Pergamon

PII: S0031-9422(98)00317-3

Phytochemistry Vol. 49, No. 8, pp. 2221–2232, 1998
© 1998 Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0031-9422/98/$ - see front matter

FOLYLPOLYGLUTAMATE SYNTHESIS IN *NEUROSPORA CRASSA*: PRIMARY STRUCTURE OF THE FOLYLPOLYGLUTAMATE SYNTHETASE GENE AND ELUCIDATION OF THE *MET-6* MUTATION

IAN J. ATKINSON, FRANK E. NARGANG and EDWIN A. COSSINS*

Department of Biological Sciences, University Of Alberta, Edmonton, Alta., Canada T6J 2E9

(*Received* 30 *September* 1997; *in revised form 13 March* 1998)

Key Word Index—*Neurospora crassa*; Ascomycetes; fungi; folylpolyglutamate synthetase; *met-6* mutant (FGSC 1330); recombinant DNA; genomic DNA sequence.

Abbreviations—C_1-metabolism = one-carbon metabolism; FGSC = Fungal Genetic Stock Center; PCR = polymerase chain reaction; 5′-RACE = rapid amplification of cDNA ends; FPGS = folylpoly-γ-glutamate synthetase; $H_4PteGlu$ = 5,6,7,8-tetrahydropteroylmonoglutamate; $H_4PteGlu_n$ = poly-γ-glutamyl derivatives (*n* ; the number of L-glutamate moieties).

Abstract—In *Neurospora crassa*, the *met-6*$^+$ gene encodes folylpoly-γ-glutamate synthetase (FPGS) which catalyzes the formation of polyglutamate forms of folate. Methionine auxotrophy of the *Neurospora crassa met-6* mutant is related to a lesion affecting this enzyme. Functional complementation of the mutant strain was achieved by introducing copies of the wild-type *met-6*$^+$ gene into mutant spheroplasts. The complementing sequences were found to be contained on a 3.5 kb *Eco*RI–*Bam*HI restriction fragment. The nucleotide sequence of the *met-6*$^+$ gene was determined and an open reading frame of 1587 bp was identified, interrupted by two introns. This open reading frame contained several AUG codons but translation beginning from either of the first two would theoretically produce a protein of appropriate size and with similarity to five other FPGS proteins. Northern blot analyses of *met-6*$^+$ transcripts revealed a 2.0 kb product. The position of the transcription stop site and an intron were identified by sequencing partial cDNA clones which were truncated at the 5′ end. DNA sequence analysis of the *met-6* mutant allele revealed a T to C transition which would result in replacement of a highly conserved serine with a proline. © 1998 Elsevier Science Ltd. All rights reserved

INTRODUCTION

Folic acid derivatives serve as coenzymes in several one-carbon transfer reactions essential for normal growth and proliferation of cells (the abbreviations for folate derivatives are those suggested by the IUPAC-IUB Commission, as summarized by Blakley and Benkovic [1]). These reactions, leading to thymidine, purine, amino acid and methyl group biosynthesis are preferentially mediated by folylpolyglutamates, the predominant intracellular form of folate [2–5]. Folylpolyglutamates are synthesized by folylpoly-γ-glutamate synthetase (FPGS; tetrahydrofolate: L-glutamate γ-ligase (ADP forming); E.C. 6.3.2.17) which catalyzes the sequential ATP-dependent addition of glutamic acid residues to tetrahydrofolates (equation (1)):

$$H_4PteGlu_n + MgATP + l-glutamate \longrightarrow H_4PteGlu_{n+1} + MgADP + Pi \quad (1)$$

Studies involving functional complementation of FPGS mutant cell lines of *E. coli* have now been used to clone several FPGS genes. In addition, transformation of mutant strains with various wild-type FPGS genes results in the expression of FPGS enzyme activity [6–10] and ability to synthesize folylpolyglutamates *in vivo* [10–12].

Polyglutamate synthesis in *Neurospora crassa* is affected by a mutation at the *met-6* locus [13–19] on chromosome 1R [20]. The resulting mutant fails to generate long-chain polyglutamates *in vivo* and displays a methionine auxotrophic phenotype [17–19, 21], which appears to be caused by a lesion affecting FPGS activity. Extracts of the *met-6* mutant strain synthesize folyldiglutamates

*Author to whom correspondence should be addressed.

(equation (2)) but not longer chain length folates [18].

$$H_4PteGlu + ATP + 1 - glutamate \longrightarrow H_4PteGlu_2 + ADP + Pi. \quad (2)$$

Preliminary biochemical and genetic evidence suggested that *Neurospora* polyglutamate synthesis might involve more than one FPGS activity [13, 18] with the diglutamate formed initially being polyglutamylated by a second enzyme to give folylhexaglutamates. A mutation originally thought to be in a second gene, *mac*, cannot form diglutamates from the monoglutamate but can form folylpolyglutamates from folyldiglutamate substrates. Subsequent work suggested that both activities are probably encoded by the *met-6*$^+$ gene [21].

This paper reports the sequence of the *met-6*$^+$ gene in three strains of wild-type genomic DNA. This information has led to the identification of a partial *met-6*$^+$ cDNA and the location of the *met-6* mutation.

RESULTS AND DISCUSSION

Transformation of polyglutamate-deficient mutants to identify the met-6$^+$ *gene*

In previous studies on the nature of FPGS in *Neurospora crassa*, cloning of the *met-6*$^+$ gene was accomplished by functional complementation of *met-6* mutant spheroplasts [21]. Cotransformation of plasmid pIA50 (containing the *met-6*$^+$ gene) with the cosmid pSV50 carrying a benomyl resistance marker [22] revealed that the methionine auxotrophic phenotype could be complemented to prototrophy by a 3.5 kb *Eco*RI/*Bam*HI fragment of genomic DNA [21]. Negative control data (pSV50 transformation of mutant spheroplasts) closely approximated spontaneous reversion rates [23], whereas prototrophic transformant colonies were generated at frequencies 800 times spontaneous rates using plasmid pIA50 DNA. Support for transformation was corroborated in this study by Southern analysis (Fig. 1). Transformations of *N. crassa* generally result in ectopic (often multiple) integration of transforming sequences [24], which is readily identified in the stable prototrophic transformant strain m6-TIA50 in Fig. 1, lane 3. The presence of additional bands large enough to contain a complete *met-6*$^+$ gene indicate integration of at least one copy of plasmid pIA50 DNA.

The common band of approximately 5.0 kb observed in 74-OR23-1*A* wild-type, the *met-6* mutant and the m6-TIA50 strains (Fig. 1, lanes 1–3) also suggest that no gross deletions or insertions of the *met-6*$^+$ DNA were present in the *met-6* mutant (Fig. 1, lane 2).

Organization of the Met-6$^+$ *gene*

To identify the *met-6*$^+$ gene, DNA sequencing was performed on both strands of CsCl-purified

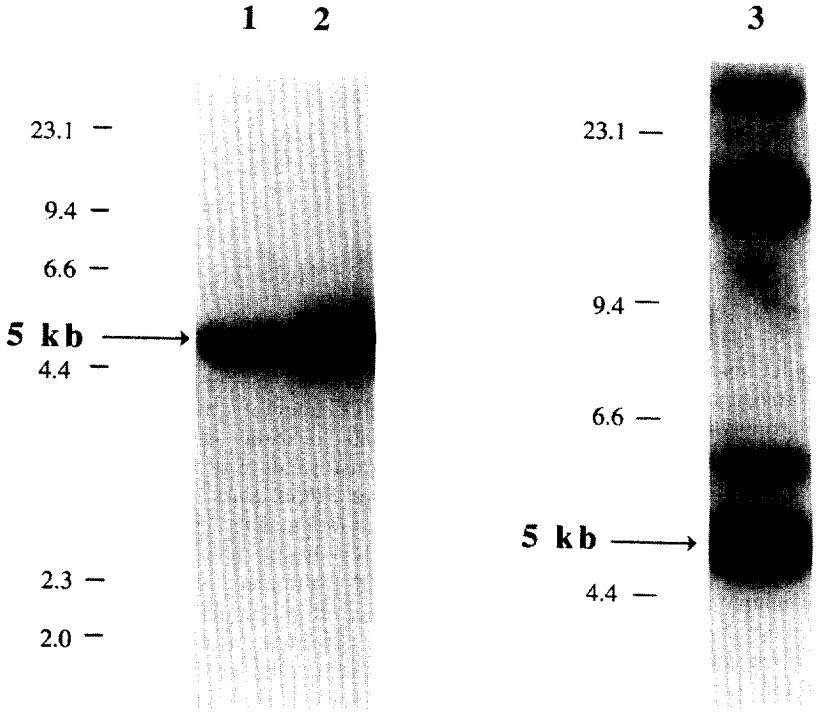

Fig. 1. Southern blot analyses of *N. crassa* genomic DNA. The *N. crassa* DNAs were digested with *Eco*RI and probed with a *met-6 specific* ^{32}P-*labeled KpnI/BamHI fragment from pIA68. Lane 1, wild-type 74-OR23-1A; lane 2, met-6 mutant; lane 3, M6-TIA50 stable transformant.*

Table 1. Conserved sequences of the longest open reading frame of the *met-6*$^-$ gene*

(A) Translation initiation site	
Potential translation start site 1	CCTG*ATG*CAC
Potential translation start site 2	CACC*ATG*GCA
N. crassa consensus sequence [25, 26]	CAMM*ATG*GCT

(B) Exon/intron boundaries

Intron 1 (hypothetical)	5′-donor site	splice branch site	3′-acceptor site
74-OR23-1*A* genomic sequence	5′ T\^GTGAGTCAACCAA	—CACTCACT–6—AT—	—CGTAG\^G 3′
N. crassa consensus sequence [25, 26]	5′ G\^GTAAGTnnYCnYY—	WRCTRACM–6—YY—	—WACAG\^ 3′
Intron 2	5′-donor site	splice branch site	3′-acceptor site
74-OR23-1*A* genomic sequence	5′ G\^GTATGTTGTCCTG	—TGCTAACC—	—5—AATAG\^G 3′
N. crassa consensus sequence [25, 26]	5′ G\^GTAAGTnnYCnYY—	WRCTRACM–6—YY—	—WACAG\^ 3′

*Nucleic acid codes recommended by the IUPAC-IUB Biochemical Nomenclature Commission.
M = adenine or cytosine; W = adenine or thymine; R = adenine or guanine; n = guanine or adenine or thymine or cytosine.

plasmid pIA50. The nucleotide sequence of the entire 3440 bp *Eco*RI/*Bam*HI fragment is shown in Fig. 2. After examination of computer translations of all six potential reading frames, two blocks of translated sequence (out of frame with one another because of the presence of introns) showed significant similarity to foreign FPGS genes.

An open reading frame of 1587 nucleotides was identified (Fig. 2) with two introns. A putative 61 nucleotide intron was identified from genomic DNA sequence at position +129 of the largest open reading frame (Fig. 2), with appropriate *Neurospora* intron/exon boundaries (Table 1; [25, 26]). Excision of the putative intron would bring the two blocks with similarity to foreign FPGS genes back into frame. A second putative intron sequence occurs at position +1333 of the largest open reading frame. The identity of this 50 nucleotide intron was confirmed by cDNA analysis (see Section 2.3, Fig. 2; Table 1).

A protein translated from the largest open reading frame (Fig. 2, beginning at ↓) would contain 528 amino acids (Fig. 3) with a deduced molecular weight of 58,672 Da. Based on size estimates of native *Neurospora* FPGS protein [21, 27], this translation could produce a protein similar in size to wild-type FPGS protein. This is also in agreement with the size of FPGS proteins from other eukaryotes which range between 60 and 70 kDa. Translation initiation at a second in-frame AUG codon (Fig. 2, ▽) could also produce a protein of appropriate size; a 1485 nucleotide open reading frame would produce a protein containing 494 amino acids and having a molecular weight of 54,646 Da.

The genomic sequence was examined for motifs characteristic of core promoter elements for transcription start sequences as well as the preferred translation start site. Computer analysis showed no obvious similarities to elements such as TATA boxes (*Neurospora* consensus sequence = TATATAA) and CAAT boxes (*Neurospora* consensus sequence = GG(C/T)CAATCT) [25, 26]. However, many fungal genes lack such elements [28]. Evaluation of the translation start site was based on similarity to the *N. crassa* translational start consensus sequence (Table 1) [25, 26]. Based on these considerations, the second in-frame AUG (Fig. 2, ▽) would appear to be the most favourable choice for the translation start sequence. However, confirmation for either translation start site remains to be achieved.

Fig. 2. Sequence of the 3440 bp plasmid pIA50 (GenBank accession number is AF005040) and subsequent sequence of the putative open reading frame. Translation start site 1 (indicated by the symbol ↓) is postulated to be the hypothetical start to the largest open reading frame of the *met-6*$^+$ gene (Table 1) at position +1 of the largest open reading frame. Translation start site 2 (indicated by the symbol ▽) is postulated to be a second hypothetical start site, 102 nucleotides smaller than the largest open reading frame of the *met-6*$^+$ gene (Table 1). The single underlined areas refer to the two primers used to amplify the wild-type or mutant PCR products of 2.2 kb from genomic template DNA. At position +400 of the largest open reading frame (indicated by the symbol ⇓ and square box), the triplet codon TCC of plasmid pIA50, 74-OR23-1*A*, Lindegren A and Emerson, giving rise to serine, is a CCC triplet codon in *met-6* mutant, giving rise to proline. This is the site of the *met-6* mutation. The 5′ end of the three cDNA clones occur at position +182 of the largest open reading frame and is postulated to be truncated and not to be the true 5′ end of the cDNA. The cDNA clone 3′ end occurs at position +1858 of the largest open reading frame of the cDNA. Both ends are indicated by asterisks (*).

```
GAATTCAAGT  GGAACCTGGC  GCCTTAGACC  GCTCGGCCAC  CTTGCCATCG  ATATTATATT  TTTATCGTGA  AAGTGGCTTT    80
TGAGGACATT  AGGGAAGATG  CTCAGCTTAA  CTAGTAGCCA  CAGCATAGAC  TACCGGTAGT  ACATAGCTGA  GGTGGGGCCA   160
GGTCGTTTTT  GGCGGGGTTC  AAACACAGTT  GAAAAAGCCA  AGATCGGGTA  TGTTACTGAC  GTCGGTGTTC  CAGAGCCGGC   240
GTGTCTAGAA  TAGAGTTCTA  ACCGAATGGG  TAAGCCAAGC  ATCGCGAACG  GAACATCCCG  TCTTCTTTTC  TTTTACCATC   320
GAAGTGACTG  TGAGTGACAA  GGTGGGTTAG  GATTGAATCA  CAGCTGATGG  GATTGGGAGT  TTTGGTATTG  TGAATGCCAC   400
TAACACCAGG  GCAGCTATCA  GACGCCTCGA  GATACTACCC  GATTATCACC  TCTTTGAAAC  CAACGGCCTG               470 (-1)
↓
ATG CAC CAT GTT TTG AGG CCC ATA GCT TTT CGA CTT GCC CTC GTC TCC CCT CTG AGG TCG CTC ACC ATC ACC CAT   545
Met His His Val Leu Arg Pro Ile Ala Phe Arg Leu Ala Leu Val Ser Pro Leu Arg Ser Leu Thr Ile Thr His   (+75)
                                    ∇
CAT CAC CTC TTC TTC ACA AAA CGC ACC ATG GCA TCC AGC GCT AGA ACA TAC AAT gtgagtcaaccaacctttttaccgctcc   626
His His Leu Phe Phe Thr Lys Arg Thr Met Ala Ser Ser Ala Arg Thr Tyr Asn                 INTRON 1      (+129)

acagaaccccaaacactcactcaccatatcgtag  GAC GCA ATC GAT GCG CTC AAC TCC CTC CAG ACC CCC TTC GCC GTC ATC   708
                                    Asp Ala Ile Asp Ala Leu Asn Ser Leu Gln Thr Pro Phe Ala Val Ile   (+177)
     *
GAA GCC CGG CGC AAA GCG GGT ATC CGT CCC GAT GCG CAC TCG GTA AAG GAA ATG CGT GCC TAC CTC GCC CGC ATC   783
Glu Ala Arg Arg Lys Ala Gly Ile Arg Pro Asp Ala His Ser Val Lys Glu Met Arg Ala Tyr Leu Ala Arg Ile   (+252)

GGC TAC TCC TCC CAA GAT CTC GAT CGC CTC AAC ATC GTC CAC GTA GCC GGT ACC AAG GGC AAG GGC GGC ACC TGT   858
Gly Tyr Ser Ser Gln Asp Leu Asp Arg Leu Asn Ile Val His Val Ala Gly Thr Lys Gly Lys Gly Gly Thr Cys   (+327)
                                                                                          ⇓
GCC TTT GTC GAC TCC ATC CTC ACG CGT CAT CAG CGG ACT CAC GGT ATC CCT AGG CGC ATC GGT CTC TTC ACC TCC   933
Ala Phe Val Asp Ser Ile Leu Thr Arg His Gln Arg Thr His Gly Ile Pro Arg Arg Ile Gly Leu Phe Thr Ser   (+402)

CCC CAT CTC ATT GCT GTC CGG GAA CGC ATC CGT ATC GAC TCC AAG CCC ATC TCC GAG GAA CTG TTT GCC CGC TAC  1008
Pro His Leu Ile Ala Val Arg Glu Arg Ile Arg Ile Asp Ser Lys Pro Ile Ser Glu Glu Leu Phe Ala Arg Tyr  (+477)

TTC TTC GAG GTC TGG GAT CGT CTC GAG ACA TCT CAG CTG GCT AAG GAT GAG GTG GAA CTG GGA AGC AAG CCC ATT  1083
Phe Phe Glu Val Trp Asp Arg Leu Glu Thr Ser Gln Leu Ala Lys Asp Glu Val Glu Leu Gly Ser Lys Pro Ile  (+552)

TAC GCT CGC TAT CTT ACC TTG ATG AGT TAC CAC GTC TAC CTG TCC GAG GGC GTC GAT GTA GCC ATT TAC GAA ACT  1158
Tyr Thr Arg Tyr Leu Thr Leu Met Ser Tyr His Ala Asn Leu Ser Glu Gly Val Asp Val Ala Ile Tyr Glu Thr  (+627)

GGC ATC GGG GGT GAA TAC GAT GCC ACC AAC GTA GTC  GAC CGT CCC GTG GTC AGC GGT ATC AGC ACC CTT GGT ATC  1233
Gly Ile Gly Gly Glu Tyr Asp Ala Thr Asn Val Val Asp Asn Pro Val Val Ser Gly Ile Ser Thr Leu Gly Ile  (+702)

GAC CAC GTC TTT GTC CTG GGT GAT ACG GTC GAC AAG ATT GCG TGG CAC AAG GCG GGT ATC ATG AAG ACG GGC AGC  1308
Asp His Val Phe Val Leu Gly Asp Thr Val Asp Lys Ile Ala Trp His Lys Ala Gly Ile Met Lys Thr Gly Ser  (+777)

CCT GCG TTC ACG ATT GAG CAG GTT CCC TCC GCA ACT CAG GTG CTC AAG GAT AGG GCG GTG GAA AAG GGC GTG GAT  1383
Pro Ala Phe Thr Ile Glu Gln Val Pro Ser Ala Thr Gln Val Leu Lys Asp Arg Ala Val Glu Lys Gly Val Asp  (+852)
```

Fig. 2.

```
GAATTCAAGT   GGAACCTGGC   GCCTTAGACC   GCTCGGCCAC   CTTGCCATCG   ATATTATATT   TTTATCGTGA   AAGTGGCTTT     80
TGAGGACATT   AGGGAAGATG   CTCAGCTTAA   CTAGTAGCCA   CAGCATAGAC   TACCGGTAGT   ACATAGCTGA   GGTGGGGCCA    160
GGTCGTTTTT   GGCGGGGTTC   AAACACAGTT   GAAAAAGCCA   AGATCGGGTA   TGTTACTGAC   GTCGGTGTTC   CAGAGCCGGC    240
GTGTCTAGAA   TAGAGTTCTA   ACCGAATGGG   TAAGCCAAGC   ATCGCGAACG   GAACATCCCG   TCTTCTTTTC   TTTTACCATC    320
GAAGTGACTG   TGAGTGACAA   GGTGGGTTAG   GATTGAATCA   CAGCTGATGG   GATTGGGAGT   TTTGGTATTG   TGAATGCCAC    400
TAACACCAGG   GCAGCTATCA   GACGCCTCGA   GATACTACCC   GATTATCACC   TCTTTGAAAC   CAACGGCCTG                  470 (-1)
↓
ATG CAC CAT GTT TTG AGG CCC ATA GCT TTT CGA CTT GCC CTC GTC TCC CCT CTG AGG TCG CTC ACC ATC ACC CAT   545
Met His His Val Leu Arg Pro Ile Ala Phe Arg Leu Ala Leu Val Ser Pro Leu Arg Ser Leu Thr Ile Thr His   (+75)
                                        ∇
CAT CAC CTC TTC TTC ACA AAA CGC ACC ATG GCA TCC AGC GCT AGA ACA TAC AAT gtgagtcaaccaaccttttaccgctcc   626
His His Leu Phe Phe Thr Lys Arg Thr Met Ala Ser Ser Ala Arg Thr Tyr Asn                 INTRON 1     (+129)

acagaaccccaaacactcactcaccatatcgtag  GAC GCA ATC GAT GCG CTC AAC TCC CTC CAG ACC CCC TTC GCC GTC ATC   708
                                    Asp Ala Ile Asp Ala Leu Asn Ser Leu Gln Thr Pro Phe Ala Val Ile   (+177)
     *
GAA GCC CGG CGC AAA GCG GGT ATC CGT CCC GAT GCG CAC TCG GTA AAG GAA ATG CGT GCC TAC CTC GCC CGC ATC   783
Glu Ala Arg Arg Lys Ala Gly Ile Arg Pro Asp Ala His Ser Val Lys Glu Met Arg Ala Tyr Leu Ala Arg Ile   (+252)

GGC TAC TCC TCC CAA GAT CTC GAT CGC CTC AAC ATC GTC CAC GTA GCC GGT ACC AAG GGC AAG GGC GGC ACC TGT   858
Gly Tyr Ser Ser Gln Asp Leu Asp Arg Leu Asn Ile Val His Val Ala Gly Thr Lys Gly Lys Gly Gly Thr Cys   (+327)
                                                                                          ⇓
GCC TTT GTC GAC TCC ATC CTC ACG CGT CAT CAG CGG ACT CAC GGT ATC CCT AGG CGC ATC GGT CTC TTC ACC TCC   933
Ala Phe Val Asp Ser Ile Leu Thr Arg His Gln Arg Thr His Gly Ile Pro Arg Arg Ile Gly Leu Phe Thr Ser   (+402)

CCC CAT CTC ATT GCT GTC CGG GAA CGC ATC CGT ATC GAC TCC AAG CCC ATC TCC GAG GAA CTG TTT GCC CGC TAC   1008
Pro His Leu Ile Ala Val Arg Glu Arg Ile Arg Ile Asp Ser Lys Pro Ile Ser Glu Glu Leu Phe Ala Arg Tyr   (+477)

TTC TTC GAG GTC TGG GAT CGT CTC GAG ACA TCT CAG CTG GCT AAG GAT GAG GTG GAA CTG GGA AGC AAG CCC ATT   1083
Phe Phe Glu Val Trp Asp Arg Leu Glu Thr Ser Gln Leu Ala Lys Asp Glu Val Glu Leu Gly Ser Lys Pro Ile   (+552)

TAC GCT CGC TAT CTT ACC TTG ATG AGT TAC CAC GTC TAC CTG TCC GAG GGC GTC GAT GTA GCC ATT TAC GAA ACT   1158
Tyr Thr Arg Tyr Leu Thr Leu Met Ser Tyr His Ala Asn Leu Ser Glu Gly Val Asp Val Ala Ile Tyr Glu Thr   (+627)

GGC ATC GGG GGT GAA TAC GAT GCC ACC AAC GTA GTC  GAC CGT CCC GTG GTC AGC GGT ATC AGC ACC CTT GGT ATC   1233
Gly Ile Gly Gly Glu Tyr Asp Ala Thr Asn Val Val Asp Asn Pro Val Val Ser Gly Ile Ser Thr Leu Gly Ile   (+702)

GAC CAC GTC TTT GTC CTG GGT GAT ACG GTC GAC AAG ATT GCG TGG CAC AAG GCG GGT ATC ATG AAG ACG GGC AGC   1308
Asp His Val Phe Val Leu Gly Asp Thr Val Asp Lys Ile Ala Trp His Lys Ala Gly Ile Met Lys Thr Gly Ser   (+777)

CCT GCG TTC ACG ATT GAG CAG GTT CCC TCC GCA ACT CAG GTG CTC AAG GAT AGG GCG GTG GAA AAG GGC GTG GAT   1383
Pro Ala Phe Thr Ile Glu Gln Val Pro Ser Ala Thr Gln Val Leu Lys Asp Arg Ala Val Glu Lys Gly Val Asp   (+852)
```

Fig. 2—*continued.*

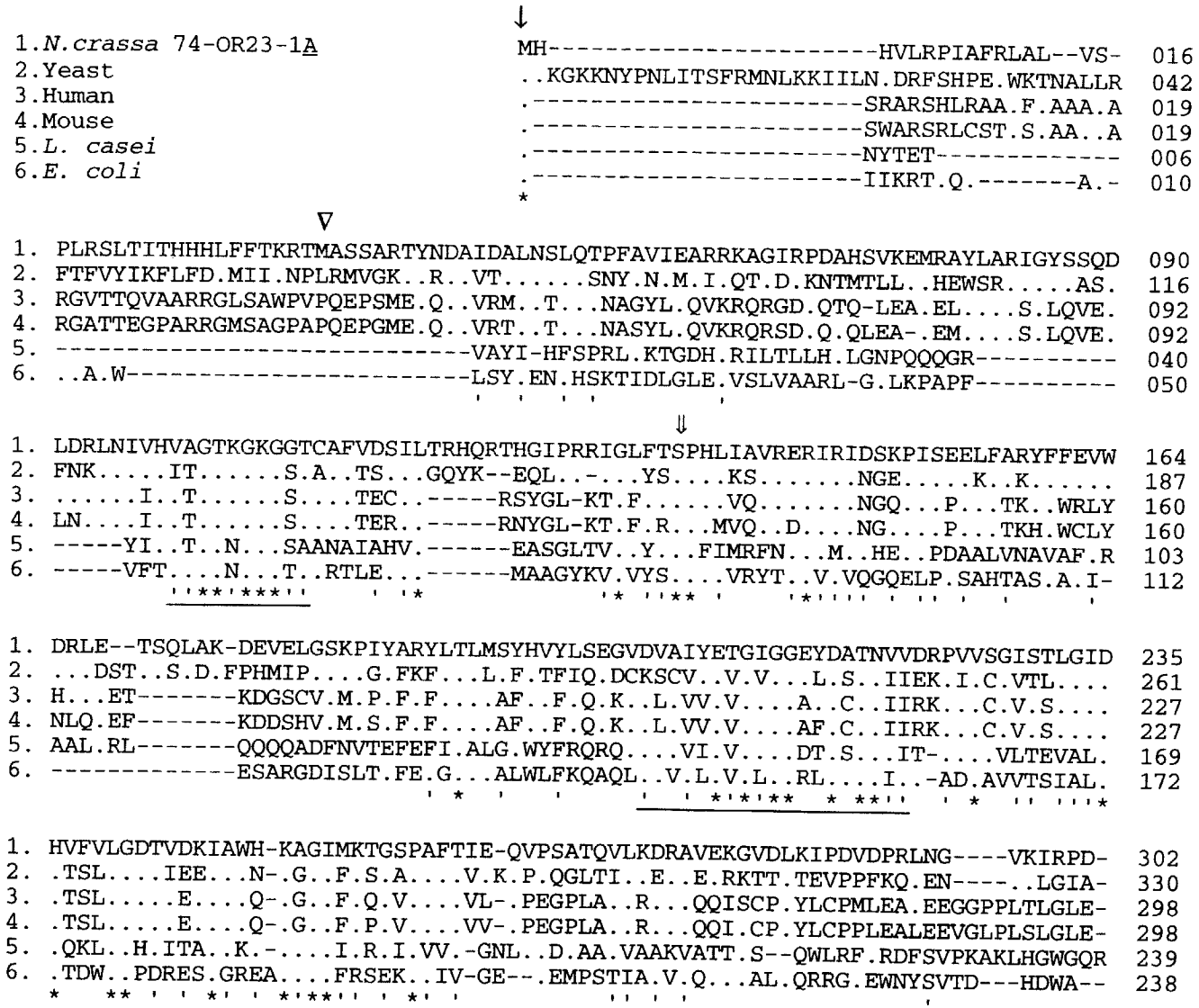

Fig. 3.

```
1. ---AVFQKKNATLAIALAETALKK---LDPSFKPGTDS-------LSP-----E-FVQGLEQVVWRGRCEVKEE  357
2. ---GE...S..S..VM..SEI.HTSNI.EEKI.CSSNA-------SI.-----.K.II..QNTK.E...Q.L.K  389
3. ---GEH.RS..A..LQ..HCW.QRQDRHGAGEPKASRPGLLWQLP.A.VFQPTSHMRL..RNTE.P..TQ.LRR  369
4. ---GAH.RS..A..LQ..HCW.ERQDHQ.IQELKVSRPSIRWQLP.A.VFRPTPHMRR..RDT..P..TQILQR  369
5. FTYEDQDGRISD.EVP.VGDYQQRNMAIAIQTAKVYAKQTEWPLT-------PQNIR...AASH.PA.L.KISD  306
6. --FSDAHGTLEN.PLP.V---PQPNAATALAAL----RASGLEVS-------ENAIRD.IASAILP..FQIVS.  296

1. DQAVWHLDGAHTVDSLKVAGRWFVEECVKKAKG----GPKVLIFNQQGRSEAVDFLDGLCNTVKSADPEGTGFS  427
2. GKN..YI.....K..MVA.ST..RD--MVRLSK----RKKI.L....S.-D.NALVNN.YSS.SPE----ITFD  452
3. GPLT.Y......AS.AQACV...RQALQGRERPSGGPEVRV.L..AT.DRDPAAL.KL.---------QPCQFD  434
4. GPLT.Y......TS.VQACVH.YRQSLERSKRTDGGSEVHI.L..ST.DRDSAAL.KL.---------QPCQFD  434
5. TPLI-VI....NP.GIN-----GLITAL.QLFSQ-PITVIAG.LADKDYAAMA.R.TA----------------  357
6. SPR.-IF.V..NPHAAE-----YLTGRM.ALPKNGRMLAVIGMLHDKDIAGTLAW.KSVVDD------------  352

1. HVIFCTNVTYATTGYKKDFVNHQYNPKDI----------ENMTQQRVFAERWSTLDP----SANVMLIPTIEE  486
2. D...T....WKSGS.SA.L.SMNTSQE.V----------.KLKV.ESLVKN.NKI.D----NRAK---THVTA  508
3. YAV..P.L.EVSSTGNA.QQ.FTVTLDQVLLRCLEHQQHWNHLDEEQASPDL..APS.EPGG..SLL.A.HPPH  508
4. YAV..P...EVSSIGNA.QQ.FTVTLDQVLLRCLQHQQHWNGLAEKQASSNL..SCG.DPAGPGSLL.A.HPPQ  508
5. ----------------------------------------------------------------------  357
6. --WY.APLE----------GPRGATAEQLLEHLGNGKSFDSVA.A------.DAAMADAKAEDT.LV------  401

1. A----------------INKARSLV-------------------DTTEGEQKVQALITGSLHLVGGALGILE  523
2. S----------------.EE.NE.I-------------------E.LYD.P-ADIFV.........L.VVFD  535
3. TCSASSLVFSCISHALQW.SQG.DPIFQPPSPPKGLLTHPVAHSGASILREAAAIHV.V.........V.KL..  582
4. PTRTSSLVFSCISHALLWISQGRDPIFQPQSLPRNLLNHPTANSGASILREAAAIHVLVTGSLHLVGGVLKLLD  582
5. ---A----FSTVYLV----------------------------------------------------------  365
6. -CGS----FHTVAHVMEV.DAR..-----------------------------------------..------  420

1. KADAL
2. RIDVK
3. P.LSQ
4. PSMSQ
5. -----
6. ----K
```

Fig. 3.—*continued.*

The putative 1587 bp open reading frame was examined for the presence of a mitochondrial targeting sequence. FPGS activity is located in the mitochondria and cytosol of eukaryotic cells and mitochondrial FPGS activity is required for normal mitochondrial folate pools and glycine metabolism [29]. Evidence from Cossins and Chan [18] showed that about 20% of recoverable wild-type *Neurospora* FPGS activity was associated with the mitochondrial fraction, which can be separated from cytosolic FPGS by precipitation with ammonium sulphate [17, 30]. However, the mutant *met-6* lacks detectable mitochondrial FPGS activity in addition to the limited cytosolic FPGS function [19]. Computer analysis using PCGENE™ software identified a putative targeting peptide rich in hydroxylated and basic amino acids and devoid of acidic amino acids, predicted in position +1 to +41 at the N-terminus of the putative protein encoded by the largest possible reading frame. Since there are two possible start sites, translation from the first start site could target the protein to the mitochondria. Further examination of the two putative AUG translational start sites may elucidate the mechanisms of FPGS cytosolic/mitochondrial distribution.

To gain additional evidence that the coding sequence identified did indeed encode *met-6*$^+$, a 2.2 kb PCR fragment of wild-type 74-OR23-1*A* genomic DNA was generated from oligonucleotide primers located just beyond the boundaries of the largest open reading frame (Fig. 2, underlined). Cotransformation of this cloned *met-6*$^+$ fragment with the cosmid pSV50 for antibiotic selection was capable of complementing mutant spheroplasts to prototrophy. Since regions of homology to foreign FPGS genes were detected in the open reading frame, encompassed by the PCR product, transformation to prototrophy with such a small fragment strongly supports the notion that the *met-6*$^+$ gene occurs within this region.

The derived amino acid sequence of the largest open reading frame was compared to five known FPGS proteins and considerable similarity was observed (Fig. 3; [6, 8, 31, 32]). The similarity was concentrated in regions which corresponded to the major conserved regions between the five foreign FPGS proteins, primarily in the N-terminus (Fig. 3). The overall percentage identity of the five proteins to the putative *met-6*$^+$ gene ranged from 20% (*E. coli*) to 43% (*S. cerevisiae*). The open reading frame contained regions with strong homology to areas described as A and B nucleotide binding sites for the *E. coli folC* gene and the human FPGS sequence deduced from cDNA [8, 32, 33]). These sites occur at amino acids 99–108 (A region) and 203–221 (B region) of the putative *met-6*$^+$ gene product and are highly conserved across FPGS enzymes (Fig. 3, underlined). All FPGS enzymes studied to date have an absolute requirement for divalent cations and ATP. One or more of the regions of similarity may represent elements which would be expected to be at or near the active site to facilitate conformation change upon binding ATP [33].

A single 2.0 kb band was observed in Northern analyses of wild type 74-OR23-1*A* RNA using a fragment of the genomic *met-6*$^+$ gene as a probe (Fig. 4). Although knowledge about the regulation of *Neurospora* FPGS expression is limited, the data presented in Fig. 4 suggest that the level of the *met-6*$^+$ transcript varied based on the age of mycelium. By examining blots of total RNA isolated from 0, 6, 12, 18 and 24 h cultures, higher transcript levels were detected in freshly germinated mycelium, with less expression seen in ungerminated conidia and cultures near stationary phase (Fig. 4).

In order to characterize the *met-6*$^+$ transcript for the positions of introns and transcription initiation–termination sites, *met-6*$^+$ cDNA clones were isolated and sequenced from two *N. crassa* 74-OR23-1*A* cDNA libraries. Three clones, each 1.7 kb long, were obtained containing poly(A) tails and sequence matching plasmid pIA50. The 5′ end of the three cDNA clones started 243 nucleotides downstream from the AUG codon of the largest open reading frame (Fig. 2, * at position +182); all three clones had the same 5′ and 3′ ends. Definitive identification of the transcriptional start site(s) was not achieved. It appears that the cDNAs isolated are truncated and do not contain the entire gene. Comparison to other FPGS sequences indicated amino acid similarity in the upstream regions beyond the 5′ end of the cDNA clones. In addition, cotransformations of *met-6*$^+$ mutant spheroplasts with each of the cDNA clones and pSV50 was unable to restore

Fig. 3. Comparison of the amino acid sequences of eukaryotic and prokaryotic FPGS proteins. The aligned sequences are: (1) 74-OR23-1*A* wild-type *Neurospora crassa* FPGS (present study), 528 residues, (2) yeast FPGS (Shane, personal communication), 540 residues, (3) human FPGS [8], 587 residues, (4) mouse FPGS [31], 587 residues, (5) *E. coli* FPGS [32], 365 residues and (6) *L. casei* FPGS [6], 423 residues. The alignment was made employing a clustal matrix (PCGENE™). Periods (·) indicate identity with the 74-OR23-1*A* sequence at a site. Asterisks (*) indicate amino acid identity among the six proteins, a position that is perfectly conserved. A quotation mark (′) indicates a position that is conserved in most cases. Underlined protein sequence indicates similarity to the A and B nucleotide binding domains. Translation start site 1 is indicated by the symbol ↓ at position +1 of the largest open reading frame. The symbol V indicates the position of translation start site 2, 102 nucleotides smaller than the largest open reading frame. The symbol ⇓ indicates the site of the *met-6* mutation.

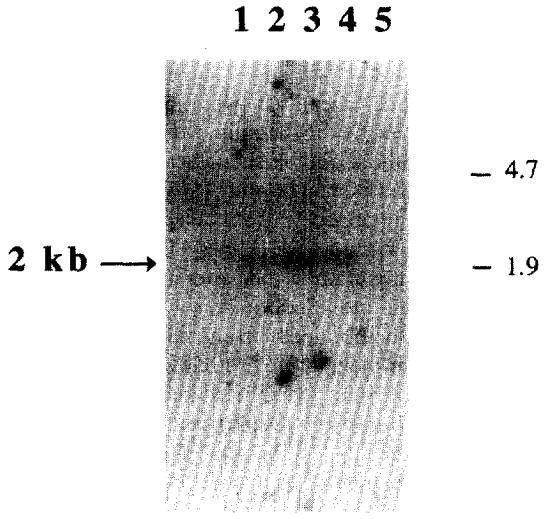

Fig. 4. Northern blot analyses of *N. crassa* total RNA. Total RNA was isolated from wild type 74-OR23-1*A* mycelia. The RNA was probed with a ^{32}P-labeled *Kpn*I/*Bam*HI fragment from pIA68 to detect *met-6*$^+$ RNA. Lane 1, 50 μg total RNA, 0 h germination prior to harvest; lane 2, 6 h germination prior to harvest; lane 3, 12 h germination prior to harvest; lane 4, 18 h germination prior to harvest; lane 5, 24 h germination prior to harvest. RNA levels were determined to be approximately the same in all lanes (not shown).

prototrophy. This lack of transformation potential is not simply the consequence of a lack of promoter sequence on cDNA function, as full length cDNAs devoid of promoter sequence can restore function in the appropriate mutants of other *Neurospora* genes ([34]; unpublished observations). Attempts to isolate the 5′ region of the cDNA using RACE were not successful. Sequence at the 3′ end of each of the three cDNA clones revealed that the transcript ended 271 nucleotides beyond the TGA stop codon of the largest open reading frame (Fig. 2, * at position +1858).

The sequence of the cDNA clones confirmed the position of the presence of the second intron shown on Fig. 2. The cDNA sequence did not extend far enough upstream to confirm the position of the first intron. The boundaries of the genomic sequence at the site of the introns appeared to have strong homology with the consensus sequence of *Neurospora* intron/exon boundaries (Table 1) [25, 26]. The remainder of the nucleotide sequence and deduced amino acid sequence corresponded to *met-6*$^+$ sequence derived from genomic DNA.

Cloning and sequencing of the mutant gene

In order to identify the mutant site in the *met-6* strain, PCR-amplification of mutant genomic DNA was performed. Products of 2.2 kb were generated by PCR using VENT DNA polymerase from oligonucleotide primers located outside the largest open reading frame (Fig. 2, underlined) and the PCR products were cloned into the vector pGEM-T. As expected, cotransformation of the cloned *met-6* mutant fragment with the cosmid pSV50 for antibiotic selection was unable to rescue *met-6* mutant spheroplasts to prototrophy.

The complete sequence of the 2.2 kb mutant clone was determined and compared with the 74-OR23-1*A* wild-type *met-6*$^+$ gene sequence of plasmid pIA50. Numerous changes were detected (Table 2). Because the origin of the mutant strain has a varied and confusing history of crosses and back crosses, sequence characterization of additional wild-type strains was performed. In all cases where the *met-6* gene was sequenced from wild-type strains, this was done directly on the PCR products without prior cloning. This analysis revealed that most of the changes seen in the *met-6* mutant gene polymorphisms as corresponding changes were present in at least one of the additional wild-type strains (Table 2). Sequence analy-

Table 2. Polymorphisms identified by sequence analysis

Position on sequence (Fig. 2)	Strains			
	74-OR23-1*A*	Emerson	Lindegren*A*	*Met-6* mutant
+63	CT*C* (Leu)	CT*A* (Leu)	CT*A* (Leu)	CT*A* (Leu)
+94*	*A*AA (Lys)	*C*AA (Gln)	*C*AA (Gln)	*C*AA (Gln)
+126	TA*C* (Tyr)	TA*T* (Tyr)	TA*T* (Tyr)	TA*T* (Tyr)
(intron 1 + 9)	*A*	*G*	*G*	*G*
(intron 1 + 14)	*C*	*T*	*T*	*T*
(intron 1 + 18)	*T*	*C*	*C*	*C*
(intron 1 + 19)	*T*	*C*	*C*	*C*
(intron 1 + 20)	*A*	*C*	*C*	*C*
+319*	G*G*C (Gly)	*A*GC (Ser)	*A*GC (Ser)	*A*GC (Ser)
+323*	A*C*C (Thr)	A*T*C (Ile)	A*T*C (Ile)	A*T*C (Ile)
+372	GG*T* (Gly)	GG*C* (Gly)	GG*C* (Gly)	GG*C* (Gly)
+378	CC*T* (Pro)	CC*C* (Pro)	CC*C* (Pro)	CC*C* (Pro)
+411	CT*C* (Leu)	CT*C* (Leu)	CT*A* (Leu)	CT*A* (Leu)
+414	AT*T* (Ile)	AT*T* (Ile)	AT*C* (Ile)	AT*C* (Ile)
+510	TC*T* (Ser)	TC*C* (Ser)	TC*T* (Ser)	TC*T* (Ser)
+677*	G*T*C (Val)	G*C*C (Ala)	G*C*C (Val)	G*T*C (Val)

Asterisks (*) indicate positions where base substitution has resulted in an amino acid change. Nucleotides at positions that vary among the different strains are italicized.

sis comparing 74-OR23-1*A*, the Lindegren *A* and Emerson wild-type strains revealed sixteen sequence differences between the three (Table 2). Of these differences, all were base substitutions. Five were present in the first intron (Table 2). Of the remaining eleven changes in the long open reading frame, seven were silent at the amino acid level, while four resulted in amino acid changes (Table 2). These four significant amino acid alterations did not produce a phenotypic change between the different wild-type strains.

Upon examination of the mutant sequence, only one change, a T to C transition at position +400 of the open reading frame (Fig. 2, ⇓ and square box) was identified to which the mutant phenotype could be attributed relative to all three wild-type strains. This alteration in the *met-6* sequence resulted in a serine to proline change at amino acid 134 (Figs 2 and 3). This occurs in a region of the protein that was highly conserved with other species, between the A and B nucleotide binding domains (Fig. 3, underlined). The site of the *met-6* mutation appears to be conserved throughout all FPGS genes studied to date. Although the mutant sequence was derived from a cloned PCR fragment generated using Vent DNA polymerase, which contains a proofreading activity, it was still conceivable that this dramatic change was due to an error that occurred during the PCR. To confirm that the Ser to Pro change was not an artifact, we also directly sequenced PCR products derived from *met-6* strain DNA without prior cloning. This analysis confirmed the existence of the Ser to Pro mutation in the *met-6* strain.

A highly similar amino acid sequence was also observed (not shown) between *Neurospora* FPGS and the *Saccharomyces cerevisiae* dihydrofolate synthetase gene (Shane, personal communication). Similarity was observed in the region of the *met-6* mutation, as well as at the A and B nucleotide binding sites. The dihydrofolate synthetase protein in *Saccharomyces cerevisiae* performs a similar reaction to FPGS, the ATP-dependent addition of an L-glutamate to dihydropteroate. Future work with *in vitro* mutagenesis of the *Neurospora met-6*$^+$ gene now appears warranted for identification of the amino acids located at the putative active site and identification of mutant proteins with altered binding and/or catalytic properties.

In the broader context of one-carbon metabolism, the characterization of FPGS is a crucial step in understanding the regulation of folate formation and utilization. The identification of the *met-6*+ gene provides a genetic tool for studying control and regulation of FPGS expression. The observation that decreased levels of human FPGS occurs in cell lines resistant to antifolates [5] is substantial evidence that further study may facilitate the identification of the molecular basis of the resistant phenotypes and may have important clinical implications both within and beyond folate metabolism.

EXPERIMENTAL

Materials

Reagent-grade chemicals were purchased from Sigma Chemical Co. (St. Louis, MO) and from the Edmonton offices of Fisher Scientific, Baxter-Canlab, BDH Chemicals and Terochem. [U-^{3}H]L-glutamate, obtained from Amersham-Searle Corporation, Arlington Heights, IL, was diluted with carrier L-glutamate to give a specific radioactivity of 25 Ci/mol. Sequenase™ DNA sequencing kits were also obtained from Amersham-Searle. Sequencing primers were purchased from New England BioLabs or were prepared by the DNA Synthesis Laboratory, Department of Biological Sciences, University of Alberta. Restriction endonucleases were purchased from Bethesda Research Laboratories (BRL), Boehringer Mannheim or New England BioLabs. Fuji X-ray film or Kodak XAR-5 was used for autoradiography.

Strains and culture conditions

Three wild-type strains of *N. crassa* (Lindegren *A*, FGSC 853; 74-OR23-1*A*, FGSC 987 and Emerson, FGSC 424) and the methionine-requiring mutant strain *met-6* (FGSC 1330) were maintained and cultured in liquid or solid Vogel's medium [35]. *Met-6* mutant cultures received supplements of 0.2 mM L-methionine and the transformant strain m6-TIA50 was grown in the presence of benomyl (0.5 $\mu g \cdot ml^{-1}$). The control transformant strain m6-pSV50 was grown on benomyl and methionine. Bacterial plasmids pUC19 and pGEM-T, as well as recombinant clones were maintained in *E. coli* strains JM 83 or XL1-Blue.

Subcloning

Cloning methods, restriction endonuclease digestion, cDNA library screening, cosmid and plasmid isolations, *E. coli* transformation and ligations, were performed as described [36–38] or according to supplier's instructions. The generation of plasmid pIA50 containing a 3.5 kb *Eco*RI/*Bam*HI digested restriction fragment capable of transforming *met6* mutant spheroplasts to prototrophy is previously described [21]. The recombinant plasmids pIA62 and pIA68 are the *Eco*RI/*Kpn*I (0.7 kb) and *Kpn*I/*Bam*HI (2.7 kb) fragments of the *Eco*RI/*Bam*HI insert of pIA50, respectively [21]. The following previously published procedures were performed: transformation of mutant *N. crassa* spheroplasts [39] with modifications [24] and *N. crassa* genomic DNA isolation [40].

Southern blot analyses

10 μg aliquots of genomic DNA from each strain were digested, electrophoresed and blotted to Biotrans nylon membrane (ICN Biomedical) following the procedure for Southern analysis [41].

Restriction fragment or PCR amplification products used as hybridization probes were excised and purified from agarose gels using sodium iodide/glass [42]. Products were either uniformly labeled using random oligonucleotide primer extension in the presence of ^{32}P-dCTP, [43, 44], or non-radioactively labeled using digoxigenin labelling (BMC instructions). Southern blots were screened using labeled 2.7 kb *Kpn*I/*Bam*HI fragment isolated from pIA68. Prehybridization, hybridization and washing of membranes were carried out at 42 or 65°C according to the ICN protocol. Autoradiography was performed for 24 to 48 h at −70°C in the presence of intensifying screens.

Polymerase chain reaction (PCR)

For rapid small scale isolation of genomic DNA template for PCR, conidial DNA extracts were prepared. Small conidial samples were suspended in sterile 1 M sorbitol, 20 mM EDTA and 3 mg/ml Lysing enzyme (Sigma) followed by a 15 min incubation at 37°C. The samples were pelleted, washed twice with 1 M sorbitol and 20 mM EDTA before resuspension in ddH_2O. After boiling for 10 min, an aliquot of the resulting preparation was used as template in the PCR reaction. PCR fragments of mutant and wild-type genomic DNA were generated using Taq DNA polymerase for direct sequencing or generated using VENT DNA polymerase for cloning and transformation. Reactions were carried out in an AutoGene Thermal Cycler or a Stratagene Robocycler 40. When PCR products were to be cloned, pGEM-T (Promega) was used as the cloning vector following standard methodology of the manufacturer.

DNA sequence determination

All DNA sequencing was determined by the dideoxy chain termination method [45] using Sequenase™ sequencing kits and [α-^{35}S]dATP. Sequencing performed on double-stranded plasmids was according to manufacturer's instructions. PCR amplification products were sequenced with modifications to the manufacturer's instructions. PCR products isolated from 0.8% agarose gels were resuspended in distilled water. Approximately 1–2 μg of DNA solution combined with 100 ng of a specific primer was denatured before annealing using immediate flash cooling in a dry ice/ethanol bath. All sequencing reactions were loaded onto 6 or 8% (w/v) polyacrylamide gels and electrophoresed. In regions where compressions were encountered, labelling and termination mixes containing dITP instead of dGTP were used to sequence the regions [46].

Preparation of RNA, electrophoresis and analysis

To inactivate RNAses, all glassware used for RNA isolation was baked at 180°C for 3 h and all solutions used were treated with 0.1% diethyl pyrocarbonate for a minimum of 12 h prior to autoclaving. Mycelium from the wild-type strain 74-OR23-1*A* was harvested by filtration, washed in distilled water and immediately frozen in liquid nitrogen. Isolation of total RNA and Northern blotting was according to standard techniques [38]. For Northern blots, 50 μg total RNA from wild-type strain 74-OR23-1*A* was electrophoresed and blotted. Following overnight transfer, the blot was baked at 80°C for 1 h and hybridized to ^{32}P-containing DNA probes of the *KpnI*/*BamHI* fragment of pIA68. Blots were hybridized, washed and exposed to X-ray film for 48–168 h at −70°C. The 5′-RACE (rapid amplification of cDNA ends) techniques were performed using a 5′-RACE kit (Bethesda Research Laboratories) following manufacturer's instructions using 10–50 μg of total RNA from wild-type strain 74-OR23-1*A*.

Isolation of met-6$^+$ *cDNA clones from the λ-ZAP library*

N. crassa cDNA libraries, made in our laboratory or generously provided by Dr Matthew Sachs, Stanford, California, were constructed in the λ-ZAP/*Eco*RI vector (Stratagene). The cDNA library clones (5×10^5 pfu per plate; 10–150 mm plates) were lifted onto Biotrans nylon membranes (ICN Biomedicals) following the protocol for plaque lifts. Products used for probes were either uniformly radioactively labeled using random oligonucleotide primer extension in the presence of ^{32}P-dCTP [43, 44] or non-radioactively labeled using digoxigenin labelling (BMC instructions). The libraries were screened using hybridization probes of the 2.7 kb *Kpn*I–*Bam*HI fragment from pIA68 or the 0.7 kb *Eco*RI–*Kpn*I fragment from pIA62. Hybridization to and washing of filters were done following the Biotrans protocol. Plaques that gave a positive hybridization signal were further purified and re-screened until single, well isolated plaques could be identified. Each phage stock was titered and subjected to *in vivo* excision. Recombinant plasmids were isolated and these clones were sequenced.

Acknowledgements—The authors are grateful to Dr Tom Schmidhauser, Southern Illinois University, for the gift of the cosmid clone pSV50-4:11:C and to R. Gary Ritzel for helpful discussions. This work was supported by grants from the National Sciences and Engineering Research Council of Canada to E. A. C. and F. E. N. GenBank Accession Number: the reported DNA sequence of pIA50 has been submitted to GenBank (accession number AF005040).

REFERENCES

1. Blakley, R. L. and Benkovic, S. J., in *Folates and Pterins*. Wiley Interscience, New York, 1984, p. xi.

2. Blakley, R. L., in *The Biochemistry of Folic Acid and Related Pteridines*, eds. A. Neuberger and E. L. Tatum. North Holland Publishing Co., Amsterdam, 1969, p. 92.
3. MacKenzie, R. E., in *Folates and Pterins*, eds. R. L. Blakley and S. J. Benkovic. Wiley Interscience, New York, 1984, p. 255.
4. Cossins, E. A., in *Folates and Pterins*, Vol. 1, eds. R. L. Blakley and S. J. Benkovic. Wiley, New York, 1984 p. 1.
5. Shane, B., *Vitamins Hormones*, 1989, **45,** 263.
6. Bognar, A. L., Osborne, C., Shane, B., Singer, S. C. and Ferone, R., *J. Biol. Chem.*, 1985, **260,** 5625.
7. Sussman, D. J., Milman, G. and Shane, B., *Somatic Cell Mol. Genet.*, 1986, **12,** 531.
8. Garrow, T. A., Admon, A. and Shane, B., *Proc. Natl. Acad. Sci. U.S.A.*, 1992, **89,** 9151.
9. Garrow, T. A. and Shane, B., in *Chemistry and Biology of Pteridines and Folates*, eds. R. L. Blakley and S. J. Benkovic. Plenum Press, New York, 1993, p. 659.
10. Osborne, C. B., Lowe, K. E. and Shane, B., *J. Biol. Chem.*, 1993, **268,** 21657.
11. Shane, B., Garrow, T., Brenner, A., Chen, L., Choi, Y.-J., Hsu, J. C. and Stover, P., in *Chemistry and Biology of Pteridines and Folates*, eds. R. L. Blakley and S. J. Benkovic. Plenum Press, New York, 1993, p. 629.
12. Lowe, K. E., Osborne, C. E., Lin, B. F., Kim, J.-S., Hsu, J.-C. and Shane, B., *J. Biol. Chem.*, 1993, **268,** 21665.
13. Selhub, J., Ph.D. Thesis, Case Western Reserve University, University Microfilms International, Ann Arbor, Michigan, 1970.
14. Sakami, W., Ritari, S. J., Black, C. W. and Rzepka, J., *Fed. Proc.*, 1973, **32,** 471.
15. Ritari, S. J., Sakami, W., Black, C. W. and Rzepka, J., *Neurospora Newslett.*, 1973, **20,** 26.
16. Ritari, S. J., Sakami, W. and Black, C. W., *Neurospora Newslett.*, 1973, **20,** 27.
17. Cossins, E. A. and Chan, P. Y., in *Folyl and Antifolyl Polyglutamates*, eds. D. Goldman, B. Chabner and J. R. Bertine. Plenum Press, New York, 1983, p. 183.
18. Cossins, E. A. and Chan, P. Y., *Phytochemistry*, 1984, **23,** 965–971.
19. Cossins, E. A. and Chan, P. Y., *Plant Cell Physiol.*, 1985, **26,** 331.
20. Perkins, D. D., Radford, A., Newmeyer, D. and Bjorkman, M., *Microbiol. Rev.*, 1982, **46,** 426.
21. Atkinson, I. J., Nargang, F. E. and Cossins, E. A., *Phytochemistry*, 1995, **38,** 603.
22. Vollmer, S. J. and Yanofsky, C., *Proc. Natl. Acad. Sci. U.S.A.*, 1986, **83,** 4869.
23. Murray, N. E., *Genetics*, 1969, **61,** 67.
24. Akins, R. A. and Lambowitz, A. M., *Mol. Cell. Biol.*, 1985, **5,** 2272.
25. Edelmann, S. E. and Staben, C., *Exp. Mycol.*, 1994, **18,** 70.
26. Bruchez, J. J. P., Eberle, J. and Russo, V. E. A., *Fungal Genet. Newslett.*, 1993, **40,** 89.
27. Chan, P. Y., Dale, P. L. and Cossins, E. A., *Phytochemistry*, 1991, **30,** 3525–3531.
28. Gurr, S. J., Unkles, S. E. and Kinghorn, J. R., in *Gene Structure in Eukaryotic Microbes*, ed. J. R. Kinghorn. IRL Press, Oxford, 1987, p. 93.
29. Lin, B.-F. and Shane, B., *J. Biol. Chem.*, 1994, **269,** 9705.
30. McDonald, D., Atkinson, I. J., Cossins, E. A. and Shane, B., *Phytochemistry*, 1995, **38,** 327.
31. Roy, K., Mitsugi, K. and Sirotnak, F. M., *J. Biol. Chem.*, 1996, **271,** 23820.
32. Bognar, A. L., Osborne, C. and Shane, B., *J. Biol. Chem.*, 1987, **262,** 12337.
33. Toy, J. and Bognar, A. L., *J. Biol. Chem.*, 1990, **265,** 2492.
34. Nargang, F. E., Kunkele, K.-P., Mayer, A., Ritzel, R. G., Neupert, W. and Lill, R., *EMBO J.*, 1995, **14,** 1099.
35. Davies, R. H. and de Serres, F. J., *Methods Enzymol.*, 1970, **17A,** 79.
36. Maniatus, T., Fritsch, E. F. and Sambrook, J., in *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbour Laboratory, Cold Spring Harbour, New York, 1982.
37. Sambrook, J., Fritsch, E. F. and Maniatus, T., in *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbour Laboratory, Cold Spring Harbour, New York, 1989.
38. Ausubel, R. A. R., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K., in *Current Protocols in Molecular Biology*. Greene and Wiley Interscience, New York, 1987.
39. Schweizer, M., Case, M. E., Dykstra, C. C., Giles, N. H. and Kushner, S. R., *Proc. Natl. Acad. Sci. U.S.A.*, 1981, **78,** 5086.
40. Schectman, M., *Fungal Genet. Newslett.*, 1986, **33,** 45.
41. Southern, E., *Methods Enzymol.*, 1979, **68,** 152.
42. Vogelstein, B. and Gillespie, D., *Proc. Natl. Acad. Sci. U.S.A.*, 1979, **76,** 615.
43. Feinberg, A. P. and Vogelstein, B., *Anal. Biochem.*, 1983, **132,** 6.
44. Feinberg, A. P. and Vogelstein, B., *Anal. Biochem.*, 1984, **137,** 266.
45. Sanger, F., Nicklen, S. and Coulson, A. R., *Proc. Natl. Acad. Sci. U.S.A.*, 1977, **74,** 5463.
46. Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F. and Petersen, G. B., *J. Mol. Biol.*, 1982, **152,** 729.