

27 rare Pro codons in Paxillin!!!

DNA Strider Seq.
 I E G R C - (PAX 38-557)
 atagaaggagatgccacacataccaggagattgccacccccgtccccccccccccgtccagcg
 aggcctcaatggcacaatcctgacccttagaccagtgccagccagcggctcccgattcaccaccagc
 agcctcagctctcatcactgtgtacggctccagtgccaaaacttccagtgctccaccctcaggacagtggt
 ggctctccgtgctccgagtggtgagggagcagctctacagctcccaacaagcagaatcagctga
 gccttaaccaccgtaatgagcacctcctggcagcaacttctgaactgaccctcgtctggaact
 gaacgctgtacagcataaccgccagctccctcagatgaggccaactcaagccccccgtctcgggg
 cctgagcccccttatggtgtccagagactaacgccccctggaggcaaaagctgggccccctgacgaaa
 gagaagcctaagcggaaatggggcggggcctggaggacgtgcccagctgtggagagctcttggat
 gaactggagagctccgtgccaccgccctcctccatcactgtaaccagggcgagatgagcagccgc
 agcgcctaccctcccaacagcagacacgcatcggcctcctcctccaccaggagctggagcagct
 gatggctcgtctcggattcaagttcatgcccagggagacagggagcagctcaccctcggggggc
 ccccgaagcccgggacagctggacagcagctggggagcctcagctgactgaacaagctggggg
 tcgccagctgcgcaaaagagctgcccggcctcaagaagccatcggcggcaggttggaccgccat
 ggggaagacgtggcaccgccagcactcgtctgcaccactgcccagggagatcggatccggaaactc
 ttcgagcgggatggacagccctactgtgaaaaggactaccacaactcttccccgcctgactactgca
 acggccccatctgataaagtgtgacagccctgaccggcagctggcaccctgaacactctctgtgac
 agrtggagcctcttggctccgaagggttccacgagaaggacgcaaggcctactgtcgaaggactact
 tcgacatgttcgaccaaagtgtggcctgcccggccatcctgagaaactatctcagccctcaaca
 cgtgtggcctcctgagcttctgtgcccgggaatgctcagccatggaacggcagctcttcgagcag
 gacgggcagccctactgtgaggtgcaactaccagagcggcggcctcgtgtgttggctgcccagaagc
 ccacaccggccctgcatcaccgcatggccaagaagttccaccggcagcctcgtctgtcctctgctc
 caagcagctcaacaaggcaccctcaaggagcagaacgacaagcctactgtcgaactgcttcccaagc
 ctctctgtag

When I mutated in IEGR, the I, G and R codons that I chose were all rare (how silly!)

Not only are there 27 rare CCC Pro codons, but many are close together (in the pro-rich regions) and many are at the beginning of the seq.

This may explain all the expression problems (including low yield)

526 codons rare = ●

MW : 57361 Dalton CAI(S.c.) : 0.073 CAI(E.c.) : 0.259

TTT phe F	2	TCT ser S	5	TAT tyr Y	2	TGT cys C	9
TTC phe F	25	TCC ser S	13	TAC tyr Y	12	TGC cys C	17
TTA leu L	1	TCA ser S	7	TAA OCH Z	-	TGA OPA Z	-
TTG leu L	2	TCG ser S	4	TAG AMB Z	1	TGG trp W	4
CTT leu L	4	CCT pro P	12	CAT his H	2	CGT arg R	-
CTC leu L	10	<u>CCC pro P</u>	<u>27</u>	CAC his H	15	CGC arg R	7
CTA leu L	-	CCA pro P	5	CAA gln Q	1	CGA arg R	2
CTG leu L	21	<u>CCG pro P</u>	7	CAG gln Q	25	CGG arg R	9
ATT ile I	1	ACT thr T	3	AAT asn N	2	AGT ser S	5
ATC ile I	11	ACC thr T	9	AAC asn N	18	AGC ser S	18
<u>ATA ile I</u>	1	ACA thr T	6	AAA lys K	6	<u>AGA arg R</u>	<u>1</u>
ATG met M	8	ACG thr T	6	AAG lys K	23	<u>AGG arg R</u>	<u>1</u>
GTT val V	2	GCT ala A	4	GAT asp D	5	GGT gly G	3
GTC val V	11	GCC ala A	23	GAC asp D	15	GGC gly G	19
GTA val V	2	GCA ala A	3	GAA glu E	8	<u>GGA gly G</u>	<u>6</u>
GTG val V	13	GCG ala A	-	GAG glu E	27	GGG gly G	15

Below are the rare codons that are supplied in the Rosetta E. coli cell line.

aua- ile 1
 agg- arg 1
 aga- arg 1
 cua- none
 ccc- **pro 27!**
 gga- gly 6

6.8% rare codons