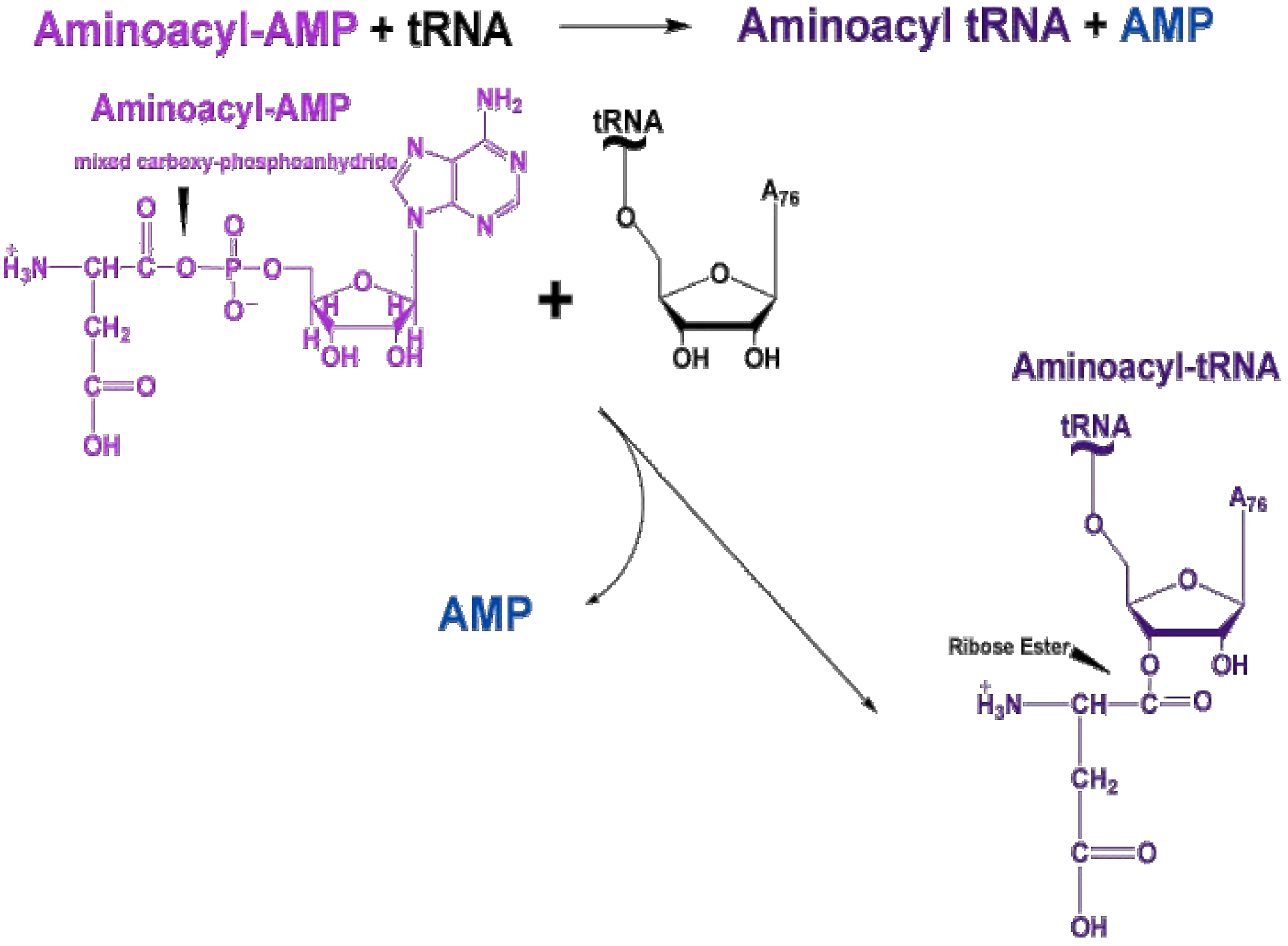
Indirect Pathway - AspRS is non-discriminating. Some organisms do not contain genes to make Asn and use Gln as source of ammonia. In that case ID pathway is the only way to obtain Asn and a direct AsnRS not found in the organism. Similar SepRS/SepCysS

Step 1: Creation of the Aminoacyl-Adenylate Complex



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 → aminoacyl-AMP + PP$_i$**

2. **aminoacyl-AMP + tRNA → aminoacyl-tRNA + AMP**

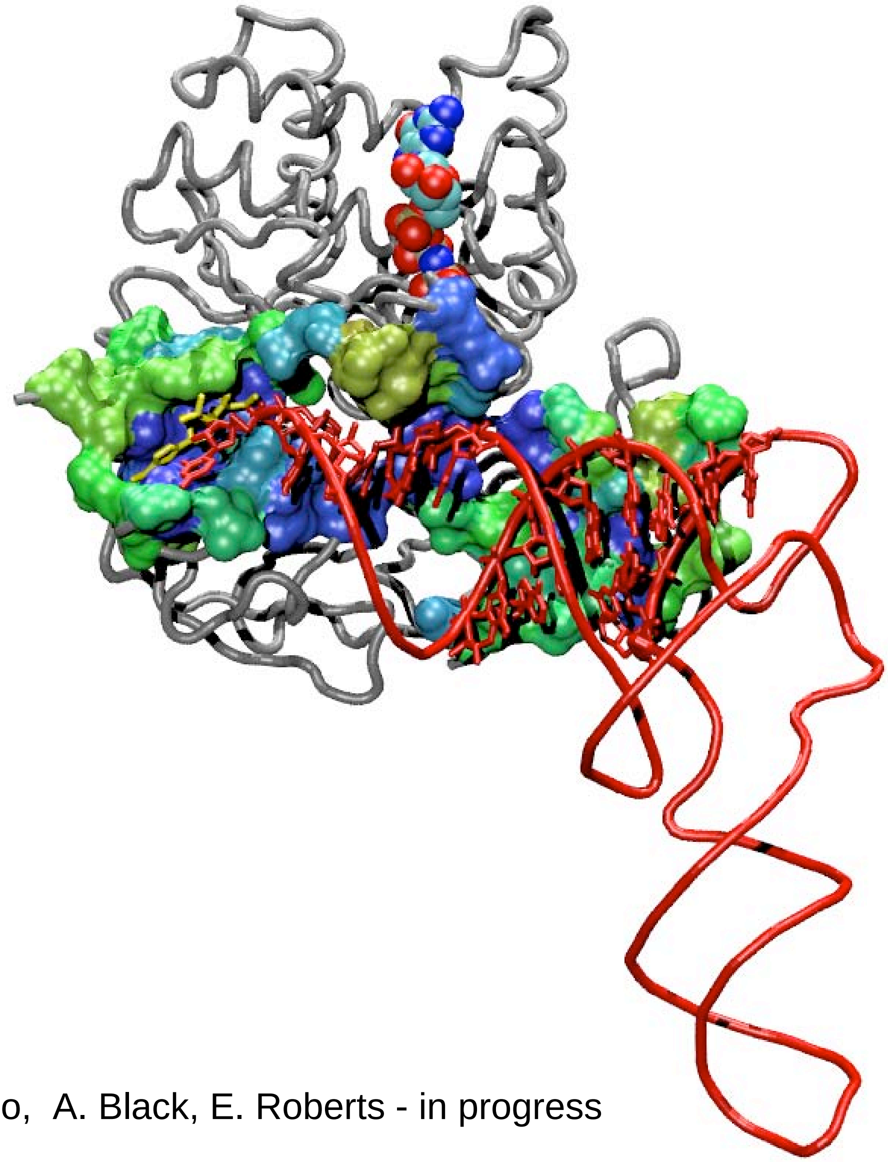
Overall Reaction:

amino acid + ATP + tRNA → aminoacyl-tRNA + AMP

+ PP$_i$

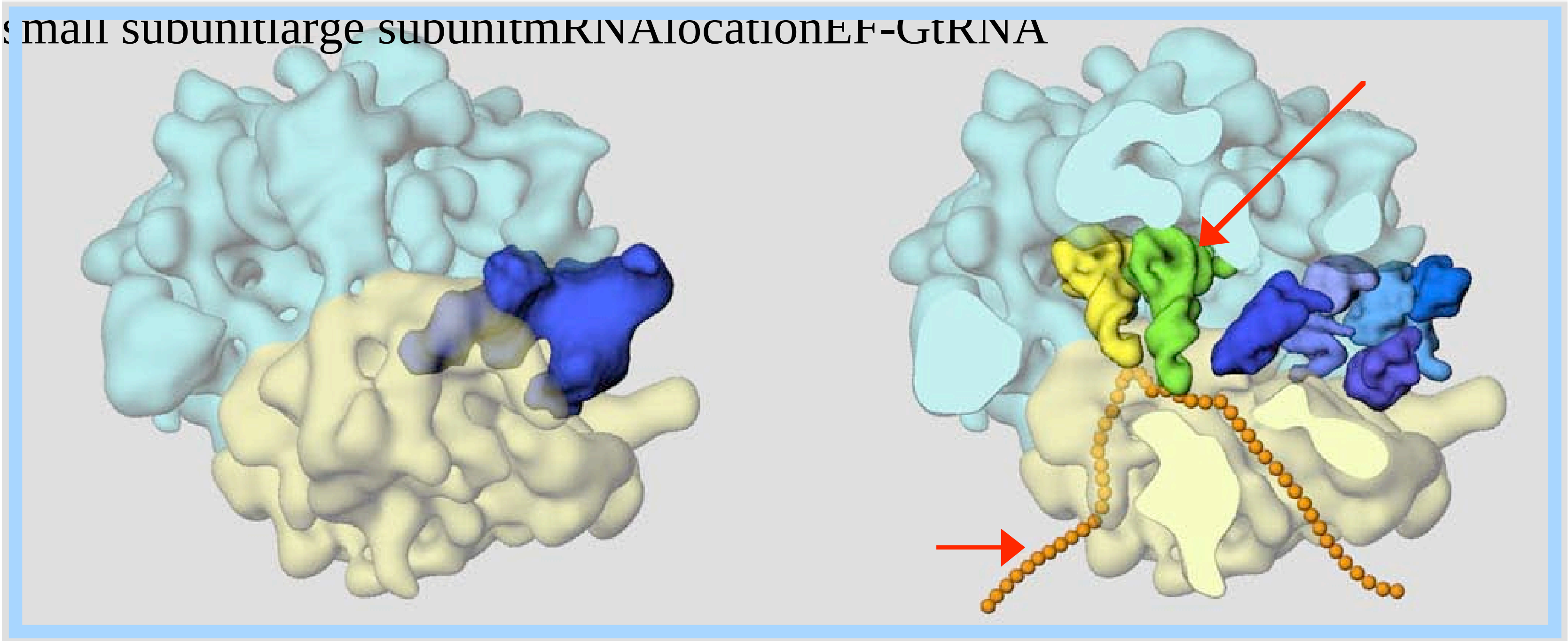
Next step: EF and Ribosome for Protein Synthesis

Evolution of Protein/RNA Interfaces: EF-Tu/tRNA Recognition



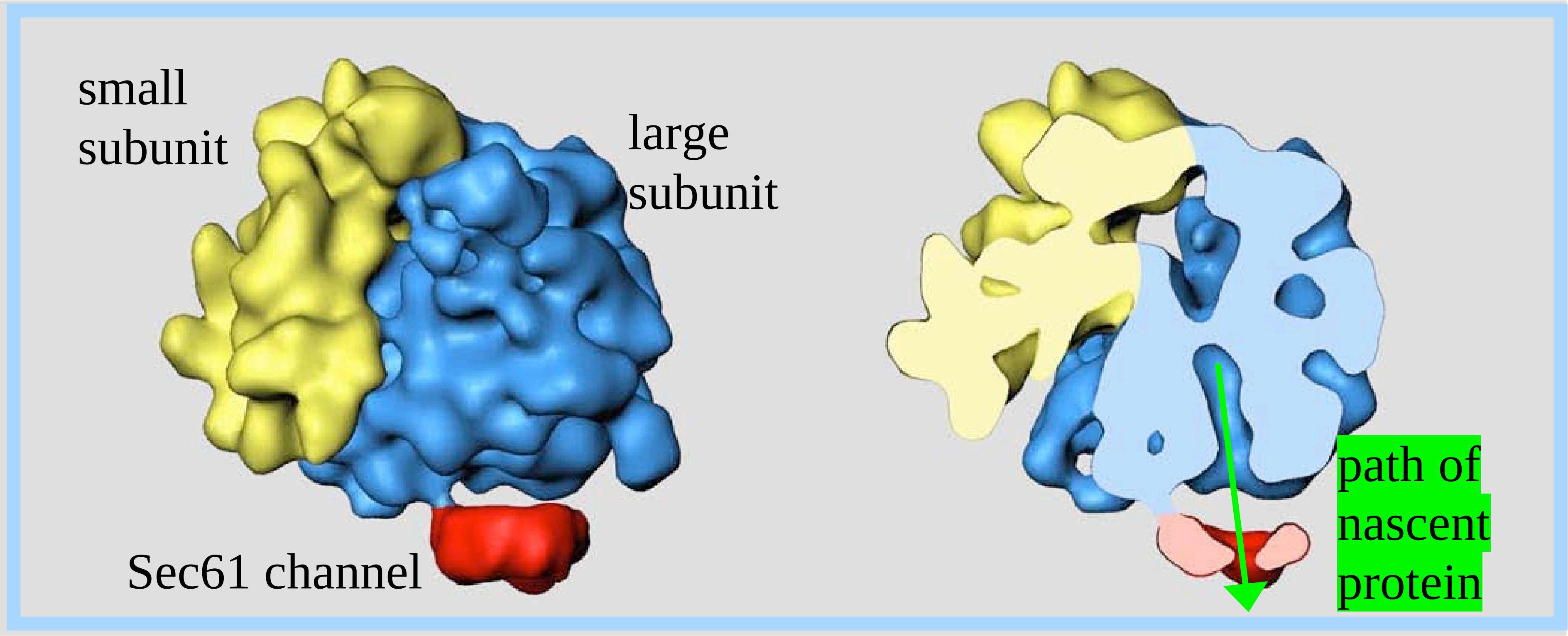
J. Eargle, A. Sethi, L. Trabuco, A. Black, E. Roberts - in progress

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



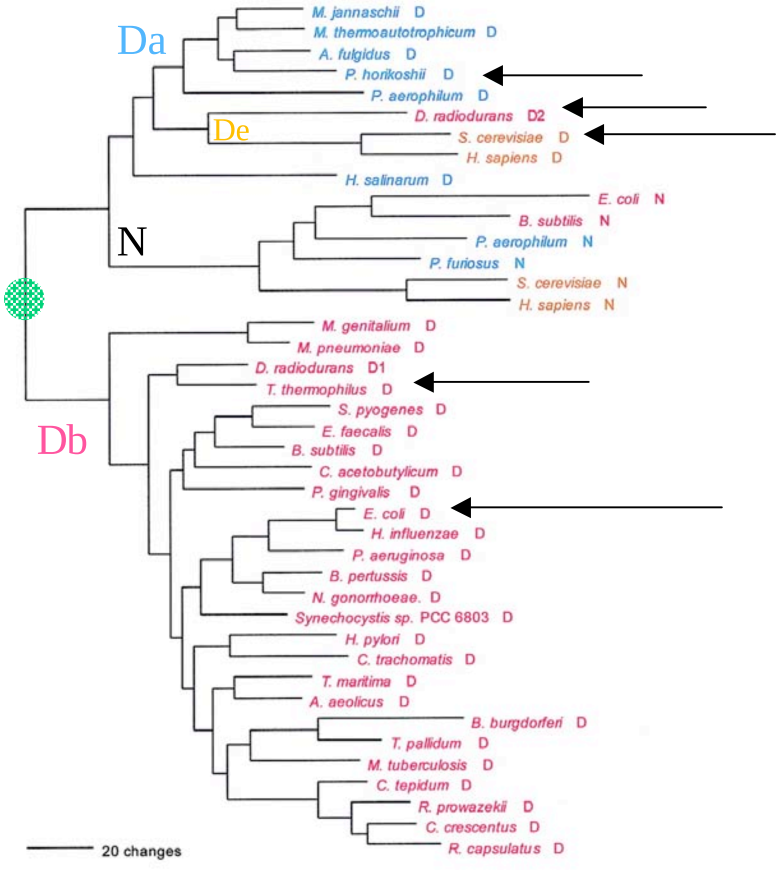
The cutaway view at right shows that the **tunnel** in the yeast large ribosome subunit, through which nascent polypeptides emerge from the ribosome, **lines up** with the lumen of the ER **Sec61 channel**.

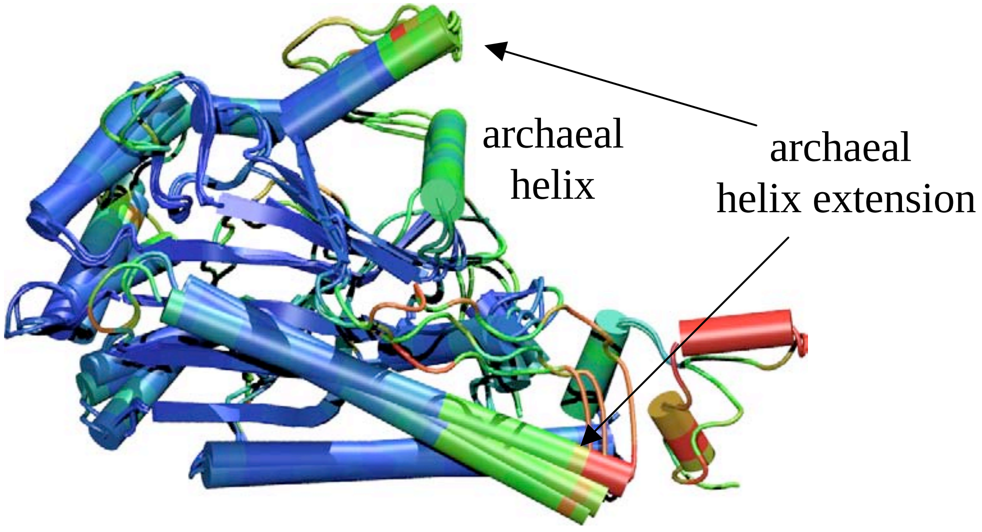
VMD Movie of Ribosomal Assembly

Horizontal Gene Transfer in Protein Structure

Sequence Phylogeny

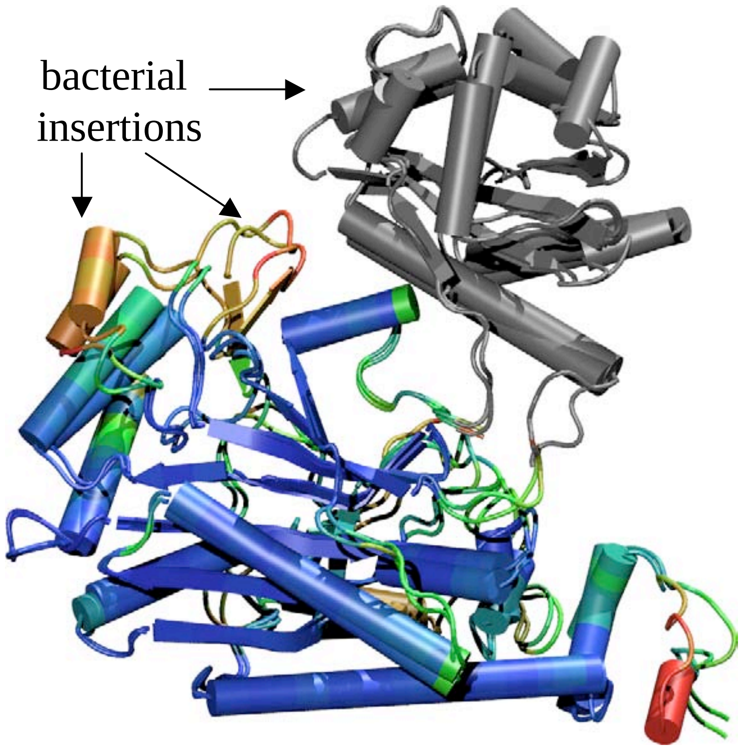
AspRS-AsnRS Group





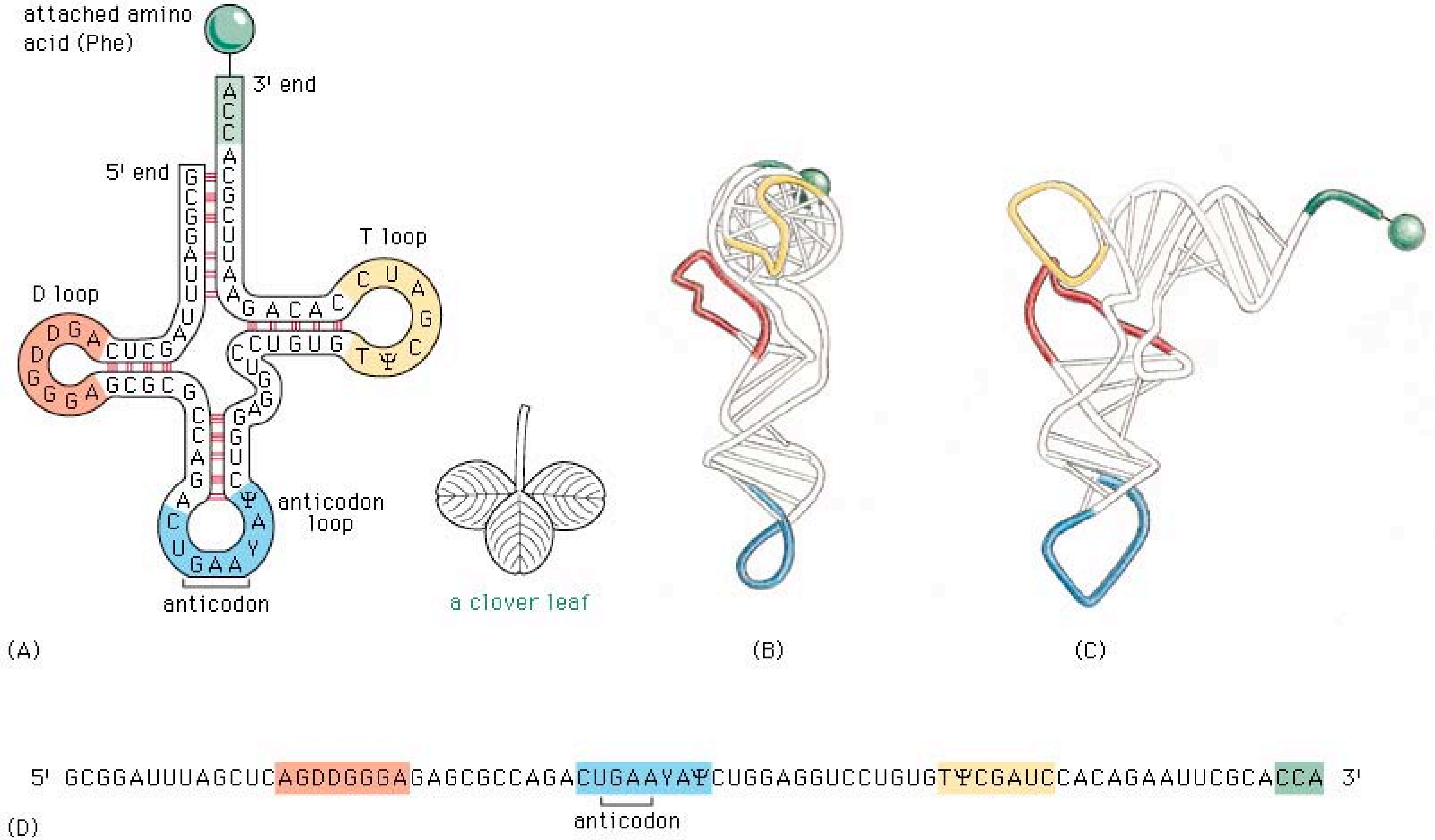
archaeal helix

archaeal helix extension



bacterial insertions

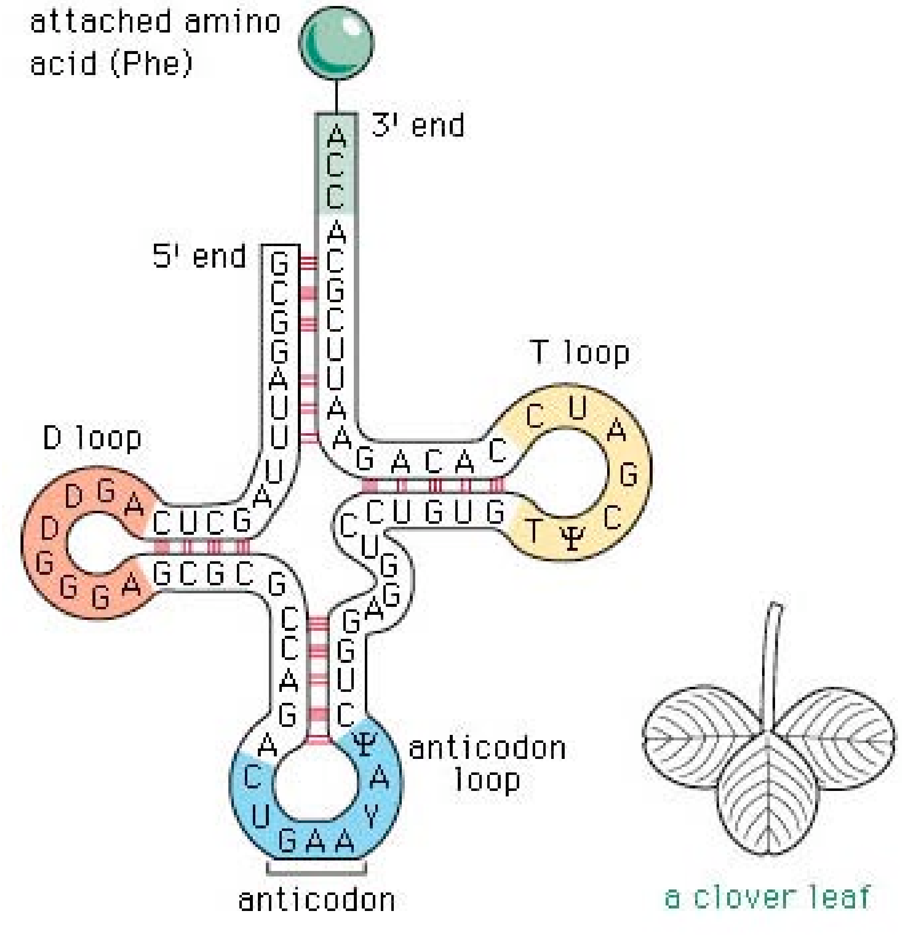
tRNA Structure



tRNA Secondary Structure

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.

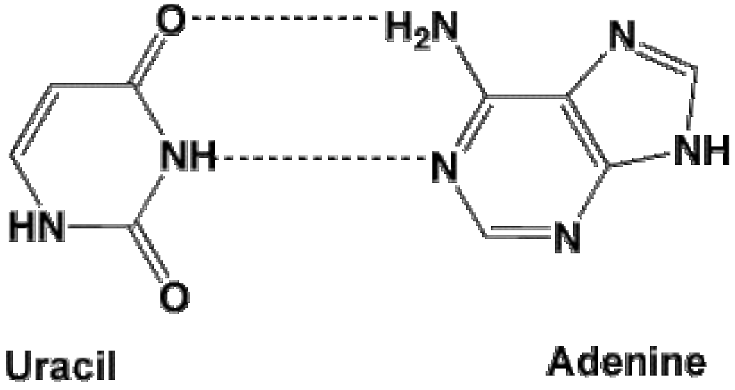


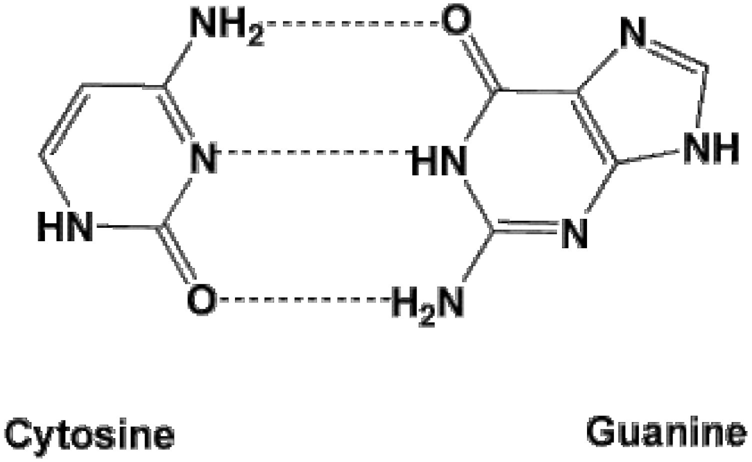
(A)

Loops occur where **lack of complementarity**, or the presence of **modified bases**, prevents base pairing.

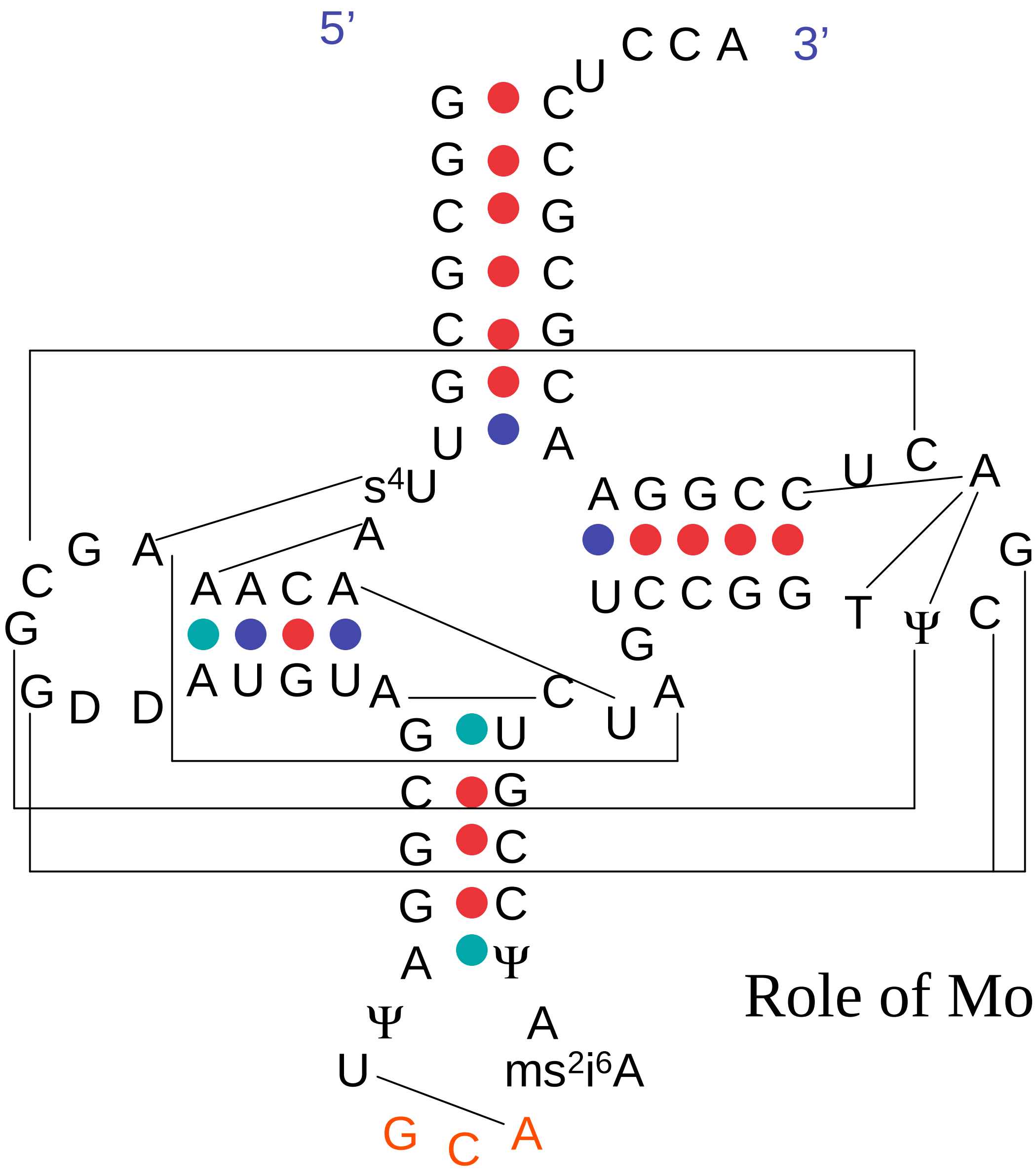
Hydrogen bonds link 2 complementary nucleotide bases on separate nucleic acid strands, or on complementary portions of the same strand.

Conventional **base pairs**: **A & U (or T)**; **C & G**.





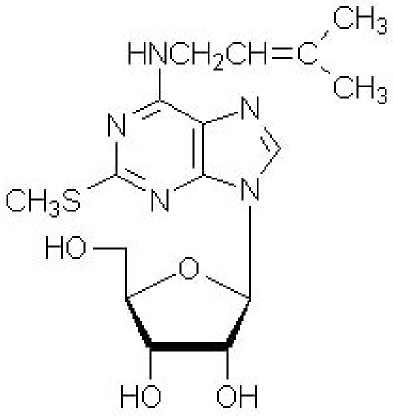
Secondary and Tertiary Interactions for tRNA^{Cys} from *E. coli*



Role of Modified Bases?

Hauenstein, S., Zhang, C.M., Hou, Y.M., Perona, J.J. *Nat. Struct. Mol. Biol.* 11:1134-1141 (2004)

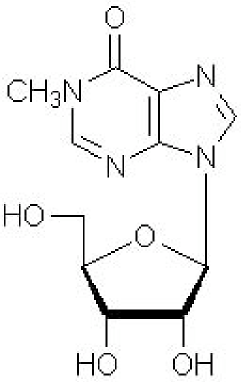
Modified Bases in tRNA^{Cys} from *E. coli*, *S. cerevisiae*, *and H. volcanii*



ms^{2}i^{6}A

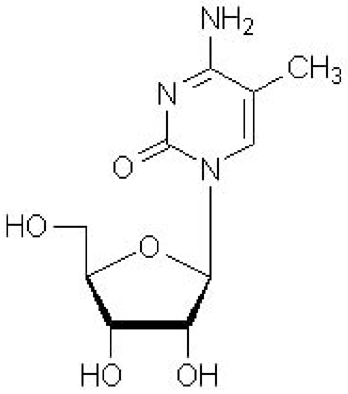
2-methylthio-*N*^{6}

-isopentenyladenosine



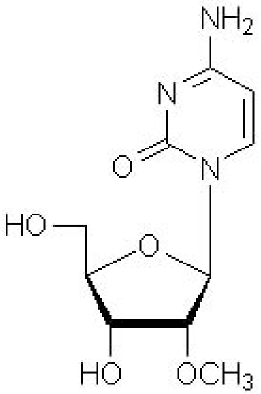
m^{1}I

1-methylinosine



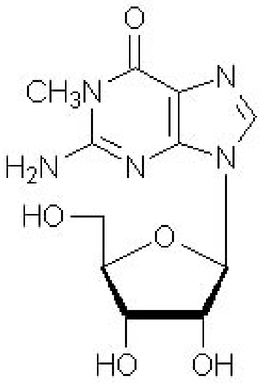
m^{5}C

5-methylcytidine



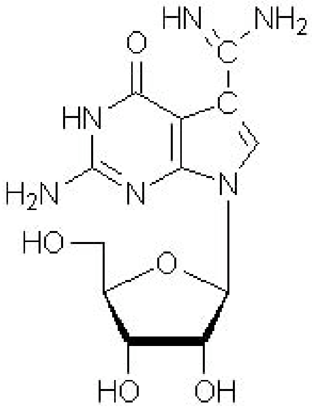
Cm

2'-*O*-methylcytidine



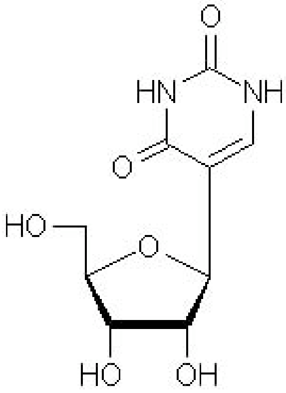
m^{1}G

1-methylguanosine



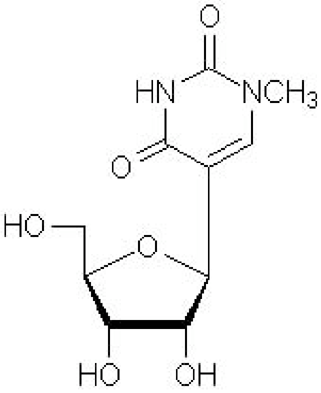
G$^+$

archaeosine



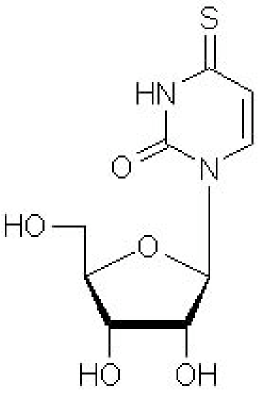
Ψ or Y

pseudouridine



m$^1\Psi$

1-methylpseudouridine



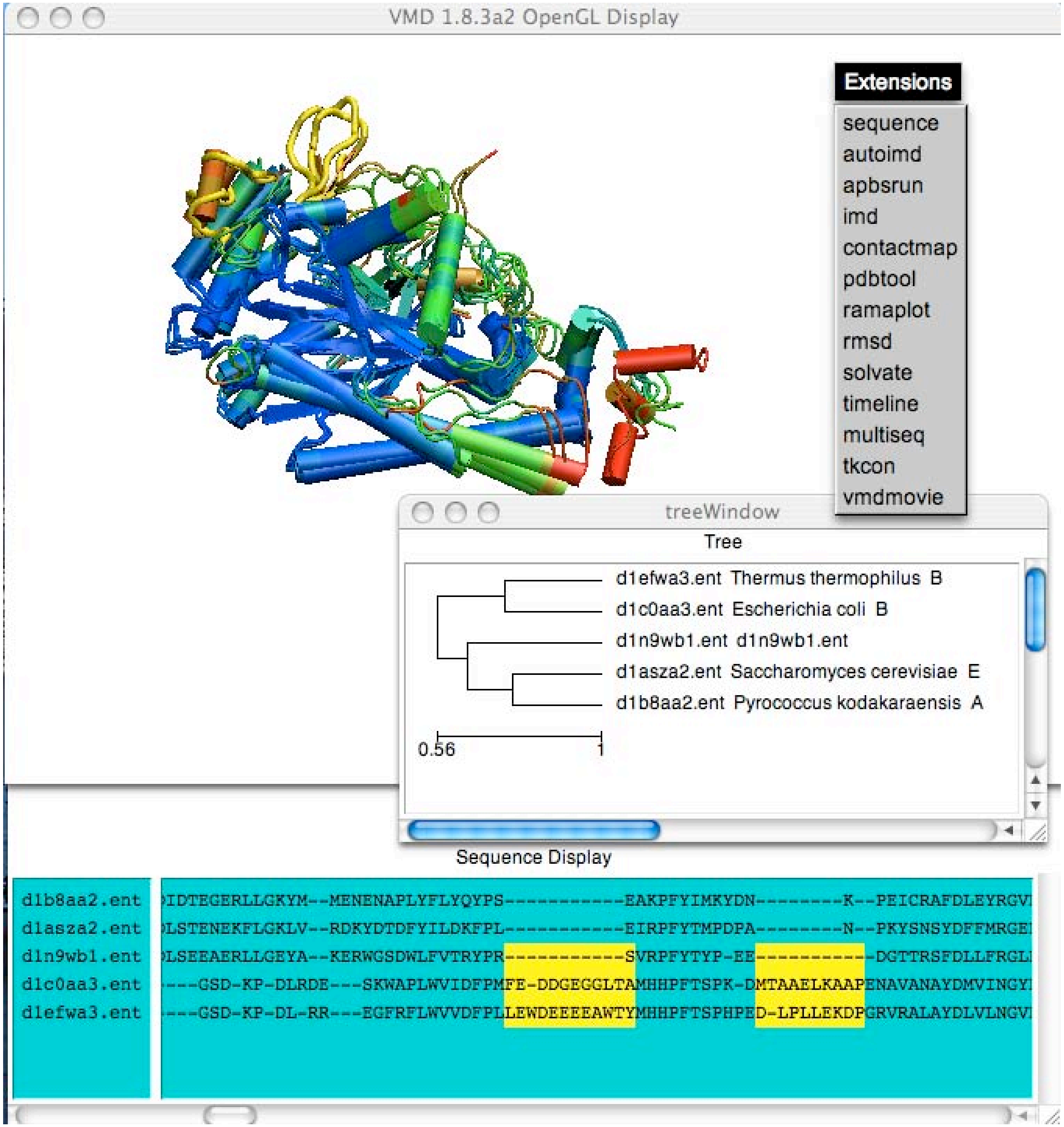
s^{4}U

4-thiouridine

Limbach, P.A., Crain, P.F., and McCloskey, J.A. *Nucleic Acids Res.* 22:2183-2196 (1994)

Multiseq

extension in VMD



d1efwa3.ent Thermus thermophilus B
d1c0aa3.ent Escherichia coli B
d1n9wb1.ent d1n9wb1.ent
d1asza2.ent Saccharomyces cerevisiae E
d1b8aa2.ent Pyrococcus kodakaraensis A

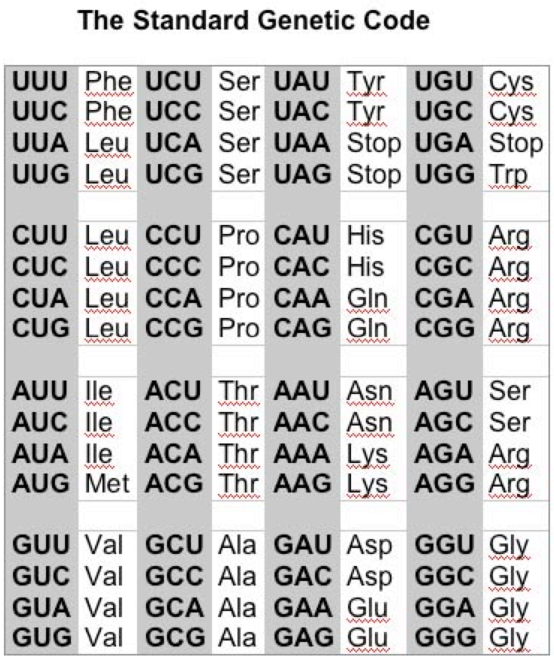
```
d1b8aa2.ent  IDTEGERLLGKYM--MENENAPLYFLYQYPS-----------EAKPFYIMKYDN-------K--PEICRAFDLEYRGV
d1asza2.ent  LSTENEKFLGKLV--RDKYDTDFYILDKFPL-----------EIRPFYTMPDPA-------N--PKYSNSYDFFMRGE
d1n9wb1.ent  LSEEAERLLGEYA--KERWGSDWLFVTRYPR-----------SVRPFYTYP-EE----------DGTTRSFDLLFRGL
d1c0aa3.ent  ---GSD-KP-DLRDE---SKWAPLWVIDFPMFE-DDGEGGLTAMHHPFTSPK-DMTAAELKAAPENAVANAYDMVINGY
d1efwa3.ent  ---GSD-KP-DL-RR---EGFRFLWVVDFPLLEWDEEEEAWTYMHHPFTSPHPED-LPLLEKDPGRVRALAYDLVLNGV
```

Genetic code

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.



The Standard Genetic Code

UUU	Phe	**UCU**	Ser	**UAU**	Tyr	**UGU**	Cys
UUC	Phe	**UCC**	Ser	**UAC**	Tyr	**UGC**	Cys
UUA	Leu	**UCA**	Ser	**UAA**	Stop	**UGA**	Stop
UUG	Leu	**UCG**	Ser	**UAG**	Stop	**UGG**	Trp
CUU	Leu	**CCU**	Pro	**CAU**	His	**CGU**	Arg
CUC	Leu	**CCC**	Pro	**CAC**	His	**CGC**	Arg
CUA	Leu	**CCA**	Pro	**CAA**	Gln	**CGA**	Arg
CUG	Leu	**CCG**	Pro	**CAG**	Gln	**CGG**	Arg
AUU	Ile	**ACU**	Thr	**AAU**	Asn	**AGU**	Ser
AUC	Ile	**ACC**	Thr	**AAC**	Asn	**AGC**	Ser
AUA	Ile	**ACA**	Thr	**AAA**	Lys	**AGA**	Arg
AUG	Met	**ACG**	Thr	**AAG**	Lys	**AGG**	Arg
GUU	Val	**GCU**	Ala	**GAU**	Asp	**GGU**	Gly
GUC	Val	**GCC**	Ala	**GAC**	Asp	**GGC**	Gly
GUA	Val	**GCA**	Ala	**GAA**	Glu	**GGA**	Gly
GUG	Val	**GCG**	Ala	**GAG**	Glu	**GGG**	Gly

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/