

Nucleotide sequence of the *dcm* locus of *Escherichia coli* K12Theodor Hanck<sup>§</sup>, Nicole Gerwin<sup>+</sup> and Hans-Joachim Fritz<sup>\*§</sup>

Max-Planck-Institut für Biochemie, Abteilung Zellbiologie Am Klopferspitz 18A, D-8033 Martinsried, FRG

Submitted June 20, 1989

EMBL accession no. X13330

The *dcm* locus (1,2) of *E. coli* K12 was isolated from phage  $\lambda$ 25D8 (3) as a 3.65 kb DNA fragment bordered by BamH I restriction sites. The nucleotide sequence of this fragment, determined by the chain termination method (4), is listed below. Open reading frames: *dcm*: 298-1713, *orf18*: 1697-2164, *orf20*: 1314-811, *orf32*: 3076-2159.

```

1  GATCGTTCG CACAGCATGC CCGCTTGTAT GATAAACGCT ACGCGCTGGA TCATTTCCAG ACTAAGTTGC TGAAGCTACC GCAAACCATG CAAMCCGCAA
101 GAGGCAAGCA GCTGGCCGAG CACAACGCCG ATTTTTTATG CGAGTTTATG CGGAAGCTCA GFGCCGAACT GCGCGGGGAG AATGAAGGTG TTGATCACAA
201 GGTGATAGAT GCGTTTTTCAT CCGCTTGAGCT AGCCGCTGGC CCTAAATGGC TGTAAATFATG TFAACCTGTC GGCCATCTCA GATGGCCGGT GAAATCTATG
301 CAGGAAAAATA TATCAGTAAC CGATTTCATC AGCACCCGGA ATGCCGCACA GGCAATGCTG GAGAAACTGC TGCAAAATTA TGATGTTAAA ACGTGGTGGT
401 CGCAGCTTAA TCGTGTAGTG GAGAATCACT GGAGCGCGGC AATTTTAAAA CGTGCCTGGC CGAATGACTC GGCATGGCAC CGTTTAAAGT AGAAAGAGTT
501 CGCCCATCTG CAAACGTTAT TACCCAACC ACCGGCACAT CATCCGCATT ATCGGTTTTG CTTTATCGAT CTATTCCCGG GAATTGGCGG CATCCGTCGC
601 GGTTTTGAAT CGATTGGCGG ACAGTGCCTG TTTACCAGCG AATGGAACAA ACATGCGGTA CGCACTTATA AAGCCAACCA TTATTGGCAT CCGCGGACCG
701 ATCATTTTAA TGAGATATC CGGCACATCA CCTCAGCCA TAAAGAAGGC GTGAGTGATG AGGCGCGCGC GGAACATATT CGTCAACACA TTCTGAAACA
801 CGATGTTTTA CTGGCCGGTT TCCCTTGTCA GCCATTTTCG CTGGCTGGC TATCGAAAAA GAACTCGCTC GGGCGGGGGC ACGGTTTTGC CTGGGATACC
901 CAGGCGACGC TGTTTTTTGA TGTGGTAGCC ATATACGAGC CGGCTCGTCC GCGCATGTTT GTGCTCGAAA ACGTCAAAAA CCTGAAAGT CACGACCCAG
1001 GTAACAGGTT CCGCATCATC ATGCAGACGC TGGACGAAT GGGCTATGAC GTGGCTGATG CAGAAGATA GATCCGAAAG TCATCGACCG
1101 CAAACATTTT CTGGCCGACG ACCGTGAAGC CATCGTGTG GTGGTTTTCT GTCCGATCT GAATCTGAAA GCGGATTTTA CCTCGCTGTA TATCAGCGAA
1201 TGTTTTCCCT GCGCAGGAGT GACGCTGGCC CAGCTGTGG ACCCGATGTT CGAGCGGAAA TATATCTGTA CCGCCGTGCT GTGGAAGTAC CTCTATCGAT
1301 ATCCGAAAAA ACATCAGGCG CCGGGTAAAG GCTTCGGTTA TGGAAATGTT TATCCGAAAC ATCCGCAAGG CGTCAACGGT ACCTGTCTG CCGGTTATTA
1401 CAAAGATGGC GGGAAATTT TAACTGATCG CGGCTGGGAT ATGGCCACGG GTGAGAAAGA CTTTGACCAT CCGCTGAATC AGCAACATCG TCCACGTCGG
1501 TTAACGCCCT GGAATGCGC CGCTTAAATG GGTTTTGAAG CGCCGGGAGA AGCGAAATTC CGTATTCGCG TTTCCGACAC TCAGGCCATC GCGCAGTTCC
1601 GTAACCTGTT GGTGCTGGCC GTCTTTGCGC CCGTGGCAA ACTGCTTGAG CCAAAAAATCA AACAGCCGTT GCGCTTGCCT CAGCAAGAGG CACAACATGG
1701 CCGACGTTCA CGATAAGGCC ACTCGCAGCA AAAATATGCG CCGCATGACC ACGGCTGATA CCGCGATAGA GAAGCCGCTC GCGCATCTGT TAACCGGGCA
1801 GGGCCTGGCA TTTCCGCTTC AGGACGCCAG TCTGCCCGGA CTCCCGGATT TTGTCGTTGA TGAATATCGC TGCCGTGAT TACCCATGTT CCGTCTCTGG
1901 CATCATCATC ACTGCTATCT GTTTAAAGTG CCTCGCAGTC GAACCGAGTT CTGGCTGGAG AAGATAGGTA AAAATTTGTA GCGCGATCGC CCGGATATCA
2001 GTGCTTTGCA GGAACCTGGC TGGCGGATAT TGATTGTCTG GGAGTGCCTG TTACGTGGCC CCGGAAAGCT CAGCGATGAA GCGCTTACCC AGCGTCTGGA
2101 AGAGTGGATC TGGCGCGAAG GTGGCCAGCT GACAGTCCAG ACCGAGGGGA TCTATTACT CGCTTATGTC GTCCCTGAATA ACTGAGCCAA CTCACGGGTT
2201 TGCCGGGAG AGATATTTT CCAAGCTGAC CAGTACCACC GCGAAGACAA TTACGCCGAG CCGCAGCCAG TCAATCTCTC ACAGTGTCTC TCCACCAGT
2301 CCGTACCCA CGAAGACCGC GACCCACGGG TTAACGTAAG CGTAGCTGTT GCGGAGAGCC GGAATGATGTA TTTCCGCAAG CCAATGGA CATAACCCACC
2401 CGATAATCGA ACCAAACAGC GCCAGTAGCC GACCCCGTAA AGGGAAGGGA GCGCCGTCAG TTTTTCACCC GCAATCATCG ACCCGATCAT
2501 TAACACCAGC CTTGCCGACA GCATCTCATC CAGCACCCCC ATCATCCCTA CAGGTAAGT AATCGCCGAG CCATAAATCT AGCCAAACCC CCACTAATC
2601 GAGCCGATTA AATCAGAAT CCGCCGCCAC GGATTGCCGC TFAAATTTCC ACCGCTATTG AGCATGATGA TTTCCGCAAG CCAATGGA CATAACCCACC
2701 ATTCAGTGT GCGCGTTTTA ATGCCAAACA GCGCGTGA GCACAGGGTA AAGAGGGGCA CCGTTGCAAC CACTACGGCC GCGATGCCGG AAGGAACATT
2801 TTGATGTTG GCAACCGTCA CATGCCATT ACCGACAGCC AGCAATAACA GGCCAATCAG CCGCGCATG AGCAGCGGAC GTAGCGGGGG GAGTTTGTGT
2901 CCGCCGAGTA GCAAAAAATG CAGCAATAAA ATACCGGCTC CCGAGGAATCG AACGCCCGCC ATCATTAAAC GAGGCCAGCT TTCCACGCCA ATCCGAATGA
3001 CAAAATAGT TGAGCCCAA ATGATATACA ACCGAAACAG CCGCCAAA AGCGGTAACA ACTGGCGGAA ACCGATAATC CCTCACGGTG GAAATAAAAA
3101 GGTGTTCAT AGTAAACGTT AAAATCATT TGTGTCGAG AGATATAATT GCACTTGATT GTTAAAAAAA TGTTGACCTG TGAAGCAAGT ATCAGAGCTG
3201 CCGTTTTTTC TTCATACTTA ACCCTTCAA CAATAAAAA GAGAGGGAAT GCTTTTTGCC GGGAGTAGTT TACTGACGTT CACTGATCAT ATCGCCACAC
3301 TACTGGACGA TATCTCCGTG ATGGGCAAAC TGGCGCGGAA GAAAACCCGC GGTGTATTAG GGGATGACTT ATCGTCAAT GCGCAACAG TTTACGGCTG
3401 GCGGGCCAA CCGGAACCTC CCGTGGCTG GCGCGTGGC AAAGGATCCG TGATTAATAA AGTGAATCTG GTGCCGCTGG CCGTGTATCT CAGTCCGTTT
3501 ATCCCGTGG CGATTACGGC TCTGTGATG ATTTGGTGGC CTTTTCTCT CTTTGAAGGA GTAGAGAAAT TGCTGCATAT GCTGGAGGCG GCTAAACATA
3601 AAGAGATCC GGGCCAGAGC CAGCAGCGTC TGGAGAACT GCGCGCCAG gatcc

```

\*To whom correspondence should be addressed

Present addresses: <sup>+</sup>Institut für Genetik und Mikrobiologie der Ludwig-Maximilians-Universität München, Maria-Ward-Str. 1a, D-8000 München 19 and <sup>§</sup>Institut für Molekulare Genetik der Georg-August-Universität, Grisebachstr. 8, D-3400 Göttingen, FRG

**References**

1. Marinus, M.G. (1973) *Mol. Gen. Genet.* **127**, 47-55.
2. Baghwat, A.S., Sohail, A. and Roberts, R.J. (1986) *J. Bacteriol.* **166**, 751-755.
3. Kohara, Y., Akiyama, K. and Isono, K. (1987) *Cell* **50**, 495-508.
4. Bankier, A.T. and Barrell, B.G. (1983) In Flavell, R.A. (ed.), *Techniques in the Life Sciences*, Elsevier Scientific Publishers Ireland Ltd., County Clare, Ireland, Vol. B5, *Nucleic Acid Biochemistry B* 508, pp. 1-34.