The nucleotide sequence of pACYC184

Ronald E.Rose

Calgene Inc., 1920 Fifth Street, Davis, CA 95616, USA
Submitted November 18, 1987 Accession no. X06403

pACYC184 is a commonly used multicopy cloning vector which was constructed by ligating restriction fragments from pSC101, Tn9, and p15A each of which have been previously sequenced (1,2,3,4,5). Despite its wide use, the complete nucleotide sequence of pACYC184 has never been reported. The sequence was completed by using oligonucleotide primers designed to span the junctions between each of the previously sequenced regions. pACYC184 is 4244 bp in length with nucleotide 1 corresponding to the EcoRI site in the original map (1). The chloramphenicol resistance (Cm) segment from Tn9 extends from the HaeII site at base 3505 to the HaeII site at base 585 with bases 219 (ATG) to 3804 encoding the Cm gene. Part of an IS1 (5) from Tn9 extends from bases 443 to 583. Bases 1494 to 3275 are derived from pSC101 with the tetracycline (Tet) resistance gene encoded by bases 1580 (ATG) to 2770. The p15A origin of replication extends from bases 581 to 1492. Three fragments; an AluI (3276) to HaeII (3368), a HaeII (3368) to HaeII (3422), and a HaeII (3422) to HaeII (3505) are located between Cm and Tet gene and are all derived from different regions of the Tet gene. During the construction of pACYC184 a precursor plasmid, pACYC175, was digested with HaeIII, AluI, and HincII, to remove extraneous DNA and to reduce the size of the plasmid. It appears that the precursor was also digested with HaeII which generated the HaeII Cm resistance segment and the HaeII fragments found between the Tet and Cm genes. The underlined sequence was determined while the rest of the sequence was taken from the published sequences of pSC101 (2), Tn9 (3,4), and p15A (5).

```
1     GAATTCCGGA TGAGCATTCA TCAGGCGGGC AAGAATGTGA ATAAAGGCCG GATAAAACTT GTGCTTATTT TTCTTTACGG TCTTTAAAAA GGCCGTAATA
101   TCCAGCTGAA CGGTCTGGTT ATAGGTACAT TGAGCAACTG ACTGAAATGC CTCAAAATGT TCTTTACGAT GCCATTGGGA TATATCAACG GTGGTATATC
201   CAGTGATTTT TTTCTCCATT TTAGCTTCCT TAGCTCCTGA AAATCTCGAT AACTCAAAAA ATACGCCCGG TAGTGATCTT ATTTCATTAT GGTGAAAGTT
301   GGAACCTCTT ACGTGCCGAT CAACGTCTCA TTTTCGCCAA AAGTTGGCCC AGGGCTTCCC GGTATCAACA GGGACACCAG GATTTATTTA TTCTGCGAAG
401   TGATCTTCCG TCACAGGTAT TTATTCGGCG CAAAGTGCGT CGGGTGATGC TGCCAACTTA CTGATTTAGT GTATGATGGT GTTTTTGAGG TGCTCCAGTG
501   GCTTCTGTTT CTATCAGCTG TCCTCCTGTT TCAGCTACTG ACGGGGTGGT GCGTAACGGC AAAAGCACCG CCGGACATCA GCGCTAGCGG AGTGTATACT
601   GGCTTACTAT GTTGGCACTG ATGAGGGTGT CAGTGAAGTG CTTCATGTGG CAGGAGAAAA AAGGCTGCAC CGGTGCGTCA GCAGAATATG TGATACAGGA
701   TATATTCCGC TTCCTCGCTC ACTGACTCGC TACGCTCGGT CGTTCGACTG CGGCGAGCGG AAATGGCTTA CGAACGGGGC GGAGATTTCC TGGAAGATGC
801   CAGGAAGATA CTTAACAGGG AAGTGAGAGG GCCGCGGCAA AGCCGTTTTT CCATAGGCTC CGCCCCCCTG ACAAGCATCA CGAAATCTGA CGCTCAAATC
901   AGTGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC GTTTCCCCTG GCGGCTCCCT CGTGCGCTCT CCTGTTCCTG CCTTTCGGTT TACCGGTGTC
1001  ATTCCGCTGT TATGGCCGCG TTTGTCTCAT TCCACGCCTG ACACTCAGTT CCGGGTAGGC AGTTCGCTCC AAGCTGGACT GTATGCACGA ACCCCCCGTT
1101  CAGTCCGACC GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGAAAGACAT GCAAAAGCAC CACTGGCAGC AGCCACTGGT AATTGATTTA
1201  GAGGAGTTAG TCTTGAAGTC ATGCGCCGGT TAAGGCTAAA CTGAAAGGAC AAGTTTTGGT GACTGCGCTC CTCCAAGCCA GTTACCTCGG TTCAAAGAGT
1301  TGGTAGCTCA GAGAACCTTC GAAAAACCGC CCTGCAAGGC GGTTTTTTCG TTTTCAGAGC AAGAGATTAC GCGCAGACCA AAACGATCTC AAGAAGATCA
1401  TCTTATTAAT CAGATAAAAT ATTTCTAGAT TTCAGTGCAA TTTATCTCTT CAAATGTAGC ACCTGAAGTC AGCCCCATAC GATATAAGTT GTAATTCTCA
1501  TGTTTGACAG CTTATCATCG ATAAGCTTTA ATGCGGTAGT TTATCACAGT TAAATTGCTA ACGCAGTCAG GCACCGTGTA TGAAATCTAA CAATGCGCTC
1601  ATCGTCATCC TCGGCACCGT CACCCTGGAT GCTGTAGGCA TAGGCTTGGT TATGCCGGTA CTGCCGGGCC TCTTGCGGGA TATCGTCCAT TCCGACAGCA
1701  TCGCCAGTCA CTATGGCGTG CTGCTAGCGC TATATGCGTT GATGCAATTT CTATGCGCAC CCGTTCTCGG AGCACTGTCC GACCGCTTTG GCCCCCGCCC
1801  AGTCCTGCTC GCTTCGCTAC TTGGAGCCAC TATCGACTAC GCGATCATGG CGACCACACC CGTCCTGTGG ATCCTCTACG CCGGACGCAT CGTGGCCGGC
1901  ATCACCGGCG CCACAGGTGC GGTTGCTGGC GCCTATATCG CCGACATCAC CGATGGGGAA GATCGGGCTC GCCACTTCGG GCTCATGAGC GCTTGTTTCG
2001  GCGTGGGTAT GGTGGCAGGC CCCGTGGCCG GGGGACTGTT GGGCGCCATC TCCTTGCATG CACCATTCCT TGCGGCGGCG GTGCTCAACG GCCTCAACCT
2101  ACTACTGGGC TGCTTCCTAA TGCAGGAGTC GCATAAGGGA GAGCGTCGAC CGATGCCCTT GAGAGCCTTC AACCCAGTCA GCTCCTTCCG GTGGGCGCGG
2201  GGCATGACTA TCGTCGCCGC ACTTATGACT GTCTTCTTTA TCATGCAACT CGTAGGACAG GTGCCGCAG CGCTCTGGGT CATTTTCGGC GAGGACCGCT
2301  TTCGCTGGAG CGCGACGATG ATCGGCCTGT CGCTTGCGGT ATTCGGAATC TTGCACGCCC TCGCTCAAGC CTTCGTCACT GGTCCCGCCA CCAAACGTTT
2401  CGGCGAGAAG CAGGCCATTA TCGCCGGCAT GGCGGCCGAC GCGCTGGGCT ACGTCTTGCT GGCGTTCGCG ACGCGAGGCT GGATGGCCTT CCCCATTATG
2501  ATTCTTCTCG CTTCCGGCGG CATCGGGATG CCCGCGTTGC AGGCCATGCT GTCCAGGCAG GTAGATGACG ACCATCAGGG ACAGCTTCAA GGATCGCTCG
2601  CGGCTCTTAC CAGCCTAACT TCGATCACTG GACCGCTGAT CGTCACGGCG ATTTATGCCG CCTCGGCGAG CACATGGAAC GGGTTGGCAT GGATTGTAGG
2701  CGCCGCCCTA TACCTTGTCT GCCTCCCCGC GTTGCGTCGC GGTGCATGGA GCCGGGCCAC CTCGACCTGA ATGGAAGCCG GCGGCACCTC GCTAACGGAT
2801  TCACCACTCC AAGAATTGGA GCCAATCAAT TCTTGCGGAG AACTGTGAAT GCGCAAACCA ACCCTTGGCA GAACATATCC ATCGCGTCCG CCATCTCCAG
2901  CAGCCGCACG CGCCGCATCT CGGGCAGCGT TGGGTCCTGG CCACGGGTGC GCATGATCGT GCTCCTTGTC TTGAGGACCC GGCTAGGCTG GCGGGGTTGC
3001  CTTACTGGTT AGCAGAATGA ATCACCGATA CGCGAGCGAA CGTGAAGCGA CTGCTGCTGC AAAAGTTCTG CGACCTGAGC AACAACATGA ATGGTCTTCG
3101  GTTTCCGTGT TTCGTAAAGT CTGGAAACGC GGAAGTCCGC TACGTGCTGC TGAAGTTGCC CGCAACAGAG AGTGGAACCA ACCGGTGATA CCACGATACT
3201  ATCACGAGA GTCAACGCCA TGAGCGGCCT CATTTCTTAT TCTGAGTTAC AACAGTCCGC ACCGCTGTCC GGTAGCTCCT TCCGGTGGGC GCGGGGCATG
3301  ACTATCGTCG CCGCACTTAT GACTGTCTTC TTTATCATGC AACTCGTAGG ACAGGTGCCG GCAGCGCCCA ACAGTCCCCC GGCCACGGGG CCTGCCACCA
3401  TACCCACGCC GAAACAAGCG CCCTGCACCA TTATGTTCCG GATCTGCATC GCAGGATGCT GCTGGCTACC CTGTGGAACA CCTACATCTG TATTAACGAA
3501  GCGCTAACGG TTTTTATCAG GCTCTGGGAG GCAGAATAAA TGATCATATC GTCAATTATT ACCTCCACGG GGAGAGCCTG AGCAAACTGG CCTCAGGCAT
3601  TTGAGAAGCA CACGGTCACA CTGCTTCCGG TAGTCAATAA ACCGGTAAAC CAGCAATAGA CATAAGCGGC TATTTAACGA CCCTGCCCTG AACGAAGGAG
3701  CGGGTCGAAT TTGCTTTCGA ATTTCTGCCA TTCATCCGCT TATTATCACT TATTCAGGCG TAGCACCAGG CGTTTAAGGG CACCAATAAC TGCCTTAAAA
3801  AAATTACGCC CCGCCCTGCC ACTCATCGCA GTACTGTTGT AATTCATTAA GCATTCTGCC GACATGGAAG CCATCACAGA CGGCATGATG AACCTGAATC
3901  GCCAGCGGCA TCAGCACCTT GTCGCCTTGC GTATAATATT TGCCCATGGT GAAAACGGGG GCGAAGAAGT TGTCCATATT GGCCACGTTT AAATCAAAAC
4001  TGGTGAAACT CACCCAGGGA TTGGCTGAGA CGAAAAACAT ATTCTCAATA AACCCTTTAG GGAAATAGGC CAGGTTTTCA CCGTAACACG CCACATCTTG
4101  CGAATATATG TGTAGAAACT GCCGGAAATC GTCGTGGTAT TCACTCCAGA GCGATGAAAA CGTTTCAGTT TGCTCATGGA AAACGGTGTA ACAAGGGTGA
4201  ACACTATCCC ATATCACCAG CTCACCGTCT TTCATTGCCA TACG 4244
```

REFERENCES

(1) Chang, A.C.Y. *et al.* (1978) J. Bacteriol. 134, 1141-1156. (2) Bernardi, A. *et al.* (1984) Nucl. Acids Research 24, 9415-9426. (3) Alton, N.K., *et al.* (1979) Nature 282, 864-869. (4) Prentki, P., *et al.* (1981) Gene 14, 289-299. (5) Selzer, G., *et al.* (1983) Cell, 119-129. (6) Ohtsubo, H. *et al.* (1978) Pro. Natl. Acad. Sci. USA 75, 615-619.