Vol. 145, No. 1, 1987 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 29, 1987 Pages 390-396

YEAST PYRUVATE CARBOXYLASE : GENE ISOLATION

C. Phillip Morris, Filip Lim, and John C. Wallace

Department of Biochemistry, University of Adelaide,
Adelaide, South Australia 5000.

Received March 31, 1987

To improve our understanding of pyruvate carboxylase (PC)(EC 6.4.1.1) structure and the evolution of the biotin-dependent carboxylases we have isolated and sequenced a yeast (*Saccharomyces cerevisiae*) genomic DNA fragment encoding PC. The identity of the cloned gene was confirmed by comparing the encoded protein with the sequence of a 26 amino acid biotin-containing peptide isolated from yeast PC. The yeast PC sequence is homologous (43% amino acid homology) to the rat PC sequence, although the carboxyl-terminus was found to be 44 residues from the biotinyl-lysine whereas in all biotin carboxylases sequenced to date the biotin is 35 residues from the carboxyl-terminus.

Pyruvate carboxylase (PC)(EC 6.4.1.1) is a widely distributed biotin-dependent enzyme found in all higher eukaryotes (other than plants), fungi and some prokaryotes. In vertebrate liver and kidney it acts as a key enzyme in the regulation of gluconeogenesis from lactate and alanine, whilst in other tissues it has an anaplerotic role in relation to the tricarboxylic acid cycle, particularly in the transport of acetyl groups from mitochondria to the cytosol in lipogenic and acetyl-choline synthesizing tissues.

PC is a member of the "family" of biotin-dependent carboxylases which are proposed to have a common evolutionary origin (1). In all eukaryotic species which have been examined the enzyme is composed of four apparently identical subunits ranging in size from 115,000 to 125,000 (2). Electron microscopic studies of PC from yeast (*Saccharomyces cerevisiae*) (3), *Aspergillus nidulans* (4), *Rhizopus arrhizus* (5), chicken (6,7), sheep (6,7) and rat (7) have shown that the subunits in the native enzyme are arranged in a tetrahedron-like structure. Each subunit contains one biotin moiety located within the active site of the enzyme and covalently attached through the ϵ-amino group of a specific lysine residue.

0006-291X/87 $1.50

The general understanding of the catalytic mechanism and potential regulatory properties of PC is more advanced than is knowledge about its structure (8, 9). Structural studies have been hampered by the lack of sequence data, which for PC is restricted to the limited peptide sequence data from the biotin attachment site of sheep, chicken and turkey PC described by Rylatt *et al.* (10) and the extremely limited nucleic acid sequence data (151 bp encoding 44 amino acids) from a human PC cDNA clone (11).

Here we report the identification and sequencing of a clone containing a fragment of yeast genomic DNA derived from the pyruvate carboxylase gene. The yeast PC sequence revealed homology with sheep, chicken, human and rat PC around the biotin attachment site. We have found that the biotin moiety is attached to the protein close to its carboxyl-terminus, which is consistent with the general arrangement seen in other biotin-containing enzymes. However in all of the biotin-containing enzymes examined so far the prosthetic group is attached 35 residues from the carboxyl-terminus whereas this distance is increased to 44 residues in yeast PC. This is the first report of any sequence data for yeast PC or any of the yeast biotin carboxylases, additionally it is the first report of a significant amount of sequence data for PC from any species.

Materials and Methods

The yeast genomic library was provided by M. Carlson and D. Botstein (12) and was prepared by subcloning a partial Sau 3A digest of yeast DNA into the yeast/*E. coli* plasmid vector YEp 24. It was screened using a [^{32}P]-labelled biotin carboxylase specific oligonucleotide probe (described below) by the high density colony screening procedure of Hanahan (13). Restriction endonucleases, T4-polynucleotide kinase and various other enzymes were purchased from New England Biolabs. [α-^{32}P]dATP, [α-^{32}P]dCTP and dideoxynucleotide sequencing kits were purchased from BRESA Pty. Ltd.

Positive colonies were purified by a further round of screening and plasmid DNA was prepared by the method of Birnboim and Doly (14). Oligonucleotide positive Eco RI restriction fragments were identified, Southern blotted (15) and subcloned into Eco RI cleaved pUC19 (16). Insert DNAs were sequenced by cleavage with the appropriate restriction endonucleases and subcloning of the isolated fragments into M13mp18 and mp19 (16) and sequenced by the dideoxynucleotide chain termination method described by Sanger (17).

Yeast pyruvate carboxylase was purified as described previously (3) and digested to completion with trypsin. The biotin-containing tryptic peptide was purified by avidin-Sepharose chromatography followed by reversed phase HPLC and sequenced using an Applied Biosystems gas phase sequencer (18).

Results and Discussion

The highly conserved sequence around the biotin attachment site of biotin carboxylases and PC in particular provides an opportunity to generate an unambiguous oligonucleotide probe composed of a mixture of four 14-mers, as shown below;

Protein	NH_2	Ala	Met	Lys*	Met	Glu	Thr	COOH
mRNA	5′	GCN	AUG	AAR	AUG	GAR	ACN	3′
Probe	3′		TAC	TTY	TAC	CTY	TG	5′

* Biotin attached in amide linkage to ϵ-NH_2 group of lysine.

N = any nucleotide, R = purine, Y = pyrimidine.

This biotin carboxylase specific probe (14-mer) was used to screen a library of genomic fragments in the vector YEp 24. Five positive clones with inserts ranging in size from 1.4 kb to 11 kb were obtained by screening $5x10^3$ recombinants. Restriction mapping and 14-mer probed Southerns revealed that three of the clones contained inserts derived from the same region of the yeast genome (data not shown).

In order to simplify further analysis and eliminate all genes other than the biotin carboxylases from these 14-mer positive clones, probe-positive Eco RI fragments were subcloned into pUC19. The three unique subclones obtained were used in Northern analysis of yeast poly(A)$^+$ mRNA. One subclone (pYPC 5) hybridized to a single mRNA species of about the right size to encode PC (data not shown) and comparable with the 4.2 kb mRNA species detected in rat liver extracts by a human PC cDNA clone (11). The majority of the 1.3 kb insert of pYPC 5 was restriction mapped and sequenced following the strategy shown in Figure 1.

The nucleotide sequence (Figure 2) was determined by subcloning the appropriate restriction fragments into M13mp19 followed by sequencing using the dideoxynucleotide chain termination method. The sequence (1184 nucleotides) contained about half coding and non-coding sequence (531 and 653 bp respectively). The remaining coding sequence (about 3 kb) awaits the isolation of a longer yeast PC clone since the sequence obtained extended to very near the end of the original YEp 24 clone insert.

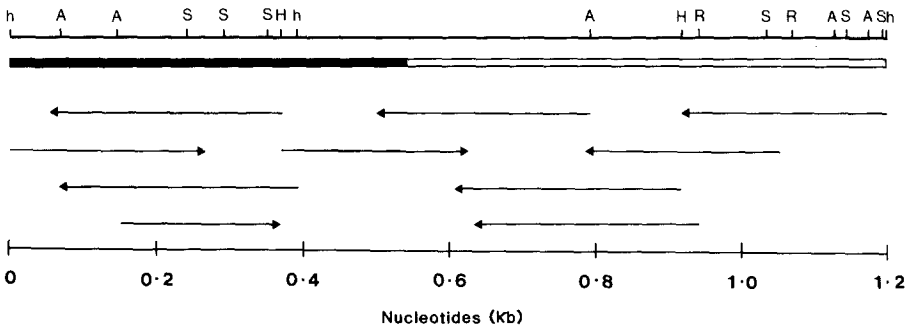

Figure 1. Partial restriction map of pYPC 5 containing part of the yeast pyruvate carboxylase gene. The extent of sequencing is indicated by the length of each arrow, and the direction of the arrow indicates the strand that was sequenced. The solid line indicates the extent of the coding region. Only restriction endonuclease cleavage sites used in the sequence analysis are indicated : Alu I (A), Hae III (H), Hha I (h), Rsa I (R) and Sau 3A (S).

The encoded peptide was identified as a biotin carboxylase since it contained the predicted sequence *Ala Met Lys Met*, corresponding to the highly conserved biotin attachment site found in all carboxylases examined to date (10, 11, 19, 20, 21). The clone was unambiguously identified as pyruvate carboxylase by the identical match obtained when the sequence around the biotin attachment site was a compared with a 26 amino acid, biotin-containing peptide isolated from purified yeast PC (Figure 2).

The sequence contained two polyadenylation sequences (AATAAA) but the correct site of transcription termination is not obvious because the sequences involved in the process of transcription termination in yeast are poorly defined (22). The termination codon TGA defines a carboxyl-terminus of the protein 44 amino acids from the lysine residue identified during amino acid sequencing of the biotin-containing peptide as the site of biotinylation in yeast PC. In this respect yeast PC differs from other biotin enzymes in that it contains an additional 9 amino acids at its carboxyl-terminus since the biocytin is positioned 35 amino acids from the carboxyl-terminus in all the biotin enzymes which have currently been sequenced including human and rat PC (11, A.I. Cassady, personal communication), human propionyl-CoA carboxylase (19), *E. coli* and chicken liver acetyl-CoA carboxylase (20, 23) and *P. shermanii* transcarboxylase (21). This suggests that the structural constraint placed on the position of the biotin in the other biotin-containing enzymes has been satisfied in a different way in yeast PC. It may be significant that there is little cross-reaction between the yeast and *Bacillus stearothermophilus* enzymes responsible

```
Tyr Pro Arg Val Tyr Glu Asp Phe Gln Lys Met Arg Glu Thr Tyr Gly Asp Leu Ser Val
TACCCAAGAGTTTATGAAGACTTCCAAAAGATGAGAGAAACGTATGGTGATTTATCTGTA 60

Leu Pro Thr Arg Ser Phe Leu Ser Pro Leu Glu Thr Asp Glu Glu Ile Glu Val Val Ile
TTGCCAACAAGAAGCTTTTTGTCTCCACTAGAGACTGACGAAGAAATTGAAGTTGTAATC 120

Glu Gln Gly Lys Thr Leu Ile Ile Lys Leu Gln Ala Val Gly Asp Leu Asn Lys Lys Thr
GAACAAGGTAAAACGCTAATTATCAAGCTACAGGCTGTGGGTGATTTGAACAAAAAGACC 180

Gly Glu Arg Glu Val Tyr Phe Asp Leu Asn Gly Glu Met Arg Lys Ile Arg Val Ala Asp
GGTGAAAGAGAAGTTTACTTTGATTTGAATGGTGAAATGAGAAAAATTCGTGTTGCTGAC 240

Arg Ser Gln Lys Val Glu Thr Val Thr Lys Ser Lys Ala Asp Met His Asp Pro Leu His
AGATCACAAAAAGTGGAAACTGTTACTAAATCCAAAGCAGACATGCATGATCCATTACAC 300

Ile Gly Ala Pro Met Ala Gly Val Ile Val Glu Val Lys Val His Lys Gly Ser Leu Ile
ATTGGTGCACCAATGGCAGGTGTCATTGTTGAAGTTAAAGTTCATAAAGGATCACTAATA 360

Lys Lys Gly Gln Pro Val Ala Val Leu Ser Ala Met Lys Met Glu Met Ile Ile Ser Ser
AAGAAGGGCCAACCTGTAGCCGTATTAAGCGCCATGAAAATGGAAATGATTATATCTTCT 420

Pro Ser Asp Gly Gln Val Lys Glu Val Phe Val Ser Asp Gly Glu Asn Val Asp Ser Ser
CCATCCGATGGACAAGTTAAAGAAGTGTTTGTCTCTGATGGTGAAAATGTGGACTCTTCT 480

Asp Leu Leu Val Leu Leu Glu Asp Gln Val Pro Val Glu Thr Lys Ala Stop
GATTTATTAGTTCTATTAGAAGACCAAGTTCCTGTTGAAACTAAGGCATGAACCGGTTAG 540

TTCTCATTTATAATGTATAATATACCCGAATCTTATTTATTTACCTTTCCTATTTTTTGA 600

CGACCAGTAAATACTAATACATAATTAGGAACAAAAGTTAAATAAAAAAAAAAAATAAT 660

TTAACGCATCCAATTAACGTGTCCTTTTTTCATCATTAATTTATCTACTATTTCGATTTA 720

AATTCCATATACAATAAATCCTAGATACATTCCCGAAAGTCATCTTTTAGCGAAACATCT 780

TCCTTGAGCTGCTAGCAGTGGGCTTAGTCCACCTGTTAGTTACTCTTGGTATACCACTAG 840

GTCTTTTCGAGGCAGGAACATGGGTCTTTCTTACTCTTCCGTTATTTGAAATTCCTGCCG 900

AAGAAACGGGCCTTCTACCAGATTGCACACCGTACCTTGGATTTAAGCCTTGATTTGGTG 960

GCTTTGCTGGTGCAGCATCCTTTTCACCCTTCCTATTTTCTAATTTTTGTCTCAAATCCA 1020

AGATCTCCTTCTTTTTCAACGCTATCACGTTGTGTAAAGTACTCCCATCTTCATCATTTG 1080

TGTTTACATTTTCATTGTGCAGTCTATCATTAATTAGCTTTAAATAGGTTTGATCGCTCA 1140

TCAATTTATCAATATAACCAGCTTGAGAATGAATGATGATCCAT 1184
```

Figure 2. Nucleotide sequence of a yeast genomic DNA fragment containing part of the pyruvate carboxylase gene and the deduced amino acid sequence. The extent of the match with the 26 amino acid peptide used to identify the gene is indicated by boxing, the lysine to which biotin is attached is indicated by a star (⋆), putative polyadenylation sequences are boxed and the sequence matching the 14-mer oligonucleotide probe is underlined – the last base did not match the probe.

for the attachment of biotin to pyruvate apocarboxylase (24), whereas other studies with biotinyl holoenzyme synthetases have shown a general lack of species or apocarboxylase specificity (25).

It is difficult to make comparisons of this data with other PC sequences because this is the first report of a significant amount of PC sequence from any species. There

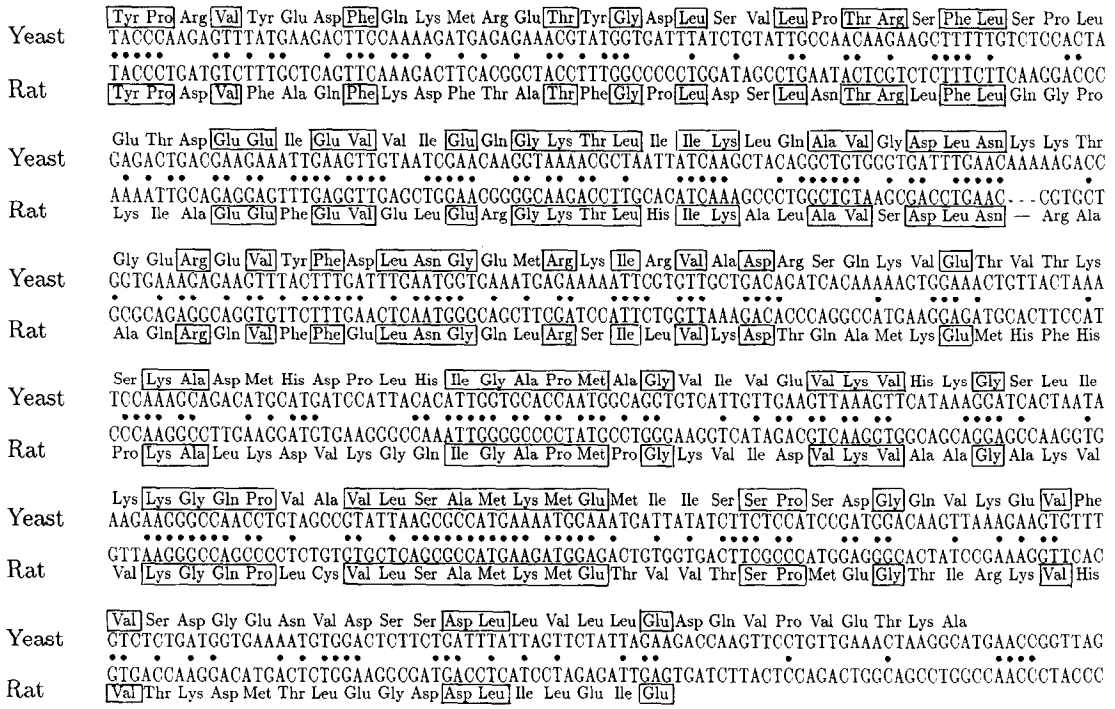

Figure 3. Comparison of the nucleotide and deduced amino acid sequences surrounding the biotin-binding site of yeast and rat pyruvate carboxylases. Homology of the nucleotide sequences is indicated by the symbol "•" spanning identical bases. Positions in the sequences where the amino acid residues match have been boxed.

is however a larger amount of as yet unpublished sequence data from rat and human PC cDNA clones (A.I. Cassady and R.A. Gravel, personal communications). Comparison of the yeast PC gene with the rat PC cDNA sequence (A.I. Cassady, personal communication, full data to be published elsewhere) revealed several blocks of sequence conservation, one of which corresponds to the site of biotin attachment. Some of the conservation that is present may be related to the recognition of the biotin attachment site by the biotinyl holoenzyme synthetase while others may serve to provide the correct cleft environment for the movement of biotin during the carboxyl transfer process at the active site of the enzyme. The sequences show a relatively low level of homology (48%) at the nucleic acid level yet the high degree of amino acid sequence conservation (43%, 57% if conservative amino acid substitutions are taken into account) indicates that strong selective forces have been acting to preserve regions of functional importance in these enzymes.

Comparison of the full yeast PC sequence with that of a higher eukaryote is likely to yield more information about structurally important regions than comparisons between

higher eukaryote sequences such as those from human and rat because the strong degree of sequence conservation shown between these PC sequences reflects their modest evolutionary separation. In support of this idea, a human PC cDNA clone (R.A. Gravel, personal communication) showed only four conservative substitutions when it was compared with the same rat PC data shown in Figure 3.

Acknowledgments

This work was supported by the Australian Research Grants Scheme grant D283 / 15800. The authors would like to acknowledge the assistance of A.I. Cassady, R.A. Gravel and G. Pure for helpful discussions. We wish to thank Ms. Yvonne Riese for her valuable technical assistance.

References

1. Obermayer, M. and Lynen, F. (1976) Trends Biochem. Sci. 1, 169–171.
2. Wallace, J.C. and Easterbrook-Smith, S.B. (1985) Pyruvate Carboxylase, pp. 65–108, Keech D.B. and Wallace, J.C. eds., CRC Press, Boca Raton, Florida.
3. Rohde, M., Lim, F. and Wallace, J.C. (1986) Eur. J. Biochem. 156, 15–22.
4. Osmani, S.A., Mayer, F., Marston, A.O., Selmes, I.P. and Scrutton, M.C. (1984) Eur. J. Biochem. 139, 509–518.
5. Mayer, F., Osmani, S.A. and Scrutton, M.C. (1985) FEBS Letts. 192, 215–219.
6. Goss, N.H., Dyer, P.Y., Keech, D.B. and Wallace, J.C. (1979) J. Biol. Chem. 254, 1734–1739.
7. Mayer, F., Wallace, J.C. and Keech, D.B. (1980) Eur. J. Biochem. 112, 265–272.
8. Attwood, P.V. and Keech, D.B. (1984) Curr. Topics Cellular Regulation 23, 1–55.
9. Barritt, G.J. (1985) (1985) Pyruvate Carboxylase, pp. 141–178, Keech D.B. and Wallace, J.C. eds., CRC Press, Boca Raton, Florida.
10. Rylatt, D.B., Keech, D.B. and Wallace, J.C. (1977) Arch. Biochem. Biophys. 183, 113–122.
11. Freytag, S.O. and Collier, K.J. (1984) J. Biol. Chem. 259, 12831–12837.
12. Carlson, M. and Botstein, D. (1983) Cell 28, 145–154.
13. Hanahan, D. and Meselson, M. (1980) Gene 10, 63–67.
14. Birnboim, H.C. and Doly, J. (1979) Nucl. Acids. Res. 7, 1531–1523.
15. Southern, E.M. (1975) J. Mol. Biol. 98, 503–517.
16. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103–119.
17. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463–5467.
18. Hunkapiller, M.W., Hewick, R.M., Dreyer, W.J. and Hood, L.E. (1985) Meths. Enzymol. 91, 399–412.
19. Lamhonwah, A., Barankiewicz, T.J., Willard, H.F., Mahuran, D.J., Quan, F. and Gravel, R.A. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 4864–4868.
20. Sutton, M.R., Fall, R.R., Nervi, A.M., Alberts, A.W., Vagelos, P.R. and Bradshaw, R.A. (1977) J. Biol. Chem. 252, 3934–3940.
21. Maloy, W.L., Bowien, B.U., Zwolinski, G.K., Kumar, K.G., Wood, H.G., Ericsson,L.H. and Walsh, K.A. (1979) J. Biol. Chem. 254, 11615–11622.
22. Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) Cell 41, 349–359.
23. Takai, T., Wada, K. and Tanabe, T. (1987) FEBS Letts. 212, 98–102.
24. Sundaram, T.K., Cazzulo, J.J. and Kornberg, H.L. (1971) Arch. Biochem. Biophys. 143, 609–616.
25. Wood, H.G. and Barden, R.E. (1977) Ann. Rev. Biochem. 46, 385–413.