DOE/PC/79912--12

DE91 000851

THIOPHENE METABOLISM BY E. COLI

Contract DE-FG22-87 PC79912

David P. Clark
Department of Microbiology
Southern Illinois University
Carbondale, IL 62901

Technical Progress Report
June 16, 1990-September 15, 1990

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

OCT 16 1990

EXECUTIVE SUMMARY

The thdF gene of E. coli, as cloned on the plasmid pKA10 has been sequenced. The sequence has been checked and edited to give an updated version. The thdF open reading frame has been translated by a computer program. The thdF gene lies near to the tnaA gene, encoding tryptophanase. We found that our thiophene degrading mutants of E. coli degraded naphthalene sulfonic acids and were poisoned by indole. One interpretation of these results is that the thd system, including thdF has reactivated an obsolete metabolic pathway for degradation of indole and related bicyclic ring compounds.

We have constructed improved strains of E. coli for use as hosts in cloning genes which allow use of organic sulfur compounds as sulfur source. These strains carry two defects in sulfur metabolism and hence, reversion is negligible. We will use these in further attempts to clone genes from Acinetobacter which confer the ability to use sulfones as sulfur source.

BACKGROUND

The objective of this project is to investigate the mechanism of degradation of sulfur containing heterocyclic molecules by mutants of Escherichia coli K-12. We previously isolated multiple mutants of E. coli which were selected for improved oxidation of furan and thiophene derivatives. We have focussed on the thdA mutation. In our subsequent research as it appears to be of central importance in thiophene oxidation. We hope that analysis of the thd genes of E. coli will lead to improvement of our thiophene metabolizing bacterial strains.

PUBLICATIONS

The manuscript submitted by Martha J. Juhl and David P. Clark has been accepted for publication in Applied and Environmental Microbiology. The proofs are attached.

RESULTS

Sequence of thdF

In our previous report we presented the DNA sequence of the 3.8 kb insert of plasmid pKA10, which contains the thdF gene. Since then we have gone over our sequencing gels and also proof read the computer file carrying this sequence. The new improved version of the sequence is attached. The number of gaps and inconsistencies in the reading frames has been substantially decreased and the thdF open reading frame is much clearer. The new printout also supplies the amino acid sequence for significant open reading frames. The thdF gene is in reading frame 4, running from approximate coordinate 2500 to 1300. [Reading frame 4 is backwards relative to the sequence as printed]. The N-terminal end of tryptophanase may be seen in reading frame 3 from 3400 onwards. As remarked before, the tna gene encoding tryptophanase has been previously sequenced and serves to locate our DNA segment to 84 min on the E. coli chromosome.

Since tryptophanase converts tryptophan to ammonia, pyruvate and indole we proposed that the original function of thdF might be the further degradation of indole derivatives - perhaps by an indole oxygenase reaction. We therefore tested our series of thd mutants for oxidation of indole and several derivatives. We included the

naphthalene - sulfonic acids (NSA) because it has been shown that naphthalene dioxygenase also oxidizes indole to indigo. As Table 1 shows, NAR11 (thdA) and NAR30 (thdACD) oxidized both 1-NSA and 2-NSA significantly, although little effect was seen with indole-2-carboxylate (I2C) or 5-methoxy-indole-2-carboxylate (MICA). In the case of indole itself, NAR11 and NAR30 were killed whereas the parent DC625, grew on the indicator medium but did not significantly oxidize the indole. One interpretation is that NAR11 and NAR30 both metabolize indole to a toxic intermediate and kill themselves. This theory needs further substantiation.

Sulfone Degradation System of Acinetobacter

We have isolated a strain of Acinetobacter calcoaceticus which is able to use a wide range of sulfones, as well as sulfonic acids, as sulfur source. These compounds are not used as carbon source.

We have attempted to clone this ability into restriction negative E. coli strains carrying a cysA mutation and which therefore cannot take up inorganic sulfate. Our first attempts used as vector the high copy number plasmid pJOE810. Since these initial attempts failed, I will not provide a lot of detail. However, we did succeed in cloning fragments of Acinetobacter DNA into our E. coli host strain. Thus the cloning procedure worked. However we did not succeed in cloning the genes we wanted. We believe the problem is that overproduction of sulfur metabolism genes on a high copy vector is deleterious. We have therefore obtained a medium copy cloning vector, based on the pSC101 replicon. We have also improved our E. coli recipient by constructing a restriction negative strain carrying both cysA and cysN::Kan defects. This greatly improves sulfur source selections and largely abolishes the appearance of cysA⁺ revertants. We will attempt a new series of cloning experiments using these improvements.

TABLE 1. Response To Indole and Related Compounds

TTCplus	DC625	NAR11	NAR20	NAR30
blank	W	P	P	P
FCA	W	R	P	R
TCA	W	R	P	RR
I2C	W	P	P	P
MICA	W	P	P	P
1-NSA	W	R	P	P
2-NSA	W	R	P	P
Indole	W	dead	P	dead(P)*

* The few survivors are pink.

Sequence Range: 1 to 3780

```

      10      20      30      40      50
      *      *      *      *      *
CTG CAG GTG CCC GAT GAC GTA TAC CGA CGC GCA GGC AAC ACG TTT ACC AAA
GAC GTC CAC GGG CTA CTG CAT ATG GCT GCG CGT CCG TTG TGC AAA TGG TTT
<Gln Leu His Gly Ile Val Tyr Val Ser Ala Cys Ala Val Arg Lys Gly Phe
<  a  a  a  a  TRANSLATION OF SEQUENCE/KISWAR  a  a  a  a

```

```

      60      70      80      90     100
      *      *      *      *      *
ACG TTT GTC CTG AAA CGT GGT GAT TAC GCT GTC AAC GTC AAC TAC AAC GTG
TGC AAA CAG GAC TTT GCA CCA CTA ATG CGA CAG TTG CAG TTG ATG TTG CAC
<Arg Lys Asp Gln Phe Thr Thr Ile Val Ser Asp Val Val Val Val His
<  a  a  a  a  TRANSLATION OF SEQUENCE/KISWAR  a  a  a  a

```

```

     110     120     130     140     150
      *      *      *      *      *
CAG AAC GCT GGC GAG AAA CCG CTG GAA ATC TCC TCG TTT GGT CAG TTG AAG
GTC TTG CGA CCG CTC TTT GGC GAC CTT TAG AGG AGC AAA CCA GTC AAC TTC
<Leu Val Ser Ala Leu Phe Arg Gln Phe Asp Gly Arg Lys Thr Leu Gln Leu
<  a  a  a  a  TRANSLATION OF SEQUENCE/KISWAR  a  a  a  a

```

```

     160     170     180     190     200
      *      *      *      *      *
CAA TCC ATC ACT CTG CCA CCG CAT CTC GAT ACC GGA AGC AGC AAC TTC GCA
GTT AGG TAG TGA GAC GGT GGC GTA GAG CTA TGG CCT TCG TCG TTG AAG CGT
<Leu Gly Asp Ser Gln Trp Arg Met Glu Ile Gly Ser Ala Ala Val Glu Cys
<  a  a  a  a  TRANSLATION OF SEQUENCE/KISWAR  a  a  a  a

```

```

     210     220     230     240     250
      *      *      *      *      *
CTG CAC ACC TTC CGT GGC GCG GCG TAC TCC ACG CCT GAC GAG AAG TAT GAG
GAC GTG TGG AAG GCA CCG CGC CGC ATG AGG TGC GGA CTG CTC TTC ATA CTC
<Gln Val Gly Glu Thr Ala Arg Arg Val Gly Arg Arg Val Leu Leu Ile Leu
<  a  a  a  a  TRANSLATION OF SEQUENCE/KISWAR  a  a  a  a

```

```

     260     270     280     290     300
      *      *      *      *      *
AAA TAC AAG TTC GAT ACC ATT GCC GAT AAC GAA AAC CTG AAC ATC TCT TCG
TTT ATG TTC AAG CTA TGG TAA CGG CTA TTG CTT TTG GAC TTG TAG AGA AGC
<Phe Val Leu Glu Ile Gly Asn Gly Ile Val Phe Val Gln Val Asp Arg Arg
<  a  a  a  a  TRANSLATION OF SEQUENCE/KISWAR  a  a  a  a

```

```

     310     320     330     340     350
      *      *      *      *      *
AAA GGT GGT TGG GTG GCG ATG CTG CAA CAG TAT TTC GCG ACG GCG TGG ATC
TTT CCA CCA ACC CAC CGC TAC GAC GTT GTC ATA AAG CGC TGC CGC ACC TAG
<Phe Thr Thr Pro His Arg His Gln Leu Leu Ile Glu Arg Arg Arg Pro Asp
<  a  a  a  a  TRANSLATION OF SEQUENCE/KISWAR  a  a  a  a

```

```

     360     370     380     390     400
      *      *      *      *      *
CCG CAT AAC GAC GGT ACC AAC AAC TTC TAT ACC GCT AAT CTG GGT AAC GGC
GGC GTA TTG CTG CCA TGG TTG TTG AAG ATA TGG CGA TTA GAC CCA TTG CCG

```

<Arg Met Val Val Thr Gly Val Val Glu Ile Gly Ser Ile Gln Thr Val Ala
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

410 420 430 440 450
* * * * *
ATC GGC CTA TCG GCT ATA AAT CTC AGC CGG TAC TGG TTC AGC CTG GTC AGA
TAG CCG GAT AGC CGA TAT TTA GAG TCG GCC ATG ACC AAG TCG GAC CAG TCT
<Asp Ala End Arg Ser Tyr Ile Glu Ala Pro Val Pro Glu Ala Gln Asp Ser
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

460 470 480 490 500 510
* * * * *
CTG GCG CGA TGA ACA GCA CCC TGT GGG TTG GCC CGG AAA TCC AGG TCA AAA
GAC CGC GCT ACT TGT CGT GGG ACA CCC AAC CGG GCC TTT AGG TCC AGT TTT
<Gln Arg Ser Ser Cys Cys Gly Thr Pro Gln Gly Pro Phe Gly Pro End Phe
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

520 530 540 550 560
* * * * *
TGG CAG CTG TTG CTC CGC ACC TGG ATC TGA CCG TTG ATT ACG GTT GGT TGT
ACC GTC GAC AAC GAG GCG TGG ACC TAG ACT GGC AAC TAA TGC CAA CCA ACA
<Pro Leu Gln Gln Glu Ala Gly Pro Asp Ser Arg Gln Asn Arg Asn Thr Thr
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

570 580 590 600 610
* * * * *
GGT TCA TCT CTC AGC CGC TGT TCA AAC TGC TGA AAT GGA TCC ATA GCT TTG
CCA AGT AGA GAG TCG GCG ACA AGT TTG ACG ACT TTA CCT AGG TAT CGA AAC
<Thr End Arg Glu Ala Ala Thr End Val Ala Ser Ile Ser Gly Tyr Ser Gln
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

620 630 640 650 660
* * * * *
TGG GTA ACT GGG GCT TCT CCA TTA TCA TCA TCA CCT TTA TCG TTC GTG GCA
ACC CAT TGA CCC CGA AGA GGT AAT AGT AGT AGT GGA AAT AGC AAG CAC CGT
<Pro Tyr Ser Pro Ser Arg Trp End End End End Arg End Arg Glu His Cys
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

670 680 690 700 710
* * * * *
TCA TGT ACC CGC TGA CCA AAG CGC AGT ACA CCT CCA TGG CGA AGA TGC GTA
AGT ACA TGG GCG ACT GGT TTC GCG TCA TGT GGA GGT ACC GCT TCT ACG CAT
<End Thr Gly Ala Ser Trp Leu Ala Thr Cys Arg Trp Pro Ser Ser Ala Tyr
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

720 730 740 750 760
* * * * *
TGT TGC AGC CGA AGA TTA GGC AAT GCG TGA GCG TCT GGG CGA TGA CAA ACA
ACA ACG TCG GCT TCT AAT CCG TTA CGC ACT CGC AGA CCC GCT ACT GTT TGT
<Thr Ala Ala Ser Ser End Ala Ile Arg Ser Arg Arg Pro Ser Ser Leu Cys
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

770 780 790 800 810
* * * * *

GCG TAT CAG CCA GGA AAT GAT GGC GCT GTA CAA AGC TGA GAA CCC GCT GGG
 CGC ATA GTC GGT CCT TTA CTA CCG CGA CAT GTT TCG ACT CTT GGG CGA CCC
 <Arg Ile Leu Trp Ser Ile Ile Ala Ser Tyr Leu Ala Ser Phe Gly Ser Pro
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

820 830 840 850 860
 * * * * *
 CGC TGC TTC CCG CTG CTG ATC CAG ATG CCA ATC TTC CTG GCG TTG TAC TAC
 GCG ACG AAG GGC GAC GAC TAG GTC TAC GGT TAG AAG GAC CGC AAC ATG ATG
 <Ala Ala Glu Arg Gln Gln Asp Leu His Trp Asp Glu Gln Arg Gln Val Val
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

870 880 890 900 910
 * * * * *
 ATG CTG ATG GGT TCC GTT GAA CTG CGT CAG GCA CCG TTT GCA CTG TGG ATC
 TAC GAC TAC CCA AGG CAA CTT GAC GCA GTC CGT GGC AAA CGT GAC ACC TAG
 <His Gln His Thr Gly Asn Phe Gln Thr Leu Cys Arg Lys Cys Gln Pro Asp
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

920 930 940 950 960
 * * * * *
 CAC GAC CTG TCG GCA CAG GAC CCG TAC TAC ATC CTG CCG ATC CTG ATG GGC
 GTG CTG GAC AGC CGT GTC CTG GGC ATG ATG TAG GAC GGC TAG GAC TAC CCG
 <Val Val Gln Arg Cys Leu Val Arg Val Val Asp Gln Arg Asp Gln His Ala
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

970 980 990 1000 1010 1020
 * * * * *
 GTA ACG ATG TTC TTC ATT CAG AAG ATG TCG CCG ACC ACA GTG ACC GAC CCG
 CAT TGC TAC AAG AAG TAA GTC TTC TAC AGC GGC TGG TGT CAC TGG CTG GGC
 <Tyr Arg His Glu Glu Asn Leu Leu His Arg Arg Gly Cys His Gly Val Arg
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1030 1040 1050 1060 1070
 * * * * *
 ATG CAG CAG AAG ATC ATG ACC TTT ATG CCG GTC ATC TTC ACC GTG TTC TTC
 TAC GTC GTC TTC TAG TAC TGG AAA TAC GGC CAG TAG AAG TGG CAC AAG AAG
 <His Leu Leu Leu Asp His Gly Lys His Arg Asp Asp Glu Gly His Glu Glu
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1080 1090 1100 1110 1120
 * * * * *
 CTG TGG TTC CCG TCA GTC TGG TGC TGT ACT ATA TCG TCA GCA ACC TGG TAA
 GAC ACC AAG GGC AGT CAG ACC ACG ACA TGA TAT AGC AGT CGT TGG ACC ATT
 <Gln Pro Glu Arg End Asp Pro Ala Thr Ser Tyr Arg End Cys Gly Pro Leu
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1130 1140 1150 1160 1170
 * * * * *
 CCA TTA TTC AGC AGC AGC TGA TTT ACC GTG GTC TGG AAA AAC GTG GCC TGC
 GGT AAT AAG TCG TCG TCG ACT AAA TGG CAC CAG ACC TTT TTG CAC CGG ACG
 <Trp End Glu Ala Ala Ala Ser Lys Gly His Asp Pro Phe Val His Gly Ala
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1180 1190 1200 1210 1220
 * * * * *
 ATA GCC GCG AGA AGA AAA AAT CCT GAT TCG GTG AGT TTT CGC TAA AAT AAG
 TAT CGG CGC TCT TCT TTT TTA GGA CTA AGC CAC TCA AAA GCG ATT TTA TTC
 <Tyr Gly Arg Ser Ser Phe Ile Arg Ile Arg His Thr Lys Ala Leu Ile Leu
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1230 1240 1250 1260 1270
 * * * * *
 GGC GGT CAG TTG ACC TGT TTT TCT TTT CGT AGG GCG GAT AAG CAC CGC GCA
 CCG CCA GTC AAC TGG ACA AAA AGA AAA GCA TCC CGC CTA TTC GTG GCG CGT
 <Ala Thr Leu Gln Gly Thr Lys Arg Lys Thr Pro Arg Ile Leu Val Ala Cys
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1280 1290 1300 1310 1320
 * * * * *
 TCC CGA CAC AAA GCA ACA GGA ACA TCA TGA GCG ATA ATG AAC CTA TCG TAG
 AGG GCT GTG TTT CGT TGT CCT TGT AGT ACT CGC TAT TAC TTG GAT AGC ATC
 <Gly Ser Val Phe Cys Cys Ser Cys End Ser Arg Tyr His Val End Arg Leu
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1330 1340 1350 1360 1370
 * * * * *
 CCC AGG CCA CGC TCC GGG ACG TGG CGG CGT TGG CAT CCT GCG CAT CTC CGG
 GGG TCC GGT GCG AGG CCC TGC ACC GCC GCA ACC GTA GGA CGC GTA GAG GCC
 <Gly Pro Trp Ala Gly Pro Arg Pro Pro Thr Pro Met Arg Arg Met Glu Pro
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1380 1390 1400 1410 1420
 * * * * *
 CTT CAA AGC CCG TGA AGT TGC CGA AAC CGT GCT GGG TAA ACT GCC TAA GCC
 GAA GTT TCG GGC ACT TCA ACG GCT TTG GCA CGA CCC ATT TGA CGG ATT CGG
 <Lys Leu Ala Arg Ser Thr Ala Ser Val Thr Ser Pro Leu Ser Gly Leu Gly
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1430 1440 1450 1460 1470
 * * * * *
 GCG CTA CAT CGA GTA TCT TCC GTA TAA AGA CGC TGT GGT TCC CTG GCC CGA
 CGC GAT GTA GCT CAT AGA AGG CAT ATT TCT GCG ACA CCA AGG GAC CGG GCT
 <Arg End Met Ser Tyr Arg Gly Tyr Leu Ser Ala Thr Thr Gly Gln Gly Ser
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1480 1490 1500 1510 1520 1530
 * * * * *
 ACT CGC ACC GGC GAA GAT GTG CTG GAA CTG CAA GGT CAT GGC GGT CCG GTG
 TGA GCG TGG CCG CTT CTA CAC GAC CTT GAC GTT CCA GTA CCG CCA GGC CAC
 <Ser Ala Gly Ala Phe Ile His Gln Phe Gln Leu Thr Met Ala Thr Arg His
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1540 1550 1560 1570 1580
 * * * * *
 ATC CTC GAC CTG CTG TCT AAA ACG CAT TCT GAC CAT TCC CGG CCT GCG GAA
 TAG GAG CTG GAC GAC AGA TTT TGC GTA AGA CTG GTA AGG GCC GGA CGC CTT
 <Asp Glu Val Gln Gln Arg Phe Arg Met Arg Val Met Gly Pro Arg Arg Phe
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1590 1600 1610 1620 1630
 * * * * *
 TTG CTC GCC CTG GTG CGA GTT TTC CGA ACG CGC GTT TCT TAA GCA TAA ACT
 AAC GAG CGG GAC CAC GCT CAA AAG GCT TGC GCG CAA AGA ATT CGT ATT TGA
 <Gln Glu Gly Gln His Ser Asn Glu Ser Arg Ala Asn Arg Leu Cys Leu Ser
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1640 1650 1660 1670 1680
 * * * * *
 TGA CTT AGC CAG GCC GAG GCG ATT GCG ATC TTG TCG ACG CAG TTC GGA ACA
 ACT GAA TCG GTC CGG CTC CGC TAA CGC TAG AAC AGC TGC GTC AAG CCT TGT
 <Ser Lys Ala Leu Gly Leu Arg Asn Arg Asp Gln Arg Arg Leu Glu Ser Cys
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1690 1700 1710 1720 1730
 * * * * *
 GCC CGT TCG CAC TAC TCG CAT GAA GAA TGA CTC GTT CTG GTA CTG CGT ATC
 CGG GCA AGC GTG ATG AGC GTA CTT CTT ACT GAG CAA GAC CAT GAC GCA TAG
 <Gly Thr Arg Val Val Arg Met Phe Ser Glu Asn Gln Tyr Gln Thr Asp
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1740 1750 1760 1770 1780
 * * * * *
 GGA TAT GAG GGA GCA TAA CCG CTG ATA GTA CGA GCG ATA GCG GCC TTT GAA
 CCT ATA CTC CCT CGT ATT GGC GAC TAT CAT GCT CGC TAT CGC CGG AAA CTT
 <Ser Ile Leu Ser Cys Leu Arg Gln Tyr Tyr Ser Arg Tyr Arg Gly Lys Phe
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1790 1800 1810 1820 1830
 * * * * *
 TCC GCG TAA GCC GAC CAC ATA GCG GTA TCG AAG AAG GAC TGA GAT ATA TAC
 AGG CGC ATT CGG CTG GTG TAT CGC CAT AGC TTC TTC CTG ACT CTA TAT ATG
 <Gly Arg Leu Gly Val Val Tyr Arg Tyr Arg Leu Leu Val Ser Ile Tyr Val
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1840 1850 1860 1870 1880
 * * * * *
 GCC CCC TCG GTT ACG GTA GCG GAA TAG GCC TAT TAA CGT ATG AAC CCC CAC
 CGG GGG AGC CAA TGC CAT CGC CTT ATC CGG ATA ATT GCA TAC TTG GGG GTG
 <Gly Gly Arg Asn Arg Tyr Arg Phe Leu Gly Ile Leu Thr His Val Gly Val
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1890 1900 1910 1920 1930
 * * * * *
 GTG CTC CAC CCC AGC GAA TTT TAC TTG GAT CCA CAG TGC AAA CGG TGC CTG
 CAC GAG GTG GGG TCG CTT AAA ATG AAC CTA GGT GTC ACG TTT GCC ACG GAC
 <His Glu Val Gly Ala Phe Lys Val Gln Ile Trp Leu Ala Phe Pro Ala Gln
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1940 1950 1960 1970 1980
 * * * * *
 ACC CAG TTC AAC GGA ACC CAT CAG CAT GTA GTA CAA CGC CAG GAA GAT TGG
 TGC GTC AAG TTG CCT TGG GTA GTC GTA CAT CAT GTT GCG GTC CTT CTA ACC

<Arg Leu Glu Val Ser Gly Met Leu Met Tyr Tyr Leu Ala Leu Phe Ile Pro
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

1990 2000 2010 2020 2030 2040
* * * * *
CAT CTG GAT CAG CAG CGG GAA GCA GCC GCC CAG CGG TTG ACC TTC GCA GCT
GTA GAC CTA GTC GTC GCC CTT CGT CGG CGG GTC GCC AAC TGG AAG CGT CGA
<Met Gln Ile Leu Leu Pro Phe Cys Gly Gly Leu Pro Gln Gly Glu Cys Ser
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2050 2060 2070 2080 2090
* * * * *
TTG TAC AGC GCC ATC ATT TCC TGG CTG ATA CGC TGT TTG TGA TCG CCC AGA
AAC ATG TCG CGG TAG TAA AGG ACC GAC TAT GCG ACA AAC ACT AGC GGG TCT
<Gln Val Ala Gly Asp Asn Gly Pro Gln Tyr Ala Thr Gln Ser Arg Gly Ser
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2100 2110 2120 2130 2140
* * * * *
CGC TCG CGC ATT GCC TGA ATC TTC GGC TGC AAC ATA CGC ATC TTC GCC ATG
GCG AGC GCG TAA CGG ACT TAG AAG CCG ACG TTG TAT GCG TAG AAG CGG TAC
<Ala Arg Ala Asn Gly Ser Asp Glu Ala Ala Val Tyr Ala Asp Glu Gly His
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2150 2160 2170 2180 2190
* * * * *
GAG GTG TAC TGG CCT TTG GTC AGC GGG TAC ATG ATG CAC CAA GCA TAA GGT
CTC CAC ATG ACC GGA AAC CAG TCG CCC ATG TAC TAC GTC GTT CGT ATT CCA
<Leu His Val Pro Arg Gln Asp Ala Pro Val His His Leu Leu Cys Leu Thr
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2200 2210 2220 2230 2240
* * * * *
GAT GAT GAT ATG AGA GCC AGT TAC CAC AAG CTG TGG ATC CAT CGG CTT TAT
CTA CTA CTA TAC TCT CGG TCA ATG GTG TTC GAC ACC TAG GTA GCC GAA ATA
<Ile Ile Ile His Ser Gly Thr Val Val Leu Gln Pro Asp Met Pro Lys Ile
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2250 2260 2270 2280 2290
* * * * *
TGC GCA CCA CGG TGA TCG GCA GTT TCG CTG GCA GAC GGG CAA TAA ATT CCG
ACG CGT GGT GCC ACT AGC CGT CAA AGC GAC CGT CTG CCC GTT ATT TAA GGC
<Ala Cys Trp Pro Ser Arg Cys Asn Arg Gln Cys Val Pro Leu Leu Asn Arg
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2300 2310 2320 2330 2340
* * * * *
GCC AGA TCT CTG CCG GAT CCA CGG CGT CTG TTG TGG TGC CAT CGA CCA TAA
CGG TCT AGA GAC GGC CTA GGT GCC GCA GAC AAC ACC ACG GTA GCT GGT ATT
<Gly Ser Arg Gln Arg Ile Trp Pro Thr Gln Gln Pro Ala Met Ser Trp Leu
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2350 2360 2370 2380 2390
* * * * *

ACA GCA CGC GGT CGG CCT GTT CAA TTT CCT GCC ACG CGT CGC TCG ATA CCA
 TGT CGT GCG CCA GCC GGA CAA GTT AAA GGA CGG TGC GCA GCG AGC TAT GGT
 <Cys Cys Ala Thr Pro Arg Asn Leu Lys Arg Gly Arg Thr Ala Arg Tyr Trp
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2400 2410 2420 2430 2440
 * * * * *
 ATA CGT TCT ACT TCG TCG CTG GCT GCA GGT AGC CCG GCG GTA TCG ATG ATA
 TAT GCA AGA TGA AGC AGC GAC CGA CGT CCA TCG GGC CGC CAT AGC TAC TAT
 <Tyr Thr Arg Ser Arg Arg Gln Ser Cys Thr Ala Arg Arg Tyr Arg His Tyr
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2450 2460 2470 2480 2490
 * * * * *
 TGC AGC GGC ATT CCG GCT ATG TGG ATA TGC GCA CGC AGC ACG TAC GCG TAG
 ACG TCG CCG TAA GGC CGA TAC ACC TAT ACG CGT GCG TCG TGC ATG CGC ATC
 <Ala Ala Ala Asn Arg Ser His Pro Tyr Ala Cys Ala Ala Arg Val Arg Leu
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2500 2510 2520 2530 2540 2550
 * * * * *
 TTC CGG CGA TGA CCA GTA GTG GCC GCT TGC GAC CCT TAC GCA CTT CAC TTG
 AAG GCC GCT ACT GGT CAT CAC CGG CGA ACG CTG GGA ATG CGT GAA GTG AAC
 <Glu Pro Ser Ser Trp Tyr His Gly Ser Ala Val Arg Val Cys Lys Val Gln
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2560 2570 2580 2590 2600
 * * * * *
 CCA GTG CGC AAT TAA GCA GAG AGC CGT TCC TGA CCA CTT CCG CAC CTG CAC
 GGT CAC GCG TTA ATT CGT CTC TCG GCA AGG ACT GGT GAA GGC GTG GAC GTG
 <Trp His Ala Ile Leu Cys Leu Ala Thr Gly Ser Trp Lys Arg Val Gln Val
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2610 2620 2630 2640 2650
 * * * * *
 GAC GAT TGG TAG AGT TTG TCT CGT ACC CGA AAC TGT GGT TGT ACC TTC CGC
 CTG CTA ACC ATC TCA AAC AGA GCA TGG GCT TTG ACA CCA ACA TGG AAG GCG
 <Val Ile Pro Leu Thr Gln Arg Thr Gly Ser Val Thr Thr Thr Gly Glu Ala
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2660 2670 2680 2690 2700
 * * * * *
 CGA AGG ACC GGA ATG TGT CCC GTT TCG CGT TGA CAA CCC TCG GAC CCG TCC
 GCT TCC TGG CCT TAC ACA GGG CAA AGC GCA ACT GTT GGG AGC CTG GGC AGG
 <Ser Pro Gly Ser His Thr Gly Asn Arg Thr Ser Leu Gly Arg Val Arg Gly
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2710 2720 2730 2740 2750
 * * * * *
 ACT TGA CGA CCG CTT CTC AAC GCA GAC CGT GTC GTC TTG CAA TTC CCA TTA
 TGA ACT GCT GGC GAA GAG TTG CGT CTG GCA CAG CAG AAC GTT AAG CCA AAT
 <Ser Ser Ser Arg Lys Glu Val Cys Val Thr Asp Asp Gln Leu Glu Cys AAT
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2760 2770 2780 2790 2800
 * * * * *
 GTG GCC CCT TAA ATG AAG TCT GCT GGA CGA CCC CGC GTA AAA GAG GTC GAA
 CAC CGG GGA ATT TAC TTC AGA CGA CCT GCT GGG GCG CAT TTT CTC CAG CTT
 <His Gly Arg Leu His Leu Arg Ser Ser Ser Gly Ala Tyr Phe Leu Asp Phe
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2810 2820 2830 2840 2850
 * * * * *
 GAC ATC AAG TAA ACC GCG CAT ACA GCC GCA TTC TGA CTG TCA GAT GCG CTT
 CTG TAG TTC ATT TGG CGC GTA TGT CGG CGT AAG ACT GAC AGT CTA CGC GAA
 <Val Asp Leu Leu Gly Arg Met Cys Gly Cys Glu Ser Gln End Ile Arg Lys
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2860 2870 2880 2890 2900
 * * * * *
 CGC TTC ATT GTT ACC GCT CCT GTT ATT CCT CAA CCC TTT TTT TAA ACA TTA
 GCG AAG TAA CAA TGG CGA GGA CAA TAA GGA GTT GGG AAA AAA ATT TGT AAT
 <Ala Glu Asn Asn Gly Ser Arg Asn Asn Arg Leu Gly Lys Lys Leu Cys End
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2910 2920 2930 2940 2950
 * * * * *
 AAA TTC TTA CGT AAT TTA TAA TCT TTA AAA AAA GCA TTT AAT ATT GCT CCC
 TTT AAG AAT GCA TTA AAT ATT AGA AAT TTT TTT CGT AAA TTA TAA CGA GGG
 <Phe Glu End Thr Ile End Leu Arg End Phe Phe Cys Lys Ile Asn Ser Gly
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

2960 2970 2980 2990 3000
 * * * * *
 CGA ACG ATT GTG ATT CGA TTC ACA TTT AAA CAA TTT CAG AAT AGA CAA AAA
 GCT TGC TAA CAC TAA GCT AAG TGT AAA TTT GTT AAA GTC TTA TCT GTT TTT
 <Ser Arg Asn His Asn Ser Glu Cys Lys Phe Leu Lys Leu Ile Ser Leu Phe
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3010 3020 3030 3040 3050 3060
 * * * * *
 CTC TGA GTG TAA TAA TGT AGC CTC GTG TCT TGC GAG GAT AAG TGC ATT ATG
 GAG ACT CAC ATT ATT ACA TCG GAG CAC AGA ACG CTC CTA TTC ACG TAA TAC
 <Glu Ser His Leu Leu Thr Ala Glu His Arg Ala Leu Ile Leu Ala Asn His
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3070 3080 3090 3100 3110
 * * * * *
 AAT ATC TTA CAT ATA TGT GTT GAC CTC AAA ATG GTT CAA TAT TGA CAA TAA
 TTA TAG AAT GTA TAT ACA CAA CTG GAG TTT TAC CAA GTT ATA ACT GTT ATT
 <Ile Asp End Met Tyr Thr Asn Val Glu Phe His Asn Leu Ile Ser Leu Leu
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3120 3130 3140 3150 3160
 * * * * *
 AAT TGT CGA TCA CCG CCC TTG ATT TGC CCT TCT GTA GCC ATC ACC AGA GCC
 TTA ACA GCT AGT GGC GGG AAC TAA ACG GGA AGA CAT CGG TAG TGG TCT CGG
 <Ile Thr Ser End Arg Gly Gln Asn Ala Arg Arg Tyr Gly Asp Gly Ser Gly
 < a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3170 3180 3190 3200 3210
* * * * *
AAA CCG ATT GAT TCA ATG TGT TCT ATT TGT TTG CTA TAT CTT AAT TTT GCC
TTT GGC TAA CTA AGT TAC ACA AGA TAA ACA AAC GAT ATA GAA TTA AAA CGG
<Phe Arg Asn Ile End His Thr Arg Asn Thr Gln End Ile Lys Ile Lys Gly
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3220 3230 3240 3250 3260
* * * * *
TTT TGC AAA GGT CAT CTC TCG TTT ATT TAC TTG TTT TAG TAA ATG ATG GTG
AAA ACG TTT CCA GTA GAG AGC AAA TAA ATG AAC AAA ATC ATT TAC TAC CAC
<Lys Ala Phe Thr Met Glu Arg Lys Asn Val Gln Lys Leu Leu His His His
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3270 3280 3290 3300 3310
* * * * *
CTT GCA TAT ATA TCT GGC GAA TTA ATC GGT ATA GCA GAT GTA ATA TTC ACA
GAA CGT ATA TAT AGA CCG CTT AAT TAG CCA TAT CGT CTA CAT TAT AAG TGT
<Lys Cys Ile Tyr Arg Ala Phe End Asp Thr Tyr Cys Ile Tyr Tyr Glu Cys
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3320 3330 3340 3350 3360
* * * * *
GGG ATC ACT GTA ATT AAA ATA AAT GAA GGA TTA TGT AAT GGA AGA CAT CGG
CCC TAG TGA CAT TAA TTT TAT TTA CTT CCT AAT ACA TTA CCT TCT GTA GCC
<Pro Asp Ser Tyr Asn Phe Tyr Ile Phe Ser End Thr Ile Ser Ser Met Pro
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3370 3380 3390 3400 3410
* * * * *
CAT GCA GGC TGC GAT GAT GCG CGG CGA CGA AGC CTA CAG CGG CAG TCG TAG
GTA CGT CCG ACG CTA CTA CGC GCC GCT GCT TCG GAT GTC GCC GTC AGC ATC
<Met Cys Ala Ala Ile Ile Arg Pro Ser Ser Ala End Leu Pro Leu Arg Leu
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3420 3430 3440 3450 3460
* * * * *
CTA CTA TGC GTT ACG CGA GTC AGT GAA AAA TAT CTT TGG TTA TCA ATA CAC
GAT GAT ACG CAA TGC GCT CAG TCA CTT TTT ATA GAA ACC AAT AGT TAT GTG
<End End Ala Asn Arg Ser Asp Thr Phe Phe Ile Lys Pro End End Tyr Val
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3470 3480 3490 3500 3510
* * * * *
CAT TCC GAC TCA CCA GGG CCG TGG CGC AGA GCA AAT CTA TAT TCC GGT ACT
GTA AGG CTG AGT GGT CCC GGC ACC CCG TCT CGT TTA GAT ATA AGG CCA TGA
<Met Gly Val End Trp Pro Arg Pro Ala Ser Cys Ile End Ile Gly Thr Ser
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3520 3530 3540 3550 3560 3570
* * * * *
GAT TAA AAA ACG CGA GCA GGA AAA AGG CCT GGA TCG CAG CAA AAT GGT GGC
CTA ATT TTT TGC GCT CGT CCT TTT TCC GGA CCT AGC GTC GTT TTA CCA CCG

<Ile Leu Phe Arg Ser Cys Ser Phe Pro Arg Ser Arg Leu Leu Ile Thr Ala
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3580 3590 3600 3610 3620
* * * * *
GTT CTC TAA CTA TTT CTT TGA TCC ACG CAG GGC CAT AGC CAG ATC AAC GGC
CAA GAG ATT GAT AAA GAA ACT AGG TGC GTC CCG GTA TCG GTC TAG TTG CCG
<Asn Glu Leu End Lys Lys Ser Gly Arg Leu Ala Met Ala Leu Asp Val Ala
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3630 3640 3650 3660 3670
* * * * *
GTA CCG TGC GTA ACG TCT ATA TCA AAG AAG CCT TCG ATA CGG GCG TGC GTT
CAT GGC ACG CAT TGC AGA TAT AGT TTC TTC GGA AGC TAT GCC CGC ACG CAA
<Tyr Arg Ala Tyr Arg Arg Tyr End Leu Leu Arg Arg Tyr Pro Arg Ala Asn
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3680 3690 3700 3710 3720
* * * * *
ACG ACT TTA AAG GCA ACT TTG ACC TTG AGG GAT TAG AAC CCG GTA TTG AAG
TGC TGA AAT TTC CGT TGA AAC TGG AAC TCC CTA ATC TTG CGC CAT AAC TTC
<Arg Ser End Leu Cys Ser Gln Gly Gln Pro Ile Leu Val Arg Tyr Gln Leu
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3730 3740 3750 3760 3770
* * * * *
AAG TTG GTC CGA ATA ACG TGC CGT ATA TCG TTG CAA CCA TCA CCA GTA ACT
TTC AAC CAG GCT TAT TGC ACG GCA TAT AGC AAC GTT GGT AGT GGT CAT TGA
<Leu Gln Asp Ser Tyr Arg Ala Thr Tyr Arg Gln Leu Trp End Trp Tyr Ser
< a a a a TRANSLATION OF SEQUENCE/KISWAR a a a a

3780
* *
CTG CAG
GAC GTC
<Gln Leu
< a

- END -

DATE FILMED

10 / 31 / 90