The URK1 gene of *Saccharomyces cerevisiae* encoding uridine kinase

Laurence Kern

Laboratoire de Génétique Physiologique du CNRS, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 rue René Descartes, F-67084 Strasbourg, France

Uridine kinase (EC 2.7.1.48) catalyses the conversion of uridine into UMP in the pyrimidine salvage pathway of *Saccharomyces cerevisiae* (1, 2). The *URK1* gene was cloned by complementation of an uridine kinase deficient strain and is included in a 2.7 kb DNA segment. A single large ORF following the ATG codon at nucleotide 774 with a stop codon at nucleotide 2276 encodes a polypeptide chain of 501 amino acid residues. Two putative ORF are located on the complementary strand. ORF1, of at least 179 amino acids, initiates upstream from the *URK1* gene at nucleotide 537. ORF2, at least 111 amino acids, ends at nucleotide 2367. The screening of the EMBL Data Library did not allow the identification of these ORF.

ACKNOWLEDGEMENTS

REFERENCES

1. Grenson,M. (1969) *Eur. J. Biochem* **11**, 249–260.
2. Jund,R. (1973) Thesis, Université Louis Pasteur, Strasbourg, France.

```
AAGCTTTGCATCATCACCAATTTCTAATCTCCTTTTTCTCATCAATTCTCTCCTTTTGAGACGATATTGTTGTAGTACGTCGTCTTCATCATCACTTAGTTCTAGCTCTAAGCTATTTT  120
 L  K  A  D  D  G  I  E  L  R  R  K  R  M  L  E  R  R  R  K  L  R  Y  Q  Q  L  V  D  D  E  D  D  S  L  E  L  E  L  S  N  K
TGCTCTTTTATTACGAAATCCTCCATTTTTAATGTCATAAAGAATTTTGTTTATCATTTTGATATCCATTTCCTTATTTTCTGGGCAAGCATTTCTCTCGATTTCATCAGGATTGAAGTT  240
 A  R  K  N  R  F  G  G  N  K  I  D  Y  L  I  K  N  I  M  K  I  D  M  E  K  N  E  A  A  L  M  E  R  I  E  H  P  N  F  N
GTTCTTGGAATAATCATCGATCATTTTTTCTACGTCAGAATCATAGTCGTCAGATCCTTCTCCATCTGCACCACCGGATACCATGCCATTCATCCTCGGATTCTTCGGCTTCCATTTCAAA  360
 N  K  S  Y  D  D  I  M  K  E  V  D  S  D  Y  D  D  S  G  E  G  D  A  G  G  I  G  H  W  E  D  E  S  E  E  A  E  M  E  F
GAAGTTTGTGACACCTCTTTTCTTCAACTCTTTTATCTTGGCCTCGTGCTCCTTTTCTTTCTGTTTTAGTTGCAATTTCTCTTTGTCAATCAGCCGACGACGGCGAATTGCTTCATCAGCTTC  480
 F  N  T  V  G  R  K  K  L  E  K  I  K  A  E  H  E  K  E  K  Q  K  L  Q  L  K  E  K  D  I  L  Q  R  R  I  A  E  D  A  E
TTCATTGACGTATACATCATCTTTGACTTCATTTTCGATTGTATTTCTTTCGGCCATCACCTCTTCAATGTTTTGTGTATCTAGAACATTGGTTTCTTGAAATTTGTCGGTATCCTCTCC  600
 E  N  V  Y  Y  V  D  D  K  V  E  N  E  I  T  N  R  E  A  M
GATATTATCACCGTAATGGCCCAGATTAATGGCGATCGATAAGATAAGAGATGTGATTTTTTTTTTTTTTTTCGTAAGGCTGTAGGCCTTGGATAACGTCTTATTTACTATTTATATGCGAC  720
ATCTTCGATAATTTCATACGTTTAAATTTCGAACTCGCATTTATTTTATTTATTATGTCCCATCGTATAGCACCTTCCAAAGAACGATCTTCATCATTTATTTCAATTTTAGACGATGAAA  840
                                   M  S  H  R  I  A  P  S  K  E  R  S  S  S  F  I  S  I  L  D  D  E  T   23
CAAGAGACACATTGAAAGCTAATGCAGTCATGGATGGTGAAGTAGATGTCAAAAAAACAAAAGGAAAAAGCTCTCGGTATATCCCACCATGGACAACTCCATATATAATAGGTATAGGTG  960
 R  D  T  L  K  A  N  A  V  M  D  G  E  V  D  V  K  K  T  K  G  K  S  S  R  Y  I  P  P  W  T  T  P  Y  I  I  G  I  G  G   63
GTGCTTCAGGTTCAGGCAAGACAAGACGTTGCTGCTAAGATTGTGTCGTCAGTACTTCCCTGGACAGTATTAATATCTTTGGATAATCTTTGCAATCCATTAGGCCCAGAGGACAGAG  1080
 A  S  G  S  K  T  S  V  A  A  K  I  V  S  S  I  N  V  P  W  T  V  L  I  S  L  D  N  F  Y  N  P  L  G  P  E  D  R  A   103
CCAGAGCCTTTAAAAATGAATACGATTTCGACGAGCCAAATGCCATCAACTTAGATTTGGCATATAAGTGCATTTTGAACTTAAAGGAGGGCAAAAGGACAAATATCCCAGTTTATAGCT  1200
 R  A  F  K  N  E  Y  D  F  D  E  P  N  A  I  N  L  D  L  A  Y  K  C  I  L  N  L  K  E  G  K  R  T  N  I  P  V  Y  S  F   143
TCGTCCACCACAATAGAGTTCCTGATAAAAATATAGTCATATACGGGGCCAGTGTGGTAGTTATCGAAGGGATCTACGCCCTTTACGATCGCCGATTGCTGGATTTGATGGACTTGAAA  1320
 V  H  H  N  R  V  P  D  K  N  I  V  I  Y  G  A  S  V  V  V  I  E  G  I  Y  A  L  Y  D  R  R  R  L  L  D  L  M  D  L  K  I   183
TTTATGTTGACGCTGATTTGGATGTCTGCTTAGCAAGAAGATTGTCGAGAGATATAGTTTCCAGAGGGAGAGATTTGGATGGTTGTATTCAACAATGGGAGAAATTTGTGAAACCAAATG  1440
 Y  V  D  A  D  L  D  V  C  L  A  R  R  L  S  R  D  I  V  S  R  G  R  D  L  D  G  C  I  Q  Q  W  E  K  F  V  K  P  N  A   223
CGGTAAAGTTTGTGAAACCAACAATGAAGAATGCAGATGCTATCATTCCATCGATGAGTGATAATGCTACAGCGGTAAATTTAATCATTAACCACATCAAGTCAAACTGGAACTAAAAT  1560
 V  K  F  V  K  P  T  M  K  N  A  D  A  I  I  P  S  M  S  D  N  A  T  A  V  N  L  I  I  N  H  I  K  S  K  L  E  L  K  S   263
CAAATGAACACTTAAGAGAGCTAATCAAATTGGGCTCTTCTCCCTTCACAAGATGTGCTTAATCGTAACATAATTCATGAATTGCCGCCCACCAACCAAGTTCTTTCGCTGCATACTATGC  1680
 N  E  H  L  R  E  L  I  K  L  G  S  S  P  S  Q  D  V  L  N  R  N  I  I  H  E  L  P  P  T  N  Q  V  L  S  L  H  T  M  L   303
TTCTAAATAAAAATCTAAATTGCGGGGACTTTGTTTTCTACTTTGACAGGTTAGCAACAATTTTGTTATCATGGGCACTTGATGACATTCCTGTAGCACATACGAACATAATTACACCTG  1800
 L  N  K  N  L  N  C  A  D  F  V  F  Y  F  D  R  L  A  T  I  L  L  S  W  A  L  D  D  I  P  V  A  H  T  N  I  I  T  P  G   343
GTGAGCATACCATGGAAAACGTTATTGCCTGTCAATTCGATCAAGTTACAGCCTGTTAATATTATTCGATCTGGCGATTGTTTTATGAAGTCTTTGAGAAAGACGATCCCCAATATCACAA  1920
 E  H  T  M  E  N  V  I  A  C  Q  F  D  Q  V  T  A  V  N  I  I  R  S  G  D  C  F  M  K  S  L  R  K  T  I  P  N  I  T  I   383
TTGGTAAATTGTTGATTCAGTCCGATTCACAAACTGGGGAAACCTCAACTGCATTGCGAATTTTTACCCCCAAATATAGAAAAGTTTGGCAAGGTTTTCTTAATGGAAGGTCAAATCATAA  2040
 G  K  L  L  I  Q  S  D  S  Q  T  G  E  P  Q  L  H  C  E  F  L  P  P  N  I  E  K  F  G  K  V  F  L  M  E  G  Q  I  I  S   423
GTGGTGGGCCATGATCATGGCCATCCAGGTGCTTTTAGATCATGGTATTGATTTGGAAAAGATTAGCGGTGGTGGTTTTATTTGGCCCACTGAAGTTGGTTAGTCACGTATATTAAATGCAT  2160
 G  A  A  M  I  A  I  Q  V  L  L  D  H  G  I  D  L  E  K  I  S  V  V  V  Y  L  A  T  E  V  G  I  R  R  I  L  N  A  F   463
TTGATAACAAAGTCAACATTTTTGCTGGTATGATCATCTCCAGAGAAAAGTTACAAAATCATCAATACAAATGGCATTGACCAGATTTTTTGATTCAAAGTATTTTGGTTGTGATTGAG  2280
 D  N  K  V  N  I  F  A  G  M  I  I  S  R  E  K  L  Q  N  H  Q  Y  K  W  A  L  T  R  F  F  D  S  K  Y  F  G  C  D  *   501
AATAATTAGAGGCAATATAAAGTAAAATTAAATAATAGTGCACGTATACATACTATATTAGTAACCTTTCATTCTCTCTATGGATAATCTTAAAAATCCATTACTTTCAATATGCCATAA  2400
                                                                    *  F  G  M  V  K  L  I  G  Y
CCTACTGTAACAATAGCGGCATAACTTAACAAACTGGCTGGTACAACCTCTCGTAATAAAAGTACCCAGGGTCAAATATCTATCACCAACTTCATCAATCATTGAAATGGCAGTCACGTTG  2520
 G  V  T  V  I  A  A  Y  S  L  L  S  A  P  V  G  R  T  I  F  T  G  V  T  L  Y  R  D  G  V  E  D  I  M  S  I  A  T  V  N
GGGAAACCAGAGGTTGGAAGACCCATAGCACTAGAGCATAATAGGCGGGCTATAACGATCAATAGTCTAGAATGATCACCTGAAGGTAGATTGGAACGATCTCACTCATTAATGGTACG  2640
 P  F  G  S  T  P  L  G  M  A  S  S  C  L  L  A  A  I  V  I  L  L  R  S  H  D  G  S  P  L  N  S  G  I  E  S  M  L  P  V
ATGATCATTGCCGCGACGGTATGGGAGACAAAAGTGGCCATTACAAGAATCACTAGACCGAAGATCC  2707
 I  I  M  A  A  V  T  H  S  V  F  T  A  M  V  L  I  V  L  G  F  I
```