The EMBO Journal Vol.1 No.6 pp.675-680, 1982

Two yeast acid phosphatase structural genes are the result of a tandem duplication and show different degrees of homology in their promoter and coding sequences

B. Meyhack*, W. Bajwa, H. Rudolph, and A. Hinnen

Friedrich Miescher-Institut, CH-4002 Basel, Switzerland

Communicated by G. Schatz
Received on 13 May 1982

We have cloned the structural genes for a regulated *(PHO5)* and a constitutive *(PHO3)* acid phosphatase from yeast by transformation and complementation of a yeast *pho3,pho5* double mutant. Both genes are located on a 5.1-kb *Bam*HI fragment. The cloned genes were identified on the basis of genetic evidence and by hybrid selection of mRNA coupled with *in vitro* translation and immunoprecipitation. Subcloning of partial *Sau*3A digests and functional *in vivo* analysis by transformation together with DNA sequence analysis showed that the two genes are oriented in the order (5') *PHO5,PHO3* (3'). While the nucleotide sequences of the two coding regions are quite similar, the putative promoter regions show a lower degree of sequence homology. Partly divergent promoter sequences may explain the different regulation of the two genes.

Key words: acid phosphatase genes/gene duplication/molecular evolution/yeast gene library/yeast promoter

Introduction

The expression of acid phosphatase activity in the yeast *Saccharomyces cerevisiae* depends on the concentration of inorganic phosphate (P_i) in the medium. At least two structural genes *(PHO3,PHO5)* and several regulatory genes *(PHO2, PHO4,PHO80,PHO81,PHO85)* control the level of acid phosphatase activity (Schurr and Yagil, 1971; Toh-e *et al.*, 1973; 1975; Ueda *et al.*, 1975). Mapping studies have revealed that the *PHO3* and *PHO5* genes are linked (Hansche *et al.*, 1978), whereas the regulatory genes are unlinked to the *PHO3,PHO5* cluster and dispersed over the yeast genome (Toh-e, 1980; Lange and Hansche, 1980). The *PHO5* gene is repressed at high P_i concentrations. In contrast, expression of the *PHO3* gene occurs constitutively, but at low levels. Repressible and non-repressible acid phosphatase activity is located in the periplasmic space and can therefore be measured with intact cells. The repressible enzyme is glycosylated and has a mol. wt. of ~410 000 (Mildner *et al.*, 1976).

Recently, Kramer and Andersen (1980) isolated from total yeast DNA two *Eco*RI restriction fragments of 5 and 7.9 kb which hydridized selectively to mRNA from cells grown at low P_i concentrations. The synthesis of the mRNA was regulated in a fashion characteristic of repressible acid phosphatase structural genes.

In the present study we have cloned the *PHO5* and *PHO3* genes by transformation of phosphatase-deficient yeast mutants. DNA sequence analysis allowed us to compare the presumptive promoter sequences and part of the coding sequences. Our data suggest that the two genes arose by gene duplication and, possibly through subsequent evolutionary divergence, acquired distinct regulatory mechanisms.

After the completion of this work Rogers *et al.* (1982) published a paper on the acid phosphatase gene family. Their results partly overlap with ours and they independently reach similar conclusions.

Results

Yeast gene library

For the purpose of cloning different genes from *S. cerevisiae* we first constructed a yeast gene library as a convenient source of genomic yeast DNA. The haploid yeast genome contains ~14 000 kb (Lauer *et al.*, 1977). To cover the whole genome by a minimal number of clones, we chose to clone into yeast cosmid vector pYcl (Hohn and Hinnen, 1980). 3000 amp^R colonies were individually grown in the wells of microtiter dishes (yeast gene library). Assuming an average insert length of 30 kb (data not shown), our library should cover the yeast genome six times, reaching a theoretical probability of >99% to have any yeast DNA sequence represented.

Complementation of a pho3,pho5 double mutation

The yeast gene library was screened for clones containing yeast DNA complementing a *pho3,pho5* acid phosphatase double mutant. Replicas of the library were grown on LB agar plates containing 50 μg/ml ampicillin. The colonies were washed off the plates and pooled. DNA was isolated from individual pools and used to transform yeast strain AH216 *(his3,leu2,pho3,pho5)* to HIS^+. Transformants were replica-plated on low P_i minimal medium and stained for acid phosphatase (see Materials and methods). Colonies with a functional *PHO5* gene stain red upon derepression of the gene on low P_i medium. By repeated subpooling of our gene library (Hinnen *et al.*, 1979) we finally isolated three independent clones exhibiting repressible acid phosphatase.

One of these clones (pG7) was further analyzed. The hybrid plasmid had a size of 42 kb, being well within the size range of cosmid clones. *Eco*RI and *Bam*HI fragments of pG7 were subcloned in pBR322*(HIS3)* and pJDB207, respectively (see Figure 1). A 5.1-kb *Bam*HI fragment, which is part of a 8-kb *Eco*RI fragment, was subcloned in the high copy number vector pJDB207. Upon transformation of yeast strain AH216 this hybrid plasmid elicited high phosphatase activity under derepressed (low P_i) conditions and low levels of activity under repressed (high P_i) conditions. We, therefore, tentatively concluded that the whole *PHO3,PHO5* gene cluster had been isolated.

Plasmid linkage of the activity complementing pho5

If the synthesis of acid phosphatase in the selected clone is caused by a gene carried on a plasmid, then a loss of the plasmid (as measured by the plasmid-borne HIS^+ marker) should be accompanied by loss of acid phosphatase from the transformed yeast strains. Plasmids can be lost under non-selective conditions. Therefore, yeast strain AH216 *(leu2, his3,pho3,pho5)* was transformed with pYcl *(HIS3)* carrying the 8-kb *Eco*RI fragment of pG7. Transformed cells were grown on YPD agar plates, replica-plated on low P_i and high

*To whom reprint requests should be sent.

© IRL Press Limited, Oxford, England. 0261-4189/82/0106-0675$2.00/0.

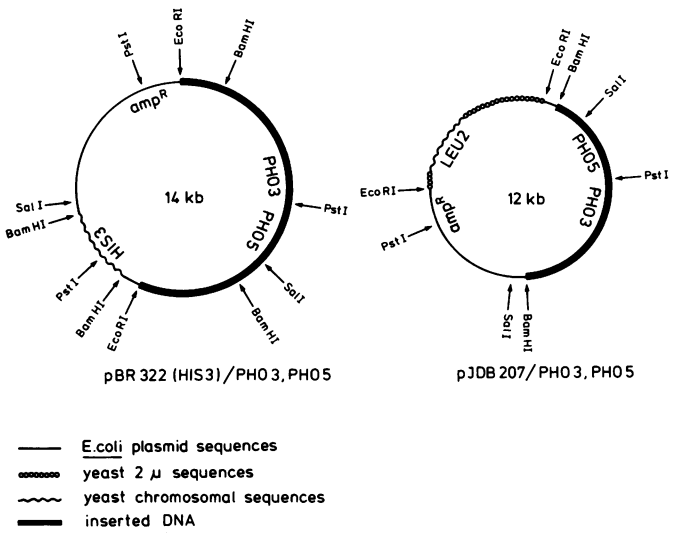

Fig. 1. Restriction map of plasmids with yeast DNA inserts containing the *PHO3, PHO5* genes. Cloning of the 8-kb *Eco*RI and 5.1-kb *Bam*HI fragments is described in Materials and methods.

Table I. Segregation of *PHO5*$^+$ and *HIS3*$^+$ in a genetic cross AH216 (pBR322 *(HIS3)/PHO5,PHO3)* x MC333[a]

	Pho$^+$: Pho$^-$		
	4:0	3:1	2:2
Expected:			
plasmid integrated at *pho5*	100	0	0
plasmid unlinked to *pho5*	17	66	17
Observed:	26	1	0
	His$^+$: His$^-$		
	4:0	3:1	2:2
Expected:			
plasmid integrated at *his3*	100	0	0
plasmid unliked to *his3*	17	66	17
Observed:	5	18	4

[a]Both parent strains used in this cross are Pho$^+$ (measured under derepressed conditions) and His$^+$. The frequency of Pho$^-$ or His$^-$ segregants (3:1 and 2:2 asci) is a measure of the linkage between the chromosomal *pho5* or *his3* locus and the *PHO5*$^+$ and *HIS3*$^+$ genes introduced by transformation.

P_i minimal medium with and without histidine and stained for acid phosphatase. Forty to fifty percent of the colonies on YPD did not grow on minimal medium without histidine, indicating loss of the plasmid. These colonies all lacked acid phosphatase activity under any growth condition. In contrast, all of the *HIS*$^+$ colonies stained strongly for acid phosphatase after growth on low P_i minimal medium but considerably less on high P_i minimal medium. The co-segregation of the His$^+$ and Pho$^+$ phenotypes indicates that the *PHO5* gene is carried on plasmid pG7. Experiments mentioned below will show that the plasmid also carries the structural gene *PHO3*.

Genetic evidence for the origin of the 8-kb EcoRI fragment

In order to determine the genetic origin of the cloned 8-kb *Eco*RI restriction fragment we recloned it into the integrating yeast vector pBR322*(HIS3)*. Plasmid pBR322*(HIS3)/PHO3 PHO5* (see Figure 1) does not contain a yeast replication origin and thus transforms yeast only at low frequency by integrating into a homologous chromosomal location (Hinnen *et al.*, 1979; Hicks *et al.*, 1979). From the extent of homology between the two possible integration sites (1.7 kb for *HIS3*, 8 kb for *PHO3,PHO5*) integration was expected to take place predominantly into or close to the *pho3,pho5* locus. Yeast strain AH216 was transformed with pBR322 *(HIS3)/PHO3 PHO5* and two of the transformants were crossed with strain MC333. In both crosses the linkage relationship between the relevant genetic markers was similar. The data for one of the crosses are given in Table I. They clearly show tight linkage of the cloned region with the *pho3,pho5* locus on chromosome II. Furthermore, the *HIS3* gene of the vector is unlinked to the *his3* locus on chromosome XV.

Hybrid selection of acid phosphatase mRNA

The original phosphatase clone (pG7) had been isolated on the basis of its ability to complement *pho5* in the recipient strain AH216. The mutation had previously been characterized genetically and evidence was presented that *PHO5* and *PHO3* represent two closely linked acid phosphatase structural genes (Toh-e *et al.*, 1975).

To provide direct evidence that the complementing DNA segment contains the structural gene for the regulated acid phosphatase we used the cloned 5.1-kb *Bam*HI DNA fragment for mRNA selection. The *Bam*HI fragment subcloned in pBR325 was bound to nitrocellulose filters and used to select complementary poly(A)$^+$ RNA from either repressed (high P_i) or derepressed (low P_i) cells of the wild-type strain S288C. RNA was eluted and translated *in vitro* in a reticulocyte system in the presence of L-[^{35}S]methionine (Pelham and Jackson, 1976). Translation products were immunoprecipitated by antibodies against purified repressible acid phosphatase. The immunoprecipitates were analyzed by SDS-polyacrylamide gel electrophoresis with subsequent fluorography and compared to the immunoprecipitated *in vitro* translation products obtained with poly(A)$^+$ RNA.

Figure 2A shows the immunoprecipitates of *in vitro* translated total RNA (lanes 2 and 4) and poly(A)$^+$ RNA (lanes 1 and 3) from derepressed (low P_i; Figure 2A, lanes 1 and 2) and repressed (high P_i; Figure 2A, lanes 3 and 4) cells of S288C. The major band obtained from low P_i mRNA has a mol. wt. of ~58 000 as judged from the mol. wt. markers. This band is most likely identical to the p60 gene product described by Bostian *et al.*, (1980). Repressed cells produce only trace amounts of translatable acid phosphatase mRNA (Figure 2A, lanes 3 and 4).

The specificity of the immunoprecipitation reaction was determined by a competition experiment using non-radioactive acid phosphatase. Increasing amounts of acid phosphatase (0.5 – 10 μg/precipitation) progressively prevent immunoprecipitation of the ^{35}S-labeled *in vitro* translation products of mRNA from low P_i cells (Figure 2A; compare lanes 5,6,7 with lane 2). The result indicates that immunoprecipitation is specific for acid phosphatase. All minor precipitation bands in Figure 2A, lanes 1 and 2 are also diluted out by the addition of non-radioactive phosphatase. suggesting that these bands are related to acid phosphatase.

Figure 2B shows the immunoprecipitates of *in vitro* translated mRNA selected by hybridization to pBR325/ *PHO3,PHO5*. The major translation product of hybrid selected mRNA from low P_i cells has a mol. wt. of ~58 000 (Figure 2B, lane 5). The size of the translation product is the same as that of total RNA or poly(A)$^+$ RNA from S288C

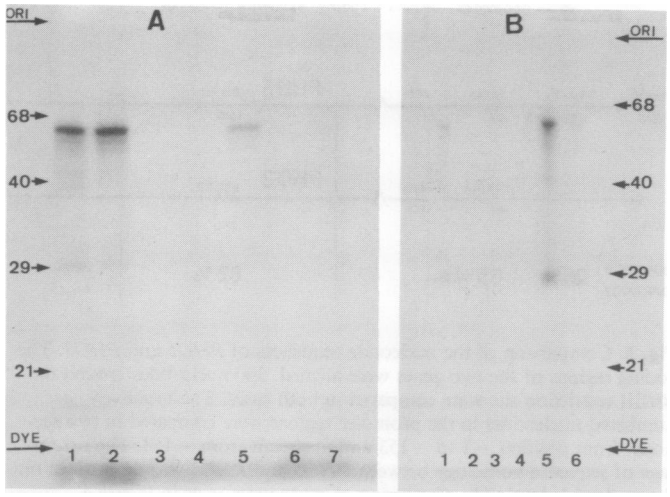

Fig. 2. Analysis of specific *in vitro* translation products. **A:** 5 μg of poly(A)$^+$ RNA from low P_i cells **(lane 1)** or high P_i cells **(lane 3)** and 400 μg of total RNA from low P_i cells **(lane 2)** or high P_i cells **(lane 4)** were used in an *in vitro* translation system. The translation products were immunoprecipitated and analysed on 12% polyacrylamide-SDS slab gels. Radioactive bands were detected by fluorography (see Materials and methods). For immunocompetition experiments the *in vitro* translation products of 400 μg each of total RNA from low P_i cells (as in **lane 2**) were immunoprecipitated in the presence of 0.5 μg **(lane 5)**, 2 μg **(lane 7)** and 10 μg **(lane 6)** of purified non-radioactive acid phosphatase. **B:** Poly(A)$^+$ RNA from low P_i cells **(lane 5)** and high P_i cells **(lane 6)** selected by hybridization to pBR325/*PHO3, PHO5* (~0.2 μg of selected RNA) was translated *in vitro*. The immunoprecipitates were analyzed as above. The following controls were included: poly(A)$^+$ RNA from low P_i cells and high P_i cells selected by hybridization to pBR325 **(lanes 3** and **4)**, 5 μg of poly(A)$^+$ RNA from low P_i cells **(lane 1)**, no RNA added **(lane 2).** Differences in the intensity of corresponding bands (Figure 2A, lane 1 compared to Figure 2B, lane 1) are due to different times of fluorography. The mol. wt. markers are bovine serum albumin (68 000), aldolase (40 000), carbonic anhydrase (29 000), and soybean trypsin inhibitor (21 000). Numbers identify the mol. wts. of the size markers x 10^{-3}. The mol. wt. of the major precipitation band was determined in a semilogarithmic plot to be 58 000. Minor bands are smaller in size and thought to be *in vitro* translation products of partly degraded mRNAs. The minor bands do not appear with freshly prepared mRNA.

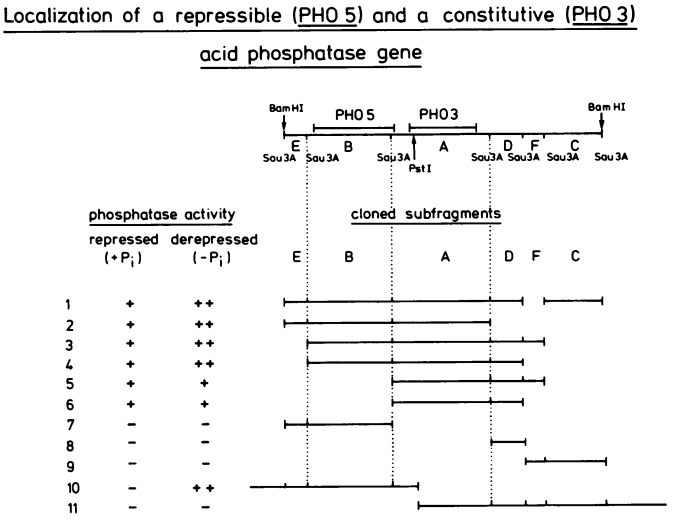

Fig. 3. A partial *Sau*3A digest of the 5.1-kb *Bam*HI fragment was cloned into the *Bam*HI site of self-replicating yeast vector pJDB207. Plasmids (No. 1–9) with different combinations of *Sau*3A fragments (A–F), as determined by restriction analysis, were used to transform the *pho3,pho5* mutant AH216. Transformed colonies were replicated on low P_i and high P_i medium plates (conditions for derepressed and repressed synthesis of acid phosphatase, respectively). Colonies were stained for acid phosphatase activity (see Materials and methods). The intensity of staining is indicated by – (no staining), + (clearly visible red staining) and + + (intense staining). Subclones No. 10, 11 contain two different *Eco*RI-*Pst*I fragments (both from the 8-kb *Eco*RI fragment) which had been cloned in pBR322, with the yeast *HIS3* gene ligated into the *Bam*HI site only subsequently (integrating vector). As a result of this analysis the approximate location of *PHO5* and *PHO3* within the 5.1-kb *Bam*HI fragment is shown.

grown on low P_i medium (Figure 2A, lanes 1 and 2). Hybrid-selected poly(A)$^+$ RNA from high P_i cells only produced a weak precipitation band with an apparent mol. wt. of 58 000 (Figure 2B, lane 6). mRNA hybridized to plasmid pBR325 (which lacks the *PHO3,PHO5* insert) did not direct the synthesis of any immunoprecipitable acid phosphatase polypeptide (Figure 2B, lanes 3 and 4).

In summary, the cloned DNA indeed selects poly(A)$^+$ RNA that directs the synthesis of a polypeptide with the same mol. wt. and similar antigenic properties as obtained with total poly(A)$^+$ RNA. This result confirms that the cloned DNA contains the structural gene for *PHO5*.

Localisation of PHO3 and PHO5 on the cloned DNA

So far we have shown that the cloned 5.1-kb *Bam*HI fragment carries the *PHO5* gene. The following experiment presents evidence that this fragment also contains the *PHO3* gene.

For the localisation of *PHO3* and *PHO5* within the *Bam* fragment we took advantage of the pattern of *Sau*3A restriction sites and a unique *Pst*I site. Digestion of the *Bam* fragment with restriction endonuclease *Sau*3A generates six fragments (A–F, Figure 3). Subcloning of a partial *Sau*3A digest into the *Bam* site of self-replicating yeast vector pJDB207 led to plasmids with different combinations of *Sau*3A fragments. These plasmids were then used to transform the *pho3,pho5* mutant AH216. Transformants were checked for acid phosphatase activity after growth on either low P_i or high P_i medium plates. Clones containing at least *Sau*3A fragments A and B (Figure 3, No. 1–4) expressed acid phosphatase at the same level as the entire 5.1-kb *Bam* fragment. Expression was regulated normally by the concentration of P_i in the medium. Clones with *Sau*3A fragment A only (Figure 3, No. 5, 6) expressed low levels of acid phosphatase, which was not influenced by the P_i concentration in the medium. This indicates that information carried by the *Sau*3A fragment A is sufficient for constitutive acid phosphatase *(PHO3)* expression. *Sau*3A fragment B (Figure 3, No. 7) alone does not lead to any expression of acid phosphatase under either repressed or derepressed conditions. However, a subclone with the complete sequence between the *Bam*HI and *Pst*I sites (Figure 3, No. 10) shows regulated, but not constitutive, synthesis of acid phosphatase. Subclones with *Sau*3A fragments C,D,F (Figure 3, No. 8, 9) have no transforming activity, fragment E alone is inactive, but seems to stimulate the regulated expression of fragment B (unpublished results).

As *Sau*3A fragment B is followed by fragment A, the structural gene for regulated acid phosphatase *(PHO5)* must be located at the 5′ end of the gene cluster followed by the constitutive gene *(PHO3). PHO5* is mostly encoded by *Sau*3A fragment B plus some additional sequences of the adjacent A fragment but contains no essential sequences beyond the *Pst*I site. *PHO3* is enclosed entirely in *Sau*3A fragment A. However, the gene product is inactive when sequences up to the *Pst*I site are missing (Figure 3, No. 11). In conclusion, the

```
                                                        GATCCG
-540.
   AAAGTTGTATTCTAACAAGAATGCGCAAATATGTCAACGTATTTGGAAGTCATCTTATGT
-480.
   GCGCTGCTTTAATGTTTTCTCATGTAAGTCGAACGTCGTCATAAACTTTCAAACGAAGGT
-420.
   AAAAGGTTCATAGCGGCTTTTTCTTTGTCTGCACAAAGAAAATTATATATTAAATTAGCA
-360.
   CGTTTTTCGCATATGACGCAACTGCACAATGCCAAAAAAAGTAAAAGCGATTAAAAGAGT
-300.
   TAATTGAATAGGCAATCTCTAAATGATCGATACAACCTTGGCACTCACACGTGGGACTAG
-240.
   CACAGACTAAATTTATGATTCTGGTCCCTGTTTTCGAAGAGATCGCACATGCCAAATTAT
-180.
   CAAATTGGTCACCTTACTTGGCAAGGCATATACCCATTTGGGATAAGGGTAAACATCTTT
-120.
   GAATTGTCGAAATGAAACGTATATAAGCGCTGATGTTTTGCTAAGTCGAGGTTAGTATGG
 -60.
   CTTCATCTCTCATGAGAATAAGAACAACAACAAATAGAGCAAGCAAATTCGAGATTACCA
  +1.                                                        . +45
   ATG TTT AAA TCT GTT GTT TAT TCA ATT TTA GCC GCT TCT TTG GCC
   MET PHE [LYS] SER VAL VAL TYR SER ILE LEU ALA ALA SER LEU ALA
                                                             . +90
   AAT GCA GGT ACC ATT CCC TTA GGC AAA CTA GCC GAT GTC GAC AAG
   ASN ALA GLY THR ILE PRO LEU GLY [LYS] LEU ALA [ASP] VAL [ASP LYS]
                                                          .+135
   ACC GGT ACC CAA AAA GAT ATC TTC CCA TTT TTG GGT GGT GCC GGA
   THR GLY THR GLN [LYS ASP] ILE PHE PRO PHE LEU GLY GLY ALA GLY
                                                          .+180
   CCA TAC TAC TCT TTC CCT AGC GAC TAT GGT ATT TCT CGT GAT TTG
   PRO TYR TYR SER PHE PRO SER [ASP] TYR GLY ILE SER [ARG ASP] LEU
                                                          .+225
   CCT GAA GGT TGT GAA ATG AAG CAA CTG CAA ATG GTT GGT AGA CAT
   PRO [GLU] GLY CYS [GLU] MET [LYS] GLN LEU GLN MET VAL GLY [ARG] HIS
```

Fig. 4. Nucleotide sequence of the *PHO5* promoter and part of the coding region. Nucleotides of the coding region, including the ATG, are numbered +1 to +225. Nucleotides upstream from the ATG are numbered −1 to −546. The (5′)*Bam*HI site is at position −546. Only the nucleotide sequence of the non-coding DNA strand is shown. The presumptive TATAA box is underlined. Amino acids deduced from the nucleotide sequence of the coding region are listed below the corresponding triplets. Charged amino acids are boxed.

300 bases between the *Sau*3A and *Pst*I sites must contain the 3′-terminal sequence of *PHO5* and some 5′-terminal information for the synthesis of *PHO3*.

Promoter sequences of PHO5 and PHO3

The localization of the two genes on the 5.1-kb *Bam*HI fragment allowed us to sequence the boundaries of the two genes and finally led to the complete *PHO5* sequence and part of the *PHO3* sequence. Figure 4 lists the nucleotide sequence of the non-coding DNA strand from the (5′) *Bam*HI site up to 225 nucleotides into the coding sequence for *PHO5*, including the *PHO5* promoter region and the putative signal sequence of the *PHO5* gene product, the regulated acid phosphatase. The TATAA-box at position −99 from the ATG is underlined. The complete nucleotide sequence of the *PHO5* structural gene will be published elsewhere. Although the total nucleotide sequence of the *PHO3* gene is not yet completed, the promoter region of *PHO3* as well as large stretches in the coding region are sequenced. The data are not confirmed yet by sequencing the complementary strand, therefore the *PHO3* sequence is not included in this paper. The data, however, allow a comparison of nucleotide sequences between *PHO5* and *PHO3* (Figure 5). In comparing the coding regions we considered in each case only 300 nucleotides around the *Bst*EII restriction site. Sequence homology for

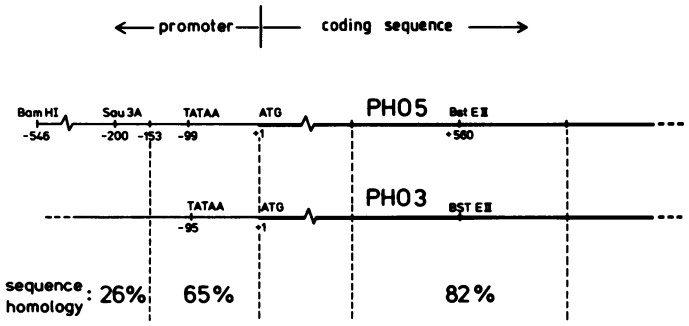

Fig. 5. Comparison of the nucleotide sequences of *PHO5* and *PHO3*. The coding regions of the two genes were aligned. 300 nucleotides around the *Bst*EII restriction site were compared in both cases. The negatively numbered nucleotides in the promoter regions were compared in two sections: from position −1 to −153 and upstream from −154. The percentage of sequence homology between *PHO5* and *PHO3* was determined on the nucleotide level. The numbering of nucleotides is as in Figure 4.

this region was 82%, both at the nucleotide and amino acid level. Sequence homology for the promoter region of both genes up to position −153 is ~65%, including a TATAA box at position −99 *(PHO5)* and −95 *(PHO3)*. No sequence homology is apparent upstream from position −154 (26% represents a random distribution).

The high degree of sequence homology (82%) of part of the coding regions for *PHO5* and *PHO3* suggests that the two genes arose by duplication of an ancestral acid phosphatase gene. Also, part of the two promoter regions shows a reasonably high sequence homology of 65%, suggesting a common origin. However, promoter sequences upstream from position −154 are clearly unrelated (Figure 5). This region of the *PHO5* promoter contains the sequence (5′)GCCAAATT(3′) at position −183 to −190 resembling the consensus sequence for a "CCAAT" box (Benoist *et al.*, 1980) as possible polymerase recognition site.

The amino acid sequence at the N-terminus of regulated acid phosphatase was deduced from the nucleotide sequence of the *PHO5* coding sequence and reveals a stretch of 20 neutral amino acids (No. 4−23) framed by charged amino acids (boxed in Figure 4). This arrangement of charged and neutral amino acids resembles that of a signal sequence, expected for exported proteins.

Discussion

We have used the yeast transformation system to isolate the yeast *PHO3,PHO5* gene cluster. According to Oshima and colleagues (Toh-e *et al.*, 1975) this cluster contains two acid phosphatase structural genes, i.e., the repressible *PHO5* gene and the constitutively expressed *PHO3* gene. Our data show that both genes are located on a single 5.1-kb *Bam*HI restriction fragment which is included within an 8-kb *Eco*RI fragment. It is very likely that the 7.9-kb *Eco*RI fragment isolated by Kramer and Andersen (1980) by differential plaque filter hybridization is identical to the 8-kb DNA region we have cloned.

We present several lines of evidence to prove that the cloned 5.1-kb *Bam*HI restriction fragment carries two distinguishable structural genes for acid phosphatase: (1) the fragment complements a *pho3,pho5* double mutation, and yeast transformation experiments with subclones of the 5.1-kb *Bam*HI fragment identify two functional units; (2) integrating transformants map at the *pho3,pho5* locus; (3) the cloned

DNA, challenged with mRNA from derepressed yeast cells, selectively hybridizes with mRNA coding for a polypeptide precipitable by antibodies raised against purified regulated acid phosphatase; (4) DNA sequence analysis reveals two adjacent coding regions with a high degree of homology; and (5) the amino acid sequence predicted from the 5'-terminal coding region is consistent with a signal sequence expected for exported proteins. In addition, the N termini of two tryptic peptides from purified regulated acid phosphatase fully agree with two amino acid sequences in the protein coding region (M.E. and A.M. Schweingruber, personal communication).

Our data support the model for transcriptional regulation of *PHO5*, presented by Toh-e *et al.*, (1978). The *PHO5* mRNA level is clearly repressed under high P_i conditions as compared to low P_i (Figure 2A, lane 3 compared to lane 1). The *PHO3* gene, on the other hand, seems to be expressed at a low level in a constitutive fashion. This conclusion is based on results from yeast transformants containing only *Sau*3A partial fragment A (Figure 3, No. 5, 6). These transformants show similar acid phosphatase activity at low P_i and high P_i conditions. However, minor differences would not have been detected by the staining assay we used in these experiments.

The *in vitro* translation product of hybrid-selected *PHO5* mRNA has a mol. wt. of 58 000 and may thus correspond to the 60 000-dalton protein described by Bostian *et al.* (1980). Our sequencing data for *PHO5* show an open reading frame of 1263 nucleotides which would correspond to a polypeptide of ~48 000 daltons. This mol. wt. is considerably smaller than the 58 000 of the *in vitro* translation product. The discrepancy would be even larger if the protein was made as a large precursor that is subsequently cleaved to the mature form. The calculated mol. wt. of the *PHO5* gene product does agree well with the presumptive coding capacity of the *PHO5* transcripts (data not shown).

DNA sequencing data for the promoters and the protein-coding sequences of *PHO5* and *PHO3* lead to the following conclusions: (i) coding sequences are relatively conserved (82% sequence homology) probably because of pressure to maintain two functional gene products; (ii) promoter sequences diverge much more (65% sequence homology), with the exception of the TATAA box which is fully conserved; and (iii) the complete loss of sequence homology upstream from position −154 could represent a DNA breaking point as a result of recombination events during gene duplication. The *PHO5* promoter sequences in this region may still contain important signals such as a recognition site for RNA polymerase ("CCAAT" box at −183 to −190).

We can only speculate about the origin of the two genes. Two main possibilities exist: (i) the original gene was constitutively expressed; after duplication the two genes evolved separately and acquired distinct regulatory features, (ii) the original gene was regulated; duplication occurred in a way which led to the separation of the regulatory sequences from the rest of the promoter. The second model can be tested experimentally by exchanging *PHO5* and *PHO3* sequences at the breaking point (position −154).

Yeast may contain more than two acid phosphatase structural genes. By Southern hybridization of total yeast DNA using the 8-kb *Eco*RI restriction fragment as a probe, we have detected a 5-kb *Eco*RI restriction fragment which gives a strong hybridization signal (data not shown). This fragment is possibly related to the 5-kb *Eco*RI fragment obtained by Kramer and Andersen (1980) and may contain a coding region for yet another acid phosphatase protein. It seems that duplication as well as transposition events have been involved in generating the yeast acid phosphatase gene family.

Materials and methods

Strains

E. coli: HB101 (r_k^-, m_k^-, *leu*$^-$, *pro*$^-$, *rec*A$^-$). *S. cerevisiae*: S288C (wild-type); 023-M15 (α, *pho5-27,pho3-1*), obtained from A. Toh-e; AH216 (a, *leu2-3, leu2-112, his3, pho5, pho3*) was constructed by crossing strain 023-M15 with well-transformable yeast strains carrying stable selectable markers for *HIS3* and *LEU2*; MC 333 (α, *leu2, met8, trp1*), obtained from M. Culbertson.

Vectors

Cosmid vector pYcl has been described by Hohn and Hinnen (1980). It consists of the bacterial plasmid pBR322 (carrying the origin of replication and genes for resistance of *E. coli* to ampicillin) joined to a portion of the yeast 2 μ plasmid (allowing autonomous replication in yeast). pYcl also carries the yeast *HIS3* gene (for selection in *his3* strains) and the cohesive end sequence *(cos)* from bacteriophage λ allowing *in vitro* packaging into λ heads (Hohn, 1979; Hohn and Hinnen, 1980). The self-replicating vector pJDB207 has been constructed by Beggs (1981) (see Figure 1). pBR322 (Bolivar *et al.*, 1977) and pBR325 (Bolivar, 1978) were used for subcloning; pBR322*(HIS3)* has been constructed in our laboratory (Hinnen and Meyhack, 1982).

Media

YPD medium contains 2% Bacto Peptone, 1% Bacto Yeast Extract (Difco) and 2% glucose. Yeast minimal medium plates are based on Difco Yeast Nitrogen Base without amino acids to which 2% glucose, the required amino acids (20 mg/l) and 2% agar were added. Low P_i and high P_i medium plates were prepared according to the recipes of the above Difco medium with 1 g/l KH_2PO_4 (high P_i) and 0.03 g/l KH_2PO_4 plus 1 g/l KCl (low P_i), respectively. Liquid low P_i and high P_i media contained 2 g/l asparagine instead of $(NH_4)_2SO_4$. LB medium contains 1% Bacto Tryptone, 0.5% Bacto Yeast Extract (Difco) and 1% NaCl.

Construction of a yeast gene bank in a cosmid vector

30 μg of total high mol. wt. yeast DNA (Olson *et al.*, 1979) from wild-type strain S288C were incubated for 30 min at 37°C with 2 units of *Eco*RI methylase (New England BioLabs) in 250 μl of RI methylation buffer as recommended by the supplier. DNA was precipitated by ethanol, resuspended in 500 μl of 25 mM Tris-HCl pH 8.5, 2 mM $MgCl_2$ (RI* buffer) (Mayer, 1978) and digested with *Eco*RI (Boehringer) until the size distribution of the DNA fragments had a maximum in the 30–50 kb range (a *Xho* digest of λ DNA provided appropriate 33-kb and 17-kb markers). The yeast DNA digested under *Eco*RI* conditions was size-fractionated on a sucrose gradient (5–20% sucrose in 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 6 h at 38 K in a SW 40 rotor). Thirty fractions of 0.4 ml each were collected from the top of the gradient. Fraction 16 contained DNA fragments of 30–40 kb in size. The DNA of this fraction (3 μg) was precipitated with ethanol and ligated for 16 h at 15°C in a total volume of 15 μl to 1 μg of cosmid vector pYcl, linearized by *Eco*RI. The DNA was packaged *in vitro* into bacteriophage λ (Hohn, 1979) and the assembled phages were used to transduce *E. coli* strain HB101. The efficiency of transduction was ~5000 ampicillin-resistant colonies/μg of pYcl vector; 3000 amp^R colonies were picked and grown individually in the wells of microtiter dishes in LB medium containing 50 μg/ml ampicillin.

Yeast transformation

Yeast strains were transformed according to published protocols (Hinnen *et al.*, 1978). Aliquots of the spheroplasts were suspended in 5 ml of regeneration agar, plated on selective yeast minimal medium plates and incubated for 4 days. To derepress synthesis of acid phosphatase, the transformants were replica-plated onto low P_i medium plates. Colonies were stained for acid phosphatase by overlayering them with staining agar [1% Difco agar in 0.1 M acetate buffer pH 4.0, 2 mg/ml Fast Blue Salt (Serva) and 0.2 mg/ml α-naphthyl phosphate (Serva)].

Genetic analysis

Genetic analysis was by standard methods (Sherman *et al.*, 1974).

Cell growth and RNA preparation

S288C was grown overnight in liquid high P_i minimal medium. Cells were spun at 2500 r.p.m. at room temperature, washed with low P_i minimal medium, centrifuged again and suspended to 10^5 cells/ml in low P_i minimal medium. Acid phosphatase activity of whole cells was assayed with p-nitrophenyl phosphate (Merck) as substrate (Toh-e *et al.*, 1973).

Total cellular RNA was prepared from yeast cells harvested in mid-logarithmic phase (10^7 cells/ml). Cells washed with 150 mM NaCl were converted to spheroplasts (Suissa and Schatz, 1982) and then lysed in extraction buffer

(150 mM NaCl, 100 mM Tris-HCl pH 8.6, 5 mM EDTA, 2% SDS) containing 100 μg/ml of proteinase K (Merck). After 30 min of gentle shaking at room temperature, an equal volume of phenol-chloroform-isoamylalcohol (50:50:1) was added, vortexed for 5 min at room temperature and spun at 10 000 r.p.m. for 10 min. The aqueous phase was re-extracted with phenol-chloroform. The phenol extraction was repeated at least three times. For the final extraction, phenol saturated with extraction buffer without SDS was used. Total nucleic acids were precipitated by ethanol, resuspended in 10 mM Tris-HCl pH 7.4 and RNA was selectively precipitated with 3 M LiCl at 0°C for at least 4 h. The LiCl precipitation was repeated until a white pellet was obtained. The RNA was desalted by precipitation with ethanol (66%, 6 h at −20°C) and finally resuspended in 10 mM Tris-HCl pH 7.4 at 300−400 A_{260} units/ml (yield: 400 A_{260} units/g wet cells). An aliquot (150 A_{260} units) of total RNA was applied to a 0.5 ml bed of oligo(dT)-cellulose (Type 2, Collaborative Research, Inc.) in 0.5 M LiCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.5% SDS. The column was washed extensively with the same buffer. Poly(A)$^+$ RNA was eluted with 10 mM Tris-HCl pH 7.4, 1 mM EDTA, 0.05% SDS and then precipitated with ethanol as above.

Selection of acid phosphatase mRNA

Twenty to thirty μg of poly(A)$^+$ RNA or 200−300 μg of total cellular RNA, from yeast cells grown under repressed or derepressed conditions were dissolved in 2 ml of 70% formamide (v/v), 0.3 M NaCl, 5 mM EDTA, 10 mM Hepes pH 7.5, 0.2% SDS ("R-loop solution") and hybridized for 4 h with gentle shaking to two nitrocellulose filters (2 x 2 cm), each containing 5 μg of the cloned *Bam*HI fragment *(PHO3,PHO5)* subcloned in pBR325, or 5 μg of pBR325 as a control. The plasmids had been linearized by *Eco*RI restriction. Each filter was washed twice at 47°C for 15 min in 5 ml of "R-loop solution", and mRNA bound to each filter was eluted twice for 30 min at 47°C with 1 ml of 90% formamide (v/v), 10 mM Hepes pH 7.5, 1 mM EDTA, 0.1% SDS. The hybrid-selected mRNA (0.4 μg) was precipitated with ethanol, washed twice with ethanol and dissolved in 10 μl of 10 mM Tris-HCl pH 7.4.

Cell-free protein synthesis and immunoprecipitation

RNA was translated for 50 min at 37°C in a cell-free reticulocyte system (Pelham and Jackson, 1976) in the presence of L-[S^{35}]methionine (1000 Ci/mmol, 3.0 x 10^7 c.p.m./ml). Specific cell-free translation products were immunoprecipitated (Suissa and Schatz, 1982) using antibodies prepared against purified acid phosphatase from derepressed yeast cells. For immunocompetition experiments, the precipitation was done in the presence of non-radioactive acid phosphatase. Immunoprecipitates were bound to glutaraldehyde-fixed *Staphylococcus aureus*, dissociated by heating to 100°C for 2 min in the presence of 0.8% SDS and 2% β-mercaptoethanol and analyzed on 12% SDS-polyacrylamide gels.

Specific *in vitro* translation products were detected by fluorography (Chamberlain, 1979). Mol. wt. markers were stained by Coomassie Blue.

Construction of recombinant plasmids

T4 DNA ligase and restriction endonucleases except *Eco*RI (Boehringer) were from New England BioLabs. These enzymes were used according to the suppliers specifications. Recombinant plasmids were constructed by using the original acid phosphatase clone (pG7) and the plasmids pBR322*(HIS3)* and pJDB207 (see Figure 1). Mixtures of corresponding restriction fragments were ligated and *E. coli* strain HB101 was transformed to ampicillin resistance. Plasmid isolation was according to Clewell (1972).

DNA sequencing

DNA sequencing was performed according to Maxam and Gilbert (1980).

Acknowledgements

Dr. A. Toh-e kindly provided us with the phosphatase mutant 023-M15. We thank Professor G. Schatz for helpful discussion, and Mrs. D. Martin for excellent technical assistance. We also want to thank Mr. V. Mysicka for his help in setting up an adequate computer program for sequencing data.

References

Beggs,J.D. (1981) in Williamson,R. (ed.) *Genetic Engineering,* **2**, Academic Press, London/NY, pp. 175-203.
Benoist,C., O'Hare,K., Breathnach,R., and Chambon,P. (1980) *Nucleic Acids Res.*, **8**, 127-142.
Bolivar,F., Rodriguez,R., Greene,P.J., Betlach,M., Heyneker,H.L., Boyer, H.W., Crosa,J., and Falkow,S. (1977) *Gene,* **2**, 95-113.
Bolivar,F. (1978) *Gene,* **4**, 121-136.
Bostian,K.A., Lemire,J.M., Cannon,L.E., and Halvorson,H.O. (1980) *Proc. Natl. Acad. Sci. USA,* **77**, 4504-4508.
Chamberlain,J.P. (1979) *Anal. Biochem.*, **98**, 132-135.
Clewell,D.B. (1972) *J. Bacteriol.*, **110**, 667-676.
Hansche,P.E., Beres,V., and Lange,P. (1978) *Genetics,* **88**, 673-687.
Hicks,J., Hinnen,A., and Fink,G.R. (1979) *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 1305-1313.
Hinnen,A., Hicks,J.B., and Fink,G.R. (1978) *Proc. Natl. Acad. Sci. USA,* **75**, 1929-1933.
Hinnen,A., Farabaugh,P.J., Ilgen,C., and Fink,G.R. (1979) in Axel,R., Maniatis,T., and Fox,M. (eds.), *Eukaryotic Gene Regulation,* vol. **14**, Academic Press, NY, pp. 43-51.
Hinnen,A., and Meyhack,B. (1982) in Henle,W., Hofschneider,P.H., Koprowski,H., Melchers,F., Rott,R., Schweiger,H.G., and Vogt,P.K. (eds.), *Current Topics in Microbiology and Immunology,* vol. **96**, Springer Verlag, Berlin/Heidelberg, pp. 101-117.
Hohn,B. (1979) in Colowick,S.P. and Kaplan,N.O. (eds.), *Methods in Enzymology,* vol. **68**, Academic Press, NY/London, pp. 299-309.
Hohn,B., and Hinnen,A. (1980) in Setlow,J.K., and Hollaender,A. (eds.), *Genetic Engineering Principles and Methods,* vol. **2**, Plenum Press, NY/London, pp. 169-183.
Kramer,R.A., and Andersen,N. (1980) *Proc. Natl. Acad. Sci. USA,* **77**, 6541-6545.
Lange,P., and Hansche,P.E. (1980) *Mol. Gen. Genet.*, **180**, 605-607.
Lauer,G., Roberts,T.M., and Klotz,L.C. (1977) *J. Mol. Biol.*, **114**, 507-526.
Maxam,A.M., Gilbert,W., (1980) in Colowick,S.P., and Kaplan,N.O. (eds.), *Methods in Enzymology,* vol. **65**, Academic Press, NY/London, pp. 499-560.
Mayer,H. (1978) *FEBS Lett.*, **90**, 341-344.
Mildner,P., Barbaric,S., Golubic,Z., and Ries,B. (1976) *Biochim. Biophys. Acta,* **429**, 274-282.
Olson,M.V., Loughney,K., and Hall,B.D. (1979) *J. Mol. Biol.*, **132**, 387-410.
Pelham,H.R.B., and Jackson,R.J. (1976) *Eur. J. Biochem.*, **67**, 247-256.
Rogers,D.T., Lemire,J.M., and Bostian,K.A. (1982) *Proc. Natl. Acad. Sci. USA,* **79**, 2157-2161.
Schurr,A., and Yagil,E. (1971) *J. Gen. Microbiol.*, **65**, 291-303.
Sherman,F., Fink,G.R., and Lawrence,C.W. (1974) *Methods in Yeast Genetics,* published by Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Suissa,M., and Schatz,G. (1982) *J. Biol. Chem.*, in press.
Toh-e,A., Ueda,Y., Kakimoto,S., and Oshima,Y. (1973) *J. Bacteriol.*, **113**, 727-738.
Toh-e,A., Kakimoto,S.I., and Oshima,Y. (1975) *Mol. Gen. Genet.*, **143**, 65-70.
Toh-e,A., Kobayashi,S., and Oshima,Y. (1978) *Mol. Gen. Genet.*, **162**, 139-149.
Toh-e,A. (1980) *Genetics,* **94**, 923-932.
Ueda,Y., Toh-e,A., and Oshima,Y. (1975) *J. Bacteriol.*, **122**, 911-922.