Cloning and Characterization of the Mitochondrial Phosphate Transport Protein Gene from the Yeast *Saccharomyces cerevisiae*[†,‡]

Anne Phelps, Christian T. Schobert,[§] and Hartmut Wohlrab*

Department of Metabolic Regulation, Boston Biomedical Research Institute, 20 Staniford Street, Boston, Massachusetts 02114, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115

Received July 17, 1990; Revised Manuscript Received September 18, 1990

ABSTRACT: We have cloned the gene of the *Saccharomyces cerevisiae* phosphate transport protein (PTP), a member of the mitochondrial anion transport protein gene family. As PTP has a blocked N-terminus, we prepared three peptides. Oligonucleotides, based on their sequences, were used to screen a Yep24-housed genomic library. A total of 2073 bases of clone Y22 code for a 311 amino acid protein (M_r 32 814), which has similarities to the anion transport proteins: a triplicate gene structure and 6 hydrophobic segments. Typical for PTP, the triplicate gene structure possesses the X-Pro-X-(Asp/Glu)-X-X-(Lys/Arg)-X-(Arg/Lys)-X (X is an unspecified amino acid) motif and the very high homology only between the first and second repeat. The 6 hydrophobic segments harbor most of the 116 amino acids that are conserved between the yeast and the beef proteins. An N-terminal-extended signal sequence, as found in the beef protein, is absent. The yeast protein has about 33% fewer basic and acidic amino acids and five fewer Cys residues than the beef protein. The protein is insensitive to *N*-ethylmaleimide since Cys-42 (beef) has been replaced with a Thr. Mersalyl sensitivity has been retained and must be due to one of its three cysteines. Among these three cysteines, only Cys-28, located in the first hydrophobic segment, is conserved between the yeast and the beef protein.

The phosphate transport protein (PTP)[1] is a major protein of the inner mitochondrial membrane. It is responsible for the transport of the bulk of the inorganic phosphate into the mitochondrial matrix where phosphate is required for steady-state oxidative phosphorylation (Wohlrab, 1986). The protein belongs to the family of mitochondrial anion transport proteins (Wohlrab, 1986; Runswick et al., 1987; Aquila et al., 1987), which includes the AAC and the UCP. Until now, efforts towards an understanding of their mechanism of transport at the molecular level have been limited to purification, reconstitution, sequencing, and chemical probe studies. We would like to extend these studies to include site-directed mutagenesis. With this objective in mind, we have now purified and sequenced a genomic clone of the PTP from *Saccharomyces cerevisiae*.

The yeast PTP has recently been purified in a reconstitutively active form (Guérin et al., 1990). Its kinetic characteristics are the same as those of the beef heart protein. However, the extremely high sensitivity of the beef protein toward *N*-ethylmaleimide is absent from the yeast protein. Nevertheless, it does retain the same high sensitivity toward mersalyl. The genomic sequence that we report in this paper permits an explanation for these inhibitor sensitivities and sets the stage for the characterization of the function of PTP at the molecular level by in vitro mutagenesis.

Materials and Methods

Preparation of Yeast PTP Peptides. *Saccharomyces cerevisiae* D273-10B (ATCC 24657) was grown aerobically in rich glycerol medium and harvested in the logarithmic phase, and mitochondria were isolated from protoplasts (Guérin et al., 1979). The yeast PTP was solubilized and purified following the procedure published for the beef heart protein (Wohlrab et al., 1986). The final preparation was alkylated with *N*-ethylmaleimide, characterized by SDS-PAGE (Figure 1), and stained for 1 h with Coomassie Blue. The protein band of lower molecular mass (Figure 1, lane E) was cut from the gel, equilibrated with 70% formic acid, and incubated with CNBr at room temperature for 40 min. The gel slices were extensively neutralized, desiccated overnight, and applied to an SDS peptide gel for electrophoresis (Kolbe & Wohlrab, 1985). The separated peptides were blotted onto an Immobilon P (Millipore) membrane, stained with Coomassie Blue, and submitted for gas-phase sequencing (Matsudaira, 1987).

Genomic Clone Isolation and Sequencing. Three sets of degenerate oligonucleotide probes were synthesized with a MilliGen/BioSearch Cyclone DNA synthesizer. The degeneracies were minimized by utilizing frequently used yeast codons of nuclear-coded mitochondrial membrane proteins: Rieske iron-sulfur protein (Beckman et al., 1987), porin (Mihara & Sato, 1985), the more highly expressed AAC-2 (Lawson & Douglas, 1988), and tabulated codons of highly expressed yeast genes (Sharp et al., 1986). They were based on peptides 1, 2, and 3 (Figure 3) with degenerate positions in parentheses: Y1, 5′GC(TCG)GA(TC)AT(TC)GC-(TCG)(TC)TGTGTCCCT3′ (Ala-129 to Leu-136); Y2, 5′ATCGA(CT)GT(TC)GT(TC)AA(GA)AC(TC)-AGAATCCAA3′ (Ile-38 to Gln-46); Y3, 5′CAAAT(CT)-AT(CT)GC(TC)GGTGAAGGTGC3′ (Gln-62 to Ala-69).

A yeast recombinant plasmid (Yep24) library (Carlson & Botstein, 1982) was screened with these oligonucleotides la-

[†] Supported by a grant from the NIH (GM33357).

[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05302.

* Address correspondence to this author at the Department of Metabolic Regulation, Boston Biomedical Research Institute.

[§] Present address: Pflanzenphysiologie, Universität Bayreuth, Postfach 3008, D-8580 Bayreuth, FRG.

[1] Abbreviations: PTP, phosphate transport protein; AAC, ADP-ATP carrier; UCP, uncoupling protein; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

0006-2960/91/0430-248$02.50/0

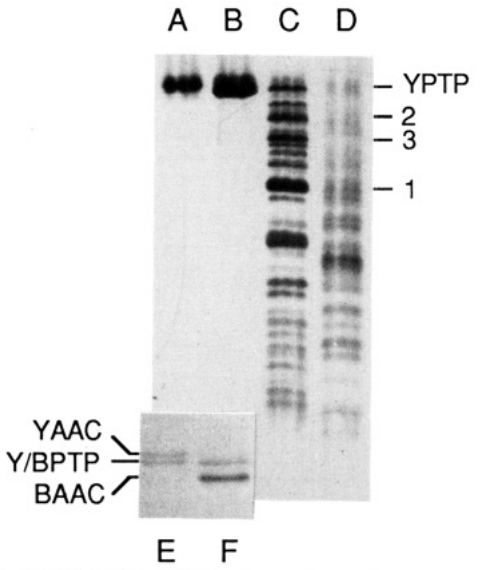

FIGURE 1: SDS–PAGE of CNBr-digested PTP from yeast and beef heart. Lanes A and C (yeast) and B and D (beef) of PTP treated in parallel with (C, D) and without (A, B) CNBr. Lanes E and F are purified preparations that contain the AAC and PTP from yeast (Y) and beef (B), respectively. The peptides have molecular masses of (1) 16 555 Da, (2) 26 204 Da, and (3) 23 803 Da.

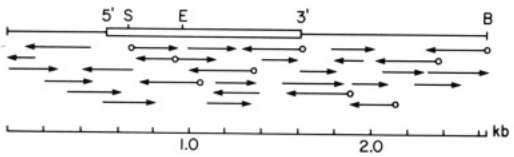

FIGURE 2: Sequencing strategy for the yeast PTP clone Y22. The arrows represent the direction and length of the sequence obtained. Information from single-stranded templates is indicated by circles; all others utilized double-stranded templates. In all cases, synthetic oligonucleotides were used as primers. The protein-coding region is represented by the closed box. Restriction sites: S (*Sal*I), E (*Eco*RI), B (*Bam*HI).

beled with T4 polynucleotide kinase and [γ-^{32}P]ATP. Four independent clones (Y22, Y33, Y181, and Y191) that hybridized to each of the three probes were isolated. Plasmids were purified, cleaved with restriction enzymes, and analyzed by Southern blotting. One of the clones (Y22) (Figures 2 and 3) contained a 2-kb *Sal*I/*Bam*HI fragment, which hybridized to each of the three probes. This fragment was subcloned into the *Sal*I/*Bam*HI site of M13mp19. A 290 bp *Sal*I/*Eco*RI fragment which hybridized to both Y2 and Y3 was subcloned into the appropriate site in both M13mp18 and M13mp19. Both fragments were sequenced by the chain termination method (Biggin et al., 1983; Bankier, 1984) with the aid of modified T7 DNA polymerase (Sequenase, U.S. Biochemical) and synthetic oligonucleotides. Double-stranded templates were also used to determine the sequence of both strands (Figure 2).

Results and Discussion

Preparation of Yeast PTP Peptides and Corresponding Oligonucleotides. Regions with an extremely high degree of amino acid conservation were found between yeast AAC-2 (Lawson & Douglas, 1988) and bovine AAC-T1 (Aquila et al., 1982). We therefore prepared oligonucleotides based on the comparable regions of beef PTP to use as hybridization probes for screening a yeast genomic library. We were unable to identify the yeast PTP using this approach. Figure 6 clearly shows that while these regions are conserved between yeast and beef AAC, homologous regions between yeast and beef

```
      -574                -554                -534                -514
ACGGTGGCTT TTTTGCCTCG GAGCGACGTT GAAAAGTGAC AGCGTCAGCC CATAGAGAAA AAAAATAAAA

      -504                -484                -464                -444
AAAGGCCGCT ACCGATATGT TTTCATTACG ACAATTAAGC AGCACAAGTG TCTGGCGCAA TGCCTGACTA

      -434                -414                -394                -374
CTGGGAAAAG TCGAAGCATT TGCACGGCGT CCATTACGAT GGTTAAGGCA GTTTACGACG AATTGCGGAG

      -364                -344                -324                -304
TCAAGTTTTA ATATTTATAG ATAGCCCTGT CACCAGTATC GGGTGTTGTT TCTACCCAAG CGCCTGCCAG

      -294                -274                -254                -234
CCTCAAACAA TACAACAGGT ATTGTTTTTG GACAGCGGAA TGCTGCGATG CAACAGTATC GCAGCATGGT

      -224                -204                -184                -164
ACTGCTTTCC TCTAGTCTTT TGTCGAATGT GAGCAAGCAG CAGCTGAGAC TCTGCCCAAC CATCGTTAGC

      -154                -134                -114                -94
CAATCAGCGT CGAGAACGGG TCCCATGCCC GGCCGCGGCC CGTTTCGAGA AGATTTTTAA GTGCGATCCA

      -84                 -64                 -44                 -24
TCCTCCCGAT CTTTTTATCT ACAAGAACCA TCGCAAAGAA CAGAGTGATC TGCAAGAAGA AGAACTACAA

      -14                 1                   20
AGATCAAAAA GTCTCATCTC ACC ATG TCT GTG TCT GCT GCT CCT GCT ATT CCA CAG
                          MET Ser Val Ser Ala Ala Pro Ala Ile Pro Gln    11

          40                  60                  80
TAC TCT GTC TCT GAC TAC ATG AAA TTC GCC CTA GCC GGC GCC ATA GGG TGC GGG
Tyr Ser Val Ser Asp Tyr MET Lys Phe Ala Leu Ala Gly Ala Ile Gly Cys Gly  29

              100                 120                 140
TCG ACT CAC TCC AGT ATG GTC CCG ATC GAT GTC GTT AAG ACG AGA ATC CAA CTG
Ser Thr His Ser Ser MET Val Pro Ile Asp Val Val Lys Thr Arg Ile Gln Leu  47
                        (2)
                  160                 180
GAA CCT ACC GTG TAC AAC AAG GGT ATG GTC GGC TCT TTC AAG CAA ATC ATC GCC
Glu Pro Thr Val Tyr Asn Lys Gly MET Val Gly Ser Phe Lys Gln Ile Ile Ala  65
                                    (3)
      200                 220                 240
GGA GAA GGT GCT GGT GCG TTG TTG ACC GGG TTT GGT CCT ACT TTG TTG GGT TAC
Gly Glu Gly Ala Gly Ala Leu Leu Thr Gly Phe Gly Pro Thr Leu Leu Gly Tyr  83

          260                 280                 300
TCT ATC CAA GGT GCC TTC AAG TTC GGT GGT TAC GAA GTG TTC AAG AAG TTC TTC
Ser Ile Gln Gly Ala Phe Lys Phe Gly Gly Tyr Glu Val Phe Lys Lys Phe Phe 101

                  320                 340
ATT GAC AAC TTG GGT TAC GAT ACC GCC TCT CGC TAC AAG AAC TCT GTT TAC ATG
Ile Asp Asn Leu Gly Tyr Asp Thr Ala Ser Arg Tyr Lys Asn Ser Val Tyr MET 119

      360                 380                 400
GGT TCT GCT GCC ATG GCC GAA TTC TTG GCA GAC ATT GCC TTG TGT CCC CTA GAA
Gly Ser Ala Ala MET Ala Glu Phe Leu Ala Asp Ile Ala Leu Cys Pro Leu Glu 137
(1)
          420                 440                 460
GCC ACT AGA ATC AGA TTG GTC TCC CAA CCA CAG TTC GCT AAT GGA CTC GTC GGA
Ala Thr Arg Ile Arg Leu Val Ser Gln Pro Gln Phe Ala Asn Gly Leu Val Gly 155

                  480                 500
GGG TTT TCC AGA ATT TTG AAA GAA GAA GGT ATC GGC TCT TTC TAC AGT GGG TTC
Gly Phe Ser Arg Ile Leu Lys Glu Glu Gly Ile Gly Ser Phe Tyr Ser Gly Phe 173

                  540                 560
ACT CCA ATT TTG TTC AAG CAA ATT CCT TAC AAC ATT GCT AAA TTC TTG GTC TTT
Thr Pro Ile Leu Phe Lys Gln Ile Pro Tyr Asn Ile Ala Lys Phe Leu Val Phe 191

          580                 600                 620
GAA CGT GCC TCC GAG TTC TAC TAC GGG TTC GCC GGT CCA AAA GAA AAA TTG TCC
Glu Arg Ala Ser Glu Phe Tyr Tyr Gly Phe Ala Gly Pro Lys Glu Lys Leu Ser 209

              640                 660                 680
TCT ACT TCC ACT ACT TTG TTG AAT TTG CTT TCC GGT TTG ACT GCT GGT TTG GCC
Ser Thr Ser Thr Thr Leu Leu Asn Leu Leu Ser Gly Leu Thr Ala Gly Leu Ala 227

                  700                 720
GCT GCC ATT GTT TCC CAG CCA GCA GAC ACT CTG TTG TCC AAG GTC AAC AAG ACT
Ala Ala Ile Val Ser Gln Pro Ala Asp Thr Leu Leu Ser Lys Val Asn Lys Thr 245

      740                 760                 780
AAG AAG GCT CCT GGC CAA TCC ACA GTC GGT TTG TTG GCT CAG TTG GCC AAA CAG
Lys Lys Ala Pro Gly Gln Ser Thr Val Gly Leu Leu Ala Gln Leu Ala Lys Gln 263

              800                 820                 840
TTA GGA TTT TTC GGT TCC TTT GCT GGG TTG CCT ACT CGT TTG GTC ATG GTT GGT
Leu Gly Phe Phe Gly Ser Phe Ala Gly Leu Pro Thr Arg Leu Val MET Val Gly 281

                  860                 880
ACC TTG ACC TCC TTG CAA TTT GGT ATT TAT GGT TCT TTA AAG AGC ACT TTG GGT
Thr Leu Thr Ser Leu Gln Phe Gly Ile Tyr Gly Ser Leu Lys Ser Thr Leu Gly 299

      900                 920                 940
TGC CCA CCA ACC ATT GAA ATT GGT GGT GGT GGT CAT TAG AAATGGTCTT GATTGATACA
Cys Pro Pro Thr Ile Glu Ile Gly Gly Gly Gly His  .                      311

          966                 986                 1006                1026
TGCATATATA TTCTCTCCTC ATCACCACCG GAGCCCCTTT TTTTTTTCAT ATTCTAAGAA AGAGTCAACA

          1036                1056                1076                1096
AAGGAAAGAA ATAAATAAAT CTGTTTAAAT GTACTCTATT CAATCACTAG TATTCTTATT CCATGTATTT

          1106                1126                1146                1166
ACTGACTTTG CAAAAGTCCA AATATCTGTA AATTAACATG CTATTATAAA TATATATATA TATATATATT

          1176                1196                1216                1236
TGAATTTTCT ATTTTTTCAT CTACAAAAGC AGGTTATAGC TCCGTTAAAC CCTTTGCTCA TTGAATTGTC

          1246                1266                1286                1306
CTTCACGGTT TTTTGGTGGA AAACACAACA TGGAACATAA CATTTAAACA TCGTTCTCAA TCCATTTCAT

          1316                1336                1356                1376
CGCAGTAACA TATTTGGTCG GGCACATACG GACACGTATA TGTTATACAG ATAAATATAG CTTAAATATG

          1386                1406                1426                1446
GCTATATTTA CACAACACAA TGCTCGTGGG TCGTTACCCG ACCATGCAGG ATCGCCAGTA TTCTTTATTT

          1456                1476                1496                1516
TTCCGCCTTC TAACATAAAG AAAAATAAAC AAAAAAGTAT TTGATCGAAA AGTAAAATAG GTAGACACCA

          1526                1546                1566                1586
CGTATTGGCG ACCCGATCTG GAGCCGTTTA GAAAGTCAAT CATCACAAGG CCTAAAGTTG CTAACCACCA

          1596                1616                1636                1656
GCCATGTCCG GAATAATTGA CGCATCTTCT GCATTAAGAA AAAGAAAGCA TTTGAAAAGA GGTATAACCT

          1666                1686                1706                1726
TCACTGTGAT GATCGTGGGC CAGTCCGGAT CTGGTAGATC GACTTTTATA AATACTTTGT GCGGTCAGCA

          1736                1756                1776                1796
AGTTGTAGAC ACTTCGACGA CAATCTTGTT ACCCACAGAT ACGTCCACAG AAATAGACTT ACAATTGAGA

          1806                1826                1846                1866
GAGGAGACGG TCGAATTAGA AGATGATGAA GGTGTCAAGA TTCAACTTAA TATCATCGAT ACTCCGGGAT

          1876                1896                1916                1936
TCGGTGATTC TCTCGACAAT TCTCCATCTT TCGAAATCAT TTCCGACTAC ATTCGCCACC AATATGATGA

          1946                1966                1986                2006
AATCTTATTG GAAGAAAGTC GTGTGAGAAG AAACCCAAGA TTTAAGGACG GCAGAGTTCA TTGTTGTCTT

          2016                2036                2056                2073
TACTTAATCA ACCCAACTGG CCACGGTTTA AAAGAGATTG ATGTGGAATT CATCAGACAG TTGGATC
```

FIGURE 3: Partial sequence of clone Y22. In the 5′ untranslated sequence are putative TATA- and CAAT-box sequences at positions −362 and −456, in-frame stop codons at positions −531, −189, and −48, and an in-frame initiation codon at −207. In the 3′ untranslated sequence are two polyadenylation signal sequences at positions 1040 and 1470. The three underlined protein segments are the partial N-terminal sequences of the peptides in Figure 1.

PTP are very short and relatively rare. Therefore, to obtain hybridization probes for screening, we proceeded to prepare peptides for protein sequencing.

The N-terminus of the yeast PTP appears to be blocked, and since our normal reconstitutively active PTP preparation from yeast and beef includes the AAC, CNBr digestion was carried out in gel slices. We digested the band of lower molecular mass (Figure 1, lane E), since we have shown, on the basis of antiserum reactivity and peptide sequence (Guérin et al., 1990), that the upper band is the AAC-2. The result is shown in Figure 1, lane C. We had attempted to remove C-terminal amino acids with carboxypeptidases A and Y to aid the identification of yeast PTP C-terminal peptides (no blocked N-terminus) among CNBr peptides. While the carboxypeptidases clearly removed C-terminal amino acids from the beef PTP (Kolbe & Wohlrab, 1985), we were not able to detect any digestion with the yeast protein. We arbitrarily chose the major bands in the CNBr digest (Figure 1, lane C, bands 1–3) for sequencing. Figure 3 shows the amino acid sequences. While we had great difficulty in obtaining significant sequence information from the blotted peptides with the ABI 477A pulsed-liquid protein sequencer, the ABI 470A gas-phase protein sequencer (Harvard Microchemistry Facility) gave excellent results with the small amounts of protein used. The corrected amino acid sequencing yields ranged from 21 pmol to less than 0.2 pmol. Band 1 (Figure 1, lane C) consisted of two peptides: peptide 1A (Figure 3) from Gly-120 to Met-279 and peptide 1B from Ala-125 to Met-279 at a ratio of 1:4. Cys-134 was not detected since the protein had been alkylated with *N*-ethylmaleimide. Oligonucleotide sequences were based on these peptides.

Cloning and Sequencing of the Yeast PTP Gene. A Yep24-housed genomic library (Carlson & Botstein, 1982) was screened with the ^{32}P-labeled oligonucleotides, and four independent clones were isolated. Clone Y22 was chosen for complete sequencing, and the sequencing strategy is shown in Figure 2. The sequenced bases of Y22 are shown in Figure 3. Perfect homology with the three yeast PTP peptides confirmed the deduced primary structure as that of the yeast PTP. Potential CAAT-box and TATA-box promoter elements can be located at positions −456 and −362, respectively. The nucleotide sequence adjacent to the initiation codon fulfills Kozak's criteria for eukaryotic initiation sites (Kozak, 1986). Two polyadenylation signal sequences (AATAAA) are located 104 and 534 bases downstream from the termination codon.

We have sequenced the following regions of the three remaining independent clones (see Figure 3 for location) in order to identify additional PTP genes: from inside the translated sequence into the 5′ untranslated sequence (111 to −86), near the middle of the translated sequence (502–301), and from within the translated into the 3′ untranslated sequence (780–989). We were not able to detect any differences among the sequences of those three regions of the clones. It seems likely that the four clones are identical and that yeast PTP, unlike yeast AAC, is translated from a single nuclear gene. This observation agrees with the mammalian system, where only a single message for PTP (Runswick et al., 1987; Wohlrab et al., 1988) but several messages and genes for AAC (Rasmussen & Wohlrab, 1986; Li et al., 1989) have been found. In addition, we have analyzed the codon usage by this yeast PTP gene and found it to correlate closely with that of the highly expressed AAC-2 and other high-level expression genes of yeast (Sharp et al., 1986) rather than that of the less frequently expressed AAC-1 gene (Lawson & Douglas, 1988).

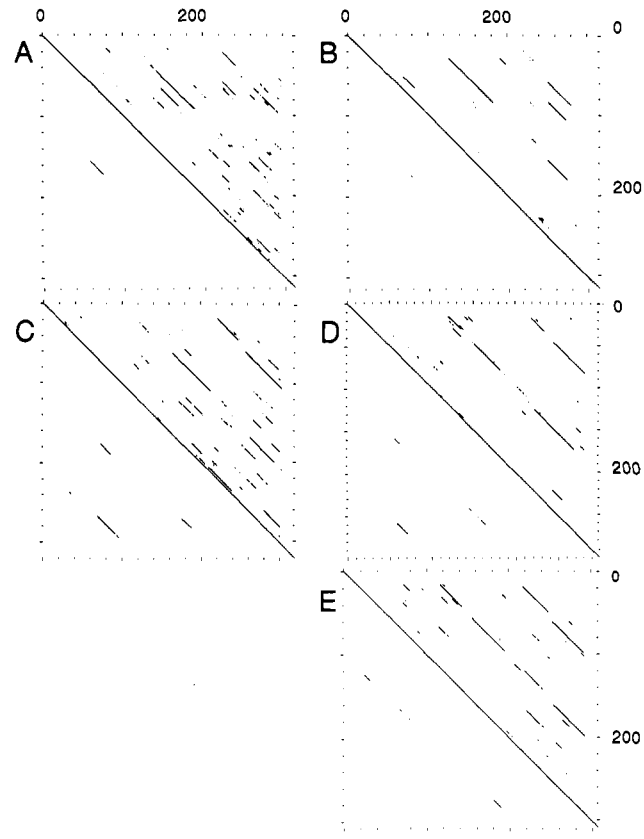

FIGURE 4: DotPlot of protein sequences against themselves. Each plot is a composite analysis at low (top part, window 30, stringency 13) and high (bottom part, window 30, stringency 20) stringency using the Genetics Computer Group program. The following proteins were analyzed: (A) yeast PTP; (B) beef PTP (Runswick et al., 1987); (C) yeast AAC-2 (Lawson & Douglas, 1988); (D) beef AAC-T1 (Aquila et al., 1982); (E) hamster UCP (Aquila et al., 1985).

This supports the conclusion that this is the only expressed PTP gene in yeast.

Triplicate Gene Structure and Six Hydrophobic Regions of the Yeast PTP. One of the key features of members of the mitochondrial anion transport protein family is the homology between the first, second, and third hundred amino acids of their protein sequence. This is demonstrated with DotPlots (Figure 4) of the PTP from yeast (A) and beef (B) against themselves at high and low stringency. Similar plots are shown for the AAC from yeast (AAC-2) and beef (AAC-T1) and the UCP from hamster brown fat mitochondria (Aquila et al., 1985). Yeast PTP clearly fits into the patterns generated by these proteins. It should be noted that at low stringency both yeast PTP and AAC (Figure 4A,C) demonstrate much more scattering of homology points than the same proteins from beef (Figure 4B,D). As the stringency is increased, homologies are present only between repeats 1 and 2 for PTP with the length of the homologous region being much longer for the yeast protein. At the same stringency, however, homologies between all three repeats can be seen for yeast and beef AAC. Again, the homologous regions for the yeast protein are clearly longer. The UCP, the third member of this protein family, has high stringency homologies only between repeats 1 and 2, and also 1 and 3, but not 2 and 3. It is not clear why these differences in homologous regions exist. The yeast PTP fits well into this new pattern that differentiates the members of this protein family.

We are left with two interesting observations: (a) the poor internal homology with repeat 3 in PTP and (b) the longer stretches of homology in the yeast proteins. The first obser-

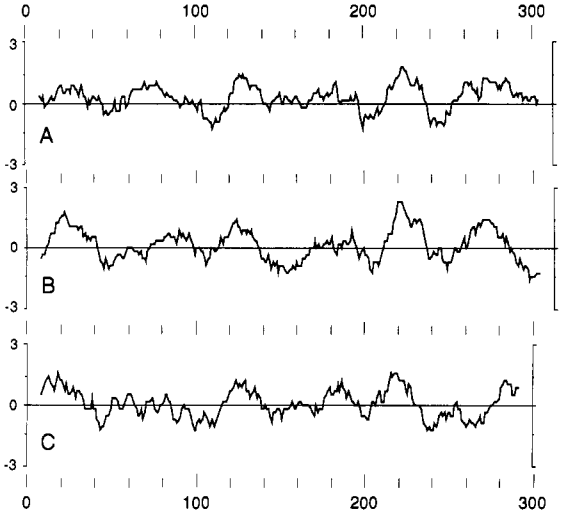

FIGURE 5: Hydropathy plots of (A) yeast PTP, (B) beef PTP (Runswick et al., 1987), and (C) beef AAC-T1 (Aquila et al., 1982). A window of 16 was used in the PepPlot program of the Genetics Computer Group.

vation suggests that an active site of PTP (third repeat) has evolved separately from the original triplicate repeat structure. The longer stretches of homology in yeast may reflect evolutionary pressure to retain high transport specificity in face of a more variable environment.

An additional characteristic common to members of the mitochondrial anion transport protein family is the presence of six hydrophobic segments which have been suggested on the basis of a hydropathy plot to represent transmembrane α-helices (Saraste & Walker, 1982). However, recent experiments by Dalbon and co-workers (Dalbon et al., 1988) place the second hydrophobic region of the beef heart AAC outside of the membrane. Figure 5 shows the hydropathy plot of the yeast PTP. Six hydrophobic segments (for detailed location, see Figure 6) can be clearly identified. The second hydrophobic segment appears to be much more hydrophobic than that of the beef AAC and thus, in contrast to the results of Dalbon and co-workers (Dalbon et al., 1988), is possibly a transmembrane α-helix. The yeast protein, contrary to that from beef, has hydrophobic and thus, most likely, membrane-embedded N- and C-termini. This may explain our observation that carboxypeptidases A and Y will digest the C-terminus of beef PTP (Kolbe & Wohlrab, 1985) but not that from yeast.

Table I: Amino Acid Composition of the Phosphate Transport Protein from Yeast (Genomic DNA) and from Beef Heart [cDNA (Runswick et al., 1987)]

	yeast	beef		yeast	beef
Gly	39	25	Gln	12	10
Leu	34	29	Glu	11	18
Ala	30	28	Arg	7	14
Ser	28	20	Met	7	9
Phe	22	21	Asn	7	5
Thr	20	14	Asp	6	8
Lys	19	24	Cys	3	8
Ile	19	14	His	2	2
Val	17	24	Trp	0	5
Pro	15	18	total amino acids	311	313
Tyr	13	17			

Conservation of Amino Acids between Beef and Yeast PTP. The amino acid compositions of yeast and beef PTP are shown in Table I. Alignment of the yeast PTP amino acid sequence with that of the bovine protein (Figure 6) shows that many of the amino acids are the same, although there are many more that are different. Notably there is a 30 and 35% lower concentration of basic and acidic amino acids, a dramatically lower number of cysteines, and an absence of tryptophan in the yeast protein. There is almost no difference in acidic and basic amino acids between yeast and beef AAC (Lawson & Douglas, 1988; Aquila et al., 1982). A most obvious feature of the yeast and beef PTP alignment is that the highest degree of amino acid conservation is in the hydrophobic segments. Of the amino acids in the hydrophobic segments, 52% are conserved in comparison to only 27% of those in the rest of the sequence. This difference is, however, even greater for the acidic and basic residues. More than 90% of those in the hydrophobic regions are identical while only 28% of those in

```
                        20      (I)         40                  60                  80      (II)        100
Yeast msVsaapaipqySvsdvmkfaLaGaIgCGsTHssmVPiDvVKtRiQlePtvYnKgmvgsFkqiiageGagaLltGfgPT-llGYSiOGafKFGgYEVFKkf
Beef   aVeeqyscdygSgrffilcgLgGiIsCGtTHtalVPlDlVKcRmQvdPqkY-KsifngFsvtlkedGfrgLakGwaPT-fiGYSlOGlcKFGfYEVFKvl
Rat      g-   ef  mkyya   f  vl   l    v                   g     i      v          l    m             a

cons.  -V---------S--------L-G-I-CG-TH---VP-D-VK-R-Q--P--Y-K-----F-------G---L--G--PT---GYS-QG--KFG-YEVFK--

                        120   (III)       140                 160                 180  (IV)           200
Yeast  fidnLGydtAsrykn-SvYmgsaAmAEFlADIALcPlEAtriRlvsQPq-fANgLvggfsrilKEEGigsFYsGftPi-lfkOIPYniaKFlvFERasEfyYgFa
Beef   ysnmLGeenAylwrt-SlYlaasAsAEFfADIALaPmEAakvRiqtQPg-yANtLrdaapkmyKEEGlkaFYkGvaPl-wmrOIPYtmmKFacFERtvEalYkFv
Rat      i     t                                    e v        n        v

cons.  ----LG---A------S-Y----A-AEF-ADIAL-P-EA---R---QP---AN-L--------KEEG---FY-G---P-----QIPY---KF--FER--E--Y-F-

                        220   (V)         240                 260                 280  (VI)           299
Yeast  g--PK--eklSststtllnllsGltAGlaaAIVSqPADtllSkvNKt-kkapgqstvgllaqlaKqLGFfGsfaGLptRlyMvGTLTsLOfgIYgSlK---stLg
Beef   v--PKprsecSkpeqlvvtfvaGyiAGvfcAIVShPADsvvSvlNK--------ekgssasevlKrLGFrGvwkGLfaRiiMiGTLTaLOwfIYdSvKvy-frLp
Rat              t a                                  q  q

cons.  ---PK-----S----------G--AG---AIVS-PAD---S--NK----------------K-LGF-G---GL--R--M-GTLT-LQ--IY-S-K-----L-

Yeast  cPPtiEigggh
Beef   rPPppEmpeslkkklgytq
Rat                    l e

cons.  -PP--E-----------
```

FIGURE 6: Alignment of the yeast PTP with the beef PTP (N-terminal extended signal sequence omitted) sequence (Runswick et al., 1987; Kolbe & Wohlrab, 1985). Amino acids of the rat liver PTP (Ferreira et al., 1989) which differ from those of the beef heart PTP are indicated. The consensus sequence between yeast and beef is shown, and the six hydrophobic segments are indicated by underlined amino acids and the Roman numerals (I–IV). The GenAlign (Intelligenetics) program was used, and the results were adjusted somewhat manually.

the rest of the sequence are conserved. The stretches of internal homology of the yeast and beef PTP are located in hydrophilic regions on the N-terminal side of hydrophobic segments 2 and 4 (Figures 4 and 6). At this time, it is not clear why amino acids in these regions are conserved.

The difference in cysteines is of great interest since it has been demonstrated that Cys-42 in the beef protein is responsible for the very high sensitivity of beef PTP to *N*-ethylmaleimide (Kolbe & Wohlrab, 1985). The yeast PTP lacks this sensitivity (Guérin et al., 1990). The mitochondrial anion transport protein associated motif X-Pro-X-(Asp/Glu)-X-X-(Lys/Arg)-X-(Arg/Lys)-X harbors Cys-42 in beef but Thr-43 in yeast (Figure 6), providing an explanation for the change in sensitivity. The yeast protein still shows a very high sensitivity to the mercurial mersalyl (Guérin et al., 1990). It must therefore react with either the conserved Cys-28, or the Cys-134, or the Cys-300. We assume that one of these is close to the active transport site(s) of PTP.

The sequence alignment of Figure 6 suggests that the yeast PTP lacks an N-terminal extended signal sequence. Support for this suggestion comes from the following: (a) the N-terminus of the yeast protein sequence aligns with the sequence of the mature beef protein (Kolbe & Wohlrab, 1985; Runswick et al., 1987); (b) the mobility of the mature yeast protein in SDS–PAGE is the same as that of the beef protein (Figure 1); and (c) the N-terminal sequence of the yeast protein is hydrophobic and thus more like the AAC (Figure 5C), which does not have an N-terminal extended sequence, than the beef PTP (Figure 5B), which does have an N-terminal extended sequence (Runswick et al., 1987; Ferreira et al., 1989). It is not clear why among all the members of the mitochondrial anion transport proteins an N-terminal-extended sequence is present only in mammalian PTP. It will be interesting to identify amino acids that are required for directing PTP into the membrane and that are present in the N-terminal region of the mature PTP in yeast and in the N-terminal extended sequence in PTP from beef and rat but probably not in the N-terminal region of the mature PTP from beef and rat. Indeed, mammalian and yeast PTPs may help define the role of signal sequences, both internal and N-terminal-extended, in directing proteins into the various membrane insertion pathways.

The evolution-directed mutagenesis that we report in this paper has provided important information on the structure and function of PTP. The availability of the cloned yeast gene will enable the chromosomal copy to be deleted and the introduction of variants to a well-defined genetic background. This will facilitate the structure/function analysis of PTP by allowing an assessment of the role of specific amino acids (many of which have been highlighted in this study) in the catalytic mechanism of the PTP.

Added in Proof

While this manuscript was in press, a yeast genomic sequence was published (Murakami et al., 1990). This sequence, which encodes a mitochondrial import receptor (Pain et al., 1990), is identical with the yeast PTP DNA sequence and extends from nucleotides −370 to 1234 (see Figure 3). The following discrepancies were noted: 459 (T), 600 (A), 804 (C), 1164/1165 (TA) insertion, 1172 (G), 1204 (C), and 1232 (C). The two catalytic functions of this protein have been discussed also in a brief communication (Meyer, 1990).

Acknowledgments

We thank Bernard Guérin for yeast mitochondria, David Botstein for making the yeast genomic Yep24 library available, Mike Cherry for running some of the computer analyses, and the technical service of U.S. Biochemical for useful double-strand sequencing suggestions.

Registry No. DNA (*Saccharomyces cerevisiae* mitochondria clone Y22 phosphate-transporting protein PTP gene), 130726-93-1; protein PTP (*Saccharomyces cerevisiae* mitochondria clone Y22 phosphate-transporting protein moiety reduced), 130726-94-2; phosphate, 14265-44-2.

References

Aquila, H., Misra, D., Eulitz, M., & Klingenberg, M. (1982) *Hoppe-Seyler's Z. Physiol. Chem. 363*, 345–349.

Aquila, H., Link, T. A., & Klingenberg, M. (1985) *EMBO J. 4*, 2369–2376.

Aquila, H., Link, T. A., & Klingenberg, M. (1987) *FEBS Lett. 212*, 1–9.

Bankier, A. T. (1984) *BioTechniques 2*, 72–77.

Beckmann, J. D., Ljungdahl, P. O., Lopez, J. L., & Trumpower, B. L. (1987) *J. Biol. Chem. 262*, 8901–8909.

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

Carlson, M., & Botstein, D. (1982) *Cell 28*, 145–154.

Dalbon, P., Brandolin, G., Boulay, F., Hoppe, J., & Vignais, P. V. (1988) *Biochemistry 27*, 5141–5149.

Ferreira, G. C., Pratt, R. D., & Pedersen, P. L. (1989) *J. Biol. Chem. 264*, 15628–15633.

Guérin, B., Labbe, P., & Somlo, M. (1979) *Methods Enzymol. 55*, 154–159.

Guérin, B, Bukusoglu, C., Rakotomanana, F., & Wohlrab, H. (1990) *J. Biol. Chem.* (in press).

Kolbe, H. V. J., & Wohlrab, H. (1985) *J. Biol. Chem. 260*, 15899–15906.

Kozak, M. (1986) *Cell 44*, 283–292.

Lawson, J. E., & Douglas, M. G. (1988) *J. Biol. Chem. 263*, 14812–14818.

Li, K., Warner, C. K., Hodge, J. A., Minoshima, S., Kudoh, J., Fukuyama, R., Maekawa, M., Shimizu, Y., Shimizu, N., & Wallace, D. C. (1989) *J. Biol. Chem. 264*, 13998–14004.

Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035–10042.

Meyer, D. I. (1990) *Nature 347*, 424–425.

Mihara, K., & Sato, R. (1985) *EMBO J. 4*, 769–774.

Murakami, H., Blobel, G., & Pain, D. (1990) *Nature 347*, 488–491.

Pain, D., Murakami, H., & Blobel, G. (1990) *Nature 347*, 444–449.

Rasmussen, U. B., & Wohlrab, H. (1986) *Biochem. Biophys. Res. Commun. 138*, 850–857.

Runswick, M. J., Powell, S. J., Nyren, P., & Walker, J. E. (1987) *EMBO J. 6*, 1367–1373.

Saraste, M., & Walker, J. E. (1982) *FEBS Lett. 144*, 250–254.

Sharp, P. M., Tuohy, T. M. F., & Mosurski, K. R. (1986) *Nucleic Acids Res. 14*, 5125–5143.

Wohlrab, H. (1986) *Biochim. Biophys. Acta 853*, 115–134.

Wohlrab, H., Kolbe, H., & Collins, A. (1986) *Methods Enzymol. 125*, 697–709.

Wohlrab, H., Bukusoglu, C., Abuerreish, G., Rasmussen, U. B., & Kolbe, H. (1988) in *Cellular Calcium and Phosphate Transport in Health and Disease* (Bronner, F., & Peterlik, M., Eds.) pp 177–182, Alan R. Liss, New York.