Volume 238, number 1, 123–128 FEB 06332 September 1988

A nuclear yeast gene (GCY) encodes a polypeptide with high homology to a vertebrate eye lens protein

Ulrich Oechsner, Viktor Magdolen and Wolfhard Bandlow

Institute for Genetics and Microbiology, Maria-Ward-Strasse 1a, D-8000 Munich 19, FRG

Received 2 August 1988

We describe here the nuclear gene for a yeast protein showing unexpectedly high homology with mammalian aldo/keto reductases as well as with ρ-crystallin, one of the prominent proteins of the frog eye lens. Although it could be proven that the gene occurs as a single copy in the haploid yeast genome, replacement of the intact by a disrupted, nonfunctional allele led to no obvious phenotype, indicating that the gene is dispensable. The gene was assigned to chromosome XV. It is transcribed in vivo into an mRNA of about 1300 bases with a coding capacity for a protein of 312 amino acids (estimated M_r 35000).

Aldo/keto reductase family; Crystallin; Nucleotide sequence; Homology; Eye lens protein evolution; (*Saccharomyces cerevisiae*)

1. INTRODUCTION

Crystallins, the major protein components of the vertebrate eye lens, comprise 20 to 60% of the wet weight of the lens, depending on the species examined. They fall into various classes named α, β, γ, δ, ϵ, ϱ and τ. While some of them are present in all vertebrate eye lenses (α, β and γ), the occurrence of others is restricted to certain animal families or is even species specific. Investigations over the past few years accumulated evidence that most crystallins are structurally and, in some cases, functionally related to proteins present in other cell types [1]. These findings led to the theory that the vertebrate lens, being of late evolutionary origin, has been constructed with proteins which originally fulfilled normal metabolic functions.

The work on yeast in our laboratory supports this hypothesis. The amino acid sequence deduced from an open reading frame on plasmid pUHC15-2 showed close homology (37% identical amino acids) with ϱ-crystallin from the European common frog [2]. In confirmation of the above theory even higher homology (about 42% identical amino acids) was found with mammalian aldo/keto reductases [3–5]. We assume the product of the gene GCY to have a related function.

Correspondence address: W. Bandlow, Institute for Genetics and Microbiology, Maria-Ward-Strasse 1a, D-8000 Munich 19, FRG

2. MATERIALS AND METHODS

2.1. *Yeast strains and growth conditions*

The *Saccharomyces cerevisiae* wild-type strain D273-10B MATα (ATCC 25657) was grown on 1% yeast extract with either 2% galactose, 3% glucose or 2% lactate as carbon sources. Strain DL1 MATα, his3, leu2, ura3 [6] and disruption mutants of GCY, named gcyD0 and gcyDD0, MATα, his3, leu2, ura3, gcy::LEU2 were grown on semisynthetic medium supplemented according to their auxotrophies. Yeast strain DczH1-1B MATα, leu1, hem1-1 [7] was grown in the absence of heme in galactose containing, semisynthetic medium supplemented with Tween 80 (2.5 mg/ml), ergosterol (12 μg/ml) and 50 μg/ml of each leucine, methionine and tyrosine. Induction of heme-dependent transcripts was done by adding 15 μM deuteroporphyrin IX to a tenfold concentrated culture of DczH1-1B and allowing 3 more hours of growth. Anaerobiosis of cells, grown in the medium described for DczH1-1B under an atmosphere of nitrogen, was tested by the expression of the ANB1 gene [8].

2.2. *Recombinant DNA methods*

Subcloning was performed using plasmids pUC12 and pUC19 and subsequent transformation into JM109 as the

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

bacterial recipient. Gene disruptions were carried out as described by Rothstein [9].

2.3. *Labelling methods and blotting procedures*

Nick-translation and fill-in reactions for labelling DNA fragments with protruding 5'-ends were made according to Maniatis et al. [10]. The double-stranded plasmid DNA sequencing method described by Chen and Seeburg [11] was used. Mapping the 5'-ends of transcripts by primer elongation and preparation of total yeast RNA was done as described previously [12]. Southern and Northern blots were performed by the original methods as described by Maniatis et al. [10]. For hybridization under stringent conditions 5 × Denhardt's, 0.2% SDS, 1 M NaCl and 50% formamide at 42°C were used. 100 μg/ml herring sperm DNA and 0.5 to 1.5 × 10^6 cpm of radioactive probes per 10 ml hybridization solution were added.

3. RESULTS AND DISCUSSION

A library of genomic yeast DNA in *E. coli* was screened using a competitive hybridization assay with duplicated colony filters. The procedure employed polyA$^+$ RNA isolated from a hem1-1 yeast mutant strain and was originally performed to favour the isolation of genes transcriptionally activated by heme [13]. One of the isolated clones contained the plasmid pUHC15-2. Various fragments, spanning the whole 3.7 kb *Bam*HI DNA insert (fig.1A), were subcloned and sequenced. Two complete major reading frames were detected, one of which was identified as the gene for yeast profilin, an actin-binding protein [14]. The other (GCY) is transcribed convergently from the opposite strand (fig.1B,C). The nucleotide sequence of GCY (fig.2) potentially codes for a protein of 312 amino acids (estimated M_r 35000). The absence of contiguous hydrophobic stretches and of presumptive secretorial or mitochondrial cleavable presequences implies a cytoplasmic location.

Genomic DNA was digested by several restriction endonucleases which do not cut within the coding sequence of GCY and hybridized in a Southern blot to a nick-translated 800 bp *Hpa*II fragment. Only single signals appeared in each lane (fig.3B) arguing that the GCY sequence is unique in the haploid yeast genome. Hybridization of chromosome blots, obtained after transverse alternating field electrophoresis, with the same fragment exhibited one signal that could be attributed to chromosome XV (not shown). This result was verified by successively probing the same blot with markers for this and other chromosomes (shown in [14]). Additional proof was obtained from tetrad analysis demonstrating genetic linkage to ADE2 [14].

To prove transcription of GCY, Northern blots, prepared from total yeast RNA, were hybridized to a 1.2 kb *Pvu*I/*Hin*dII fragment. The autoradiogram shown in fig.4A exhibits one signal in each lane which corresponds to an RNA of about

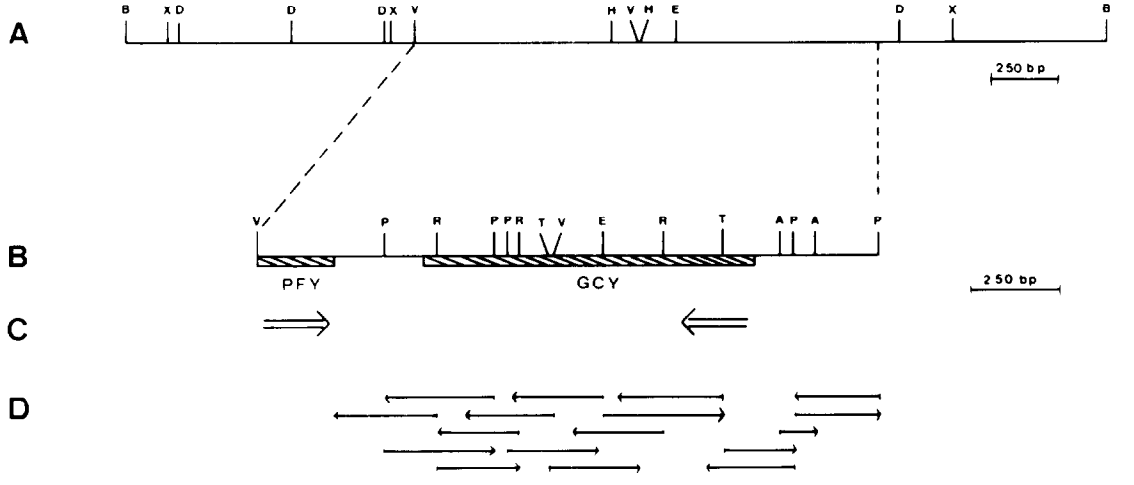

Fig.1. (A) Restriction map of the genomic yeast DNA inserted into pUHC15-2; (B) detailed restriction map of the 1.7 kb region comprising the 3'-part of PFY and the complete GCY gene; (C) arrows indicate the direction of transcription of PFY and GCY; (D) sequencing strategy for the fragment shown in B. Abbreviations used: A, *Alu*I; B, *Bam*HI; D, *Hin*dII; E, *Eco*RV; H, *Hin*dIII; P, *Hpa*II; R, *Rsa*I; T, *Taq*I; V, *Pvu*II; X, *Xho*I.

```
-355   CCGGGGAAGAACAAGGAAGGGCGGTCTTTTCTCCCTCATTGTCATAGCAAGGTCATTTCGCCTTCTCAGAAAGGGGTAG

-276   AATCAATCTAGCACGCAGATTGCAAACACGGCTTAATAATATGCCTATCAGGCATTCACCCGTGTGACGAATCGCACACC

-196   GCTGCTCTCCTTAATTCCCTAGAGTAGAAACCGAGCTTTCAGGAAAAGACTACGGCAGTAAAGAATTGCTTTACTGGGCG

                                                                    o     o       ●●
-116   TATAAAACCGGGAGAATCAAGACATTCTAATGACTTGATTCAGGATGAGAGCTTAATAGGTGCATCTTAGCAAGCTAAAA

                                                  M   P   A   T   L   H   D   S   T   K   I
 -36   TTTGGACAGCTCTCATTACTAAATTAAGATAGAAAA ATG CCT GCT ACT TTA CAT GAT TCT ACG AAA ATC

        L   S   L   N   T   G   A   Q   I   P   Q   I   G   L   G   T   W   Q   S   K
  34    CTT TCT CTA AAT ACT GGA GCC CAA ATC CCT CAA ATA GGT TTA GGT ACC TGG CAG TCG AAA

        E   N   D   A   Y   K   A   V   L   T   A   L   K   D   G   Y   R   H   I   D
  94    GAG AAC GAT GCT TAT AAG GCT GTT TTA ACC GCT TTG AAA GAT GGC TAC CGA CAC ATT GAT

        T   A   A   I   Y   R   N   E   D   Q   V   G   Q   A   I   K   D   S   G   V
 154    ACT GCT GCT ATT TAC CGT AAT GAA GAC CAA GTC GGT CAA GCC ATC AAG GAT TCA GGT GTT

        P   R   E   E   I   F   V   T   T   K   L   W   C   T   Q   H   H   E   P   E
 214    CCT CGG GAA GAA ATC TTT GTT ACT ACA AAG TTA TGG TGT ACA CAA CAC CAC GAA CCT GAA

        V   A   L   D   Q   S   L   K   R   L   G   L   D   Y   V   D   L   Y   L   M
 274    GTA GCG CTG GAT CAA TCA CTA AAG AGG TTA GGA TTG GAC TAC GTA GAC TTA TAT TTG ATG

        H   W   P   A   R   L   D   P   A   Y   I   K   N   E   D   I   L   S   V   P
 334    CAT TGG CCT GCC AGA TTA GAT CCA GCC TAC ATC AAA AAT GAA GAC ATC TTG AGT GTG CCA

        T   K   K   D   G   S   R   A   V   D   I   T   N   W   N   F   I   K   T   W
 394    ACA AAG AAG GAT GGT TCT CGT GCA GTG GAT ATC ACC AAT TGG AAT TTC ATC AAA ACC TGG

        E   L   M   Q   E   L   P   K   T   G   K   T   K   A   V   G   V   S   N   F
 454    GAA TTA ATG CAG GAA CTA CCA AAG ACT GGT AAA ACT AAG GCC GTT GGA GTC TCC AAC TTT

        S   I   N   N   L   K   D   L   L   A   S   Q   G   N   K   L   T   P   A   A
 514    TCT ATA AAT AAC CTG AAA GAT CTA TTA GCA TCT CAA GGT AAT AAG CTT ACG CCA GCT GCT

        N   Q   V   E   I   H   P   L   L   P   Q   D   E   L   I   N   F   C   K   S
 574    AAC CAA GTC GAA ATA CAT CCA TTA CTA CCT CAA GAC GAA TTG ATT AAT TTT TGT AAA AGT

        K   G   I   V   V   E   A   Y   S   P   L   G   S   T   D   A   P   L   L   K
 634    AAA GGC ATT GTG GTT GAA GCT TAT TCT CCG TTA GGT AGT ACC GAT GCT CCA CTA TTG AAG

        E   P   V   I   L   E   I   A   K   K   N   N   V   Q   P   G   H   V   V   I
 694    GAA CCG GTT ATC CTT GAA ATT GCG AAG AAA AAT AAC GTT CAA CCC GGA CAC GTT GTT ATT

        S   W   H   V   Q   R   G   Y   V   V   L   P   K   S   V   N   P   D   R   I
 754    AGC TGG CAC GTC CAA AGA GGT TAT GTT GTC TTG CCA AAA TCT GTG AAT CCC GAT CGA ATC

        K   T   N   R   K   I   F   T   L   S   T   E   D   F   E   A   I   N   N   I
 814    AAA ACG AAC AGG AAA ATA TTT ACT TTG TCT ACT GAG GAC TTT GAA GCT ATC AAT AAC ATA

        S   K   E   K   G   E   K   R   V   V   H   P   N   W   S   P   F   E   V   F
 874    TCG AAG GAA AAG GGC GAA AAA AGG GTT GTA CAT CCA AAT TGG TCT CCT TTC GAA GTA TTC

        K   *
 934    AAG TAA TTGTTTTTGCGTGTTTCTCGTATGATTGTAATATGTAGATAAATTAAACATAAGTATATCAAATGTCGATAT

1012    TCTTAAGACTGTTTCGTAAATATGTTAAATACCGGATGTGACTACATATAAAAAGTGCGCCTTTTACAGTGATAATTAGC

1092    GTCAATCAATAAAATATTTGTAAAATAGGAAGCGAAATCCCGCAGATGCCAGAATAGGTGGTGTATAAGGCAAGAAAACT

1172    TTACCTGCATAAATTA
```

Fig.2. Nucleotide sequence of the region shown in fig.1C and deduced amino acid sequence of GCY. The TATA-like element (underlined), transcription termination signals (overlined), major (●) and minor (○) mRNA initiation sites are indicated.

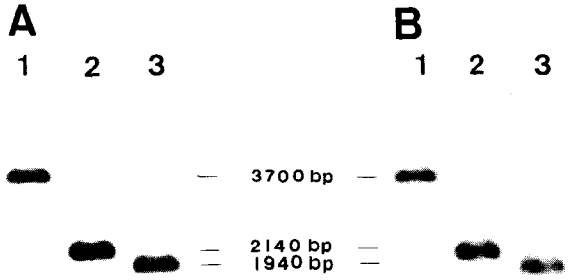

Fig.3. Southern blot hybridization to examine genomic copy number of GCY. (A) pUHC15-2 and (B) genomic yeast DNA restricted with *Bam*HI (lanes 1), *Xho*I (lanes 2) and *Hin*dII (lanes 3).

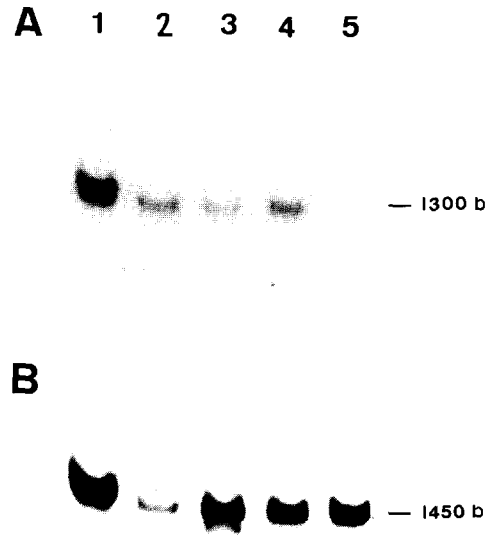

Fig.4. Northern blot hybridized with GCY (A) and Actin (B), respectively. RNA from strain DczH1-1B, hem1-1, after 0 h (lane 1) and after 3 h (lane 2) of growth in the presence of deuteroporphyrin IX. RNA from wt D273-10B grown on galactose with (lane 3) and without O_2 (lane 4). RNA from the disruption mutant gcyDD0 grown on galactose (lane 5).

1300 bases. The RNA in lane 2 was extracted after induction of heme-dependent transcription by adding deuteroporphyrin IX to the cultures [15]. It can be seen that the RNA from uninduced cells (lane 1) exhibits an even stronger signal as compared to that from induced cells, due to a higher RNA concentration in lane 1. Hybridization of the same blot with a probe from the constitutively expressed actin gene (fig.4B) reveals that the relative signal strength of GCY mRNA is about equal in both lanes. According to the assumption that heme is the mediator of oxygen availability [15], also RNA extracted from anaerobically (lane 4) and aerobically (lane 3) grown wild-type yeast was analyzed and revealed no difference in GCY mRNA content either. Both results imply that the procedure for isolation of clone UHC15-2 was not selective for GCY and bears no significance for regulation of GCY transcription by heme.

Determination of the 5′-ends of the mRNA was achieved by annealing a 58 bp *Hin*fI/*Rsa*I fragment, located 21 to 78 bp downstream of the AUG start codon, to total RNA derived from galactose-grown yeast and by its elongation with murine reverse transcriptase. Two major and two minor transcription initiation sites were found as indicated in fig.2. A potential TATA element is present (underlined in fig.2) 52–67 bp upstream of the 5′-ends of the mRNAs. The 3′-terminus of the transcripts has not yet been determined, but several canonic termination signals [13], one of them on the opposite strand, could be found (overlined in fig.2).

Further attempts aimed at the identification of the function of the GCY gene product: as a first approach, the gene was disrupted by replacement of the internal 850 bp *Eco*RV/*Hin*dII fragment, encoding the N-terminal half of the presumptive polypeptide, by a 2.3 kb *Hpa*I fragment carrying the LEU2 marker gene. The haploid yeast strain DL1 was transformed with the linear construct, carrying the disrupted gcy::LEU2 gene on a 3.2 kb *Hin*dIII/*Bam*HI fragment. LEU transformants were selected and tested for correct displacement of GCY by genomic Southern blotting. The blot was probed with both, the original 1.8 kb *Hin*dIII/*Bam*HI fragment of pUHC15-2 (fig.5A) and a LEU2-derived fragment (fig.5B). It can be seen that the length of the genomic fragment in strain gcyDD0 is increased by the insertion as expected, when compared to the wild-type (fig.5A). After hybridization with a LEU2-specific probe, a signal appears at a position identical with the wild-type

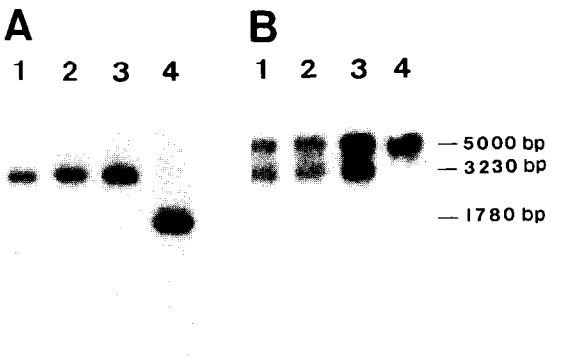

Fig.5. Southern blot of genomic DNA from DL1 (lanes 4) and three independently isolated disruption mutants of the gcyDD0 type (lanes 1,2,3), restricted with *Bam*HI plus *Hin*dIII, hybridized with a GCY- (A) and a LEU2-specific probe (B).

LEU2 fragment in addition to the one characteristic for the disruption (fig.5B). Lack of transcription of the disrupted gene in one of the gcyDD0 mutants could be proven by Northern blot analysis (fig.4, lane 5). Besides the deletion, a similar disruption mutant was constructed by insertion of the 2.3 kb *Hpa*I LEU2 fragment into the *Eco*RV site in the GCY coding region (gcyD0). The growth behaviour of both types of disruption mutants and the wild-type was compared. No differences were found, concerning growth rates on lactate, galactose, glucose or minimal media and adaptation rates determined after a change of carbon sources even after long term shift experiments (20 days). As the disruption mutants are viable in the haploid state and exhibit no obvious phenotype it must be concluded that the gene in question is not important in yeast under the metabolic conditions tested.

Comparison of the nucleotide and the deduced amino acid sequences with published data from a combined sequence data base (Pfeiffer, F., personal communication) revealed extensive homology with three mammalian enzymes: aldose reductase from rat eye lens [3], aldehyde reductase from human liver [4] and prostaglandin F synthase from bovine lung [5]. These enzymes comprise one representative from each of the three subgroups of the aldo/keto reductase family [16]. They exhibit largely overlapping substrate specificities. About 60% homology (identical plus homologous amino acids) of GCY is observed with all three of them (fig.6), so that it is not possible to assign GCY to

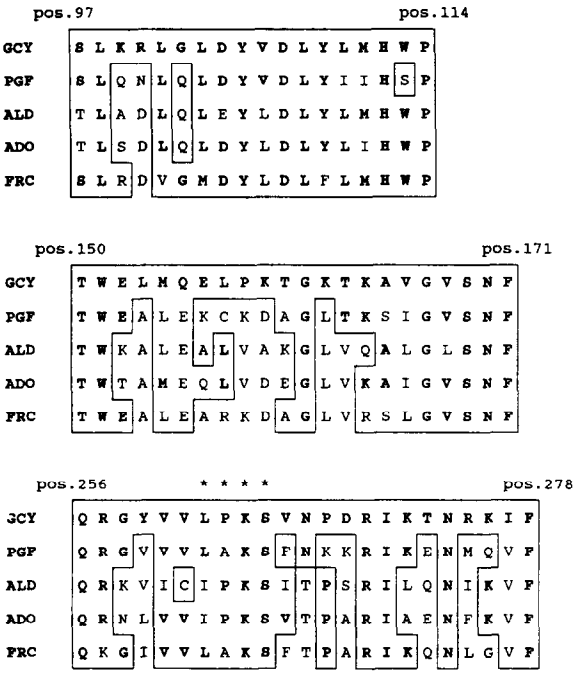

Fig.6. Alignment of three homologous blocks from the amino acid sequences of GCY, ζ-crystallin (FRC), aldose- (ADO) and aldehyde reductase (ALD), and prostaglandin F synthase (PGF). Block 3 contains the presumptive reaction center of mammalian aldo/keto reductases [19], marked by asterisks. Position numbers refer to the deduced amino acid sequence of GCY. Identical (bold characters) and homologous amino acids are boxed. Homologous amino acids are P, A, G, S, T; L, M, I, V; H, K, R; W, Y, F; Q, N, E, D.

one of the subgroups. Although the mammalian enzymes are well characterized with respect to their activities and substrate preferences, there is little idea about their cellular functions [16]. Some evidence exists that they may be involved in the manifestation of diabetic tissue damages [17].

Surprisingly, the deduced amino acid sequence of GCY was also pronouncedly homologous with ϱ-crystallin from the frog eye lens [2]. The GCY gene product shares 59% identical or homologous amino acid positions with ϱ-crystallin (fig.6). At a first view this relationship between a yeast and an eye lens protein does not make sense. But recently, the exploration of eye lenses yielded results that do not further isolate this highly specialized tissue from other body cells. In different animals the fiber cells, building up the lens body, seem to use a variety of proteins for lens construction which are identical or closely related to enzymes that

fulfill quite ordinary metabolic functions in other cells: e.g. ϵ-crystallin is identical with LDH-B4 from bird heart muscle [18]; δ- and τ-crystallins are homologous to argininosuccinate lyase and enolase, respectively [1]. In the further course of evolution some of the genes for these proteins were pertained [18], whereas others were duplicated and modified in their primary structure in such a way as to allow their products to fulfill better their function as lens proteins at the expense of their original enzymic characters. These proteins are highly overexpressed in lenses – a single crystallin may comprise up to 40% of the weight of the lens. An aldo/keto reductase may be the ancestor of ϱ-crystallin [3,5]. An aldose reductase occurs in its active form in rat lenses, where it reaches concentrations up to 0.2%. In *Rana temporaria* such a protein may have been recruited as a crystallin by overexpression, perhaps after gene duplication.

The discovery of a yeast protein homologous to an eye lens protein could shed light on the evolution of lens proteins in higher eukaryotes: A highly specialized tissue, like the eye lens, may have developed using suitable cellular proteins (for instance an aldo/keto reductase) that are present even in lower eukaryotes like yeast in an ancestorial form. It may fulfill its original function in this organism which, however, seems to be not essential. Perhaps in frog similar 'unimportance' was one parameter that allowed the recruitment of this protein as a crystallin specific for this species. Further analysis of GCY and the identification of the natural role of its product in yeast may clarify this hypothesis.

Acknowledgements: The skillful technical help by G. Strobel is gratefully appreciated. Strain DczH1-1B, hem1-1, was obtained from J. Rytka, Warsaw, plasmid pYActI from H. Domdey, Martinsried, and plasmid pYeCYC1(2,4), containing the ANB1 gene, was a gift from R. Zitomer, Albany.

REFERENCES

[1] Wistow, G. and Piatigorsky, J. (1987) Science 236, 1554–1556.
[2] Tomarev, S.I., Zinovieva, R.D., Dolgilevich, S.M., Luchin, S.V., Krayer, A.S., Skryabin, K.G. and Gause, G.G. (1984) FEBS Lett. 171, 297–302.
[3] Carper, D., Nishimura, C., Shinohara, T., Dietzchold, B., Wistow, G., Craft, C., Kador, P. and Kinoshita, J.H. (1987) FEBS Lett. 220, 209–213.
[4] Wermuth, B., Omar, A., Forster, A., Di Francesco, C., Wolf, M., Wartburg, J.-P., Bullock, P. and Gabbay, K.H. (1987) in: Enzymology of Carbonyl Metabolism 3: Aldehyde Dehydrogenase, Aldo-keto Reductase and Alcohol Dehydrogenase (Weiner, H. and Flynn, T.G. eds) pp.297–307, Alan R. Liss, New York.
[5] Watanabe, K., Fujii, Y., Nakayama, K., Ohkubo, H., Kuramitsu, S., Kagamiyama, H., Nakanishi, S. and Hayaishi, O. (1988) Proc. Natl. Acad. Sci. USA 85, 11–15.
[6] Van Loon, A.P.G.M., Van Eijk, E. and Grivell, L.A. (1983) EMBO J. 2, 1765–1770.
[7] Sledziewsky, A., Rytka, J., Bilinski, T., Hörtner, H. and Ruis, H. (1981) Curr. Genet. 4, 19–23.
[8] Lowry, C.V. and Zitomer, R.S. (1984) Proc. Natl. Acad. Sci. USA 81, 6129–6133.
[9] Rothstein, R.J. (1983) Methods Enzymol. 101, 202–211.
[10] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[11] Chen, E. and Seeburg, P.H. (1985) DNA 4, 165–170.
[12] Strobel, G., Magdolen, V., Oechsner, U., Huh, H.-S. and Bandlow, W. (1988) Curr. Genet., in press.
[13] Magdolen, V., Oechsner, U. and Bandlow, W. (1987) Curr. Genet. 12, 405–411.
[14] Magdolen, V., Oechsner, U., Müller, G. and Bandlow, W., submitted.
[15] Schmalix, W., Oechsner, U., Magdolen, V. and Bandlow, W. (1986) Biol. Chem. Hoppe-Seyler 367, 379–385.
[16] Wermuth, B. (1985) in: Enzymology of Carbonyl Metabolism 2: Aldehyde Dehydrogenase, Aldo-keto Reductase and Alcohol Dehydrogenase (Flynn, T.G. and Weiner, H. eds) pp.209–230, Alan R. Liss, New York.
[17] Kador, P.F., Robison, W.G. and Kinoshita, J.H. (1985) Annu. Rev. Pharmacol. Toxicol. 25, 691–714.
[18] Wistow, G.J., Mulders, J.W.M. and De Jong, W.W. (1987) Nature 326, 622–624.
[19] Morjana, N.A., Lyons, C. and Flynn, T.G. (1987) Fed. Proc. 46, 2154.