

Appendices

Appendix A:

Codon Preference tables (www.kazusa.or.jp/codon)

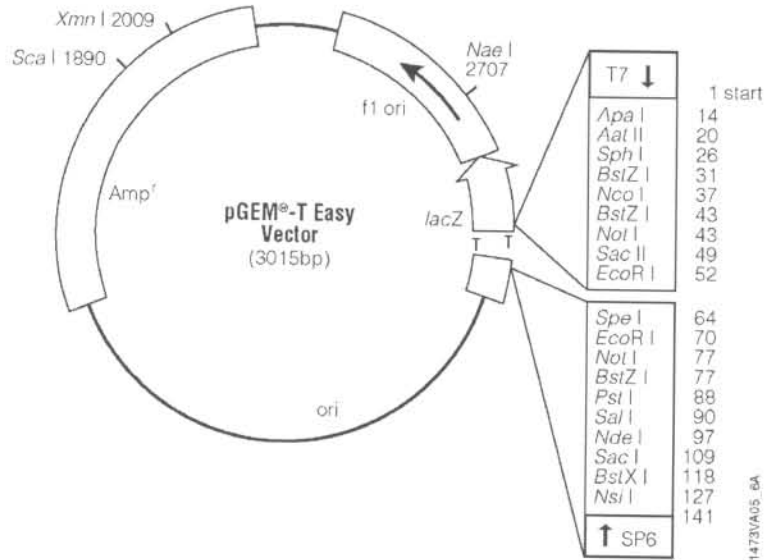
E.coli K12 strain codon preference table (1608122 codons)

Phe	UUU	22.4 (35982)	Ser	UCU	8.5 (13687)	Tyr	UAU	16.3 (26266)	Cys	UGU	5.2 (8340)
	UUC	16.6 (26678)		UCC	8.6 (13849)		UAC	12.3 (19728)		UGC	6.4 (10347)
Leu	UUA	13.9 (22376)		UCA	7.2 (11511)	Stop	UAA	2.0 (3246)	Stop	UGA	0.9 (1468)
	UUG	13.7 (22070)		UCG	8.9 (14379)		UAG	0.2 (378)	Trp	UGG	15.3 (24615)
	CUU	11.0 (17754)	Pro	CCU	7.1 (11340)	His	CAU	12.9 (20728)	Arg	CGU	21.0 (33694)
	CUC	11.0 (17723)		CCC	5.5 (8915)		CAC	9.7 (15595)		CGC	22.0 (35306)
	CUA	3.9 (6212)		CCA	8.5 (13707)	Gln	CAA	15.4 (24835)		CGA	3.6 (5716)
	CUG	52.7 (84673)		CCG	23.2 (37328)		CAG	28.8 (46319)		CGG	5.4 (8684)
Ile	AUU	30.4 (48818)	Thr	ACU	9.0 (14397)	Asn	AAU	17.7 (28465)	Ser	AGU	8.8 (14092)
	AUC	25.0 (40176)		ACC	23.4 (37624)		AAC	21.7 (34912)		AGC	16.1 (25843)
	AUA	4.3 (6962)		ACA	7.1 (11366)	Lys	AAA	33.6 (54097)	Arg	AGA	2.1 (3337)
Met	AUG	27.7 (44614)		ACG	14.4 (23124)		AAG	10.2 (16401)		AGG	1.2 (1987)
Val	GUU	18.4 (29569)	Ala	GCU	15.4 (24719)	Asp	GAU	32.2 (51852)	Gly	GGU	24.9 (40019)
	GUC	15.2 (24477)		GCC	25.5 (40993)		GAC	19.0 (30627)		GGC	29.4 (47309)
	GUA	10.9 (17508)		GCA	20.3 (32666)	Glu	GAA	39.5 (63517)		GGA	7.9 (12776)
	GUG	26.2 (42212)		GCG	33.6 (53988)		GAG	17.7 (28522)		GGG	11.0 (17704)

Plasmodium falciparum 3D7 strain codon preference table (2776356 codons)

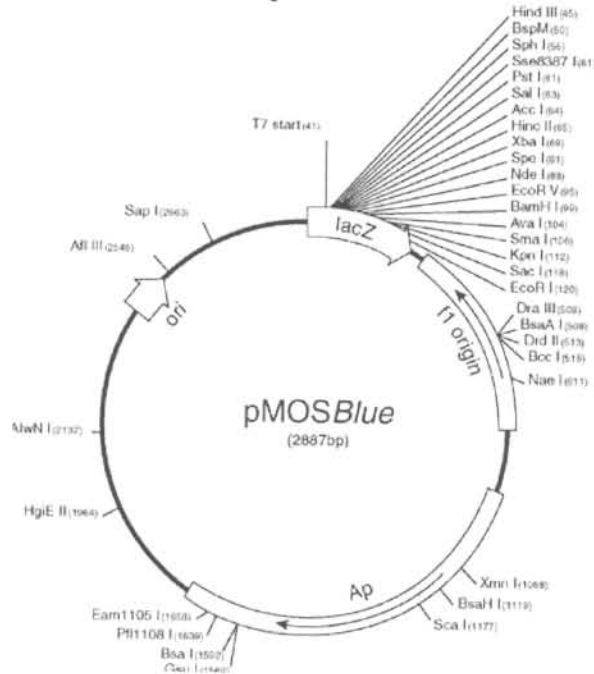
Phe	UUU	36.3 (100866)	Ser	UCU	14.8 (41088)	Tyr	UAU	50.5 (140208)	Cys	UGU	15.1 (42007)
	UUC	7.1 (19624)		UCC	5.1 (14067)		UAC	6.1 (16953)		UGC	2.2 (6245)
Leu	UUA	47.3 (131340)		UCA	16.6 (46169)	Stop	UAA	0.9 (2582)	Stop	UGA	0.3 (810)
	UUG	10.3 (28684)		UCG	3.0 (8225)		UAG	0.1 (379)	Trp	UGG	4.9 (13545)
	CUU	8.6 (23932)	Pro	CCU	8.0 (22094)	His	CAU	20.8 (57660)	Arg	CGU	3.0 (8270)
	CUC	1.8 (4959)		CCC	2.0 (5664)		CAC	3.5 (9631)		CGC	0.4 (1149)
	CUA	6.0 (16735)		CCA	9.2 (25658)	Gln	CAA	24.2 (67221)		CGA	2.4 (6674)
	CUG	1.4 (3983)		CCG	0.9 (2611)		CAG	3.7 (10280)		CGG	0.3 (746)
Ile	AUU	35.8 (99371)	Thr	ACU	10.6 (29480)	Asn	AAU	22.6 (340414)	Ser	AGU	20.3 (56431)
	AUC	6.2 (17289)		ACC	4.8 (13257)		AAC	19.8 (54952)		AGC	3.9 (10940)
	AUA	50.0 (138720)		ACA	21.7 (60278)	Lys	AAA	95.6 (265519)	Arg	AGA	16.0 (44311)
Met	AUG	21.7 (60328)		ACG	3.8 (10464)		AAG	21.4 (59397)		AGG	4.3 (11924)
Val	GUU	15.6 (43447)	Ala	GCU	8.3 (23103)	Asp	GAU	55.6 (154404)	Gly	GGU	12.0 (33208)
	GUC	2.4 (6713)		GCC	2.1 (5802)		GAC	8.6 (23893)		GGC	1.3 (3600)
	GUA	16.0 (44474)		GCA	8.5 (23527)	Glu	GAA	62.3 (172861)		GGA	12.7 (35265)
	GUG	4.8 (13274)		GCG	1.0 (2894)		GAG	10.4 (29008)		GGG	2.8 (7749)

Appendix B: pGEM T Easy Vector map



Position of the A-T cloning site is indicated in the multiple cloning cassette within the *lacZ* gene. The T7- and SP6 promoter regions indicate vector primer positions respectively. The ampicillin resistance gene (Amp^r) is also indicated.

Appendix C: pMOSBlue vector map



Cloning takes place at the linearised *EcoRV* site, which is also dephosphorylated. Phosphorylation of the insert during the *pk* reaction facilitates blunt ended cloning in the subsequent step. The vector also confers ampicillin resistance to identify transformed cells. Inactivation of the *lacZ* gene by the cloned insert also results in a white phenotype of recombinant clones. The plasmid furthermore contains *lacZ* Δ M15 and thus tetracycline resistance is observed in recombinant strains. This eliminates false white colonies that have lost the plasmid.

Appendix D: List of oligonucleotide sequences

Oligonucleotide name	Sequence (5' to 3' direction)
Fwd1	CGCGGACATATGGAGAAGAACCAGAACGATAAAAGCAACAAAAACGATATTATTCACATGAACGATAAAAGC
Fwd2	AACAACCTTTATTGATAAGAACGATGAACATGATATGAGCGATATTCTGCATAAAAATTAATAATGAGGAGAAGAAA
Fwd3	TGCCTGGAAGCTGCTGTATAAAAACCCATGCGCTGAAACTGGGCTGGATAACCCGAAGAAGCTGAACGAAAG
Fwd4	CCATCCATATTGCTGGCACCAACGGCAAAGGCAGCGTGTGCTATAAAAATTTATACCTGCCTGAAGATCAA
Fwd5	AGCCCGCATATTTTTAGCCTGCGCGAACGCATTATTGTGAACGATGAACCGATTAGCGAAAAAGAACTGATTC
Fwd6	CAAAGCGAAGAAGCTGTATATTAACCCGAGCTTTTTTGAATTTATTACCCTGGTGGCGTTTTCTGCATTTTTTAAAC
Fwd7	CGATTATTGAAACCGGCATTGGCGGCCGCTGGATGCCACCAACATTCTGACCAAACCGGAAGTGATTGTGATTAC
Fwd8	CGATAACCTGCCGATTATTTGCAACGAAAAAATGGCATTTTTTAAAAAAGATGCGAACGTGGTGATTG
Fwd9	TGATAAAGCGAAAGAGCTCAACTGCACCATTCATACCGTTGTGCCGGAACCGCGCGGCGAACGCTATAACGAAGAA
Fwd10	CTGCGCACCCCTGGAAATCTGAACATTAGCATTGATTATTTTTCTGAAAAGCATTATTCCGATTAAACCGCCGCT
Fwd11	GATTGAGCATATTAAGAAGAAGTTCAGCCAGACAACCTGGAACATAACGTGCAGTATCCGCTGGCGGTGATTG
Fwd12	TGATCGCCTGTGCACCGATATCAACTATTTTTATAAAGGCCAGAACATTTCGCATTTGCATTAGCATTAC
Fwd13	CATCCGTTTATTGCGCAGTTTGGCGATACCCTGAAAGATATTTTTTATCTGCCGAGCCTGAACGAACGCACCTA
Fwd14	TGAACAACGAAGAAGAAATTA AAAACGAAATTAAGAAGCTGATTCTGAGCAGCAGCAAAAAAGTGGGCAA
Fwd15	ATTAACGAAGAAGATGCGCTGAAACTGTATAAACCGGGCTGCATTCCGCTGATTATTA AAAACCGCTTTCTG
Fwd16	GCTGGTTTGTGGTACCTTCTTCGTGTTTGATGAAGTGTGTAACGTGTTTGATATTATAGCGATATGCAGGA
Rev1	GTTTCATCGTTCTTATCAATAAAGTTGTTAATGTTGTTTTATCATAGTTGCCGCTTTTATCGTTCATGTGAATAAT
Rev2	GTTTTATACAGCAGTTCAGGCATTTCGCTATAGCTTTAATTTCTTCATATTTCTTCTCCTCATTATTAATTTTAT
Rev3	TGGTGCCAGCAATATGGATGGTTTTATATTTATCGCACGGATGGCCAAAGCTTTCGTTTCAGCTTCTTCGGGT
Rev4	CAGGCTAAAAATATGCGGGCTGCTAAACAGTCCCCTTTAAATTTTTGATCTTCAGGCAGGTATAAATTT
Rev5	GTTAATATACAGCTTCTTCGCTTTGTTTCAGCACTTCGTTCCAGCATGAATCAGTTCTTTTTTCGCTAATCG
Rev6	CCAATGCCGTTTTCAATAATCGCATAATCCACTTTTTTGTTTAAAAAATGCAGAAACGCCAC
Rev7	GCAAATAATCGGCAGGTTATCGCCAGAATGTTGAGATGATCATAGCCAATGCTGGTAATCACAATCACTTCC
Rev8	GAGCTCTTTCGCTTTATCAAACAGTTTTTATAAATCGCCAGCTCGGGCCAATCACCACGTTTCGCATCT
Rev9	TAATGTTTCAGAATTTCCAGGGTGCAGCAGCAATGCGGCTGTTTTCTTCGTTATAGCGTTTCG
Rev10	GAACCTCTTCTTAATATGCTGAATCTGTTTCGGTTCGCCAGATACTGAATGCGCAGCGGCGGTTAATCGGAAT
Rev11	TATCGGTGCACAGGCGATCAATCGCGGTTTCGTTATGGCCACATCCAGAATCACCGCCAGCGGATAC
Rev12	CAAACCTGCGCAATAAACGGATGAAACACGCTCAGGTTGCGCGTTTTGGTAATGCTAATGCAAATGCGAATG
Rev13	CGTTTTTAATTTCTTCTTCGTTGTTTCAGCATTCCACAATTTCTTCAAATCATAGGTGCGTTTCGTTTCAGGCT
Rev14	TTCAGCGCATCTTCTTCGTTAATGTTGCCCTGTTTTTCATGCGCCAGCCATTTGCCACTTTTTTGCTGCTG
Rev15	AGAAGGTACCACAAACCAGCAGAATGCTGTTATCTTTGCAACACTCCAGAAACGCGTTTTTAATAATCAGC

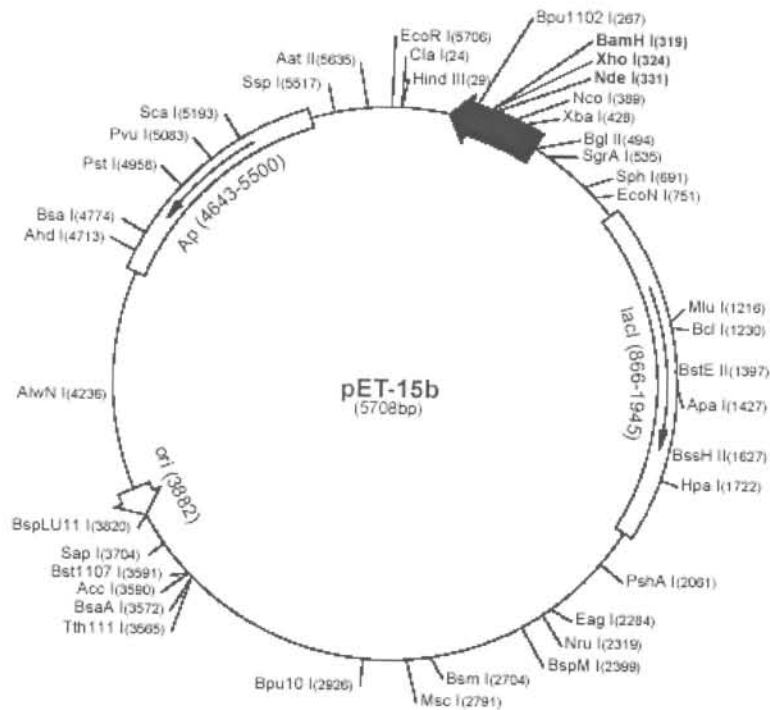
Oligonucleotide name	Sequence (5' to 3' direction)
Rev16a	CCGGATCCTTACACCAGGCTCGGTTTCGTTTCATAAAAATGGTATCCTGCATATCGCTATGAATATCA
Rev16b	CTCGGATAATGTTTCACCAGGCTCGGTTTCGTTTCATAAAAATGGTATCCTGCATATCGCTATGAATATCA
dhfs2f1	AATTTAAAGTGGGACTGTTTAGCAGCCCGCATATTTTTAGCCTGCGCGA
dhfs2r1	TGCCTAATCGGTTTCATCGTTAACAATAATGGCTTCGCGCAGGCTAAAAATATGC
dhfs2f2	AACGATGAACCGATTAGCGAAAAAGAACTGATTCATCTGGTGAACGAAGTGCTG
dhfs2r2	AGCTCGGGTTAATATACAGCTTCTTCGCTTTGTTTCAGCACTTCGTTACCAGAT
dhfs2f3	GCTGTATATTAACCCGAGCTTTTTTGAATTATTACCCTGGTGGCGTTTCTGCA
dhfs2r3	AATAATCGCATAATCGACTTTTTTGTTTAAAAAATGCAGAAACGCCACCAGGGT
dhfs2f4	CAAAAAAGTCGATTATGCGATTATTGAAACCGGCATTGGCGGCCCGCTGGATGCCA
dhfs2r4	GTAATCACAATCACTTCCGGTTTGGTCAGAATGTTGGTGGCATCCAGGCGGCCG
dhfs2f5	CCGGAAGTGATTGTGATTACCAGCATTGGCTATGATCATCTGAACATTCTGGGC
dhfs2r5	GTTGCAAATAATCGGCAGGTTATCGCCAGAATGTTTCAGATGAT
dhfs3f1	TGCGCACCCCTGGAAATCTGAACATTAGCATTGATTATTTTCTGAAAAGCATTATTCC
dhfs3r1	GGTCGCCAGATACTGAATGCGCAGCGGCGGTTTAAATCGGAATAATGCTTTTCAGAAAA
dhfs3f2	GCGCATTCAGTATCTGGCGACCGAACAGATTAGCATATTAAGAAGAAGTTCAGCCCA
dhfs3r2	CACCGCCAGCGGATACTGCACGTTATGTTCCAGGTTGTCTGGGCTGAACTTCTTCTTAA
Sequencing primer	Sequence (5' to 3' direction)
U19	AGGGTTTTCTCAGTCACGA
T7 promoter	TAATACGACTCACTATAGGG
Sp6 promoter	ATTTAGGTGACACTATA

Appendix E: Vector systems used for recombinant expression

- N-terminal His₆-tagged expression: pET15b vector
(www.novagen.com)

pET-15b sequence landmarks

T7 promoter	463-479
T7 transcription start	452
His*Tag coding sequence	362-380
Multiple cloning sites (<i>Nde</i> I - <i>Bam</i> HI)	319-335
T7 terminator	213-259
lacI coding sequence	(866-1945)
pBR322 origin	3882
<i>bla</i> coding sequence	4643-5500



- C-terminal His₆-tagged expression: pET22b vector
(www.novagen.com)

pET-22b(+), sequence landmarks

T7 promoter	361-377
T7 transcription start	300
<i>pelB</i> coding sequence	224-289
Multiple cloning sites (<i>Nco</i> I - <i>Xba</i> I)	158-225
His*Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	764-1843
pBR322 origin	3277
<i>bla</i> coding sequence	4038-4895
Π origin	5027-5482

