

Problem #4. The four polypeptides in the example consist of 24, 10, 3 and 25 amino acids, and the mRNA sequence contains $195 = ((24 + 10 + 3 + 25) + 3) \times 3$ nucleotides. It appears probable that three nucleotides (a triplet) denote one amino acid or are a separator between polypeptides (in reality a signal to terminate synthesis). However, since there are $4^3 = 64$ possible triplets (all but two of which are present in the example) and only 20 different amino acids, some triplets have the same meaning.

	... U C A G ...
U ...	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
C ...	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
A ...	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 → ?	AAG → Lys	AGG → Arg
G ...	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 → ?

All mRNA sequences start with AUG → Met.

AUG UUA ACG UUC UAA AUG UGG GGG GGA CAC CAG
 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓

- (a) *Met-Leu-?Thr-Phe* STOP *Met-Trp-?Gly-Gly-His-Gln.* The sequence contains both nucleotide triplets that were absent from the example, so we cannot be sure in the answer, but it will be confirmed when we have solved the problem to the end.

(b) *Met-Lys-Cys-Ile* ← AUG $\left\{ \begin{array}{l} \text{AAA} \\ \text{AAG} \end{array} \right\} \left\{ \begin{array}{l} \text{UGU} \\ \text{UGC} \end{array} \right\} \left\{ \begin{array}{l} \text{AUU} \\ \text{AUC} \\ \text{AUA} \end{array} \right\}$ ($1 \times 2 \times 2 \times 3 = 12$ possibilities).

- (c) A root XY is strong if XYA, XYG, XYC and XYU encode the same amino acid (UC, CC, CG, GC). A root is weak if this is not the case (UU, CA, AG, GA).