

# Identification of a gene encoding GMP synthetase from a *Neurospora crassa* cDNA library by bacterial complementation

Jeffrey A. Smiley<sup>1</sup> and David K. Asch<sup>2</sup>. Departments of Chemistry<sup>1</sup> and Biological Sciences<sup>2</sup>, Youngstown State University, Youngstown, OH 44555

We report the isolation and identification of a gene encoding GMP synthetase from a *Neurospora crassa* cDNA library. Phage infection of the purine-requiring *Escherichia coli* strain SØ3834 using the NO<sub>3</sub><sup>-</sup> induced cDNA phage library from the Fungal Genetics Stock Center resulted in colonies able to grow on minimal media with no added purine source. A plasmid, termed pGMPS1, was isolated from one of these colonies and shown to reproducibly support growth of strain SØ3834 in the absence of purines in the media. Identification of this gene as one encoding GMP synthetase is confirmed by DNA sequencing and comparison to the known *guaA* gene from yeast.

Nucleotide biosynthetic enzymes are common targets for anti-nucleic acid drug therapy, and deficiencies of the corresponding genes in humans are implicated in a number of inherited metabolic disorders. GMP synthetase (EC 6.3.5.2) has implications in anticancer and immunosuppressive therapies; additionally, this enzyme is used industrially in the semi-enzymatic synthesis of GMP (Fujio *et al.* 1997, *Bioscience, Biotechnology, and Biochemistry* 61: 840-845). The enzyme is a member of the amidotransferase family, which utilizes glutamine hydrolysis as a source of ammonia for a subsequent enzymatic step, and catalyzes the replacement of O<sub>2</sub> in XMP with an amino group to yield GMP.

To extend the collection of GMP synthetase genes from different organisms, we attempted isolation of the gene from *Neurospora crassa* using bacterial complementation. The auxotrophic *Escherichia coli* strain, SØ3834 (Chang *et al.* 1991, *Biochemistry* 30: 2273-2280), lacks both genes for GMP synthetase (*guaA*) and adenosine deaminase (*add*), and requires a purine source that can be metabolized to GMP for growth. In *guaA*<sup>-</sup> *add*<sup>-</sup> strains, the purine source can be 2,6-diaminopurine, since this compound can be enzymatically ribosylated and subsequently deaminated by adenosine deaminase to guanosine, but in SØ3834, 2,6-diaminopurine is incapable of providing the purine requirement. Thus, colonies in the complementation screen that arise on media containing 2,6-diaminopurine could contain either the *guaA* or *add* genes, but those that arise on purine-free media could only contain *guaA*.

From colonies arising in a complementation screen of about 10<sup>5</sup> phage [NO<sub>3</sub><sup>-</sup> induced cDNA phage library, Fungal Genetics Stock Center; *Fungal Genetics Newsletter*, 45 (Supplement), 8], one colony yielded plasmid DNA from an alkaline lysis preparation that, when reinserted into the host strain, provided the ability to grow on purine-free media. This plasmid was presumed to contain the GMP synthetase gene and was designated pGMPS1. Sequencing of the inserted region within the vector was carried out at the Automated DNA Sequencing Facility, University of North Carolina, Chapel Hill, using the M13 (-20) primer and primers designed from newly acquired sequence information. An open reading frame of 1629 nucleotides was identified; when translated into the amino acid sequence, a high degree of similarity was found with the yeast GMP synthetase gene (duJardin *et al.* 1994, *Gene* 139:127-132). Of the 543 amino acids from the deduced sequence, 354 (65%) have identical counterparts in the yeast sequence (Figure 1). The high degree of sequence similarity with the yeast GMP synthetase gene and the ability of the gene to support growth of *E. coli* strain SØ3834 in the absence of purines confirm the identification of a GMP synthetase gene from an *N. crassa* library.

We thank Dr. Per Nygaard, Department of Biological Chemistry, University of Copenhagen, for providing *E. coli* strain SØ3834. Contributions to this work from Maria T. Kovach and Jennifer R. Shook are gratefully acknowledged.

Figure 1. Amino acid sequence comparison of GMP synthetases. Yeast sequence is shown in light type, sequence from *N. crassa* is bold. Vertical dashes between sequences indicate identical amino acids.

|            |            |            |            |            |            |     |
|------------|------------|------------|------------|------------|------------|-----|
| MAA        | GEQV       | SNMFDTILVL | DFGSQYSHLI | TRRLREFNIY | AEMLPCTQKI |     |
|            |            |            |            |            |            |     |
| MAAATLGEVP | TKAFDTILTL | DFGSQYTHLI | TRRMRELNIL | SEMLPCTTKI |            |     |
| SELGWTPKGV | ILSGGPYSVY | AEDAPHVDHA | IFDLNVPILG | ICYGMQELAW | 97         |     |
|            |            |            |            |            |            |     |
| ADLDYKPKGV | ILSGGPYSVY | EDGAPHVDPA | VFELGVPILG | ICYGMQEIWA | 100        |     |
| INGKQVGRG  | DKREYGPATL | KVI        | DDSNS      | LFKGMNDS   | TVWMSHGDK  |     |
|            |            |            |            |            |            |     |
| RASPENVIAG | VHREYGHNSL | KALKGDDAHV | DRLFAGLEDG | MRVWMSHGDK |            |     |
| LHGLPTGYKT | IATSDNSPYC | GIVHETKPIY | GIQFHPEVTH | STQGKTLLKN | 191        |     |
|            |            |            |            |            |            |     |
| LGALPEGFHT | VAVSDNSEYA | AIAHKTKPIY | GLQFHPEVTH | SQNGTQLLKN | 200        |     |
| FAVDLCHAKQ | NWTMENFIDT | EINRIRKLVG | PTAEVIGAVS | GGVDSTVASK |            |     |
|            |            |            |            |            |            |     |
| FAVDICGCAQ | NWTMARFLDQ | EIARIRDLVG | PEGQVLGAVS | GGVDSTVAAK |            |     |
| LMTEAIGDRF | HAILVDNGVL | RLNEAANVKK | TLVEGLGINL | MVVDASEEFL | 291        |     |
|            |            |            |            |            |            |     |
| LMKEAIGDRF | WAVLVNNGVM | RLDECEQVER | DLKQHLGINL | TVIDASKDFL | 300        |     |
| SKLKGVTDPK | KKRKIIGNTF | IHFVEREAEK | IKP        | KDGK       | EIQFLLQG   |     |
|            |            |            |            |            |            |     |
| EGLKGLHDPE | QKRKFIGGKF | IDVFEAEAQK | IEEAAKSGK  | GTKIGFFLQG |            |     |
| TLYPDVIESI | SFKGPSQTIK | THHNVGGLLE | NMK        | LKL        | IEPLRELFKD | 382 |
|            |            |            |            |            |            |     |
| TLYPDVIESL | SFKGPSATIK | THHNVGGLPE | RMTNGQGLQL | IEPLRSLYKD | 400        |     |
| EVRHLGELLG | IPHDLVWRHP | FPGPGIAIRV | LGEVTKEQVE | IARKADNIYI |            |     |
|            |            |            |            |            |            |     |
| EVRELGRITL | IHEELVMRHP | FPGPGIAVRI | LGEVTEEKVR | IARQADHIFI |            |     |
| EEIKKAGLYN | QISQAFACLL | PVKSVGVMGD | QRTYDQVIAL | RAIETTDGMT | 482        |     |
|            |            |            |            |            |            |     |
| SEIRKAGLYD | QISQAYAAVD | PSRAVGVMGD | KRVYGYIIIL | RAVTTTDMGT | 500        |     |
| ADWFFFEHSF | LKKVASRIVN | EVDGVARVTY | DITSKPPATV | EWE        | 525        |     |
|            |            |            |            |            |            |     |
| AEAFNFPWDF | LQRMNRIVN  | EVNGVCRVTY | DITSKPPGTI | ELE        | 543        |     |