

CLONING AND DISRUPTION OF THE *cefG* GENE ENCODING ACETYL
COENZYME A: DEACETYLCEPHALOSPORIN C O-ACETYLTRANSFERASE
FROM *ACREMONIUM CHRYSOGENUM*

Akio Matsuda, Hiroyosi Sugiura, Kenji Matsuyama, Hideyuki Matsumoto,
Sigeaki Ichikawa and Ken-ichi Komatsu

Pharmaceutical Research and Development Department, Asahi Chemical Industry Co., Ltd.,
2-1 Samejima, Fuji 416, Japan

Received May 26, 1992

SUMMARY: Acetyl CoA: deacetylcephalosporin C o-acetyltransferase(DCPC-ATF) catalyses the final step in the biosynthesis of cephalosporin C (CPC) in *Acremonium chrysogenum*. The gene encoding DCPC-ATF, *cefG*, has been isolated from an *A. chrysogenum* genomic library using a DCPC-ATF cDNA probe. Nucleotide sequence analysis revealed that *cefG* contains two short introns of 79bp and 65bp. The gene was found to be closely linked to the *cefEF* gene encoding deacetoxycephalosporin C synthetase / deacetylcephalosporin C synthetase, which catalyses the preceding two steps in the pathway. The two genes are separated by a 1114 bp segment from which they are divergently transcribed. Introduction of the cloned *cefG* gene to *A. chrysogenum* resulted in an increased level of DCPC-ATF activity. A plasmid carrying a *cefG* gene interrupted in the coding region by a selectable marker for resistance to hygromycin B was constructed and used to disrupt the *cefG* locus in *A. chrysogenum*. The *cefG*-disrupted strains were found to lack the ability to produce CPC, and accumulated its precursor, deacetylcephalosporin C in the culture broth. Southern hybridization analysis confirmed that the disruption resulted from a gene replacement event at the *cefG* locus.

© 1992 Academic Press, Inc.

Cephalosporin C biosynthesis in *Acremonium chrysogenum* proceeds through six steps in which five enzymes involved (1). The last enzyme of the pathway which converts deacetylcephalosporin C (DCPC) to cephalosporin C (CPC) is acetyl CoA: deacetylcephalosporin C o-acetyltransferase (DCPC-ATF) and coded by the named *cefG* gene (1,2). Recently, the enzyme has been purified to near homogeneity from *A. chrysogenum* IS-5 and shown to be composed of two nonidentical subunits with molecular weights of 27000 and 14000, respectively (submitted elsewhere). Subsequently a cDNA encoding DCPC-ATF has been isolated from an *A. chrysogenum* cDNA library and expressed in a heterologous host, *S. cerevisiae*, with a functional DCPC-ATF activity. The amino acid sequence deduced from the cDNA revealed the existence of a precursor protein which is processed to the two constituent subunits of DCPC-ATF. The primary structure of DCPC-ATF precursor protein showed a significant similarity with those of acetyl CoA: homoserine o-acetyltransferase from other fungi (3). The availability of a cloned cDNA provided a probe for the *cefG* gene. In this paper we report the cloning and sequence of the *cefG* gene from *A. chrysogenum* IS-5. We also provide evidence that *cefG*

is closely linked to *cefEF* which was previously identified (4,5) on chromosome II in *A.chrysogenum*. Furthermore we describe the disruption of *cefG* locus by one step gene replacement method using the cloned gene.

MATERIALS AND METHODS

Cloning, hybridization and sequencing. Genomic DNA of *A.chrysogenum* strain IS-5 (3) was extracted according to the method of Johnston et al.(6). An *A.chrysogenum* genomic library was constructed by inserting the products of a partial MboI digest of chromosomal DNA in the BamHI site of λ EMBL3 and screened by using 32 P-labeled 1.25 Kb EcoRI fragment (DCPC-ATF cDNA insert) from pCCS1(3) as a probe. Plaque and Southern hybridizations were essentially carried out as described by Maniatis et al (7). The desired DNA fragments were subcloned in M13 vectors and sequenced by using the dideoxy chain termination method (8).

Plasmid construction. A 3.8Kb BglII-BamHI fragment from pATF1(see result and Fig.1) containing the entire *cefG* gene was treated with the Klenow fragment to make blunt ends and then inserted into the SmaI site of pACTHY83 to yield pTATF1, a plasmid for introduction of extra copies of the *cefG* gene. The pACTHY83, a plasmid for *A.chrysogenum* transformation, carries a hygromycinB resistance (HYB^R) expression cassette comprising bacterial hygromycinB phosphotransferase gene from pLG83 (9), and a putative promoter and terminator of *A.chrysogenum* actin gene. The actin gene was isolated from the above mentioned library by heterologous hybridization using the third exon region of human β actin gene as a probe (manuscript in preparation). pDATF1, a plasmid for disrupting the *cefG* locus in *A.chrysogenum*, was constructed as follows. A 3.7Kb SpeI fragment containing the entire *cefG* gene was isolated from pATF1 and inserted into the XbaI site of pUC18 to obtain pSATF1. A 3.2Kb SpeI-XbaI fragment from pACTHY83 containing the HYB^R expression cassette was inserted into the unique Eco47III site (cut at codon ARG₂₃₄ of the DCPC-ATF precursor protein) of pSATF1, which had been converted to the XbaI site by attachment of a corresponding linker, yielding pDATF1.

Transformation, culture condition and assays. Preparation of protoplasts from *A.chrysogenum* and transformation were substantially carried out according to the method of Skatrud et al (10). All transformants were selected on the plate containing 50 μ g/ml hygromycinB. Cultivation of *A.chrysogenum* including transformants was carried out in a 500ml Erlenmeyer flask containing 30ml of a fermentation medium (consisting of beet molasses, 3%; defatted soy bean, 4%; corn steep liquor, 1%; ammonium acetate, 0.5%; ammonium sulfate, 0.7%; calcium sulfate, 0.8%; calcium carbonate, 1.5%; HCl-hydrolysate of starch, 6%; and methyl oleate, 3%) at 25°C for 6 days on a rotary shaker (220 r.p.m.). Cell free extracts were prepared from mycelia grown in the medium for 3 days and assayed for DCPC-ATF activities and protein concentration as previously described (3). The amount of cephalosporins (DCPC and CPC) in the culture broth was determined by a high performance liquid chromatography system using ZORBAXBp NH₂ column.

RESULTS AND DISCUSSION

Cloning of the *cefG* gene. When 1.5×10^4 λ EMBL3 recombinant phages from an *A.chrysogenum* genomic DNA library were screened with a cDNA probe containing the entire coding sequence of DCPC-ATF(3), six positive clones were obtained. A 7Kb BamHI fragment which is common to three clones and strongly hybridizes to the cDNA probe was subcloned into the BamHI site of pUC18 to obtain pATF1 and mapped for several restriction site (Fig. 1). Further Southern analysis localized the hybridizing sequence to a 2.6Kb SacI-StuI fragment on the insert within pATF1, which was subsequently sequenced according to the strategy as shown in Fig. 1. The complete nucleotide sequence of the fragment and the translated amino acid sequence are depicted in Fig. 2. Comparison of cDNA and genomic DNA sequences revealed that the *cefG* coding region is interrupted by two short introns consisting of 79 and 65 bps, respectively. Both introns contain a typical eukaryotic 5'GT and

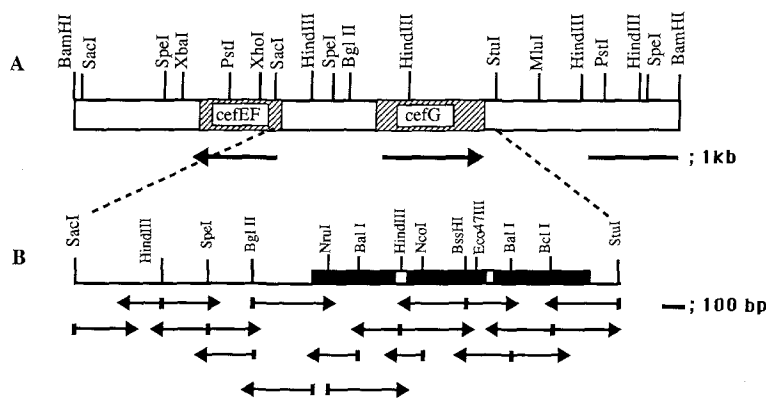


Fig.1. Restriction maps and sequencing strategy for the *cefG* gene. (A) Physical map of a 7Kb BamHI insert on pATF containing the *cefEF-cefG* gene cluster in *A.chrysogenum*. (B) Restriction map and sequencing strategy of a 2.6Kb SacI-StuI fragment containing the entire *cefG* gene. The coding regions and introns are indicated by solid boxes and open boxes, respectively. The arrows below the map indicate the direction and extent of sequencing.

3'AG splice site and internal sequence which show strong homologies with those identified (11) in other fungi as being important in lariat formation and subsequent splicing, as underlined in Fig. 2 . The sequence of the coding region was completely identical to that of the cDNA and encoded a DCPC-ATF precursor protein of 385 amino acid residues. A 3.8Kb BglII-BamHI fragment containing the entire

-1180 GAGCTCGGTGAGGACCTTGGCGCTCTTGAGGTCGTCGAGACGAAAGACGGGACCTTGGAGTCTATGTTGATGCTGTGCTTTGAGCGATGACTTGAGA
-1080 GAGTAGCGTGGAGGAAGTTCTGCAAGAGGAATTTAAGGATTACAGATCCAGTCGAGACGAAACGTTGTCAAAGCGGTATATATATCTCAAACCCACCTCGTAGCTTACGCCG
-960 GGAACCTCCTTTTATGACAACCTGCTACTTACGCCGTAAGTGACGCCCTGCTTCCCTCAGCCTTGGCCGCACACGTCATGTAGCATTTGTAACCCACGAGTGTCTTGTGAAGTTTGTGCA
(3'TCACTGC 5')

-840 CGAATCATAAGAGCCATCGAGTTCTCTCTCGTCTTGGTTCGACGAGAATATGTATCGTGCATGGTCCCTGATCGTCGAGACGCCCATGGAATCGTGCAAGCCTTAATCTCCGTAT
-720 AAGCTTCCCATTCGGACAAGATTGGCATGATGTGGATGCGGGCTCTTTTAATAAGGACCTTCTTAACCGATGTCGAGAGTGCCTAGGACGGGTCCATGTGCATACACGACGACCC
-600 CGACCTCCTATTAGGAGCATGAGGACGACAAAATGCGAAGCAGCATCAAAATGCATGCAACGTCGAGTTGTGGGCTACTCGCCTTCTGATTTCGAAGCCCTCGCGGATCCACAT
-480 TACTAGTAGCTTGGATAAACAAGCAAGTTTCGCCGCCAAAGGGCTGCCCGGATCCGATTTCGATGCCATGTACATCAAGTCGGAATGGTGTCTCCGTTTCCCGCTGGGTGAGAGG
-360 CGAAGGAGTAGTTGACCAATGCGACGCCACCCAGAGCCGACAGGCTTTATCGGATGTTGCTTCGATCCGATGCTATCCCGCGGCGCTAGATCTTGCTAATACGAGTCGGAGAGTTACT
-240 TTCCGGGCTTATGCGGACGGGCGGCCGCTCGATGCGGCCAAGGCTTGTCTGTCATGTATAGATGCTGCGCTCGGCCAAAGTGCGCCGCTCTAAGAGCGGACCTTTTCCCGCGAGTCT

-120 TCCCGCATCCCGCAGCGGGCGTCACTTTCCTGCGCCTCGCTCCTTGTGTAACCTACCTATATCTTCATCCCGCAATGCTGCGGATAGGCTACCTACAGCCACACGTCGCCACG
→ *cefG* start
1 ATGTCGCCTCAGATCGCAATCGCTTCGAGGCTTCGCTAGATGCCAAGACATAGCCAGAATATCGCTCTTACACTGGAATCTGGCGTCACTCTTCGCGATGTACCGCTGGCATACAA/
M S P Q I A N R F E A S L D A Q D I A R I S L F T L E S G V I L R D V P V A Y K
121 TCGTGGGTCCGATGAATGCTCAAGGATAACTGCTCATCGCTGCCACACCTTGACGACGAGCGGCCATGTACCTCGTGTGGGCCACACTGTTTGCCAAAGGACGGGCTTTCGAT
S W G R M N V S R D N C V I R C H T L T S S A H V T S W W P T L F G Q G R A F D
241 ACCTCTGCTACTTTCATCTGCTAAATATCTCGGGAGCCCTTGGGAGTGTGGACCATGTTACCGGACCCGATGAGAAAGCCGCGCCGTCACGGGCGCAAGTTTCCTCGC
T S R Y F I I C L N Y L G S P F G S A G P C S P D P D A E G Q R P Y G A K F P R
361 ACGACGATTCAGATGATGTG/GTAGGTAAGGCGACCGATCCAGCTTGTCTCAATATCGAGTGGTCAGGACAATCCAGGCTAAGCTTCCGTGTCCAAAAG/TATTCTATCGCCAGGT
T T I R D D V R T I H R Q V
479 CTCGACAGGTAGCGCTCAGCAAAATGCTGCGGTAGTCGGCGCATCCATGGGTGGAATGCACACTCTGGAATGGGCTTCTTGGTCCGAGTACGTGCGAAAGATTGTGCCCATCGC
L D R L G V R V Q I A A V V G A S M G M H T L E W A F F G P E Y V R K I V P I A
599 ACATATCGGCTCAGAGCGGTGTCGCGAGCTTGGTTGAGACACAGAGCGAGTGCATATGATGACCCAAAGTACCTGGACGGGAGTACGACGTACAGACCCAGCCTGTCGCGGGG
T S C R Q S G W C A A W F E T Q R Q C I Y D D D P K Y L D G E Y D V D D Q P V R G
719 CTCGAAACGCGCGCAAGATTGCGAATCTCAGCTACAAGAGCAAACTGCGATGGACGAGCGCTTCCATATGGCTCCAGGATCCCAAGCGG/ GTGAGTTTATAGATGCTTCCGCTCGG
L E T A R K I A N L T Y K S K P A M D E R F H M A P G V Q A G
837 TCGATGCTCAGAGCTAATCAGACCGAACCCGCTGTAG/ GCGGAAATATCAGACGCCAGGATGCGAAGAAGGAAATCAACGGCACAGACGCGCAACAGCCGCTGCTGCCACGCC
R N I S S Q D A K K E I N G T A D S G N S H R A G Q P
955 ATTGAAGCGTATCTTCTATCTCCGCTACCGGCCAGAAAGTTTGGCGGAGCTTCGAGCGCAACTGCTACATCGCCATGACACTCAAGTTGACACCCAGACATCAGCAGAGGCGCG
I E A V S S Y L R Y Q A Q K F A A S F D A N C Y I A M T L F D T H D I S R G R
1075 GCAGGATCAATCCGAGGCTCTGGCAATGATTACACAACAGCGCTTGATCATTGGCCAGGTCAGACGCTTGTACTCTGTTTTCGCGGACAGCTTTCAGATGGGCGCGAGTATCCCAAC
A G S I P E A L A M I T Q P A L I I C A R S D G L Y S F D E H V E M G R S I P N
1195 AGTCGCTTTTGGCGTGGACAGCAATGAGGCTCATGACTTCTTTGTAATGGAAGCGGACAAAGTTAATGATGCGGTCAGAGGATTCCTCGATCAGTCATTAATGTGAGGCTATGGAGGT
S R L C V V D T N E G H D F F V M E A D K V N D A V R G F L D Q S L M .

1315 GTCAGCCTGCCGCTGCGGCTACTTGCAGGGTGATCGATGACTCTCAGATAGTCTCCATGTGAGTATGGATTTCGGCTGTTTCCGCTCGGATATAGGCACTCTCAGGCCATCTCGCAGTA
1435 GGTATCAGAACAGCAGCTGAGGCT

Fig.2. Nucleotide sequence of the *cefG* gene of *A.chrysogenum* IS-5 including 5' and 3' untranslated regions. The nucleotide sequence is numbered from the first nucleotide of the presumed initiation codon of the *cefG* gene. The start codons of *cefEF* and *cefG* gene are overlined with an arrow, respectively. The underlined or broken-underlined sequences are discussed in the text. The intron -exon boundaries are indicated by slashes. The putative cleavage positions of DCPC-ATF precursor are indicated by vertical arrows.

Table.1. DCPC-ATF activities of *A.chrysogenum* IS-5 transformants with pTATF1 or pACTHY83

Strain (introduced plasmid)	DCPC-ATF activities (10 ³ U/mg-protein)
IS-5 (pACTHY83) 1	3.1
IS-5 (pACTHY83) 2	2.5
IS-5 (pTATF1) 1	9.4
IS-5 (pTATF1) 2	8.4
IS-5 (pTATF1) 3	9.7

coding region for DCPC-ATF was subcloned from pATF1 into pACTHY83, a vector which carries the hygromycinB resistance marker, to yield pTATF1 as described in Materials and Methods. This plasmid and vector pACTHY83 as a control were introduced into *A.chrysogenum* IS-5 and transformants selected for resistance to hygromycin B were assayed for DCPC-ATF activity. As shown in Table1, all of the examined transformants with pTATF1 exhibited 3-4 fold higher DCPC-ATF activities as compared to those with pACTHY83. The result indicates that the cloned *cefG* is a functional gene. It also suggests that the 272bp 5'-untranslated region extending from the initiation codon to the BglIII site contains some sequences required for the expression of *cefG*. In this region there are two pyrimidine tracts (underlined in Fig.2), a frequently observed fungal promoter element that has been shown to be involved in correct initiation of transcription (12), while neither canonical TATAA box nor CAAT box are observed. Interestingly, a sequence, 5'CGTCACT3', starting at 16 bp downstream the second pyrimidine tract is identical to the sequence around transcription start points of

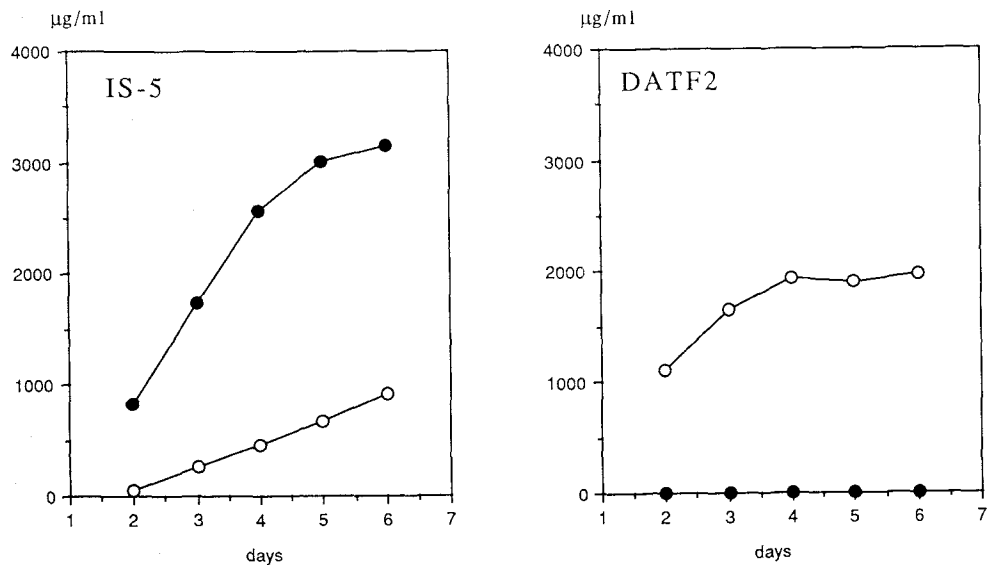


Fig.3. Production of CPC (●●●) and DCPC (○●○) by a *cefG* disrupted strain DATF2 and parental strain IS-5.

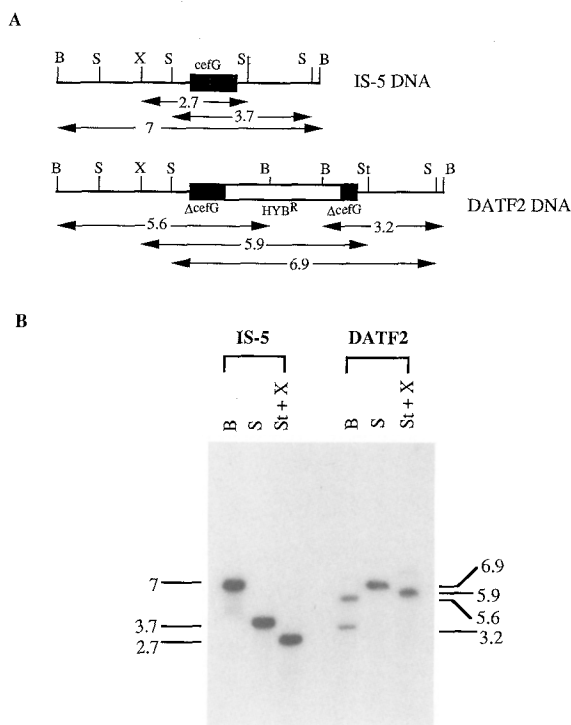


Fig.4. Disruption of the *cefG* gene. (A) Partial maps around the *cefG* gene with or without gene disruption. B, BamHI; S, SpeI; X, XhoI; St, StuI. The expected sizes of fragments hybridized to the cDNA probe are shown in kilobases. (B) Blot analysis of DNA from a *cefG* disrupted strain DATF2 and the parental strain IS-5. Genomic DNA were digested with BamHI (B), SpeI (S), or StuI plus XhoI (St + X), respectively. Samples were fractionated in 1% agarose gel and gel blot was hybridized with a 32 P-labeled 1.25kb EcoRI fragment. The sizes of fragments are indicated in kilobases.

A. chrysogenum *pcbC* gene encoding isopenicilline N synthetase (13) : the latter also occurs downstream from its own pyrimidine tracts. Furthermore similar sequences have been found at 5'flanking regions of two other CPC biosynthetic genes from *A. chrysogenum*, 5'CGTCATA3' in *pcbAB*(14) and 5'CGTCACT3' in *cefEF* gene (underlined in Fig.2), suggesting that the sequence element may play a role in a coordinate regulation on the expression of the CPC biosynthetic genes. It is noteworthy that pentanucleotide 5'CGTCA3', completely conserved among these four sequences, is known to be a core motif of CRE (cAMP-responsive element) in higher eukaryotes (15).

Linkage of *cefG* and *cefEF*. Interestingly, the sequencing analysis of a region located further upstream the *cefG* gene revealed a linkage between the gene and the *cefEF* gene encoding deacetoxycephalosporin C synthetase / deacetylcephalosporin C synthetase, the enzyme for the fourth and fifth steps in CPC biosynthesis. The sequence complementary to that (broken-underline in Fig. 2.) between -1180 and -1015 relative to the initiation codon of the *cefG* gene is completely identical to that around the N-terminal coding region of the *cefEF* gene (4). Furthermore an oligonucleotide probe corresponding to the C-terminal region of the *cefEF* gene was found to strongly hybridize to a 2Kb SpeI fragment within *pATF1* (data not shown). The results indicate that the *cefG* gene is closely linked to the *cefEF* gene on *A. chrysogenum* chromosome II (5), and both genes are separated by an intergenic

region of 1114 bp from which they are divergently transcribed. The linkage was also confirmed by Southern blot analysis of *A.chrysogenum* DNA using both genes as probes.

Disruption of *cefG* gene. In order to disrupt the *cefG* locus in *A.chrysogenum* a plasmid, designated pDATF1, was constructed as described in Materials and Methods. The plasmid contains the *cefG* gene interrupted in the coding region by a HYB^R expression cassette. The plasmid was linearized with *SacI* and used for transforming *A.chrysogenum* IS-5 to HYB resistance. 10 transformants were selected randomly and grown in a shake flask to test for their ability to produce CPC, because the disruption of *cefG* is expected to result in the loss of CPC productivity. Two strains, termed DATF2 and DATF8, respectively, were found to lack the CPC productivity and accumulate exclusively a precursor DCPC from the early period of cultivation. Typical time courses of CPC and DCPC production by DATF2 and parental strain IS-5 are shown in Fig. 3. As expected, no DCPC-ATF activity was detected in the cell free extracts of the two CPC-negative strains. DNA was isolated from the two strains and IS-5 and subjected to Southern hybridization analysis with the DCPC-ATF probe after digestion with *Bam*HI, *Spe*I, or *Stu*I plus *Xho*I. DNA from strain DATF2 gave the hybridization pattern expected for replacement of the endogenous *cefG* locus with the disrupted gene on pDATF1 by a double cross-over event (Fig. 4). These results confirm the identity of the cloned gene and indicate that the *cefG* gene is the only source for the DCPC-ATF activity in *A.chrysogenum*. The *cefG*-disrupted strains were also found to be methionine prototroph, supporting our previous assumption (3) that the *cefG* gene is not identical to *Met2* gene encoding homoserine α -acetyltransferase in *A.chrysogenum*, although the primary structure of DCPC-ATF showed a significant homology with those of MET2 encoded protein from other fungi. In the course of the study, we have isolated several clones which hybridized to the DCPC-ATF probe under a low stringent condition. Preliminary studies suggest that the clones contains *A.chrysogenum* MET2 gene (manuscript in preparation). Further analysis of the gene should provide us with interesting information for an evolutionary relationship between MET2 and *cefG* genes.

ADDENDUM

During the preparation of this paper, Martín and his colleagues (J.Bacteriol. 174, 3056-3064, 1992) reported the sequence of the *cefG* gene from *A.chrysogenum* C10. There are remarkable differences between our understanding and theirs on 5'-untranslated region and the translated polypeptide. A part of these discrepancies will be solved by our results of the enzyme purification (submitted elsewhere) which indicate that DCPC-ATF is composed of two nonidentical subunits, presumably derived from a single precursor polypeptide.

REFERENCES

1. Ingolia, T.D. and Queener, S.W. (1989) Med. Res. Rev. 9, 245-264.
2. Fujisawa, Y. and Kanzaki, T. (1975) Agr. Biol. Chem. 39, 2043-2048.
3. Matsuda, A., Sugiura, H., Matsuyama, K., Matsumoto, H., Ichikawa, S. and Komatsu, K. (1992) Biochem. Biophys. Res. Commun. 182, 995-1001.
4. Samson, S.M., Dotzlaf, J.F., Slisz, M.L., Becker, G.W., Van Frank, R.M., Veal, L.E., Yeh, W.K., Miller, J.R., Queener, S.W. and Ingolia, T.D. (1987) Bio/Technology 5, 1207-1214.
5. Skatrud, P.L. and Queener, S.W. (1989) Gene 79, 331-338.
6. Johnstone, I.L., Hughes, S.G. and Clutterbuck, A.J. (1985) EMBO J. 4, 1307-1311.
7. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

8. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
9. Gritz, L. and Davies, J. (1983) *Gene* 25, 179-188.
10. Skatrud, P.L., Queener, S.W., Carr, L.G. and Fisher, D.L. (1987) *Curr. Genet.* 12, 337-348.
11. Ballance, D.J. (1986) *Yeast* 2, 229-236.
12. Punt, P.J., Dingemanse, M.A., Kuyvenhoven, A., Soede, R.D.M., Pouwels, P.H. and Van den Hondel, C.A.M.J.J. (1990) *Gene* 93, 101-109.
13. Smith, A.W., Ramsden, M., Dobson, M.J., Harford, S. and Peberdy, J.F. (1990) *Bio/Technology* 8, 237-240.
14. Gutiérrez, S., Díez, B., Montenegro, E. and Marín, J.F. (1991) *J. Bacteriol.* 173, 2354-2365.
15. Riabowol, K.T., Fink, J.S., Gilman, M.Z., Walsh, D.A., Goodman, R.H. and Feramisco, J.R. (1988) *Nature* 336, 83-86.