JOURNAL OF VIROLOGY, Sept. 1986, p. 759–760
0022-538X/86/090759-02$02.00/0

Vol. 59, No. 3

Nucleotide Sequence and Protein Overproduction of Bacteriophage T4 Thioredoxin

DAVID M. LEMASTER
Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511

Received 18 February 1986/Accepted 10 June 1986

The bacteriophage T4 thioredoxin gene was cloned and physically mapped to 47.6 kilobases from the reference *Bam*HI site. The DNA sequence is consistent with that reported from earlier protein sequence studies. The gene was subcloned into a lambda p_L overexpression vector which allowed for the isolation of approximately 5 mg/liter.

The bacteriophage T4 *nrdC* gene encodes a small dithiol-disulfide couple redox protein, thioredoxin. It has been mapped between genes 49 and *rI* (15), which is the most poorly characterized region of the T4 genome. Although numerous in vitro transcripts have been found in this region, recent genetic-physical correlations have not located any genes in the 15-kilobase (kb) stretch between genes 49 and *tk* (3). This note describes the mapping, cloning, and sequence of the *nrdC* gene as well as a closely adjacent open reading frame.

Additional interest in T4 thioredoxin arises from the considerable similarity between its tertiary structure and that of *Escherichia coli* thioredoxin (2). This is in marked contrast to the virtual absence of sequence homology (12). As a result, the two proteins are believed to offer a useful comparative system for the study of protein dynamics. *E. coli* thioredoxin has already proven suitable for such studies, as evidenced by nuclear magnetic resonance (4, 7), circular dichroism (9), fluorescence (6, 14), and cleavage and reconstitution (5, 13) experiments. Unfortunately, the solution properties of the T4 thioredoxin are far less well known. This is primarily due to the difficulties involved in purifying adequate amounts of the material from its natural source. To provide larger amounts of the protein as well as render labeling and mutagenesis experiments more tractable, we have subcloned the T4 thioredoxin gene into a lambda p_L induction vector for overproduction.

The *nrdC* gene was mapped to the *Pst*I-*Xba*I fragment extending from 45.7 to 49.1 kb by the Southern blot technique with a ^{32}P-labeled synthetic oligonucleotide (A-A-PUR-T-G-PYR-G-T-PUR-T-A-PYR-T-G-PYR-G-A), which corresponds to the active-site protein sequence, with degenerate mixtures added for the third codon positions. The 5.3-kb *Xba*I fragment covering this region was preparatively isolated by gel electrophoresis and cut with *Pst*I and *Kpn*I. The gene was found to reside on the 1.9-kb *Kpn*I-*Xba*I fragment which was cloned into the plasmid pUC19. Restriction analysis of this clone indicated that the gene was located near the *Kpn*I terminus. A 490-base-pair (bp) *Hae*III fragment extending from the 5′ flanking region of the gene to the *Hae*III site adjacent to the polylinker sequence was cloned in both orientations into the M13 derivative mp9 and sequenced by the dideoxynucleotide method. The 620-bp *Rsa*I fragment covering the 5′ flanking region was partially sequenced as well. The DNA sequence of the thioredoxin gene and the flanking regions is given in Fig. 1. The predicted protein sequence identically matches the sequence previously reported (12). The T-A-A sequence which is only one base pair away from the initiator codon serves to terminate an open reading frame, at minimum, 300 bp long. Unfortunately, the paucity of genetic data for this region of the T4 map makes it impossible to suggest an identity for the gene product which may be encoded. Furthermore, a survey of the National Biomedical Research Foundation protein data bank failed to reveal any strongly homologous sequences. There is also a short 105-bp open reading frame starting at the 3′ terminus of the thioredoxin coding region.

The 490-bp *Hae*III fragment was cloned into the *Hpa*I site of pKC30 (10) to yield pDL51. The plasmid pKC30 carries the lambda p_L promoter upstream from the *Hpa*I site. When pDL51 is transformed into M5219 (8), a strain carrying only part of the lambda functions, including the temperature-sensitive repressor C_{I857}, T4 thioredoxin was produced as roughly 5% of the total protein. This higher level of synthesis facilitated simplification of the previously published purification procedure (1, 11).

A culture was grown in rich medium to mid-log phase at 35°C, shifted to 43°C, and grown for 3 h. The cells were ruptured by sonication in 50 mM sodium phosphate (pH 7.0) containing 5 mM 2-mercaptoethanol. Streptomycin sulfate and ammonium sulfate precipitations were performed as described earlier (1) by using narrower (35 to 70%) ammonium sulfate fractionation. The $(NH_4)_2SO_4$ pellet from 20 to 25 g (wet weight) of cells was suspended in 50 mM ammonium acetate (pH 6.1) and fractionated on an Ultrogel AcA54

```
     gttcatataaaaata tgctatcagaagtaa gtgatattatataca agcataaagaacaaa
-119 ----------+---- ----+----------+ ----------+---- ----+---------+  -60

     cgggccaaagtcttg attacgacgagactt ttaaacaagtactag gaattaaggaataat
 -59 ----------+---- ----+----------+ ----------+---- ----+---------+    0

     atgtttaaagtatat ggttatgatagcaac attcataaatgtgtg tattgcgataatgca
   1 ----------+---- ----+----------+ ----------+---- ----+---------+   60
     MetPheLysValTyr GlyTyrAspSerAsn IleHisLysCysVal TyrCysAspAsnAla

     aaacgtcttttgact gtgaagaaacagccg tttgaatttatcaac attatgccggaaaaa
  61 ----------+---- ----+----------+ ----------+---- ----+---------+  120
     LysArgLeuLeuThr ValLysLysGlnPro PheGluPheIleAsn IleMetProGluLys

     ggtgtttttgatgat gagaaaattgctgag cttctgactaaacta ggtcgtgatactcaa
 121 ----------+---- ----+----------+ ----------+---- ----+---------+  180
     GlyValPheAspAsp GluLysIleAlaGlu LeuLeuThrLysLeu GlyArgAspThrGln

     atcggcttaacaatg cctcaggtatttgct cctgatggaagtcat attggtggatttgac
 181 ----------+---- ----+----------+ ----------+---- ----+---------+  240
     IleGlyLeuThrMet ProGlnValPheAla ProAspGlySerHis IleGlyGlyPheAsp

     caattgcgggaatac tttaaatgatagaat taaatgaacaaatta tttttctaggcgatg
 241 ----------+---- ----+----------+ ----------+---- ----+---------+  300
     GlnLeuArgGluTyr PheLys

     gaacagaaggcgacc tggaatataaactct atgaatacatgattt ggttactaaagcaga
 301 ----------+---- ----+----------+ ----------+---- ----+---------+  360

     aggtatagactttgt tgtgtctaacccgta tggcgagaacacagt tgtaatcggtggtac
 361 ----------+---- ----+----------+ ----------+---- ----+---------+  420
```

FIG. 1. Nucleotide sequence of bacteriophage T4 *nrdC* gene. The predicted protein sequence of T4 thioredoxin is given below the nucleotide sequence.

column (2.5 by 90 cm) equilibrated in the same buffer. The pooled fractions were then loaded onto a Biorex 70 column equilibrated in the acetate buffer, and the T4 protein was eluted with a gradient to 0.8 M ammonium acetate. After dialysis and lyophilization, approximately 25 mg of T4 thioredoxin was recovered, which appeared homogeneous by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

This work was supported by project grant PCM8305203 from the National Science Foundation and Public Health Service program grant GM22778 from the National Institutes of Health.

I thank F. M. Richards and T. Platt for the use of their laboratory facilities and for their advice during this work. I also thank T. Lin and E. Spicer for providing the phage DNA and N. Grindley for offering the use of his Biosearch SAM-1 oligonucleotide synthesizer. The assistance of M. Bannon and J. Mouning in preparing the manuscript is also appreciated.

LITERATURE CITED

1. **Berglund, O., and B. M. Sjöberg.** 1970. A thioredoxin induced by bacteriophage T4. J. Biol. Chem. **245**:6030–6035.
2. **Eklund, H., C. Cambillau, B. M. Sjöberg, A. Holmgren, H. Jörnvall, J. O. Höög, and C. I. Bränden.** 1984. Conformational and functional similarities between glutaredoxin and thioredoxins. EMBO J., **3**:1443–1449.
3. **Gram, H., H. D. Liebig, A. Hack, E. Niggemann, and W. Rüger.** 1984. A physical map of bacteriophage T4 including the positions of strong promoters and terminators recognized in vitro. Mol. Gen. Genet. **194**:232–240.
4. **Holmgren, A., and G. Roberts.** 1976. Nuclear magnetic resonance studies of redox-induced conformational changes in thioredoxin from *Escherichia coli*. FEBS Lett. **71**:261–265.
5. **Holmgren, A., and I. Slaby.** 1979. Thioredoxin-C′: mechanism of noncovalent complementation and reaction of the refolded complex and the active site containing fragment with thioredoxin reductase. Biochemistry **18**:5591–5599.
6. **Kelley, R. F., and E. Stellwagen.** 1984. Conformational transitions of thioredoxin in guanidine hydrochloride. Biochemistry **23**:5095–5102.
7. **LeMaster, D. M., and F. M. Richards.** 1985. ^{1}H-^{15}N Heteronuclear NMR studies of *Escherichia coli* thioredoxin in samples isotopically labeled by residue type. Biochemistry **24**:7263–7268.
8. **Remaut, E., P. Stanssens, and W. Fiers.** 1981. Plasmid vectors for high-efficiency expression controlled by the p_L promoter of coliphage lambda. Gene **15**:81–93.
9. **Reutimann, H., B. Straub, P. L. Luisi, and A. Holmgren.** 1981. A conformational study of thioredoxin and its tryptic fragments. J. Biol. Chem. **256**:6796–6803.
10. **Rosenberg, M., Y. S. Ho, and A. Shatzman.** 1983. The use of pKC30 and its derivatives for controlled expression of genes. Methods Enzymol. **101**:123–138.
11. **Sjöberg, B. M.** 1972. Studies on the structure of T4 thioredoxin. J. Biol. Chem. **247**:8058–8062.
12. **Sjöberg, B. M., and A. Holmgren.** 1972. Studies on the structure of T4 thioredoxin. J. Biol. Chem. **247**:8063–8068.
13. **Slaby, I., and A. Holmgren.** 1979. Structure and enzymatic functions of thioredoxin refolded by complementation of two tryptic peptide fragments. Biochemistry **18**:5584–5591.
14. **Stryer, L., A. Holmgren, and P. Reichard.** 1967. Thioredoxin. A localized conformational change accompanying reduction of the protein to the sulfhydryl form. Biochemistry **6**:1016–1020.
15. **Yeh, Y. C., and I. Tessman.** 1972. Control of pyrimidine biosynthesis by phage T4. Virology **47**:767–772.