

The complete nucleotide sequence of the *tdc* region of *Escherichia coli*

Herbert P. Schweizer and Prasanta Datta

Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109-0606, USA

Submitted March 10, 1989

EMBL accession no. X14430

The nucleotide sequence of a 6,294 bp *EcoRI* fragment of the *E. coli* chromosome encompassing the *tdc* operon was completed. The sequence reported here overlaps with previously determined sequences from nucleotides 1734 (formerly 1) to 6025 (formerly 4282)(1,2,3). This sequence contains three open reading frames: *tdcR*, bases 1670 (GTG) to 1374; and two reading frames of unknown function: *orfA*, bases 1115 (ATG) to 579 and a truncated *orfB*, bases 554 (ATG) to 1. All gene products have been identified in a minicell system (unpublished data).

```

GAATCAATT TAAAGGATT TAAATAGTA TTAGCCATAT GTTATCTGT GAAATCATAA TAATCTAAAT CGACATTACG TAAGTCCAGA 90
GAGSATAGAT CAACGATATC TCCAGTCCCA CAACATAAAA TTACGTTTAA CCTCCTGATT TCATTGTAAA TATCTTTTAA TGGCAATTA 180
ATATTGATAT TACCTGTGGA TATTTCTCGG TTTTTAACG GTTCATTAAT GGTAAAGTCA CAAGTATTGT CCGATCGTTG AATAATTTGG 270
CGTTATAAC CTAAAAAATC CAGACGAAGA GACGAGACAT TATAAAGAAT TACATTTTGT AAGTACGCTA ATTCTTTTGG AAAATCACCA 360
GTGATAATGC ATTTCCACC AGGTGTGACG GTGTTAAGTA AAGTGACCAT TTCAATAGAG GTCTGTTTTA GTTTTTCTAT ACCTGGTGTCT 450
AGTGGTTCGA GGTATTTCCT TGTGGGCTT TCAGGAAGGT CATTCAATTT TCGACGGACT AAATCACTAC TAATATCATT ACCAATTTGAT 540
GAAACAGGGA ACATTATTTT ATCCCTTATA AATAATCAAT AAAATCATT ATTAATAATG TCGGGTGA AAAGAGAAAT TTGCCAACAA 630
TCTGAAGTAA TGCTGGTTTT ATTAAGGTGC ACAAGGGGGG TTAAGATGGC TTCTTTTCGG GTGAATTTTG CACAATACT AGAATCGTC 720
TCCTTAATAC ATTCAGCTC TTTTTATTG ACGGCTGAGC AAGTACATTG TTCACAAAAA GGTGCAAAAG TGCTCATCAT TAATGTTGT 810
ATGATATAAT CATCATGTTC TCCTGTAGAA ATATCTATAA TGAATGGACT CATATTTTCC GGTGGGATTC GGATCATGAC AGAAGGAACA 900
TCGCCATGAT AGAAATTCAA TGCAGTATT GCAGATAATA CTCTCCGACG TTCATTACTT ATACTGTAAA TTTTATTAAC GTCATTGGTG 990
GAGACAAATT CTGCAAACTT TCTTTTTACC GTCCAATAAT CAATATTTCT ATTGTGATTG TTAACCACTG CGTGCATTGG CGGGATTGAA 1080
GGGTGAAAAA TATGCCAAT TGGGAAGCCT TTCATTGATG CCTGCCCTGT CATCATGACA AACAAATAA GATGCTTATT TTCTCGCAG 1170
GGTTCTTTCC CACTATCAAA AAACGCTCAT TTACCCGGCG TTACAGTAAA CGAAGCAACA TTCTGTGGCC TACGCCGATC AAAAGCTCTG 1260
TGAAAAATTC CTTTGTGTC TTTGTCTGAC CTTATCCGCG CSATCAGCAA TTATTCGTTT TGTGATCCTC AACACAAAAA GCTATTCGTC 1350
GCCTATGAGG GGAGAGAAT TTATGGAAAA CGATTGCATT CACAAAAATGA CGTCCGTGCA GAAGAGAAAA AGTTGACGTT AATCAATTTG 1440
TCGAGATGAT CAATACATGT ATTCGAAAT TTTAAATATT GATAAAATAA ATGTTTGTTC TCCAACAAAT AGCGTGA AAAATGTGACAT 1530
GCCGCAATAT TTACTACTGA AAATGGTTTA TTTATAATAA AATTTGASAA CTGCTCATT AAATTTAAT CAAATTTTGC CTGGTAATTA 1620
TCCGGTAATT GCTTGAATA TGGTCTCAGC CCCTTTTTTG TATTAACCAAC ATAACGATGA TGTATCACC ATAAAAATG GTTATCCCCG 1710
TCATTTTTTT TGACAAAAAT CAGGGTTTTAT GCTGATTTTT ATACTTTAAC TTGTTGATAT TTA AA 1775 ...6020 G TCGACTTTGG 6030
TCCGATGTC TGGTCCCGC CCAACCAAC CAGAGCTGG GTGACCTGGA ACGCGTAGTG AATAAAGAGT CGGGATTATT AGGATTTTCC 6120
GGTCATTCTT CGGATTTACG TGTCTGGGA AAAAGCCTGG CATGAAGGTC ACGAACCGCG GCAACTGSCA ATTAACCTT TTGTTCCACC 6210
AATTGCCCGT CATATTGCCG GACAGCGAGC TTCATTACGT CGCTGGATG GAATTATATT CACCGGCGGA ATAGGAGAGA ATTC 6294

```

**ACKNOWLEDGEMENTS**

This work was supported by NIH grant GM21436.

**REFERENCES**

- (1) Goss, T.J. & Datta, P. (1985) *Mol. Gen. Genet.* 201, 308-314.
- (2) Datta, P. et al. (1987) *Proc. Natl. Acad. Sci. USA* 84, 393-397.
- (3) Goss, T.J. et al. (1988) *J. Bacteriol.* 170, 5352-5359.