The EMBO Journal vol.4 no.8 pp.2093–2099, 1985

Identification and molecular analysis of a third *Aspergillus nidulans* alcohol dehydrogenase gene

G.L.McKnight, H.Kato[1], A.Upshall, M.D.Parker, G.Saari and P.J.O'Hara

ZymoGenetics, Inc., 2121 N. 35th Street, Seattle, WA 98103, USA, and [1]Tsukuba Research Laboratories, Eisai Co. Ltd., 5-1-3 Tokodai, Toyosato-machi, Tsukuba-gun, Ibaragi 300-26, Japan

Communicated by N.Fiil

An *Aspergillus nidulans* functional cDNA encoding an alcohol dehydrogenase (ADH) was isolated by its ability to complement an *adh1* mutation in *Saccharomyces cerevisiae*. Alignment of the cDNA and cloned genomic DNA sequences indicated that the ADH gene contains two small introns. The presence of ethanol in the growth medium was shown to result in ADH mRNA accumulation presumably due to transcriptional induction of the gene. However, ADH mRNA accumulation was at most only partially repressed by the presence of glucose. The ADH gene characterized here is designated *ADH3* since it is distinct from the *alcA* gene which encodes ADH I and appears distinct from the gene which encodes ADH II. We demonstrated that the first intron in the *A. nidulans ADH3* gene was not efficiently spliced in *S. cerevisiae* whereas the promoter region was utilized weakly. We also present a comparison of the primary structure of *A. nidulans* ADH III with the alcohol dehydrogenases of *S. cerevisiae* and *Schizosaccharomyces pombe*.

Key words: cDNA/introns/regulation/yeast expression

Introduction

In *Aspergillus nidulans* the relatively low transformation efficiency and the integrative nature of the transformation events (Ballance *et al.*, 1983; Tilburn *et al.*, 1983; Yelton *et al.*, 1984) makes gene cloning with genomic DNA libraries a rather time-consuming approach, although the use of cosmid vectors has streamlined the recovery of the complementing gene (Yelton *et al.*, 1985). Attempts to isolate *A. nidulans* genes by functional complementation in *Saccharomyces cerevisiae* or *Escherichia coli* have only been successful for *argB* (Berse *et al.*, 1983) and *trpC* (Yelton *et al.*, 1983). A major problem in cloning genes by functional complementation in heterologous systems lies in the dual requirements of promoter function and intron splicing. Such limitations can be overcome by transformation experiments with cDNA pools in which the coding sequences are properly oriented in an expression vector (McKnight and McConaughy, 1983). We report here the application of this strategy to clone an alcohol dehydrogenase (ADH) cDNA from *A. nidulans* by its ability to complement a mutation in the *ADH1* gene of *S. cerevisiae*. Using the cDNA as a probe we have cloned the corresponding gene and present here the cDNA and chromosomal DNA sequences. Comparison of the nucleotide sequences indicates that the gene contains two introns and we demonstrate that the gene appears to be transcriptionally regulated by ethanol. We designate this gene *ADH3* and the encoded isozyme ADH III based on comparison with the cloned *alcA* gene which encodes ADH I (Lockington *et al.*, 1985; Doy *et al.*, 1985) and with the regulatory pattern of the ADH II isozyme (Sealy-Lewis and Lockington, 1984). We additionally show that *S. cerevisiae* is unable to splice efficiently the first intron of the *A. nidulans* gene but can weakly utilize the promoter.

Results

Isolation of the ADH cDNA and gene

S. cerevisiae contains three isozymes of ADH. ADH I is expressed during growth on glucose, ADH II is expressed during growth on non-fermentable carbon sources such as ethanol and ADH III is associated with the mitochondria. *S. cerevisiae* strain 500-11 lacks functional ADH I isozyme due to a mutation in the structural gene *ADH1* and also lacks the functional ADH II isozyme because the *ADH2* gene is not expressed due to a mutation in the positive regulatory locus *ADR1* (Ciriacy, 1979). *S. cerevisiae* strains which lack functional ADH I and ADH II isozymes are unable to grow in the presence of the respiratory inhibitor antimycin A (Williamson *et al.*, 1980). Therefore, transformants of strain 500-11, which express functional ADH, are readily identified by growth in the presence of antimycin A. *S. cerevisiae* strain 500-11 was transformed with 8 μg of the cDNA plasmid pool using a modification of the described procedure (Beggs, 1978) and plated on synthetic minimal medium lacking tryptophan. After 2 days of growth the plates were overlayed with antimycin A to select for ADH activity. One Adh^+ colony was obtained which exhibited mitotic instability for resistance to anti-

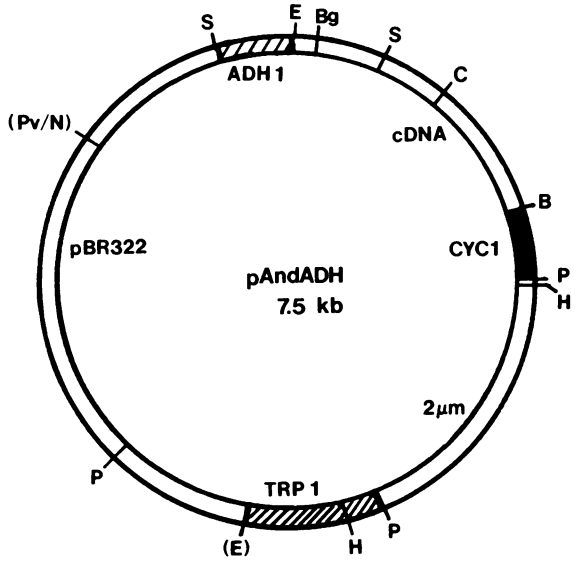

Fig. 1. Restriction endonuclease map and structure of plasmid pAndADH. The *S. cerevisiae ADH1* promoter, *CYC1* terminator, *TRP1*$^+$ selectable gene and 2 μm replication origin are indicated. The pBR322 portion, which contains the selectable *bla* gene and bacterial replication origin, has been altered by removal of the *Eco*RI site (E) and deletion of the *Pvu*II-*Nru*I segment (Pv/N). The restriction endonuclease sites shown are B, *Bam*HI; Bg; *Bgl*II; C, *Cla*I; E, *Eco*RI; H, *Hind*III; P, *Pst*I and S, *Sph*I.

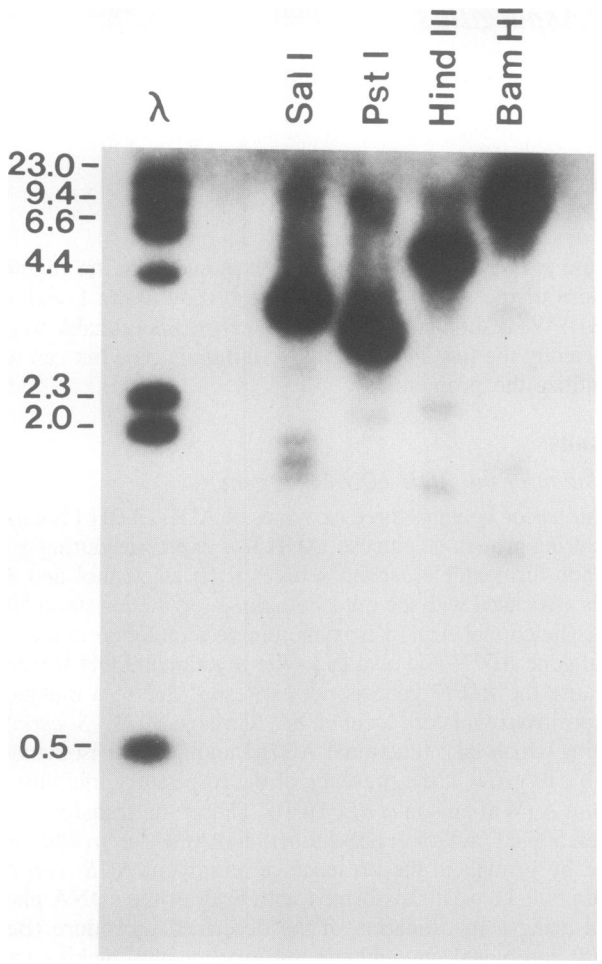

Fig. 2. Southern blot analysis of *A. nidulans* chromosomal DNA. 10 μg of DNA was digested with each of the indicated restriction endonucleases and electrophoretically separated on a 0.8% agarose gel, blotted onto nitrocellulose and hybridized with the ADH cDNA. End-labelled *Hind*III fragments of λ DNA were used as size markers.

mycin A after growth on non-selective medium. Plasmid DNA was isolated from this transformant and recovered in *E. coli* following transformation of competent MC1061 (Casadaban and Cohen, 1980). This plasmid DNA transformed the Adh$^-$ yeast strain 500-11 to antimycin A resistance (Adh$^+$) at high frequency and was designated pAndADH (Figure 1). The 1.35-kb cDNA insert in pAndADH is flanked by the poly-linker restriction sites *Eco*RI and *Sst*I on the 5′ side and *Xma*I and *Bam*HI on the 3′ side. The cDNA insert was excised by digestion with *Eco*RI and *Bam*HI, nick-translated and used to probe nitrocellulose blots of digested *A. nidulans* genomic DNA. The results shown in Figure 2 indicate that the cDNA is homologous to *Pst*I and *Sal*I fragments of 3.4 and 4.0 kb, respectively. The cDNA additionally hybridized to multiple restriction fragments in DNA digested with *Bam*HI, *Hind*III, *Pst*I or *Sal*I (Figure 2). The *A. nidulans* genomic DNA library in bacteriophage λ (Orr and Timberlake, 1982) was screened with the cDNA insert and four λ plaques were purified which hybridized to the nick-translated cDNA probe. The λ DNAs were isolated and the genomic DNA fragments were excised by *Eco*RI digestion and inserted into pUC19. These plasmid DNAs were digested with *Pst*I and with *Sal*I and hybridized to the cDNA probe. The inserts from two of the four λ DNAs contained *Pst*I and *Sal*I hybridizing fragments of 3.4 and 4.0 kb, respectively. The 3.4-kb *Pst*I fragments were inserted into pUC19

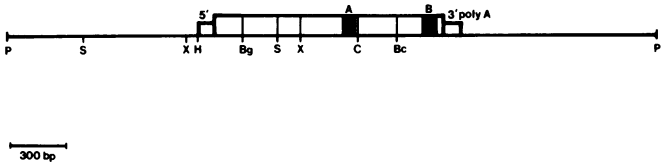

Fig. 3. Restriction endonuclease map and structure of the *ADH3* gene from *A. nidulans*. The coding region is shown as a large open box with the intervening sequences A and B indicated in black. The 5′- and 3′-non-coding regions are indicated by smaller boxes. The restriction endonuclease sites shown are Bc, *Bcl*I; Bg, *Bgl*II; C, *Cla*I; H, *Hind*III; P, *Pst*I; S, *Sph*I and X, *Xba*I.

and designated pM019 and pM020. Plasmid pM020 was chosen for subsequent analysis and the restriction map of the 3.4-kb *Pst*I fragment from pM020 is shown in Figure 3.

Structure of the cDNA and gene

The cDNA contained a 1056-bp coding sequence flanked by 5′- and 3′-non-coding regions of 76 and 113 − 114 bp, respectively. In addition, the cDNA contained a 3′ tail of ~100 deoxyadenosines and was attached to a 5′ tail of 15 nucleotides [$(C)_{14}G$]. The chromosomal nucleotide sequence was in total agreement with the cDNA except for the presence in the gene of two introns A and B (Figure 4). Both introns contained the higher eukaryotic putative 3′ splice signal CTGAT (Keller and Noon, 1984) located 15 and 22 bp upstream of the 3′ splice site. Moreover, both introns lack in-frame termination codons and are in-frame with the distal portion of the coding region. The 5′ end of the gene contained a TATAAAT sequence located 37 bp upstream of the 5′ end of the cDNA. In contrast, the 3′ end of the gene lacked both the higher eukaryotic polyadenylation signal elements AATAAA and CAYTG (Berget, 1984) as well as the consensus *S. cerevisiae* transcription termination/polyadenylation signals (Zaret and Sherman, 1982; Henikoff *et al.*, 1983). The codon usage for this gene, shown in Figure 5, indicates that among the 61 possible coding triplets 59 are utilized, which results in an absence of the extreme codon bias observed in highly expressed genes from *S. cerevisiae* (Bennetzen and Hall, 1982a) and *Schizosaccharomyces pombe* (Russell and Hall, 1983).

Comparison of the fungal ADH proteins

The inferred amino acid sequence of the *A. nidulans* ADH protein was aligned with the amino acid sequences of the *S. cerevisiae* isozymes ADH I (Bennetzen and Hall, 1982b) and ADH II (Russell *et al.*, 1983) and the *Schizosaccharomyes pombe* ADH enzyme (Russell and Hall, 1983). The alignment, shown in Figure 6, indicates the *A. nidulans* ADH isozyme contains two sites of insertion of two amino acids each relative to the *S. cerevisiae* isozymes. The amino-terminal methionine was included in the alignment even though mature ADH enzymes generally contain an acetylated serine residue at the amino terminus. All of the fungal ADH enzymes have identical amino acids at positions previously identified (Eklund *et al.*, 1976; Jornvall, 1977) as (i) an active site pocket (Figure 6, positions 46, 55, 56, 57, 58 and 120); (ii) ligands for binding co-enzyme (Figure 6, positions 45, 95, 173, 204 and 209); (iii) ligands for the Zn atom bound at the active site (Figure 6, positions 44, 67 and 158); (iv) cysteines which could bind a second atom of Zn but apparently do not, unlike ADH enzymes from higher eukaryotes (Figure 6, positions 100, 103, 106 and 114); and (v) glycines which are structurally important (Figure 6, positions 66, 77 and 88). The *S. cerevisiae* isozymes differ in the preferred direction of the metabolic reaction they catalyze. The ADH I isozyme preferen-

```
                                30                                  60                                  90
CGTTTAGCTGCCCCGGCAAAAGCAGGAATTTCGGACGATGTCCTCGTCGGCTTGTGTGGATTCGTTTCGTTGTGGACCCATCCCCTGTATAGTGTTGCGTTGTGTTGCACTG

      120   Xba I                   150                                 180        Hind III          # 210
TTATAGCCAGTGTCTAGAGGCAGTGCTGCTACACCCCTCATTCTCTTGTGGTATAAATGGCAAGAGAGCTCTCACGAAGCTTGTGAATTAGCTTGTCTCTGTACATCCTCT

                  240                                 270                                 300
CAGTTAATAGTTTATTACCTCTTCTATCAATCTCATTTTACACACTCTCATCCCAAGG ATG TCT GTC CCC GAA GTG CAA TGG GCC CAA GTG GTC GAG
                                                          Met Ser Val Pro Glu Val Gln Trp Ala Gln Val Val Glu

            330                                 360                                 390
AAG GCA GGC ACT CCT CCG GTT TAC AAA CAG GTT CCC GTT CCA AAA CCT GGA CCA GAC GAG ATT TTG GTC AAG ATG CGA TAT TCG
Lys Ala Gly Thr Pro Pro Val Tyr Lys Gln Val Pro Val Pro Lys Pro Gly Pro Asp Glu Ile Leu Val Lys Met Arg Tyr Ser

                     420Bgl II                          450                                 480
GGC GTC TGC CAT ACA GAT CTT CAC GCC ATG AAG GGC GAC TGG CCT CTT CCT TCA AAG ATG CCA CTG ATA GGC GGC CAT GAG GGC
Gly Val Cys His Thr Asp Leu His Ala Met Lys Gly Asp Trp Pro Leu Pro Ser Lys Met Pro Leu Ile Gly Gly His Glu Gly

                               510                                 540                                 570
GCT GGT GTC GTC GTT GCT AAG GGA GAA TTA GTC AAG GAC GAA GAT TTC AAG ATT GGT GAC AGA GCC GGC ATC AAA TGG CTT AAT
Ala Gly Val Val Val Ala Lys Gly Glu Leu Val Lys Asp Glu Asp Phe Lys Ile Gly Asp Arg Ala Gly Ile Lys Trp Leu Asn

                                    600Sph I                       630
GGC TCC TGC CTC TCC TGC GAA ATG TGC ATG CAA GCC GAC GAA CCT CTT TGC CCT CAC GCC TCG CTA TCC GGA TAC ACC GTC GAT
Gly Ser Cys Leu Ser Cys Glu Met Cys Met Gln Ala Asp Glu Pro Leu Cys Pro His Ala Ser Leu Ser Gly Tyr Thr Val Asp

   660                                 690                                 720  Xba I
GGC ACA TTC CAG CAA TAC ACT ATC GGA AAG GCT GCA TTG GCT TCT AAG ATC CCG GAT AAC GTG CCT CTA GAC GCG GCG GCT CCT
Gly Thr Phe Gln Gln Tyr Thr Ile Gly Lys Ala Ala Leu Ala Ser Lys Ile Pro Asp Asn Val Pro Leu Asp Ala Ala Ala Pro

            750                                 780                                 810
ATC CTC TGC GCT GGG ATT ACC GTA TAC AAA GGA CTG AAA GAG TCC GGA GCG CGT CCA GGC CAG ACT GTC GCT ATC GTT GGT GCG
Ile Leu Cys Ala Gly Ile Thr Val Tyr Lys Gly Leu Lys Glu Ser Gly Ala Arg Pro Gly Gln Thr Val Ala Ile Val Gly Ala

                     840                                 870                                 900
GGA GGT GGC CTG GGT TCT CTG GCA CAA CAG TAT GCC AAA GCA ATG GGG TTA CGC ACC ATC GCT ATT GAT TCC GGT GAC GAG AAG
Gly Gly Gly Leu Gly Ser Leu Ala Gln Gln Tyr Ala Lys Ala Met Gly Leu Arg Thr Ile Ala Ile Asp Ser Gly Asp Glu Lys

                               930                                 960                                 990
AAA GCC ATG TGC GAG CAA TTA GGA GCT GAG GTATTTTACCTACTCTACCGCATTCGACAAGCCGAACCCATCCGGAGAGCAGAGCTGATGTTGCGGCATTGA
Lys Ala Met Cys Glu Gln Leu Gly Ala Glu                                 IVS A

          1020Cla I                           1050                                1080
CAG GTC TTT ATC GAT TTC AGC AAA TCC GCC GAC GTC GTC GCA GAT GTC AAA GCT GCC ACG CCG GGC GGC CTA GGC GCC CAC GCC
    Val Phe Ile Asp Phe Ser Lys Ser Ala Asp Val Val Ala Asp Val Lys Ala Ala Thr Pro Gly Gly Leu Gly Ala His Ala

                     1110                                1140                                1170
GTC ATC CTC CTC GCC GTC GCC GAA AAG CCC TTC CAA CAA GCA ACC GAG TAC GTT CGC TCC CAC GGC TCC GTC GTC GCA ATC GGC
Val Ile Leu Leu Ala Val Ala Glu Lys Pro Phe Gln Gln Ala Thr Glu Tyr Val Arg Ser His Gly Ser Val Val Ala Ile Gly

                               1200                                1230Bcl I                           1260
TTG CCA GCC AAT GCA TTT CTC AAG GCA CCC GTG TTC ACA ACT GTC GTC CGC ATG ATC AAC ATC AAG GGA AGT TAT GTC GGA AAC
Leu Pro Ala Asn Ala Phe Leu Lys Ala Pro Val Phe Thr Thr Val Val Arg Met Ile Asn Ile Lys Gly Ser Tyr Val Gly Asn

                                    1290                                1320
CGC CAG GAC GGC GTT GAG GCG TTA GAC TTC TTT GCC AGG GGC CTT ATC AAG GCG CCA TTC AAG AAG GCA CCG CTG CAA GAT TTG
Arg Gln Asp Gly Val Glu Ala Leu Asp Phe Phe Ala Arg Gly Leu Ile Lys Ala Pro Phe Lys Lys Ala Pro Leu Gln Asp Leu

  1350                                1380                                1410                                1440
CCA CAG ATT TTT GAG TTG ATG G   GTGTGTTTCCCCTCTTTTGTGTTCTCCCCTCGACGTTAATATTGATGCTGATGGTGGTTGGAATGGCAG GG CAA GGG
Pro Gln Ile Phe Glu Leu Met Gly                                         IVS B                                Gln Gly

                               1470                                1500                                1530
AAG ATT GCG GGT CGT TAC GTC TTG GAG ATT CCC GAG TGA ACGGGCAGGAAGCAGAAATTCACCTACGCGGATCAGGATGGAAACAAGTAGGAAAATTAC
Lys Ile Ala Gly Arg Tyr Val Leu Glu Ile Pro Glu END

            1560                                1590  **                            1620                                1650
AGAATCAGCAGCCACTTTGGGAGTACACGAGATGCAAAGGCTGGAGTTTCATTTCGGTGAGACCTTTACTCGCTTAGTTCCTGTACCGACCCTGCTCCGAGCTCCTGTAGTA

TGGCGTCTGGGGTAAGCCAGG
```

Fig. 4. Nucleotide sequence of the *ADH3* gene from *A. nidulans*. The 5′ end of the cDNA is indicated by # and the polyadenylation site is shown by **. The polyadenylation site cannot be precisely determined due to the presence of an A in the genomic sequence at this location. The intervening sequences A and B are designated IVS A and IVS B, respectively. The TATA sequence upstream of the 5′ end of the cDNA is overlined and the sequence common to both IVS A and IVS B, which encompasses the putative 3′ splice signal GCTGAPy, is underlined.

tially catalyzes the conversion of acetaldehyde to ethanol whereas the ADH II isozyme preferentially catalyzes the conversion of ethanol to acetaldehyde (Wills, 1976). Seven positions differing between the yeast ADH I and ADH II isozymes have been predicted (Russell *et al.*, 1983) to be responsible for the difference in direction of the reaction catalyzed by these enzymes (Figure 6, positions 16, 171, 214, 232, 270, 275 and 288). Among these positions the *A. nidulans* ADH III isozyme matches only the yeast ADH II isozyme, at the positions 171, 270 and 275. This pattern suggests the *A. nidulans* ADH III isozyme preferentially catalyzes the conversion of ethanol to acetaldehyde.

Transcriptional regulation by ethanol

We investigated the transcriptional regulation of the gene by comparing the blot hybridization pattern of poly(A)$^+$ RNA with the cDNA insert from pAndADH (Figure 7). We compared poly-(A)$^+$ RNAs prepared from hyphae grown in 0.1% fructose, 10 mM urea medium which differed only in the absence (lane

1) and presence (lane 2) of 1% ethanol and in the presence of both 1% ethanol and 1% glucose (lane 3). The results shown in Figure 7 are consistent with the induction of transcription by the presence of ethanol and the lack of absolute transcriptional repression by the further presence of glucose. Additional RNA blots (data not shown) have demonstrated that ADH mRNA is at most only partially repressed by glucose. We have recently resolved the RNA signal into two RNA bands which hybridize to the ADH cDNA and which are co-regulated (data not shown). These RNAs may represent overlapping ADH transcripts with heterogeneous termini or unprocessed and processed forms of ADH RNA which differ in the presence or absence of the introns A and B. We have also probed samples of these poly(A)$^+$ RNA preparations with a cloned *argB* DNA fragment (Figure 7) and demonstrated that *argB* mRNA levels are similar in the absence (lane 4) and presence (lane 5) of 1% ethanol and are elevated in the presence of both ethanol and glucose (lane 6).

	U		C		A		G	
U	Phe	4	Ser	3	Tyr	3	Cys	0	U
		7		8		6		7	C
	Leu	4		1	Stop	0	Stop	1	A
		6		2		0	Trp	3	G
C	Leu	5	Pro	8	His	2	Arg	2	U
		5		5		4		4	C
		3		7	Gln	10		1	A
		5		4		6		0	G
A	Ile	7	Thr	4	Asn	2	Ser	1	U
		12		4		3		1	C
		1		3	Lys	9	Arg	1	A
	Met	10		1		17		1	G
G	Val	7	Ala	10	Asp	8	Gly	7	U
		21		15		10		18	C
		1		10	Glu	6		10	A
		4		7		12		4	G

Fig. 5. Codon utilization of the *ADH3* gene from *A. nidulans*.

A. nidulans ADH promoter and intron splicing signals do not function efficiently in S. cerevisiae

We examined the capability of *S. cerevisiae* to utilize efficiently the *A. nidulans* promoter and to splice efficiently intron A by replacing segments of pAndADH with segments of chromosomal DNA, transforming strain 500-11 with the hybrid plasmids and testing the transformants for antimycin A resistance and ADH enzyme activity in cell-free extracts. The *Sph*I-*Sph*I fragment of pM020, which contains ~700 bp of DNA upstream of the ADH-coding region and presumably the complete promoter (see Figure 3), was used to replace the *Sph*I-*Sph*I fragment of pAndADH, which contains the *S. cerevisiae ADH1* promoter (see Figure 1), and this plasmid was designated pM042-5. The *Bgl*II-*Cla*I fragment of pM020 containing intron A (see Figure 3) was used to replace the *Bgl*II-*Cla*I cDNA fragment of pAndADH (see Figure 1) and this plasmid was designated pM038. The ADH enzyme activities in cell-free extracts and the relative resistance to antimycin A of strain 500-11 and transformants containing pAndADH, pM038 and pM042-5 are compared in Table I. The results suggest that the *A. nidulans* promoter does function in *S. cerevisiae* but only at ~4% of the level of the yeast *ADH1* promoter. We have further demonstrated that the ADH activity and resistance to antimycin A of pM042-5/500-11 are plasmid-dependent by demonstrating co-segregation of these phenotypes with the Trp$^+$ plasmid marker (data not shown). The results also indicate that the presence of intron A completely eliminates ADH enzyme activity, presumably due to the lack of proper splicing by *S. cerevisiae*.

The level of expression in *S. cerevisiae* of the *A. nidulans* ADH cDNA from pAndADH was compared with the level of expression of the *S. cerevisiae ADH1* cDNA inserted into the identical vector and designated pScADH1. Comparison of the ADH enzyme activities indicated that pAndADH yielded 225-fold lower activity than did pScADH1, as shown in Table I. Northern blot analysis of poly(A)$^+$ RNA isolated from pAndADH/500-11 demonstrated the presence of high levels of *A. nidulans* ADH mRNA of the expected size (data not shown) which indicated that transcription and mRNA stability are not grossly abnormal.

```
                          20                  40                  60                  80
A.n.      MSVPEVQWAQVVEKAGTPPV--YKQVPVPKPGPDEILVKMRYSGVCHTDLHAMKGDWPLPSKMPLIGGHEGAGVVVAKGELVKDEDFKIGDR
S.p.       TI DK L A FHTH G ENVKFEE   AE  Q  V  NIK T        LQ      A              KV AG TR--L
S.c.(I)     I  T KGVIFYESHGKLE--H DI     KAN L INVK          WH      V L  V         GM  N  G--W    Y
S.c.(II)    I  T K IIFYESNGKLE--H DI     K N L INVK          WH      T L  V         GM  N  G--W    Y

                     100                 120                 140                 160                 180
A.n.      AGIKWLNGSCLSCEMCMQADEPLCPHASLSGYTVDGTFQQYTIGKAALASKIPDNVPLDAAAPILCAGITVYKGLKESGARPGQTVAIVGAG
S.p.      V V  M S  GN  Y  K G TI   IQ          H C AN TH TI  ES   EV    M     C RA    KVG  EWIC P
S.c.(I)            MA  Y ELGN SN    D     H  S    ATAD VQ AH  QGTD AQV           A  SANLMA HW   S  A
S.c.(II)           MA  Y ELGN SN    D     H  S  E ATAD VQ AH  QGTD AEV           A  SANL A HWA  S  A

                                              #
                          200                 220                 240                 260
A.n.      GGLGSLAQQYAKAMGLRTIAIDSGDEKKAMCEQLGAEVFIDFSKSADVVADVKAATPGGLGAHAVILLAVAEKPFQQATEYVRSHGSVVAIG
S.p.         H  V       AM VV   T  D AELVKSF     L  K E  MIEA   CTN —   GTLV STSP SYE  AGFA PGSTM TVS
S.c.(I)         V        Y VLG  G EG EELFRSI G      T EK I GA LK  D —   G  NVS S AAIEAS R   AN TT LV
S.c.(II)        V        Y VLG  G PG EELFTSL G      T EK I S  V  N —   GI NVS S AAIEAS R C AN T  LV

                                                                        #
                     280                 300                 320                 340
A.n.      LPANAFLKAPVFTTVVRMINIKGSYVGNRQDGVEALDFFARGLIKAPFKKAPLQDLPQIFELMGQGKIAGRYVLEIPE
S.p.      M  G K G DI WLT K LK C  H     I SI   EYVS    V PYY VQPFST  DVYR  HEN      I  DLSK
S.c.(I)   M  G KCCSD  NQ  KS S V        A TR           V S I VVG ST  E Y K EK Q V     VDTSK
S.c.(II)     G KCSSD  NH  KS S V        A TR           V S I VVG SS  E Y K EK Q       VDTSK
```

Fig. 6. Comparison of the amino acid sequence of fungal ADH polypeptides. A.n., *A. nidulans* ADH III; S.p., *Schizosaccharomyces pombe*; S.C.(I), *S. cerevisiae* ADH I; S.C.(II), *S. cerevisiae* ADH II; #, location of intervening sequences.

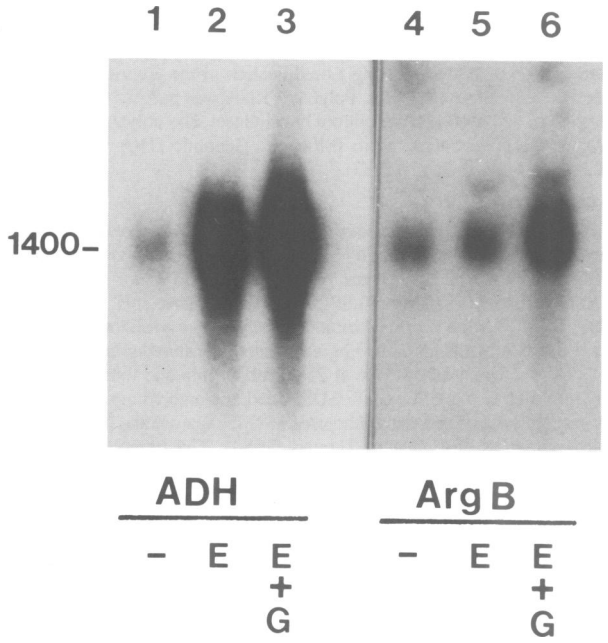

Fig. 7. Northern blot analysis of *A. nidulans* poly(A)$^+$ RNA. Poly(A)$^+$ RNA was isolated from cells grown in 0.1% fructose, 10 mM urea medium **(lanes 1** and **4)**, supplemented with 1% (v/v) ethanol **(lanes 2** and **5)** and with 1% (v/v) ethanol and 1% (w/v) glucose **(lanes 3** and **6)**. 10 μg of each poly(A)$^+$ RNA sample was glyoxylated and electrophoretically separated on a 1.0% agarose gel, blotted to nitrocellulose and hybridized with ADH cDNA **(lanes 1 – 3)** and with the *argB XbaI* fragment (Berse *et al.*, 1983) **(lanes 4 – 6)**.

Table I. Relative ADH expression levels in yeast transformants

Strain	ADH activity	Antimycin A resistance
500-11	0	–
pAndADH/500-11	234	+
pM042-5/500-11	10	±
pM038/500-11	0	–
pScADH1/500-11	52 760	++

The ADH enzyme activities are expressed in milliunits/mg protein and were determined from total crude extracts using ethanol as a substrate (Williamson *et al.*, 1980). The values shown are the average from two independent transformants.

The low ADH enzyme activity expressed from pAndADH, relative to pScADH1, may represent differences in translational efficiency due to differences in codon usage and differences between the enzymes in specific activity, substrate specificity and stability.

Discussion

Gene identity

The *A. nidulans alcA* gene has been shown to code for an ethanol-induced ADH (Pateman *et al.*, 1983), which according to Creaser *et al.* (1985) is the only ADH species present in hyphae grown in the absence or presence of ethanol. In contrast, Sealy-Lewis and Lockington (1984) have concluded that two ADH species exist in *A. nidulans*, where the ethanol-induced ADH I polypeptide is encoded by the *alcA* gene and the ADH II polypeptide is ethanol-repressed. Both groups agree that ADH enzyme activity is glucose-repressed (Pateman *et al.*, 1983; Sealy-Lewis and Lockington, 1984). A comparison of the restriction map of the region surrounding the cloned *alcA* gene (Lockington *et al.*, 1985; Doy *et al.*, 1985) with Figure 3 indicates these genes are clearly different. The gene encoding ADH II also appears distinct from the *ADH3* gene characterized here because the mRNAs encoding ADH II and ADH II activity are normally repressed by ethanol (Sealy-Lewis and Lockington, 1984). The Southern blot hybridization analysis, shown in Figure 2, indicates that the genome of *A. nidulans* contains additional DNA sequences which hybridize to the ADH cDNA. The *alcA* gene (Lockington *et al.*, 1985; Doy *et al.*, 1985) probably corresponds to the 4.0-kb cross-hybridizing *Bam*HI fragment shown in Figure 2.

Gene structure and regulation

We have cloned an *A. nidulans* ADH cDNA by functional complementation in *S. cerevisiae* and have used this cDNA to clone the corresponding *A. nidulans ADH3* gene. A comparison of the amino acid sequence of the polypeptide encoded by this gene with the ADH polypeptides of *S. pombe* and *S. cerevisiae* (Figure 6) provided further evidence that the cloned gene encodes a functional ADH enzyme. A comparison of the cDNA and genomic DNA sequences (Figure 4) revealed the presence of two introns which are 75 and 69 bp in length. The introns are homologous at nine of 10 positions in a region which encompasses the putative 3′ splice signal CTGAT. Based on sequence comparisons of the 3′ splice signals present in the *ADH3* gene with those present in the glucoamylase gene of *A. niger* (Boel *et al.*, 1984) and *A. awamori* (Nunberg *et al.*, 1984) and the exocellobiohydrolase I gene of *Trichoderma reesei* (Shoemaker *et al.*, 1983) we observe GCTGAPy as a 3′ splice signal consensus sequence. The 5′ junction sequence GTATTT of intron A differs from the typical sequence GTPuNGT found in the introns of *S. cerevisiae* genes (Teem *et al.*, 1984) and in the introns of the glucoamylase (Boel *et al.*, 1984; Nunberg *et al.*, 1984) and the exocellobiohydrolase I (Shoemaker *et al.*, 1983) genes. Both intron A and intron B lack in-frame termination codons which suggests the possibility of differential splicing and the generation of polypeptides with different activities. We have been unable to determine whether the two introns separate structural domains of the protein, as has been found with the introns of the maize *ADH1* gene (Branden *et al.*, 1984), because the three-dimensional structure of the fungal ADH proteins are unknown and because the ADH proteins of fungi and higher eukaryotes differ considerably in amino acid sequence and polypeptide length. The relationship of the 5′ end of the cDNA to the transcription initiation site is currently unknown. However, the location of a typical TATAAAT sequence 37 bp upstream from the 5′ end of the cDNA suggests that the *ADH3* promoter more closely resembles promoters of higher eukaryotes than of *S. cerevisiae*, which have TATA sequences generally located 60 – 100 bp upstream of transcription initiation sites (Russell, 1983). The region flanking the translational initiator codon is generally conserved in yeast genes and has the structure (−3)ANNATGNNT(+6) (Ammerer *et al.*, 1981). The *ADH3* gene also has this structure surrounding the initiator ATG. The polyadenylation signal in the *ADH3* gene has not yet been identified but appears distinct from the signals employed by higher eukaryotes and *S. cerevisiae*. Comparison of the sequences flanking the polyadenylation sites of the *ADH3* and *trpC* (Mullaney *et al.*, 1985) genes of *A. nidulans* and the glucoamylase genes of *A. niger* (Boel *et al.*, 1984) and *A. awamori* (Nunberg *et al.*, 1984) does not yield any consensus sequences in this region.

The results shown in Figure 7 suggest that the *ADH3* gene is transcriptionally induced by ethanol but is apparently not tran-

scriptionally repressed by glucose. Further work has demonstrated the *ADH3* gene is at most only partially repressed by glucose. Despite the presence of cross-hybridizing sequences, we believe that the Northern blot hybridization analysis, shown in Figure 7, reflects only those RNAs equivalent to the cDNA probe since the RNA blots were washed under stringent conditions (see Materials and methods). In addition, the co-regulation of the two recently resolved RNA bands suggests both RNAs are homologous to the cDNA probe and may represent overlapping transcripts with heterogeneous termini or processed and unprocessed transcripts. The apparent transcriptional induction of the *ADH3* gene by ethanol agrees with the predicted catalytic preference of the encoded enzyme for the conversion of ethanol to acetaldehyde. The physiological role of ADH III is unknown because *alcA ADH3*$^+$ strains are deficient in the utilization of ethanol. The ADH III isozyme may display an unusual substrate specificity and have a role in the detoxification of molecules other than ethanol. The low level of *A. nidulans* ADH III activity in *S. cerevisiae*, relative to the *S. cerevisiae* ADH I activity, may reflect in part differences in the utilization of ethanol as a substrate.

Aspergillus regulatory signals in S. cerevisiae

The *A. nidulans ADH3* promoter apparently functions in *S. cerevisiae* in glucose medium but only at 4% of the level of the *S. cerevisiae ADH1* promoter (Table I). This effect could be due to a difference between the promoters in the rate of transcriptional initiation in *S. cerevisiae* and/or a difference in the level of translatable *ADH3* mRNA, which is due to different lengths of the 5'-non-coding region in the mRNA. Because the RNA polymerase II of *S. cerevisiae* extends an unusually large distance between the TATA box and the transcriptional initiation site (Russell, 1983), the 5'-non-coding region of the *ADH3* mRNA transcribed from the *A. nidulans* promoter may be unusually short and thus poorly translated. The presence of intron A completely eliminated ADH enzyme activity. The intron presumably cannot be spliced by *S. cerevisiae* because the yeast 3' splicing signal TACTAAC (Langford *et al.*, 1983) is not present. Similar results have been obtained by Innes *et al.* (1985) who have demonstrated that the *A. awamori* glucoamylase promoter does not function in yeast and that the introns are inefficiently or incorrectly spliced in yeast. The regulatory signals for transcription initiation and for intron splicing apparently differ between *Aspergillus* and *Saccharomyces*.

Materials and methods

Strains, plasmids and bacteriophage

The wild-type *A. nidulans* strain FGSC4 was obtained from the Fungal Genetics Stock Center, Arcata, CA; the *S. cerevisiae* strain 500-11 (*MATo adh1-11 adr1-1 trp1 leu2 ura1*) from E.T.Young, University of Washington; the *E. coli* strain used was MC1061. The yeast expression vector pYcDE8 was constructed by G.McKnight at the University of Washington and is similar to pMAC561 (McKnight and McConaughy, 1983). Genomic DNA clones were recovered from a bacteriophage λ Charon 4 library containing *Eco*RI-linkered *Hae*III and *Alu*I partially-digested *A. nidulans* genomic DNA (Orr and Timberlake, 1982). The bacteriophage vectors M13mp18 and M13mp19 (Norrander *et al.*, 1983) were used for DNA sequencing (Sanger *et al.*, 1977) of subcloned restriction fragments. Computer analyses of the sequences were performed using an IntelliGenetics program package.

Enzymes and chemicals

Restriction endonucleases and T4 DNA ligase were obtained from Bethesda Research Laboratories; terminal deoxynucleotidyl transferase and RNase H from PL Biochemicals; placental ribonuclease inhibitor from Promega Biotech Co.; AMV reverse transcriptase from Seikagaku America, Inc.; oligo(dT)-cellulose from Collaborative Research, Inc.; ^{32}P- and ^{35}S-labelled nucleotides from New England Nuclear.

Preparation of nucleic acids from A. nidulans

Total RNA was extracted from shake flask cultures grown at 37°C for 14 h following inoculation with spores, using a modification of the procedure described by Timberlake and Bernard (1981). Poly(A)$^+$ RNA was purified by two cycles of adsorption to and elution from oligo(dT)-cellulose. The poly(A)$^+$ RNA used in cDNA synthesis was a gift from P.Russell. Genomic DNA was extracted as described in Tilburn *et al.* (1983).

Blot hybridizations

Genomic DNA digests were electrophoresed in 0.8% agarose and blotted onto nitrocellulose paper as described by Southern (1975). The immobilized DNA was hybridized with the nick-translated ADH cDNA at 65°C for 18 h in 2 x SSC (1 x SSC is 0.15 M NaCl and 0.0015 M sodium citrate, pH 7.0), 0.1% each of SDS, Ficoll, bovine serum albumin and polyvinyl pyrrolidone, 20 mM Tris-HCl pH 8.0, 1 mM EDTA and 50 μg/ml sheared, denatured salmon sperm DNA. The DNA blot was washed once at 25°C for 15 min and three times at 65°C for 30 min each in 3 x SSC, 0.2% SDS, dried and autoradiographed. Poly(A)$^+$ RNA was glyoxylated and electrophoresed in 1.0% agarose and blotted onto nitrocellulose paper as described in Thomas (1980). The immobilized RNA was hybridized at 42°C for 18 h in 50% formamide, 5 x SSC, 0.2% each of Ficoll, bovine serum albumin and polyvinyl pyrrolidone, 50 mM sodium phosphate pH 7.0 and 250 μg/ml of sheared, denatured salmon sperm DNA. The RNA blot was washed four times at 25°C for 5 min each in 2 x SSC, 0.1% SDS and twice at 50°C for 15 min each in 0.1 x SSC, 0.1% SDS, dried and autoradiographed.

Preparation of the cDNA plasmid pool and plasmid recovery

The system used was similar to that previously described in McKnight and McConaughy (1983). The vector-primer and linker fragments used in the preparation of the cDNA plasmid pool were prepared from the yeast expression plasmid pYcDE8. The vector-primer fragment contained ~100 deoxythymidylate residues and the linker fragment contained ~15 deoxycytidylate residues. The cDNA synthesis reaction was incubated at 42°C for 1 h in a reaction mixture containing 3.4 μg of the vector-primer (0.85 pmol) and 2 μg of poly(A)$^+$ RNA in 100 mM Tris-HCl pH 8.5, 50 mM KCl, 10 mM $MgCl_2$, 10 mM dithiothreitol and 1 mM each of the four deoxyribonucleoside triphosphates in a volume of 10 μl containing 20 units of ribonuclease inhibitor and 10 units of reverse transcriptase. Homopolymer tails of ~15 deoxyguanylate residues were added to the 3' termini of the cDNAs. The oligo(dG)-tailed cDNA-vector DNA was digested with 10 units of *Sph*I for 2 h in a volume of 20 μl, annealed with 0.9 pmol of the linker fragment and then adjusted to 1 ml of ligase buffer containing 9 units of T4 DNA ligase and incubated at 12°C for 3 h. The ligation reaction was adjusted to 100 mM KCl and 20 units of RNase H were added and the mixture was incubated at 22°C for 1 h. Competent *E. coli* MC1061 (Dagert and Ehrlich, 1979) were transformed with aliquots and transformants were selected on the basis of ampicillin resistance. Approximately 40 000 transformants were pooled and amplified by growth and plasmid DNA was isolated (Birnboim and Doly, 1979).

Acknowledgements

We thank P.Russell for a gift of poly(A)$^+$ RNA, M.Irani for assistance with enzyme assays, M.Parker for assistance with the computer, I.McCullough for preparation of the manuscript and J.Holly for artwork.

References

Ammerer,G., Hitzeman,R., Hagie,F., Barta,A. and Hall,B.D. (1981) in Walton, A.G. (ed.), *Recombinant DNA, Proceedings of the Third Cleveland Symposium on Macromolecules*, Elsevier, Amsterdam, pp. 185-197.

Ballance,D.J., Buxton,F.P. and Turner,G. (1983) *Biochem. Biophys. Res. Commun.*, **112**, 284-289.

Beggs,J.D. (1978) *Nature*, **275**, 104-109.

Bennetzen,J.L. and Hall,B.D. (1982a) *J. Biol. Chem.*, **257**, 3026-3031.

Bennetzen,J.L. and Hall,B.D. (1982b) *J. Biol. Chem.*, **257**, 3018-3025.

Berget,S.M. (1984) *Nature*, **309**, 179-182.

Berse,B., Dmochowska,A., Skrzypek,M., Weglenski,P., Bates,M.A. and Weiss, D. (1983) *Gene*, **25**, 109-117.

Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513-1523.

Boel,E., Hansen,M.T., Hjort,I. and Fiil,N.P. (1984) *EMBO J.*, **3**, 1581-1585.

Branden,C.I., Eklund,H., Cambillau,C. and Pryor,A.J. (1984) *EMBO J.*, **3**, 1307-1310.

Casadaban,M.J. and Cohen,S.N. (1980) *J. Mol. Biol.*, **138**, 179-207.

Ciriacy,M. (1979) *Mol. Gen. Genet.*, **76**, 427-431.

Creaser,E.H., Porter,R.L., Britt,K.A., Pateman,J.A. and Doy,C.H. (1985) *Biochem. J.*, **225**, 449-454.

Dagert,M. and Ehrlich,S.D. (1979) *Gene*, **6**, 23-28.

Doy,C.H., Pateman,J.A., Olsen,J.E., Kane,H.J. and Creaser,E.H. (1985) *DNA*, **4**, 105-114.

Eklund,H., Branden,C.I. and Jornvall,H. (1976) *J. Mol. Biol.*, **102**, 61-73.

Henikoff,S., Kelly,J.D. and Cohen,E.D. (1983) *Cell*, **33**, 607-614.
Innes,M.A., Holland,M.J., McCabe,P.C., Cole,G.E., Wittman,V.P., Tal,R., Watt,K.W.K., Gelfand,D.H., Holland,J.P. and Meade,J.H. (1985) *Science (Wash.)*, **228**, 21-26.
Jornvall,H. (1977) *Eur. J. Biochem.*, **72**, 443-452.
Keller,E.B. and Noon,W.A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7417-7420.
Langford,C.J., Nellen,W., Niessing,J. and Gallwitz,D. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1496-1500.
Lockington,R.A., Sealy-Lewis,H.M., Scazzochio,C. and Davies, R.W. (1985) *Gene* **33**, 137-149.
McKnight,G.L. and McConaughy,B.L. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4412-4416.
Mullaney,E.J., Hamer,J.E., Roberti,K.A., Yelton,M.M. and Timberlake,W.E. (1985) *Mol. Gen. Genet.*, **199**, 37-45.
Norrander,J., Kempe,T. and Messing,J. (1983) *Gene*, **26**, 101-106.
Nunberg,J.H., Meade,J.H., Cole,G., Lawyer,F.C., McCabe,P., Schweickart,V., Tal,R., Wittman,V.P., Flattgard,J.E. and Innis,M.A. (1984) *Mol. Cell. Biol.*, **4**, 2306-2315.
Orr,W.C. and Timberlake,W.E. (1982) *Cell*, **33**, 607-614.
Pateman,J.A., Doy,C.H., Olsen,J.E., Norris,U., Creaser,E.H. and Hynes,M. (1983) *Proc. R. Soc. London Ser. B.*, **217**, 243-264.
Russell,D.W., Smith,M., Williamson,V.M. and Young,E.T. (1983) *J. Biol. Chem.*, **258**, 2674-2682.
Russell,P. (1983) *Nature*, **301**, 167-169.
Russell,P.R. and Hall,B.D. (1983) *J. Biol. Chem.*, **258**, 143-149.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
Sealy-Lewis,H.M. and Lockington,R.A. (1984) *Curr. Genet.*, **8**, 253-259.
Shoemaker,S., Schweickart,V., Ladner,M., Gelfand,D., Kwok,S., Myambo,K. and Innis,M. (1983) *Biotechnology*, **1**, 691-696.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503-517.
Teem,J.L., Abovich,N., Kaufer,N.F., Schwindinger,W.F., Warner,J.F., Levy, A., Leer,R.J., van Raamsdonk-Duin,M.M.C., Mager,W.H., Planta,R.J., Schultz,L., Friesen,J.D., Fried,H. and Rosbash,M. (1984) *Nucleic Acids Res.*, **12**, 8295-8312.
Thomas,P. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201-5205.
Tilburn,J., Scazzachio,C., Taylor,G.G., Zabicky-Zissman,J.H., Lockington,R.A. and Davies,R.W. (1983) *Gene*, **26**, 205-221.
Timberlake,W.E. and Bernard,E.C. (1981) *Cell*, **26**, 29-37.
Williamson,V.M., Bennetzen,J., Young,E.T., Nasmyth,K. and Hall,B.D. (1980) *Nature*, **283**, 214-216.
Wills,C. (1976) *Nature*, **261**, 26-29.
Yelton,M.M., Hamer,J.E., de Souza,E.J., Mullaney,E.J. and Timberlake,W.E. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7576-7580.
Yelton,M.M., Hamer,J.E. and Timberlake,W.E. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1470-1474.
Yelton,M.M., Timberlake,W.E. and Van den Hondel,C.A.M.J.J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 834-838.
Zaret,K.S. and Sherman,F. (1982) *Cell*, **28**, 563-573.

Received on 19 April 1985; revised on 17 May 1985