Cloning, molecular characterization and chromosome localization of the inorganic pyrophosphatase (PPA) gene from *S.cerevisiae*

Lee F.Kolakowski Jr, Manfred Schloesser and Barry S.Cooperman

University of Pennsylvania, Department of Chemistry, Philadelphia, PA 19104, USA

Received September 30, 1988; Accepted October 24, 1988

Accession no. X13253

ABSTRACT

The gene for *Saccharomyces cerevisiae* inorganic pyrophosphatase, *PPA*, has been cloned by hybridization of "long" oligonucleotide probes with both cDNA and genomic *S. cerevisiae* libraries. The nucleotide sequence of 1612 bp from a genomic subclone that includes the entire coding region gives a deduced amino acid sequence that has nine differences (out of a total of 286 residues) from the previously published amino acid sequence that was determined directly. The codon usage in *PPA* is as expected for a "highly expressed" yeast gene. The upstream region contains a poly dA/dT sequence that might comprise a constitutive promoter. The *PPA* gene appears to be present in a single copy within the *S. cerevisiae* genome and has been localized to chromosome II.

INTRODUCTION

Saccharomyces cerevisiae inorganic pyrophosphatase (PPase) (E.C. 3.6.1.1) is among the best characterized simple phosphoryl transfer enzymes (1). It is a homodimer with a published sequence containing 285 amino acids per subunit (2). A crystal structure at 3Å resolution has also been published (3,4) and work on a 2.6Å structure is nearing completion (D. Voet, personal communication). In addition, numerous kinetic (5-8), spectroscopic (9-14) and chemical modification (15-17) studies have resulted in detailed proposals for its active site structure and catalytic mechanism (3,14,17).

The *in vivo* role of PPase is also of considerable interest. Ermakova *et al.* (18) and ourselves (N. Ahmad and B.S. Cooperman, in preparation) have shown that total PPI levels in yeast cells can get quite high (up to 50mM). The relationship of PPase to PPI levels remains to be determined, but is of interest because of the potential role internal PPI concentrations may play in regulating the flux through biosynthetic reactions (e.g. RNA, DNA, oligosaccharide biosynthesis, tRNA charging). While little is known about the regulation of PPase expression in yeast, in *E. coli* a clear relationship has been established between nucleotide pools, rates of DNA synthesis and PPase levels (19,20).

In this report, we describe the cloning, molecular characterization and chromosome localization of the PPase gene, *PPA*, from *Saccharomyces cerevisiae*. This represents an important step toward our dual goals of preparing site-specific mutants to deepen our understanding of the active site structure and mechanism of action of this enzyme and of

establishing the nature of the linkage between *in vivo* PPase levels and the levels of PPi and other phosphorus metabolites.

MATERIALS AND METHODS

Restriction enzymes were obtained from Bethesda Research Laboratories (BRL), Boehringer Mannheim Biochemicals (BMB), and New England Biolabs (NEB) and were used as specified by the supplier. T4 DNA ligase was obtained from either BRL or NEB. T4 polynucleotide kinase was obtained from Pharmacia. Calf intestinal phosphatase was obtained from BMB. All materials for DNA sequencing except for specific oligonucleotide primers were part of the Sequenase Kit from United States Biochemicals. Oligonucleotide primers for DNA sequencing were prepared by the University of Pennsylvania Cancer Center - DNA Synthesis Service. All materials for unidirectional deletions were part of the Erase-A-Base Kit from Promega. Radioisotopes including [γ - 32 P]dATP, [α - 35 S]dATP, and [α - 32 P]dCTP were obtained from Amersham Corp. Long oligonucleotide probes 1 (a 53-mer) and 2 (an 86-mer) were kindly provided by Dohn Glitz (UCLA) and Naomi Thompson (Wistar Institute), respectively.

Southern Blot Analysis

Yeast genomic DNA from strains KT743, Red Star (commercial yeast), and Budweiser (commercial yeast) was prepared (21), and digested with restriction enzymes. Digested DNA was transferred from agarose electrophoretic gels to nitrocellulose filters in 20X SSC (3.5 M NaCl, 0.3 M Na Citrate pH 7.0). Filters were baked at 80°C for 1.5 hours, then prehybridized for 3 hours in 1 M NaCl, 0.1M Tris pH 8.0, 1X Denhardt's (0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin (Pentax V fraction, Sigma), and 2 mM NaPPi. Probe (either 1 or 2) was labeled with [γ - 32 P]dATP, using T4 polynucleotide kinase, and added to the prehybridization mix. Following incubation overnight at 50°C, filters were washed at 55°C (probe 1) or 60°C (probe 2) for 2 hours and then exposed with a Dupont Cronex Intensifying Screen to Kodak XAR-5 film for 18-72 hours at -70°C.

Screening of cDNA and Genomic DNA Libraries

The *S. cerevisiae* λ gt10 cDNA library made from strain X2180 was a gift from R.A. Young (Whitehead Institute, M.I.T.), and the λ gt11 genomic DNA library from the same strain was a gift from M. Snyder (Yale University). The cDNA library was made using the S1 nuclease procedure (22) and cloned into the EcoRI site of the λ gt10 *imm434* gene. The genomic DNA library was prepared by shearing yeast genomic DNA to an average size of 5 kbp and then cloned into the EcoRI site of λ gt11 *lacZ* gene. Plaque hybridizations were performed on both the genomic and cDNA libraries as described by Benton and Davis (23). Probe preparation and hybridization conditions were the same as in the Southern analysis, except that washes with probe 2 were performed at 52°C. All filters were probed in duplicate to minimize false positives. All positives were rescreened to isolate single plaques. A general review of these methods may be found in reference 22.

Confirmation of Positives

Positive phage from both the cDNA and genomic libraries were prepared on small-scale following Blattner *et al.* (24) except that 0.2 ml of STE (0.3 M Tris HCl pH 9.0, 0.15 M EDTA, and 1.5% SDS) was added to 0.6 ml of clarified lysate, then heated for 15 minutes at 70°C. DNA was prepared on a larger scale from phage which contained a hybridization positive 0.7kbp EcoRV band on Southern blot analysis (22,25).

Subcloning of PPA Fragments

The EcoRI inserts from positive λ gt10 and λ gt11 phage were isolated and ligated (26) into EcoRI digested, dephosphorylated pUC118 (27). Ligations were used to transform JM105 competent cells, prepared as described by Hanahan (28), and plated on XIA plates (0.004% X-gal, 0.1 mM IPTG, 100 μ g/ml Ampicillin). Recombinant transformants were confirmed by restriction enzyme digest analysis on small-scale plasmid preparations prior to large scale plasmid preparations (29).

Sequencing of cDNA and Genomic Clones

All DNA sequencing was performed using the chain termination method (30) on plasmid DNA prepared by alkaline denaturation (31). Sequences were obtained using the M13 universal primer (BRL 15-mer), the M13 reverse primer (NEB 16-mer), and specific oligonucleotide primers. Unidirectional deletions performed as described by Henikoff (32) were used for some of the cDNA sequencing. Sequence analysis was performed using the Staden Software Package from the MRC (33) and the NIH Bionet Computing Resource operated by Intelligenetics.

Yeast Chromosome Analysis

CHEF gels of yeast chromosomes were run according to Chu *et al.* (34) for 17 hours with a 60-second switching interval and 8 hours with an 80-second switching interval.

RESULTS AND DISCUSSION

Design of Oligonucleotide Probes

We employed two "long" [32 P]-labeled single-stranded oligo-DNA probes (35,36) to clone the PPA gene from *S. cerevisiae*. The design of these probes was based on the published amino acid sequence for yeast PPase (2) and the known codon bias for highly expressed *S. cerevisiae* proteins (37). Probe 1 is a 53-mer corresponding to amino acids 81-98. This peptide is the longest one in PPase having a clearly favored codon bias. Probe 2 is an 86-mer corresponding to amino acids 186-214, the oligopeptide having the second most biased sequence in PPase. In addition, probe 2 is useful in being closer to the carboxy-terminus, whereas probe 1 is closer to the amino-terminus.

PROBE 1 5-AACTGTTTCCACACACGGTTACATCCACAACACGGTGCTTTCCACAAAC-3'

PROBE 2 5-GCGTACTTCTTGTCTTAGCTTCACCAGAGAAAGCGAATTG
 GTTTTCTGGCTTACCGTCTGGGATCTTGTAGATTCTGAACCATTC-3'

Southern blot analysis of yeast genomic DNA

Southern blot analysis of restriction enzyme digests of total genomic DNA of *S. cerevisiae*,

Table 1. Fragment Size (kbp) Hybridizing to probes 1 and 2 in Southern blots of restriction enzyme digests of genomic DNA

Enzyme	Size	Enzyme	Size	Enzyme	Size	Enzyme	Size
ClaI	1.8	EcoRV	0.7	BamHI	10.0	HindIII	10.0
EcoRI	18.0	SalI	19.0	PvuI	20.0	PstI	13.0
XbaI	16.0	AccI	5.5	SphI	5.0	PvuII	15.0
XhoI	7.0	Fnu4HI	0.7	—	—	—	—

using [^{32}P]-labeled probes 1 and 2, showed that both probes bind to a single restriction fragment (Table 1), thus providing strong evidence that the *PPA* gene is present in a single copy within the yeast genome. In addition, comparison of the hybridization patterns obtained with genomic DNA from different strains of *S. cerevisiae* showed considerable restriction fragment length polymorphism. A 0.7kbp EcoRV restriction fragment contained within a 1.8kbp ClaI fragment bound both probes quite specifically, and proved quite useful as a diagnostic tool for the presence of the *PPA* gene in screens of the cDNA and genomic libraries (Figure 1).

As determined by washing hybridized filters at progressively higher temperatures, the T_m for probe 2 hybridization to the *PPA* gene (72°C) was approximately 15°C higher than the T_m for probe 1 hybridization.

Screening of *S. cerevisiae* cDNA and genomic libraries

Screening of $\sim 9 \times 10^5$ λ gt10 recombinant phage in a cDNA library with probe 2 led to the identification of four positives. Secondary screening gave single, pure plaques (Figure 2). Southern blot analysis showed that each of the clones contained EcoRI inserts of length 1-1.3kb. and that each contained the 0.7kbp EcoRV fragment that hybridized strongly to both probes 1 and

Figure 1: Southern blot of genomic and isolated lambda phage DNA. Samples 1-4 were digested with EcoRI (RI) and EcoRV (RV). Genomic DNA pooled from ClaI and EcoRV digests were run in lane G. Probe was the 86-mer prepared as described in the Materials and Methods section.

Figure 2: A secondary plaque screening of a λ gt10 cDNA library. The many positives shown are from a single positive in the first screen.

2 (Figure 1). Large-scale phage preparation (25) of each clone yielded sufficient DNA insert for subcloning into EcoRI digested, dephosphorylated pUC118 plasmid, yielding three unique subclones, pFK6, pFK7, and pFK8.

An essentially identical procedure was used to isolate three positives on screening of $\sim 2 \times 10^5$ λ gt11 phage (15% recombinant) in the genomic library. Southern blots of the 3-4kbp inserts showed the presence of the 0.7kbp EcoRV fragment. Large scale preparation of DNA from the largest clone and subcloning into pUC118 yielded subclones pMI3 and pMI4 that contained a 4kbp insert in two different orientations relative to the multiple cloning site in pUC118.

Figure 3: The restriction map and sequencing strategy of cDNA and genomic clones. Restriction enzymes are: E - EcoRV, C - ClaI, D - DraI, M - MboI, Nc - NcoI, N - NlaIV, Cf - CfoI, B - BlnI. Arrows labeled with a 'p' were extensions with specific primers and 'd' were extensions from deletions.

[illegible]

Table 2. Differences between the amino acid sequence of PPase^a and the sequence deduced from the PPA Gene

AA#	PPase	PPA	Possible Cause
-1	None	Met	N-Terminal Amino Peptidase activity
40	Asp	Asn	Protein Sequencing error or Strain Difference (single base change)
71	Asn	Asp	Protein Sequencing error or Strain Difference (single base change)
72	ThrLys	ThrLysLys	Protein Sequencing error
117	Asn	Asp	Protein Sequencing error or Strain Difference (single base change)
123	Gln	Glu	Protein Sequencing error or Strain Difference (single base change)
136	Glu	Gln	Protein Sequencing error or Strain Difference (single base change)
186	Asp	Asn	Protein Sequencing error or Strain Difference (single base change)
224	Asn	Asp	Protein Sequencing error or Strain Difference (single base change)
266	Pro	Leu	Protein Sequencing error or Strain Difference (single base change)

^aprepared from Red Star Yeast (ref. 2)

open reading frame beginning with ATG at position 390 is fully consistent with the amino-terminal amino acid sequence TYTT reported by Cohen *et al.* (2), except for the methionine in position one which is presumably cleaved off by N-terminal aminopeptidase activity in *S. cerevisiae*. The predicted carboxy-terminal sequence SGSV is also in accord with the results of Cohen *et al.* and in the gene is followed by a TAA triplet, which is the most common termination codon in "highly expressed" yeast genes (37,38). The full gene-deduced amino acid sequence is, in general, in very good agreement with the amino acid sequence of Cohen *et al.* (2). The differences are summarized in Table 2, the most important being the insertion of an additional lysine between amino acid residues 73 and 74 (LysGly) as numbered according to Cohen *et al.* (2). This corresponds to deoxynucleotides 612-614 in Figure 4. A detail of the DNA sequencing gel in this region is presented in Figure 5. The inserted Lys comes at a position for which the original amino acid sequencing data is weakest, and we believe the original work is simply in error (note that LysLys sequences are frequently difficult to determine). Thus, the proper number of amino acid residues per subunit in *S. cerevisiae* is 286, rather than 285 as described by Cohen *et al.* (2). Also noteworthy are the large number (7 in all) of Asp/Asn and Glu/Gln differences. It is possible that all seven of these changes are due to protein sequencing errors, with the three in the direction carboxylic acid (amino acid sequence) to carboxamide (gene-deduced sequence) attributable to hydrolysis of the carboxamide during amino acid sequence determination. Protein sequencing errors might also account for the four changes in the opposite direction, since with the technology available when the PPase amino acid sequence was determined, in the middle 1970's (2), residues were often identified as Asn or Gln rather than Asp or Glu on the basis of quite low recoveries of the carboxamide sidechains (R.L. Heinrikson, private communication). It is also possible that some or all of these differences are due to strain differences, since the paired residues are obviously isosteric and are interconvertible with single-base changes in the DNA.

The codon usage for the PPA gene is very good accord with codon usage for "highly expressed" genes in *S. cerevisiae* (37), as demonstrated in Table 3. Here it should be noted that

Figure 5: The presence of a lysine residue (*-Lys) not present in the reported protein sequence (2) was deduced from the region of DNA sequence shown.

probes 1 and 2 had only 4 and 6 mismatches, respectively, with the *PPA* gene, and that for each probe the longest correct match was 17 bases.

5' and 3' Flanking Sequences of the *PPA* Gene

Several sequences in the upstream region (bases 1-391) resemble the consensus sequences for regulatory or structural elements of interest. **First**, the canonical sequence "TATA" has been shown to be important for the transcription of messenger RNA in *S. cerevisiae* and other eukaryotes. While there are no exact TATA sequences, there are sequences beginning at bases 4, 109, 255, 291, and 322 that are TATA-like (39). The TATA-like sequence at base 4 is most probably too far away to be functional. **Second**, Struhl *et al.* (40,41) have identified a region, termed poly dA/dT, that in *S. cerevisiae* genes *HIS3*, *PET56*, and *DED1* seems to function as a constitutive promoter. Beginning at base 233 there is a stretch of 17 bases, of which 14 are T. This region might comprise a constitutive promoter for the *PPA* gene. Consistent with the notion that *PPA* is constitutive are the results of Ermakova *et al.* (18) and ourselves (Ahmad and Cooperman, in preparation) showing that, over the course of a yeast growth curve, PPase

Table 3. Comparison of *S. cerevisiae* Codon Bias with Observed Bias^a

AA	CODON	%	Obs	AA	CODON	%	Obs	AA	CODON	%	Obs	AA	CODON	%	Obs
F	TTT	18	21	S	TCT	54	53	Y	TAT	18	13	C	TGT	100	88
F	TTC	82	79	S	TCC	31	36	Y	TAC	82	87	C	TGC	0	12
L	TTA	22	13	S	TCA	0	4	*	TAA	100	87	*	TGA	0	8
L	TTG	61	75	S	TCG	0	2	*	TAG	0	5	W	TGG	100	100
L	CTT	0	2	P	CCT	11	12	H	CAT	33	26	R	CGT	17	11
L	CTC	0	1	P	CCC	5	1	H	CAC	67	74	R	CGC	0	0
L	CTA	17	7	P	CCA	84	86	Q	CAA	86	95	R	CGA	0	0
L	CTG	0	2	P	CCG	0	1	Q	CAG	14	4	R	CGG	0	0
I	ATT	63	45	T	ACT	42	49	N	AAT	6	14	S	AGT	0	3
I	ATC	37	53	T	ACC	53	45	N	AAC	94	86	S	AGC	15	2
I	ATA	0	2	T	ACA	5	5	K	AAA	21	19	R	AGA	67	87
M	ATG	100	100	T	ACG	0	1	K	AAG	79	81	R	AGG	17	2
V	GTT	79	55	A	GCT	55	68	D	GAT	54	42	G	GGT	94	95
V	GTC	7	41	A	GCC	36	28	D	GAC	46	58	G	GGC	6	4
V	GTA	7	1	A	GCA	9	3	E	GAA	84	92	G	GGA	0	0
V	GTG	7	3	A	GCG	0	1	E	GAG	16	8	G	GGG	0	1

^a The columns labeled % and Obs. are the percent usage of that codon in the *PPA* gene and the observed percentage of that codon in "highly expressed" genes (37), respectively.

levels stay almost constant while PPI levels vary over a wide (>10-fold) range. Third, a consensus sequence RRYRR (R=purine, Y=pyrimidine) has recently been described for the mRNA initiation sites of "highly" transcribed genes (42). Five such sequences are found in the *PPA* gene upstream of base 392 at bases 69, 139, 216, 227, and 361.

The combination of these potential regulatory sequences that most closely resembles the organization of regulatory elements in better characterized *S. cerevisiae* genes (40) includes the polydT stretch at base 233, the TATA-like sequence at base 255, and the mRNA initiation site at base 361.

There are no sequences homologous to enhancer or UAS elements in the 5' region. There is a 20 base pair sequence at bases 40-60 which could possibly form a stem-loop structure as proposed for some UAS elements (43), but this structure is predicted to be weak since it has many non-Watson-Crick pairings.

In the downstream region of the *PPA* gene (bases 1254-1612) there are two potential polyadenylation signals (1454 and 1513) which match the consensus polyadenylation signal in eukaryotes of AATAAA (44,45). In addition, there is a region (1487-1495) which resembles the consensus sequence TTTTATA for transcription termination as described by Bennetzen and Hall (46).

Chromosomal Location of *PPA* Gene

The *PPA* gene was localized to chromosome II using a nick-translated (12) 0.7kbp EcoRV fragment as a probe (specific radioactivity 2×10^9 cpm/ μ g), as shown in Figure 6. CDC28 gene (gift of S. Reed, Scripps Institute), previously localized to chromosome II (47), was used to confirm identification. Importantly, no hybridization was detected to any other chromosome, thus supporting the conclusion that *PPA* is present as a single copy gene. Computer analysis of

Figure 6: A CHEF gel of *S. cerevisiae* chromosomes localizing the PPA gene to chromosome II. Chromosome assignments are taken from Carle and Olson (51). Lanes A, B, and C correspond to yeast strains X2180-1A, JC482, and 1380 respectively.

the entire sequence in Figure 4 revealed no overlap with any published yeast DNA sequences. However, the sequence of bases 753 to 1612 in Figure 4 is virtually identical to an unpublished partial sequence of the open reading frame labeled *SMT1* that is immediately adjacent to copy-I of the histone H3 and H4 genes (48) (kindly provided by M. Mitchell Smith of the University of Virginia). Identifying *PPA* and *SMT1* as the same gene also localizes copy-I of the histone H3 and H4 genes to chromosome II.

Other PPase-like genes

In addition to the *PPA* gene reported here, other PPase or PPase-like genes have recently come to our attention. Lahti *et al.* (49) have cloned and expressed at high levels the 175 amino acid PPase from *E. coli*. In addition, a gene from *K. lactis* (*IPP*), located upstream of the histone H3 and H4 genes, has been cloned and sequenced (50). It gives a gene-deduced amino acid sequence that is 83% identical to the directly determined protein sequence of *S. cerevisiae* PPase (2) and almost certainly corresponds to the *PPA* gene described here.

ACKNOWLEDGEMENTS

This work was supported in part by NIH grant AM-13212. L.F.K. was supported in part by NIH predoctoral training grant fellowship 5T32-GM-07229. M.S. is supported by a German Academic Exchange Service fellowship 312-402-579-8. We thank K. Tatchell, J. Ecker, and J. Smith for helpful discussions and suggestions, M. Stark and M.M. Smith for making unpublished data available to us, M. Mitchell for synthesis of DNA oligomers, and J. Ecker for use of his CHEF

gel apparatus. Yeast strains X2180, JC482, and 1380 are from the Yeast Genetic Stock Center, Department of Biophysics and Medical Physics, University of California, Berkley, CA 94720; J. Cannon; and J. Ecker respectively.

REFERENCES

1. Cooperman, B.S. (1982) *Methods Enzymol.* **87**, 526-548.
2. Cohen, S.A., Sterner, R., Keim, P.S., and Heinrichson, R.L. (1978) *J. Biol. Chem.* **253**, 889-897.
3. Kuranova, I.P., Terzyan, S.S., Voronova, A.A., Smirnova, E.A., Vainstein, B.K., Höhne, W., and Hansen, G., (1983) *Bioorg. Khim.* **9**, 1611-1619.
4. Terzyan, S.S., Voronova, A.A., Smirnova, E.A., Kuranova, I.P., Nekrasov, Yu. V., Arutyunyun, Vainstein, B.K., Höhne, W., and Hansen, G. (1984) *Bioorg. Khim.* **10**, 1469-1482.
5. Springs, B., Welsh, K.M., and Cooperman, B.S. (1981) *Biochemistry*, **20**, 6384-6391.
6. Welsh, K.M., Jacobyansky, A., Springs, B., and Cooperman, B.S. (1983) *Biochemistry* **22**, 2243-2248.
7. Gonzalez, M.A., Webb, M.R., Welsh, K.M. and Cooperman, B.S. (1984) *Biochemistry* **23**, 797-801.
8. Knight, W.B., Fitts, S.W., and Dunaway-Mariano, D. (1981) *Biochemistry* **20**, 4079-4086.
9. Banerjee, A. and Cooperman, B.S. (1983) *Inorg. Chim. Acta*, **79(B7)** 146-148.
10. Knight, W.B., Dunaway-Mariano, D., Ransom, S.C., and Villafranca, J.J.: (1984) *J. Biol. Chem.* **259**, 2886-2895.
11. Banerjee, A., LoBrutto, R., and Cooperman, B.S. (1986) *Inorganic Chemistry*, **25**, 2417-2424.
12. Hamm, D.J., and Cooperman, B.S. (1978) *Biochemistry*, **17**, 4033-4040.
13. Welsh, K.M., Armitage, I.M., and Cooperman, B.S. (1983) *Biochemistry*, **22**, 1046-1054.
14. Welsh, K.M., and Cooperman, B.S. (1984) *Biochemistry*, **23**, 4947-4955.
15. Bond, M.W., Chiu, N.Y., and Cooperman, B.S. (1980) *Biochemistry*, **19**, 94-102.
16. Komissarov, A.A., Makarova, I.A., Sklyankina, V.A., and Avaeva, S.M. (1985) *Bioorg. Khim.* **11**, 1504-1509.
17. Gonzalez, M.A., and Cooperman, B.S. (1986) *Biochemistry*, **25**, 7175-7185.
18. Ermakova, S.A., Mansurova, S.E., Kalebina, T.S., Lobakova, E.S., Selyach, I.O., and Kulaev, I.S. (1981) *Arch. Microbiol.*, **128**, 394-397.
19. Heinonen, J., and Kukko, E. (1977) *Chem-Biol. Interact.*, **17**, 113-116.
20. Kukko, E., and Sundell, L. (1983) *Folia Microbiol. (Prague)*, **28**, 145-148.
21. Sherman, F., Fink, G.R. and Hicks, J.B., (1986) *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, L.I., N.Y.
22. Huynh, T.V., Young, R.A., and Davis, R.W. (1985) in *DNA Cloning: Volume I*, D.M. Glover (ed.) IRL Press, Washington DC, 49-78.
23. Benton, W.D., and Davis, R.W. (1977) *Science*, **195**, 180-182.
24. Blattner, F.R., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Richards, J.E., Slightom, J.L., Tucker, P.W. and Smithes, O. (1978) *Science*, **202**, 1279-1284.
25. Davis, R.W., Botstein, D. and Roth, J.R., (1980) *Advanced Bacterial Genetics*, Cold Spring Harbor Laboratory.
26. Maniatis, T., Fritsch, E.F. and Sambrook, J., (1982) *J. Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory.
27. Viera, J. and Messing, J. (1987) *Methods Enzymol.*, **153**, 3-11.
28. Hanahan, D. (1985) in *DNA Cloning: Volume I*, D.M. Glover (ed.) IRL Press, Washington DC, , 109-135.
29. Birnboim, H.C. (1983) *Methods Enzymol.*, **100**, 243-255.
30. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
31. Haltiner, M., Kempe, T. and Tjian, R. (1985) *Nucl. Acids Res.*, **13**, 1015-1026.

32. Henikoff, S. (1987) *Methods Enzymol.*, **155**, 156-165.
33. Staden, R., (1986) *Nucl. Acids Res.*, **14**, 217-231.
34. Chu, G., Vollrath, D. and Davis, R.W. (1986) *Science*, **234**, 1582-1585.
35. Jaye, M., Salle, H., Schamber, F., Balland, A., Kohli, V., Findeli, A., Tolstoshev, P. and Lecocq, J., (1983) *Nucl. Acids Res.*, **11**, 2325-2335.
36. Ullrich, A., Berman, C.H., Dull, T.J., Gray, A. and Lee, J.M. (1984) *EMBO J.*, **3**, 361-364.
37. Sharp, P.M., Tuohy, T.M.F. and Mosurski, K.R. (1986) *Nucl. Acids Res.*, **14**, 5125-5143.
38. Smith, D.H., Brutlag, D., Friedland, P. and Kedes, L.H. (1986) *Nucl. Acids Res.*, **14**, 17-20.
39. McNeil, J.B. and Smith, M. (1986) *J. Mol. Biol.*, **187**, 363-378.
40. Struhl, K., Chen, W., Hill, D.E., Hope, I.A. and Oettinger, M.A. (1985) *C.S.H. Symp. Quant. Biol.*, **50**, 489-503.
41. Struhl, K. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 8419-8423.
42. Rudolph, H. and Hinnen, A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1340-1344.
43. Huet, J., Cottrelle, P., Cool, M., Vignais, M.L., Thiele, D., Marck, C., Buhler, J.M., Sentenac, A. and Fromageot, P. (1985) *EMBO J.*, **4**, 3539-3547.
44. Gallwitz, D., Perrin, F. and Seidel, R., (1981) *Nucl. Acids Res.*, **9**, 6339-6350.
45. Proudfoot, N.J., Cheng, C.C. and Brownlee, G.G. (1976) *Progr. Nucleic Acids Res. Mol. Biol.*, **19**, 123-134.
46. Bennetzen, J.L. and Hall, B.D. (1982) *J. Biol. Chem.*, **257**, 3018-3025.
47. Mortimer, R.K. and Schild, D. (1985) *Microbiol. Rev.*, **49**, 181-212
48. Smith, M.M. and Murray, K. (1983) *J. Mol. Biol.*, **169**, 641-661
49. Lahti, R., Pitkäranta, T., Valve, E., Ilta, I., Kukko-Kalske, E., and Heinonen, J. (1988) *J. Bacteriol.*, in press
50. Stark, M.J.R. and Milner, J.S. (1988) 14th Int. Conf. on Yeast Genetics and Molecular Biology, Helsinki, Abstract #L53
51. Carle, G.F. and Olson, M.V. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3756-3760.