

## Solution to the DNA sequence Problem.

The key to this problem is to remember that there are two possible RNA sequences that could be produced from any segment of DNA: one sequence would be transcribed from the DNA strand as given, and one that could be transcribed from the complementary DNA strand.

Both possible RNA strands are given below. Which is the correct one? The one that gives an open reading frame 94 amino acids in length! As you will see, only the sequence transcribed from the opposite DNA strand meets this requirement.

PS: There something very unusual about the end of this gene. All 3 "STOP" codons, in-frame, are present, ensuring that this protein definitely will be terminated. There's a reason for this unusual degree of care. This is a coat protein for the viral surface. If a supressor mutation interferes with the reading of its stop codon, a fusion protein will be produced resulting in a viral coat protein fused to a viral DNA polymerase. Very bad news, hence the extra "care" to make sure that this protein is properly terminated.

**Kenneth R. Miller    Professor of Biology    Brown University**

RNA as it would be transcribed from the opposite DNA strand: (STOP codons underlined)

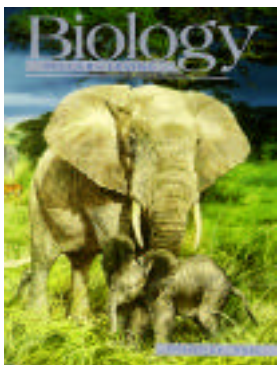
5' -  
UAGAGCCCUCAACCGGAGUUU  
GAAGCAUGGCUUCUAACUUA    *Bases #1-126*  
 CUCAGUUCGUUCUCGUCGACA  
 AUGGCGGAACUGGCGACGUGA  
 CUGUCGCCCAAGCAACUUCG  
 CUAACGGGGUCGCUGAAUGGA  
 -----  
 UCAGCUCUAACUCGCGUUCAC    *Bases #127-252*  
 AGGCUUACAAAGUAACCUGUU  
 GGCGUUCGUACUAAAUAUGG  
 AACUAACCAUUCCAAUUUUCG  
 CUACGAAUUCGACUGCGAGC  
 UUAUUGUUAAGGCAAUGCAAG  
 -----  
 GUCUCCUAAAAGAUGGAAACC    *Bases # 253-348*  
 CGAUUCCCUCAGCAAUCGCAG  
 CAAACUCCGGCAUCUACUAAU  
AGACGCCGGCCAUUCAAACAU  
GAGGAUUACCC - 3'

RNA transcribed from the same DNA strand which has been sequenced: (STOP codons underlined)

3' -  
 AUCUCGGGAGUUGGCCUCAA  
 CUUCGUACCGAAGAUUGAAU    *Bases #1-126*  
GAGUCAAGCAAGAGCAGCUGU  
 UACCGCCUUGACCGCUGCACU  
 GACAGCGGGGUUCGUUGAAGC  
GAUUGCCCCAGCGACUUACCU  
 -----  
 AGUCGAGAUUGAGCGCAAGUG    *Bases #127-252*  
 UCCGAAUGUUUCAUUGGACAA  
 CCGCAAGCAUGAAUUUAUACC  
 UUGAUUGGUAAGGUAAAAGC  
 GAUGCUUAAGGCUGACGCUCG  
AAUAACAAUUCCGUUACGUUC  
 -----  
 CAGAGGAUUUUCUACCUUUGG    *Bases # 253-348*  
 GCUAAGGGAGUCGUUAGCGUC  
 GUUUGAGGCCGUAGAUGAUUA  
 UCUGCGGCCGGUAAGUUUGUA  
 CUCCUAAUGGG - 5'

Direction of Translation:

Direction of Translation:



For Additional Teaching Resources, including a guide to exploring the human genome, visit our website:

**[www.millerandlevine.com](http://www.millerandlevine.com)**

*Sincerely,*  
 Ken Miller and Joe Levine  
 authors of *BIOLOGY* (Prentice Hall)

## Sequence Decoded:

5' UA	GAG	CCC	UCA	ACC	GGA	GUU	U-
-GA	AGC	AUG	GCU	UCU	AAC	UUU	A-
		<i>met</i>	<i>ala</i>	<i>ser</i>	<i>asn</i>	<i>phe</i>	
-CU	CAG	UUC	GUU	CUC	GUC	GAC	A-
<i>thr</i>	<i>gln</i>	<i>phe</i>	<i>val</i>	<i>leu</i>	<i>val</i>	<i>asp</i>	
-AU	GGC	GGA	ACU	GGC	GAC	GUG	A-
<i>asn</i>	<i>gly</i>	<i>gly</i>	<i>thr</i>	<i>gly</i>	<i>asp</i>	<i>val</i>	
-CU	GUC	GCC	CCA	AGC	AAC	UUC	G-
<i>thr</i>	<i>val</i>	<i>ala</i>	<i>pro</i>	<i>ser</i>	<i>asn</i>	<i>phe</i>	
-CU	AAC	GGG	GUC	GCU	GAA	UGG	A-
<i>ala</i>	<i>asn</i>	<i>gly</i>	<i>val</i>	<i>ala</i>	<i>glu</i>	<i>trp</i>	
-UC	AGC	UCU	AAC	UCG	CGU	UCA	C-
<i>ile</i>	<i>ser</i>	<i>ser</i>	<i>asn</i>	<i>ser</i>	<i>arg</i>	<i>ser</i>	
-AG	GCU	UAC	AAA	GUA	ACC	UGU	U-
<i>gln</i>	<i>ala</i>	<i>tyr</i>	<i>lys</i>	<i>val</i>	<i>thr</i>	<i>cys</i>	
-GG	CGU	UCG	UAC	UUA	AAU	AUG	G-
<i>trp</i>	<i>arg</i>	<i>ser</i>	<i>tyr</i>	<i>leu</i>	<i>asn</i>	<i>met</i>	
-AA	CUA	ACC	AUU	CCA	AUU	UUC	G-
<i>glu</i>	<i>leu</i>	<i>thr</i>	<i>ile</i>	<i>pro</i>	<i>ile</i>	<i>phe</i>	
-CU	ACG	AAU	UCC	GAC	UGC	GAG	C-
<i>ala</i>	<i>thr</i>	<i>asn</i>	<i>ser</i>	<i>asp</i>	<i>cys</i>	<i>glu</i>	
-UU	AUU	GUU	AAG	GCA	AUG	CAA	G-
<i>leu</i>	<i>ile</i>	<i>val</i>	<i>lys</i>	<i>ala</i>	<i>met</i>	<i>gln</i>	
-GU	CUC	CUA	AAA	GAU	GGA	AAC	C-
<i>gly</i>	<i>leu</i>	<i>leu</i>	<i>lys</i>	<i>asp</i>	<i>gly</i>	<i>asn</i>	
-CG	AUU	CCC	UCA	GCA	AUC	GCA	G-
<i>pro</i>	<i>ile</i>	<i>pro</i>	<i>ser</i>	<i>ala</i>	<i>ile</i>	<i>ala</i>	
-CA	AAC	UCC	GGC	AUC	UAC	UAA	U-
<i>ala</i>	<i>asn</i>	<i>ser</i>	<i>gly</i>	<i>ile</i>	<i>tyr</i>	<span style="border: 1px solid black; padding: 1px;">STOP</span>	
-AG	ACG	CCG	GCC	AUU	CAA	ACA	U-
<span style="border: 1px solid black; padding: 1px;">STOP</span>							
-GA	GGA	UUA	CCC	- 3'			
<span style="border: 1px solid black; padding: 1px;">STOP</span>							



Note the 3 in-frame STOP codons!