

Nucleotide sequence of the *Escherichia coli* uridine phosphorylase (*udp*) geneLeslie Walton, Cynthia A. Richards<sup>1</sup> and Lynn P. Elwell

Departments of Molecular Genetics and Microbiology and <sup>1</sup>Experimental Therapy, Burroughs Wellcome Co., Research Triangle Park, NC 27709, USA  
Submitted July 5, 1989

EMBL accession no. X15689

The uridine phosphorylase (*udp*) gene from *E. coli* has been localized on a recombinant plasmid, pVMK27 (1). Sequencing of this gene was performed by the Sanger chain termination method (2) on both strands of a 2,480 bp *Pst*I-*Hinc*II fragment derived from pVMK27 cloned into the phagemid cloning vector pBS+ (Stratagene, San Diego, CA). The sequence revealed an open reading frame (ORF) of 759 bases which codes for a polypeptide of 253 amino acids (Mr 27,162 Da). The identity of this ORF as the *udp* gene was confirmed by N-terminal amino acid analysis of purified uridine phosphorylase (UDP).

Fig.1 shows that distal to the coding region is a DNA sequence (underlined) which may form a hairpin loop (-11.8 kcal/mol) thereby causing transcription termination or pausing at this site. A putative ribosome binding site is shown (asterisks) at nucleotides 152-155 (Fig.1).

Previous reports on UDP from *E. coli* have suggested that this enzyme has 4-fold (3), 6-fold (4) or 8-fold (5) symmetry. These sequence data support the conclusion that UDP is a hexamer, composed of six identical subunits with an aggregate Mr of 162,972 Da.

```
TGCAGAAATGAAGGATTTATGTGATTTGCATCACTTTTGGTGGTAAATTTATGCAACGCATTTGGCTCATGGTGATG 80
AGTATCACGAAAAATGTTAAACCCCTTCGGTAAAGTGTCTTTTTCCTTCTCTGACTAAACCCGATTACAGAGGAGTGTCT 160
*****
ATATGTCCAAGTCTGATGTTTTTCATCTCGGCCCTCACTAAAAACGATTACAAAGGGCTACGCTTGCCATCGTCCCTGGC 240
M S K S D V F H L G L T K N D L Q G A T L A I V P G
GACCCGGATCGTGTGGAAAAGATCGCCGGCTGATGGATAAGCCGGTTAAGCTGGCATCTACCCCGAATTCCTACTCTG 320
D P D R V E K I A A L M D K P V K L A S H R E F T T W
CGGTGCAGACCTGGATGTAACCTGTTATCGTCTGCTTACCCGGTATCGCCGGCCGCTTACCTCTATTGCTGTGGAAG 400
R A E L D G K P V I V C S T G I G G P S T S I A V E
AGCTGGCAGACCTGGGCATTCGCACCTTCTCGGTATCGGTACAACGGCGCTATTCAGCCGATATTAATGTGGGTGAT 480
E L A Q L G I R T F L R I G T T G A I Q P H I N V G D
GTCTCGGTTACCCGGCGTCTGTCGGTCTGGATGGCGGAGCCCTGCACTTCGCACCCGTTGAATTCGCCGGCTGTCGCTGA 560
V L V T T A S V R L D G A S L H F A P L E F P A V A D
TTTCGAATGTACGACTGGCTGGTTGAAGCTGCGAAATCCATTGGCCGCAACTCACGTTGGCGTGACAGCTTCTTCTG 640
F E C T T A L V E A A K S I G A T T H V G V T A S S
ATACCTTCTACCCAGGTGAGAACGTTACGATACTTACTCTGTCGCGTAGTTCGTCACTTTAAAGGTTCTATGGAAGAG 720
D T F Y P G Q E R Y D T Y S G R V V R H F K G S M E E
TGCCAGCCGATGGCGTAAATGAACTGAAATGGAATCTGCAACCCCTGTGACCATGTGTGCAAGTCAGGCCCTGCGTC 800
W Q A M G V M N Y E M E S A T L L T M C A S Q G L R A
CGGTATGGTAGCGGGTGTATCGTTAACCACCCAGCAAGAGATCCCGAATGCTGAGACGATGAAACAAACCCGAAAGCC 880
G M V A G V I V N R T Q Q E I P N A E T M K Q T E S
ATCGCGTGAATAATCGTGGAAAGCGCCGCTGCTGCTGTAATTTCTTCTCTCTGCTGTAAGCCGACGGCTTCGCGCC 960
H A V K I V V E A A R R L L U
TTTGTATTTTTCGGTAGCGCCTCGCAGGAAATGCCCTTCCAACCTGGACGTTTGTACAGC 1020
```

**Acknowledgments:** We are grateful to Drs. Valerie Maples and Sidney Kushner for plasmid pVMK27 and to Dr. David Klapper for amino acid analysis.

**REFERENCES**

- Aldea, M., Maples, V. and Kushner, S. (1988) *J. Mol. Biol.* 200, 427-438.
- Sanger, F., Nicklen, S. and Coulson, R. (1977) *Proc. Natl. Acad. Sci. USA.* 74, 5463-5467.
- Vita, A. and Magni, G. (1983) *Anal. Biochem.* 133, 153-156.
- Cook, W. et al. (1987) *J. Biol. Chem.* 262, 2852-2853.
- Leer, J., Hammer, J., Jaspersen, K. and Schwartz, M. (1977) *Eur. J. Biochem.* 75, 217-224