

Supplemental Figure 11: Selected PU elements in *Escherichia coli*

Site	Ctx	PU element and flanking region
PU-Y(G) (75)		
2289267B	C	GACAAAGTGCCTTAGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAATTGTGTAAATTCAATA
1814188F	C	TACATAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAACCTGTGCAAAATTCATA
3068004F	PA	TATCAATGTGCGCTTTGTCAATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAACCTGTGCAAAATTCATA
698627F	PA	GACAAAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAAGCATGAAAATTCATA
3637211F	C	GACATAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAACATGCAAAATTCATA
814851B	C	GACAAAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAACGCGCAAAATTCATA
714531B	PA	GACAAAGTGTTCGTTGTTTTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAATCGCTTAAATTCATA
4407155F	PP	GACAAAAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACATAATCAGCAAAATTCATA
3267788B	C	TGTTGAGTGCCTTTATTCATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAAGCATGCAAAATTCATA
2712349B	C	GACAAAGTGCCTTTATTCATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
3510569B	PP	GGCAAAAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAATCGTGCAAAATTCATA
3674187B	C	GGCAAAAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
2682183B	C	GTCAATGTGCTCGTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAATCGTGCAAAATTCATA
2547532B	C	GACCAAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAATCGTGCAAAATTCATA
609384B	PP	GACAAAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAATCGTGCAAAATTCATA
2943961B	PP	GACTAAGTGCCTTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCATGCAAAATTCATA
4125954F	PA	GACAACGTGCTCGTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
4078022F	C	GACAACGTGCTCGTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
138743B	C	GACAACGTGCTCGTTGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
707104B	PA	TTGGATTGTTGCTGTTTCTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAAGCATGCAAAATTCATA
5616B	C	AGCGTTAAGTTTGCAGTCAATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAAGCATGCAAAATTCATA
1952432F	I	CCGTAAGGCTCGTATTACTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAAGCATGCAAAATTCATA
3875588F	PA	CCCCCTTTCTTATAGCTTCTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAAGCATGCAAAATTCATA
356687F	C	GTTACAGCGAGCTTAGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATAGGCACTGAATCTGATAG
356887F	C	TGTCTGCCAGCTGTAGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATAGGCACTGAATCTGATAG
4552508B	PA	GGTGCCTTAATCAAAATTTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAACGGGCAATTTACGGCC
338960F	I	CGCTGTGACCTACATGTTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
631318B	PA	TAACCATGTTGCTGAAAGAGTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
4233424B	I	CGCTGCGCATCAGGCATCGGTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
757654F	PP	CGCTGCGCATCAGGCACCAAGTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
2840424F	I	AGATTTAGAGTAGGTAACAGTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAATCGTGCAAAATTCATA
507786B	C	ATCAGGCAATGGTGTCTCGGATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGAGATGAAGCGATGTTTGG
2234651F	C	ATCCGGCATTTGGCTCCTGTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACAAAACCATCAGATCATGAT
2131435F	PA	GGCTTATGAGCTGTTTGTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGGGGCGGTGCGAATGCAGGC
339053F	PP	CATCAGGCATTGTGACCATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
339239F	PP	CATCAGGCATTGTGACCATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
4455188B	C	CATCAGGCATTGTGACCATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
4612462B	C	ATCAGGCATTGTGACCATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
4612361B	C	ATCAGGCATTGTGACCATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
4612260B	I	ATCAGGCATTGTGACCATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
410335B	C	TCAGGCATCGGAGCACTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
4216439B	PA	ATCAGGCATCGGAGCACTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
74234F	PP	ATCAGGCATCGGAGCACTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
374335F	PP	ATCAGGCATCGGAGCACTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
374133F	PP	TATGCCCATCCAGCAGCACTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
376692B	C	TGCGTTGATAATAGCAGCACTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
376592B	C	CATCAGGCATGCGGAGCACTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
4025572B	C	TTTGAGCGGGTTTAACTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
2671777F	I	TCGATCGGCGTTAGGAAAGATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
740156F	C	GTCAGGCAATCGAGCCAGATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
3096414F	I	CTGGCTGAGTAACTTGTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
1083937B	PA	ATCAGGCATCAGTGACGATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
410439B	C	TCTCGGTGCTCATTACTTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
3734206F	C	ATCAGGTAACGCTGCGTGTTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
2660390F	PA	CGCTGCGCATCAGGCAACGGTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
2660481F	PA	CGCTGCGCATCAGGCAACGGTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
111526B	C	CATCAGGCATCTTGTGCAAAATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
173535B	C	AGTTGTGATCCAGCATGTTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
3080805B	C	CGGGTAAAGCTGCCATTATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
72167B	C	GATTACGGGTTAGTATGACATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
2595686F	PA	GTCTGGTGTGCGGATGTTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
248281B	C	ATCCAGCATTTGGTGACAACGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
844833F	PA	TGACAACGTGCGGGCTTTGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
2116525B	PP	CGCTGCGCATCAGGCACTGAATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG
4631137F	C	ATCAGGCATCGTGACCGAATGCCGGATGCGGCGT- GA-ACGCCCTTATCCGGCCTACGATGCGGCACCAACATGAG

1112684B PA TTTTATTGGCTCGTTGTTT **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCCAC**AAAGTCTTGG **AAAA**TCAATA
4482386B PP GCGTCAGGCAAGGCTGTTAT **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCTAC**ACAGCACTGAACTCGTAGGC
2441790F C TAATCTGCTCTCGCAGTGTT **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCTAC**GTGTGACGTGTAGGCCTGAT
4604619B PP AATTATCTCGCCGAATACCG **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCTAC****AAAA**CCTGCACGTTAGCCCT
3137623B C CATCAGGCATTGTGCACGAT **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCCTAC**ATGGCTTGTGCGACGAACA
3214646F D TCTCAGGCATCATTACTCAA **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCCTAC**GTGTGAGATGAGTCCACGTT
2566240F PA CGGCAATTGCACGCGCCAC **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCTAC**ATGTGCAATCCCGTAGGCTG
3229353F PP CGGCAACTGCACGCGCCAC **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCTAC**ATGTGTGTTCCCGTAGGCTG
1550074F C GACAACGTAGGCTTTGTTCA **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCCTAC****AAAA**CCCTTCAAATCCAATAG
1550252F C GACAACGTAGGCTTTGTTCA **TGCCGGATGCGGCGT** -GA-ACGCCTTATCCG **CCCTAC****AAAA**CCCTTCAAATCCAATAG
consensus * * * * * _ * * * * *

PU(A) (55)

4371265B C GATAACGTGATCGTTGGTCA **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**AAATTCTCAAGCATTCAATC
3596465B C ATATCGAGTACGTTGCTTCA **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**AAGATCCAAAGAAATCAGTA
1067621B C GACAACGTACGCATTGTTCA **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**CAGTGCAAAATTCATA
831556B C GACAAGTGCTTGTGTGAAAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**GCATGCAAAATTCATA
2536536F C **AAAAAA**GTGCGCTTTGTTTAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**TCGTGCTAATTCAATA
3328461F C GACAAAGTGCCTTTGTTTAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**TCGTGCAAAATTCATA
898898F C GACAAAGTGCCTTTGTTTAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**TCGTGCAAAATTCATA
1999971F PA TGACAAATTGCGCTTTATGT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**TCGTGCAAAATTCATA
2653035B PP GTTAACGGCGTAACTAATA **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**AAATCTGCAAAATTCATA
4315969B PP GTTTGCAGAAATTGAGCCGG **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**AAATGCGCTCCCGTAGGTC
2833117B C GTGTTGCATTTGGACACCAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**ACATGCGTCCCGTAGGTC
216035B C CATCAGGCGTTGAGCACCAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**TCGTGCAAAATTCATA
247438B I ATCAGGCACCTGCGTACAAC **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**TCGTGCTTCCCGTCAA
4101559F C CATCAGGCATCCTGCTCAAA **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**AAATGCGAAGTTTAACTCCG
3982222F C GCGCGCGCGCTGCGGTTGAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**ATGATCGTGCAAAATTCATA
4458468B PP TTTACTCTGAGCTAAGCAAA **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**ATTTGGCAATTGCTCCAGGT
4283305B PP GCGTCGCATCAGGCATTGAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**ATTCGCGAAGGTTACCCGA
1133853B PA ACAAAGTGCATATGCTTCT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**AAGTGTGCG **AAAA**TCGATA
25742B PA CTTTATATTAAAGGATTTT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**AGTTTCGAGCAATCTGTAG
4247415B PA ATAATGACAACCTGT **TTTT**AT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**ATTTGACAGCGTTGTAGGC
4061563B C TATTTCTTCTGATGTGAAAC **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTGATACGG **TTTT**CGTTG
2345161F I TGTAAGATCTGATATTGAGA **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGCTCGGTTGTAGGCCTGA
3232692B PP GCGTAAGATTGGGAGTTACG **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTACGAGCAATCTGTAG
805151B C AGTAAGAGCAATTAACATAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTGATGCGATTGTAGG
4025473B C ATCAGGCATCATGCACCAAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTGCGACTATTTGTAGG
4025473B C ATCAGGCATCATGCACCAAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTGCGACTATTTGTAGG
2314884F C TATCTGGCATTGTCACCGAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTCCCATTTATTCACAA
4323953F PA CATCAGGCAGTCGGCACCATT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGAGGGTGCGGGAATTTGTA
4324253F PA CATCAGGCAGTCGGCACCATT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGAGGGTGCGGGAATTTGTA
4323853F PA CATCAGGCAGTCGGCACCATT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGAGGGTGCGGGAATTTGTA
4324153F PA CATCAGGCAGTCGGCACCATT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGAGGGTGCGGGAATTTGTA
4324053F PA CATCAGGCAGTCGGCACCATT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGATGGTGCGGGAATTTGTA
374436F PP ATCAGGCATCTGCGCAGGAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGATGGCGCGGAATTTGTA
2338363B C CACTGTCTAGTCGGCACCATT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGGAGCATTTGTAGGCCCTG
507988B C ATTTGGCATCGACGATGAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTGAGGACAGGTTTGTGTA
507887B C ATCAGGCAATGGTGCTCGGAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTGAGGACAGGTTTGTGTA
550680B C GTTAAGTGGTCTCTATTCT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**CGATCCGGTACCCATTGTAG
2509408B C AGCGTCGCATCTGGCAGTAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**CGGTTAAAGAGCCGGTTTGT
2566140F PA CCCTGTTAGTTGTTATTTAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**ATGTGCAATCCCGTAGGCTG
3229253F I AAGCAATAATTTGCAACTG **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**ATGTTGCAATCCTCTGCGCCG
836830F C GCGTCGCATCAGGCATCTGT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GAATTACTTATAAATTACCG
3390260F PA GCGTCGCATCAGGCAACGG **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTATGTTCCGTTTGTAGG
3390169F PA GCGTCGCATCAGGCAACGG **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTATGTTCCGTTTGTAGG
2428930F PA TTGCGGCATCGGCAATGCG **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTGATGTTCCGTTGAGGC
4294024B C AGCGTCGCATCAGGCGTTGAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**TCACGCTAAATCAGGC
2792071B PA CTTCTGTTTACGACAGAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**ATCCCCGTAGGCGGACAG
4006400B C GAGGCATGCGT **AAAA**TAAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**CTGTAAGACGCTGCTGGCAT
3295051B C TTCTGATTTGAGGG **AAAA**CAAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GGTTGATGTTCCGTTGAGGC
1689492B C GACAAAGTGCCTG **TTTT**AT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC****AAAA**CTCATAAATTCAAAG
3590566F PA ACAGATTGCTCATCGTTTCA **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**AAGATCGTGCAAAATTCATA
4216537B PA CCTCATCAGGAGCAGAGAAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GTTCGGCAACGGCTGTAGGC
3201145F I TGAATGAGTATTGGGCA **TTTT**AT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GTTCACGACAGTTGTAGGC
1532886F I **AAAA**TAATTTATGTCAGGT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**ATTAGCCCGGTTGTAGACCT
3672383F I **AAAA**CGCAGCCTT **AAAA**TAAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**TAG **TTTT**ATTCGCGCTGGCTC
2314998F C CATCTGGCATTGCACTGAAT **TGCCGGATGCGGCGT** -AA-ACGCCTTATCCG **CCCTAC**GAATCCCGCATTTCTGAC
consensus * * * * * _ * * * * *

PU-Z1 (47)

2652971F	PA	TTCAACGTTTCGGGGCGCAAA	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AAGGGATTTCGCAATTTGTTG
3772238F	PP	ATGGAGAAGGTGGGGTTAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AGGATGCATCACAATTTGTT
3148669B	C	ATCCGGCATTAGGTGCTCAA	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AAAACTATTGCAACATGTT
844904B	PP	CGTCTCTTAATAAGACACAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AGGATATCTGGCAACTTATT
3875658B	PP	TGGCGCTAACCATGCGAGCG	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GAAATCTTGCAATGTATTG
3596399F	C	TTTGATTTTAAAGTGATAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AATGTGTGTGCAATTTACT
3637281B	C	ACTGATTACAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATAATCTCTGCAATCTATTG
707039F	PP	AACAGTTTCAAACGCTAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATGATCTCTGCAATATATTG
4135861B	PP	CGGGAGTTTGAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GCAATCTCTGCAATATATTG
1112619F	I	CTTCGCAATAAAACTGAGT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATCGTTCCTGCAATTTATTG
3738966F	I	GGCTGAAAGTCACTAAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATGATTTCTGCAATATATTG
430171F	I	AAGTGATTAAACGCTAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GTTATTTTCAGCAATATATTG
2943897F	PA	CCTCAGCCTAAACGCATAAA	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATAACTCTGCAATATTTTG
138678F	C	ACTGATGACAAACGCAATC	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GCTAGCTGTGCAATCCATTG
3267723F	C	ACTGATGGCAACCGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GCCATCTCTGCAATATATTG
831491F	C	CTGATGAACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GTGAACCTGCAATATATTG
1550321B	C	CTGATGACAAACGCAACAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ACGTTCACTGCAATCTATTG
1550321B	C	ACTGATGACAAACGCAACAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ACGTTCACTGCAATCTATTG
983633B	C	ACTGATGACAAACGCAACAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GCGTCCCCTGCAATATTTTG
3068074B	PP	TTTTTCGCCAAAGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ACGTATCCTGCAATATATTG
989688B	C	ACTGATGACAAAGGTAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATTGCCATGCAATGTATTG
4468443B	C	TCTGATGACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GTGATTCCTGCAATTTATTG
2289202F	C	ATTTATGACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GTGGATCGATCAATTTATTG
2547467F	C	AGTGCTGACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GTGGAGCCCAATATATTG
2712284F	C	AGTGCTGGCAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GTGGCTCATGCAATATATTG
3561584F	C	ACTGATGACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GTGGTTTATGCAATATATTG
714466F	PP	ACGTATGACAAACACAGAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GAGGATGGTGCAATATATTG
698697B	PP	TCTGATGACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GCACTCTGCAATATATTG
489232B	C	GTTGCGGGCAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ACCGCTCTGCAATATATTG
2536606B	C	ACTGATGACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AAGGTTTCTGCAATATATTG
1689428F	C	TCTGAAACCAACAGAGAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	CTTTCTGCAACACTTTGA
609319F	PA	ACTGATGACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATGGTCTGCAATATATTGA
3510512F	PA	AGTGATGATAAAGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATTTCTTGCAATATGTGCA
3010528B	C	ACTGATGACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AAGTGCAATGCAATCTATTG
2682118F	C	GATGATCATAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ACGGAGATTGCAATATATTG
356938B	C	CACAGGTCTGCGGCTGGAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GAGTTCACTGCTCTACGTAG
356738B	C	CACAGATCAGCGGCTGGAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GAGTTCACTGCTCTACGTAG
339105B	PA	CATCCGGCAGTGGTGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AGATTCTTGCGCCATTTCGTA
339012B	PA	CATCCGGCAATGGTGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AGATTCTTGCGCCATTTCGTA
339198B	PA	CATCCGGCAATGGTGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AGATTCTTGCGCCATTTCGTA
339290B	PA	GGTTATGACCAACGCCAGAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AATTCTTGCGCCATTTCGTA
814787F	C	ACTGATGACAAACGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATTTCTTGCGCCATTTCGTA
856818F	C	GCTGATGACAAATGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GCACTCTGCAATATATTG
3328531B	C	CTGATGATAAACGCCAAATC	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AAGAGAATTGCAATATGTTG
3099748B	PP	GGCGATGACAAATGCAAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	AAAGATGCACGATCGAGTAG
4126023B	PP	AGAGCGCGCGGAGTGGAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	GTCAATATTGCAATTTATTGA
3674123F	I	TCGTTGTAATAAGAGTGAAAT	TGCC	GATGCG	-CTA---CGC-	TTATC	GGCCTAC	ATTTCATACAATATATTGA
consensus		*****_***_***_*****						

PU-Z2(G) (28)

2302472F	C	CATCCAACGCCCGAGCCGGT	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	ACCGCTGTGAAAGTGC	GGCAC
2302585F	C	CATCCAACGCCCGAGCCGGT	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	ACCGCTGTGAAAGTGC	TCCAC
2302811F	C	CATCCAACGTCGAGCCGGT	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	ACCGCTGTGAAAGTGC	TCCAC
2302698F	C	CATCTAACGCCCAAGCCGGT	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	ACCGCTGTGAAAGTGC	TCCAC
2302924F	C	CATCTAACGCCCAAGCCGGT	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	ACCGCTGTGAAAGTGC	TCCAC
2428979B	PP	ACGCCGAATTAATACGGTCT	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	GCAGGGGTAGAACC	GATAGT
2345294B	PA	GCA TTTT ACGCCGATCCTG	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAATCCGAGCCGT	AGGCCGG
4025497F	C	ATCCGGCAATTGGTGCATGA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	ATTGGTGCCGGAT	CGGTAGA
3040418F	C	ATCCGGCAGTTGTGCTCCGA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAAGGCATACCCAT	TACGCC
111434F	C	CATCCGACATTCGCACACGA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AGGGATTCTA	AACTCATTGA
500672F	C	ATCTGGCATTGGCTCTCGA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	ATAT TTTT CATAT TTTT	ACAT
66672B	PP	AAAGCGCATCCGGCATTCAA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	GCCTGCGA TTTT	GTAGGCC
66587B	PP	AAAGCGCATCCGGCATTCAA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	ACCGTGCGA TTTT	GTAGGCC
5544F	C	GGTAACAATCCGGCATTCA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	GTTAATTCTGCAAT	TATATTG
2303037F	C	CATCTAACGCCCGAGCCGGT	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	CCACATCCGGCAT	ACCATGC
3040306F	C	TTATCGCATCAGGCATTTAT	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	GATTTGATGCCT	TTATCACA
2345206B	PA	ACGCTGCATCCGGCACCCGG	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
900002F	PA	CGGCACAGGCACCGTGCTGA	TGTC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
216059F	C	CATCCGGCAATGGTGCTCAA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
247462F	PP	ATCCGGCAGTTGTACGCAGG	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	GAGCCCGCAAT	ATTTGCGA
248204F	C	ATCCGGCAGTTGTACGCAGG	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
2806180F	PP	ATCCGGCGTTCTGTGCACAA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
4486528B	C	AGTCAGGCGAGATTTGTCTT	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
4038866F	C	TGCATTGGCGCAG AAAAAA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
2833068F	C	GTGGGCTACTGTGCCT AAAA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
3648714F	PP	AG AAA GTGCGCTTTATTTA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
173483F	C	AGCGTTTACGCGACAACGGA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
4604560F	PA	GGTATGCATGCGGGGCGAGA	TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC	AAACCGAGCCGT	AGGCCGGA
consensus			TGCC	GATGCG	CG	TGGCGCGT	CTTATC	GGCCTAC		

PU-X (12)

4092687B	PP	GAGTCGCCGCGGGTAAACAG	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
3112513F	PA	ACCCGGGTGCGCTTATTTCA	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
2217626F	PA	ATATTGGTCCGGCGTTTCC	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
4092319B	PP	GCGTCGCATCAGGCATTGAT	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
4151415F	PA	GCACTGGAGTTTGGCAACAG	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
4151538F	PA	GCACTGGAGTTTGGCAACAG	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
4151661F	PA	GCACTGGAGTTTGGCAACAG	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
4448947B	PA	ATCAGGCTTCAAGTGCACCAT	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
1942222F	C	AATCTATCGCTGAACACATT	TGTCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
361087B	PP	GAAGTCGCTTAAGCAATCAAT	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
3010458F	C	AAGTAAGTGCCTTTGTTGA	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
4324353F	PA	CATCAGGCATCTGGCACCGTT	TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	GGGTCGGCAACAGTTGTAGG
consensus			TGCCGGATGCGGCG	-GA-	TCGCCCTTATCCGGCCTAC	

PU-Z2(T) (12)

4324205B	PP	CATCCGGCAACGGTGCCGAC	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AAATTCCCGCACCCCTCCGTA
4324005B	PP	CATCCGGCAACGGTGCCGAC	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AAATTCCCGCACCCCTCCGTA
4323905B	PP	CATCCGGCAATGGTGCCGAC	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AAATTCCCGCACCCCTCCGTA
4324105B	PP	CATCCGGCAACAGTGCCAAC	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AAATTCCCGCACCCCTCCGTA
4324305B	PP	CATCCGGCAACGGTGCCGAC	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AAATTCCCGCACCCCTCCGTA
757612B	I	ACGCCGCATCCGGCACTGGT	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	GGTTTACGCATTACGTTGCA
4247348F	PP	ACGCCGCATTCGGCAACCAA	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AACGGCTGTCAAATGTAGGC
3201194B	C	ACGCCGCATCCGGCAATCAA	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AACTGTCTGTTGACATGTAGC
1083884F	PP	ATAAGGCATTTGTGCGCAGA	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AATCAGTGCGGGTTTGGTAG
4216463F	PP	ATCCGGCAATCGGTGCACGA	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AGCCGTTGCCGAACGTAGGC
2537608F	PA	TCTCTATCCATACGTGGTGT	TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AGGTTACAAACCTTGCCATT
3096463B	C	ACCCCGCATCCGGCAATCAA	TGTC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	AACCGCTGCCACATGTCCGG
consensus			TGCC	GATGCG	CG	TTGGCGCGT	CTTATC	GGCCTAC	

PU-Z2(GA) (12)

2116549F	PA	ACGCCGCATCCGGCATTCA	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	AGGTCCCGAGCACAGAACCG
2116469F	PA	TTTT GTTTCGTTGCCGCT TTTT	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	AAGACCCGAGCACAGAACCG
353915F	PP	CATCCGACACCATGCTCAGA	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	CCACTG TTTT ACACCGATA
4243135F	PA	ACGCCTTATCCGACAACAAC	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	ATACGTTTCGG TTTT GTAGG
3706324F	PA	CCGACAATGGCCGGCACATT	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	GTCGAGTCGAACCTCCCTGC
2314950B	C	CATCCGGCATTCAAGTGCAAA	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	TTTAATGTAATTTGTTGAAA
4547786B	C	AGCAAAGTGCCTTTGTTTA	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	AAAAGGGCTAACGTGCAGGTT
244133F	PP	ATCCGGCATTGTGCGCTGAT	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	AAATCTGTACGCGAACCCTA
2579663F	C	GGTGATTGCCGGATGCTGAT	TGCCGGATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	GGCCTAC	AGCGTCTTATCCGGCCTACA
4315919F	PA	TTTT CGCCTCGTGGTTCGGG	TGCCGGATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	GGCCTAC	GGGGAGCGCATTGTAGGCC
66757B	PP	TAGCCGCATCCGGTATGTAA	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	ACGCTGCGAT TTTT GTAGGCC
762005F	PP	TGCGCTTCTCCGGCAATTGA	TGCC	GATGCG	CGCT	GACCGG	CTTATC	GGCCTAC	GGGACCACCAATGTAGGTCG
consensus			* * ***** *						

Supplemental Figure 11: Sequences and contexts of selected PU elements of *Escherichia coli* K12. All instances of PU elements with 24-nt segments with copy number of at least 10 (Table 2 in the main text) in the genome of *E. coli* are shown. The elements were expanded to their full conserved lengths of 30 to 37 nt, highlighted in blue. Conventions are as described in Supplemental Figure 1, except that the coordinate listed under **site** is the low order coordinate of the sequence shown (including the sequence flanking the SDR element). In addition, runs of at least four thymidines or four adenosines are rendered in blue or green letters, respectively. All 18 instances where the element is given as inside the gene actually overlap the 3' end of the gene.