

## **Sense/Antisense Open Reading Frames and the Amino Acid Composition of Ribosomal Proteins**

**Robert Huether**<sup>1</sup>, William L. Duax<sup>1,3</sup>, Lukas Habegger<sup>1</sup>, Vladimir Pletnev<sup>2</sup>, Sonjay Connare<sup>1</sup>, <sup>1</sup>*Structural Biology Dept. Hauptman-Woodward Medical Inst., 73 High St., Buffalo, NY 14203-1186.* <sup>2</sup>*Inst. Bioorg. Chem., RAS Moscow.* <sup>3</sup>*SUNY, Dept. of Structural Biology, Buffalo, NY 14260.* E-mail: [bhuether@hwi.buffalo.edu](mailto:bhuether@hwi.buffalo.edu)

By combining information on three dimensional structures and conservation of amino acid sequences in families of ancient proteins it is possible to trace the evolution of the genetic code and the amino acid composition of proteins. We discovered a pattern of multiple open reading frames (ORFs) and amino acid bias in *Streptomyces coelicolor*. Of particular interest was the high incidence of sense/antisense ORFs (SASORFs) and the absence of Trp and Cys residues in ribosomal proteins. The L1 proteins from the 50S ribosomal subunit were selected for further analysis. Of 125 sequences of L1 proteins in the SWISS-PROT TrEMBL database, 50% are missing Trp, 64% are missing Cys and 35% are missing both. Those from archaea rarely have Trp and/or Cys residues whereas those from eukaryote usually have both. Comparison of the amino acid sequences of the 125 50SL1 proteins reveals that Cys is not conserved in any sequence position at greater than 8%. In 37 structures a Trp residue resides in a common position on the surface of L1 in bacteria but not archaea. These finding suggest that SASORFs, severe codon bias and absence of Trp and Cys residues are hallmarks of ancient enzymes that have been little altered by millions of years of evolution or lateral genes transfer.

Research supported by NIH Grant No. DK26546

**Keywords:** open reading frames, codon bias, ribosome