Volume 282, number 1, 209–213 FEBS 09654 April 1991

ADONIS 001457939100362T

Sequence conservation in the α and β subunits of pyruvate dehydrogenase and its similarity to branched-chain α-keto acid dehydrogenase

Isaiah D. Wexler, Sullia G. Hemalatha and Mulchand S. Patel

Departments of Biochemistry and Pediatrics, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

Received 16 January 1991; revised version received 20 February 1991

Amino acid sequence comparison of 8 α and 6 β subunits of the α-keto acid dehydrogenase (E_1) component of the pyruvate dehydrogenase complex and branched-chain α-keto acid dehydrogenase complex from multiple species was performed by computer analysis. In addition to 2 previously recognized regions of homology in the α subunit, a 3rd region of extensive homology was identified in $E_1\alpha$, and may be one of the sites involved in subunit interaction. $E_1\beta$ contains 4 regions of extensive homology. Region 1 contains 10 amino acids that are homologous to a 10-amino acid stretch in *Escherichia coli* E_1. Regions 2 and 3 have sequence homologies with other dehydrogenases suggesting that these regions may be involved in catalysis.

Pyruvate dehydrogenase; Pyruvate dehydrogenase complex; Branched-chain α-keto acid dehydrogenase complex; α-Keto acid dehydrogenase complex; Amino acid sequence comparison

1. INTRODUCTION

The mammalian pyruvate dehydrogenase complex (PDC) is a mitochondrial multienzyme complex that catalyzes the conversion of pyruvate to acetyl-CoA [1,2]. The complex consists of 3 catalytic and 2 regulatory components, and a protein X component. In plant [3], eukaryotic, and some prokaryotic species [4], the first catalytic component, pyruvate dehydrogenase, (E_1) (pyruvate: lipoamide 2-oxidoreductase, EC 1.2.4.1), is composed of 2 non-identical α and β subunits which form a heterotetramer. E_1, which requires thiamin pyrophosphate as a cofactor, catalyzes the oxidative decarboxylation of pyruvate according to the following 2 partial reactions.

$$CH_3COCO_2^- + E_1 \cdot TPP + H^+ \rightarrow E_1 \cdot CH_3C(OH)=TPP + CO_2 \quad \text{(i)}$$

$$E_1 \cdot CH_3C(OH)=TPP + E_2LipS_2 \rightarrow E_1 \cdot TPP + E_2lip(SH)\text{-}SCOCH_3 \quad \text{(ii)}$$

Based on limited data, it has been postulated that reaction **(i)** is carried out by the α subunit, and reaction **(ii)** by the β subunit [5]. $E_1\alpha$ activity is regulated by phosphorylation-dephosphorylation by a specific E_1 kinase and phosphatase [1]. Sequence comparison of multiple thiamine-dependent enzymes has allowed for the identification of a putative thiamin pyrophosphate binding motif in $E_1\alpha$ [6]. Chemical modification of amino acid residues has identified critical amino acids in the E_1 heterotetramer that are critical for enzyme function [7]. Recent studies have shown that $E_1\beta$, but not $E_1\alpha$, is protected from trypsin digestion when bound to dihydrolipoyl acetyltransferase (E_2), the 2nd catalytic component of PDC, suggesting that $E_1\beta$ binds to E_2 [8]. Structure-function analysis has been facilitated by the cloning of both the α [4,9–11] and β [4,12,13] subunits of PDC from multiple species. However, the critical amino acid residues in the functional domains of $E_1\alpha$ and $E_1\beta$ remain unidentified. In the present study, we have used computer analysis to identify several regions of the α and β subunits that are highly conserved not only in all the known sequences of the 2 subunits of PDC from various species, but also in those of the branched-chain α-keto acid dehydrogenase complexes (BCKDC) in both prokaryotic and eukaryotic species. Similarities with other proteins suggest that one of the conserved domains may play a critical role in the oxidative decarboxylation of the α-keto acids.

Correspondence address: M.S. Patel, Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA. Fax: (1) (216) 3684544.

2. MATERIALS AND METHODS

All the published sequences known for both $E_1\alpha$ and $E_1\beta$ PDC and BCKDC were analyzed. The sequence for PDC $E_1\beta$ of the nematode, *Ascaris suum* [13] was kindly provided by Keith Johnson (University of Toledo, Toledo, OH). Comparison of sequences was done using Prosis software (Hitachi America, Brisbane, CA) using a best fit algorithm. Final alignment was done visually based on a comparison of all the sequences analyzed for a specific subunit. Comparison with other protein sequences was done by searching the National Biomedical Research Foundation (NBRF) data bank updated to September 1990 using the FastA program (Genetics Computer Group, University of Wisconsin). For reference, amino acid number-

Published by Elsevier Science Publishers B.V.

ing referred is based on the mature $E_1\alpha$ (chromosome X) and $E_1\beta$ of human PDC. Throughout this paper, an asterisk (*) denotes identical amino acids and a plus (+) indicates conserved amino acids.

3. RESULTS AND DISCUSSION

Four $E_1\alpha$ PDC and 4 $E_1\alpha$ BCKDC sequences were analyzed, including both eukaryotic and prokaryotic species (Fig. 1). There were 2 regions of extended homology, as identified previously including the thiamin pyrophosphate binding motif [6] located between human $E_1\alpha$ amino acids 156–203, and a region encompassing phosphorylation sites 1 and 2, spanning amino acids 263–274. Also present is a previously unrecognized region of high homology located between amino acids 217–261. Within this 45 amino acid stretch, 10 of the amino acids are identical in all 8 sequences. There is 33% identity and 48% homology among at least 7 of the 8 sequences (Fig. 1). No function has been attributed to this region. This region appears to be unique to α-keto acid dehydrogenases which have both α and β subunits, and it is not homologous to any sequences in *E. coli* PDC E_1 [14] or the E_1 components of *E. coli* and yeast α-ketoglutarate dehydrogenase complexes [15,16], which are dimers composed of only a single E_1 polypeptide. Neither is there significant homology between this region and amino acid sequences from other proteins listed in the NBRF protein sequence data bank. We have found a human $E_1\alpha$ point mutation located in this region [17]. This mutation affects protein stability but not catalytic activity [18]. Since previous studies have shown that α and β require each other's presence for stability in both PDC and BCKDC [19,20], this mutated amino acid residue may disrupt a critical site of interaction for the 2 subunits located in this homologous region thereby rendering both subunits unstable. The fact that this highly conserved region is only found in PDC and BCKDC which contain α and β subunits, and is not found in single protein E_1 catalytic units (in distinction to the thiamin pyrophosphate binding motif) is also suggestive that this region may play a role in subunit interaction. There is another region of a lesser degree of homology, spanning amino acid residues 61–90 (Fig. 1). In this region, 6 residues are identical and 6 other residues are conserved in all 8 of the sequences compared. The significance of this homology is not yet known.

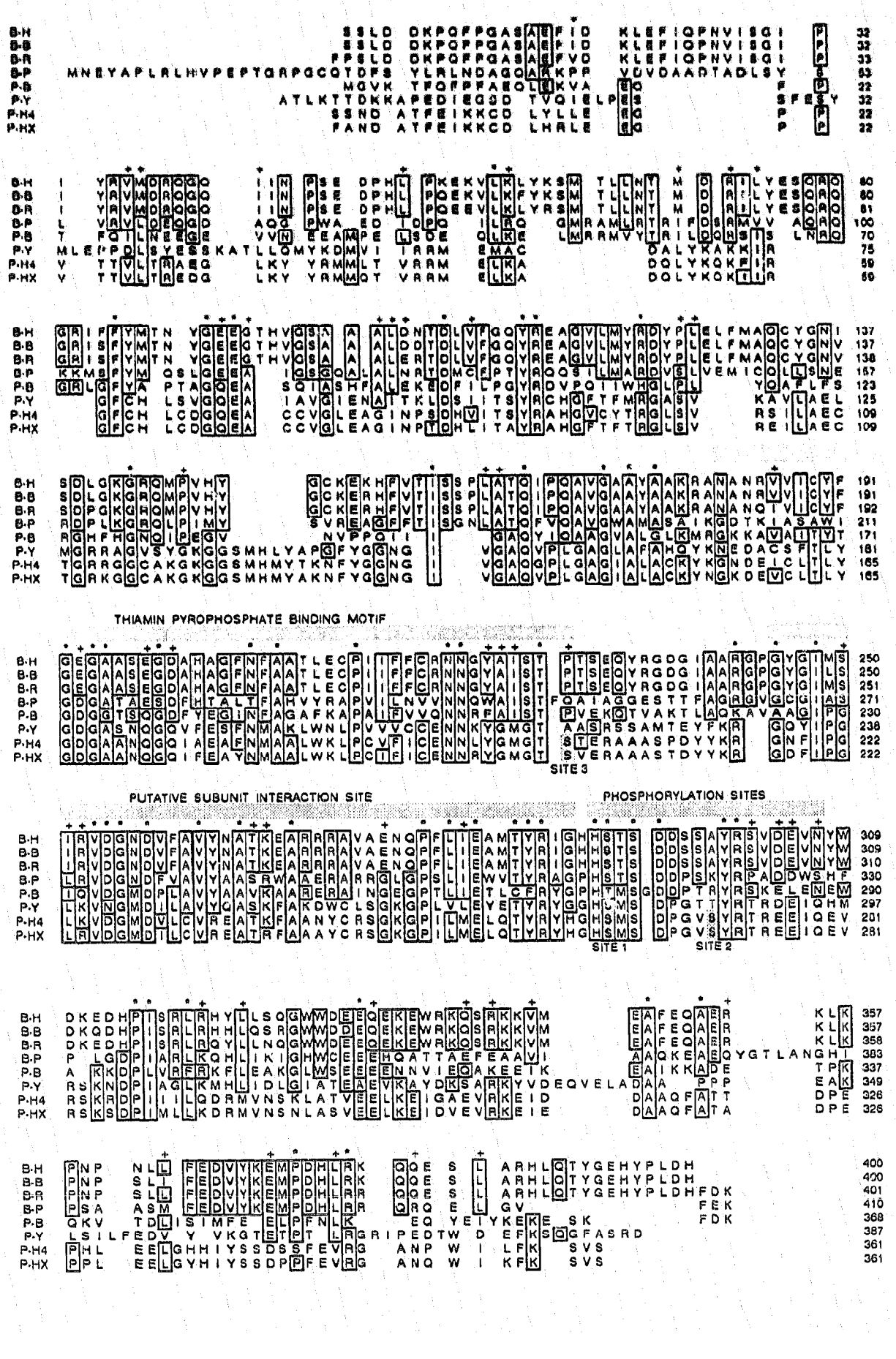
Six $E_1\beta$ sequences were also compared and included human [12], nematode *Ascaris suum* [13], and *Bacillus stearothermophilus* $E_1\beta$ [4] sequences as well as human [21], bovine [22], and *Pseudomonas putida* BCKDC $E_1\beta$ [23] (Fig. 2). There is greater sequence homology among β than α subunit sequences. There are 4 regions of extensive homology (Fig. 2). Region 1 consisting of 35 amino acids spanning amino acid residues 25–59 (34% identity and 69% homology among at least 5 of the 6 sequences) contains 10 residues (human $E_1\beta$ amino acid residues 41–50) which are homologous to a region in *E. coli* E_1 (amino acid residues 672–681) [14] as shown below. There is 50% identity and 70% homology in this region with the underlined identical amino acid residues being conserved in all the species that were compared (see below and Fig. 2).

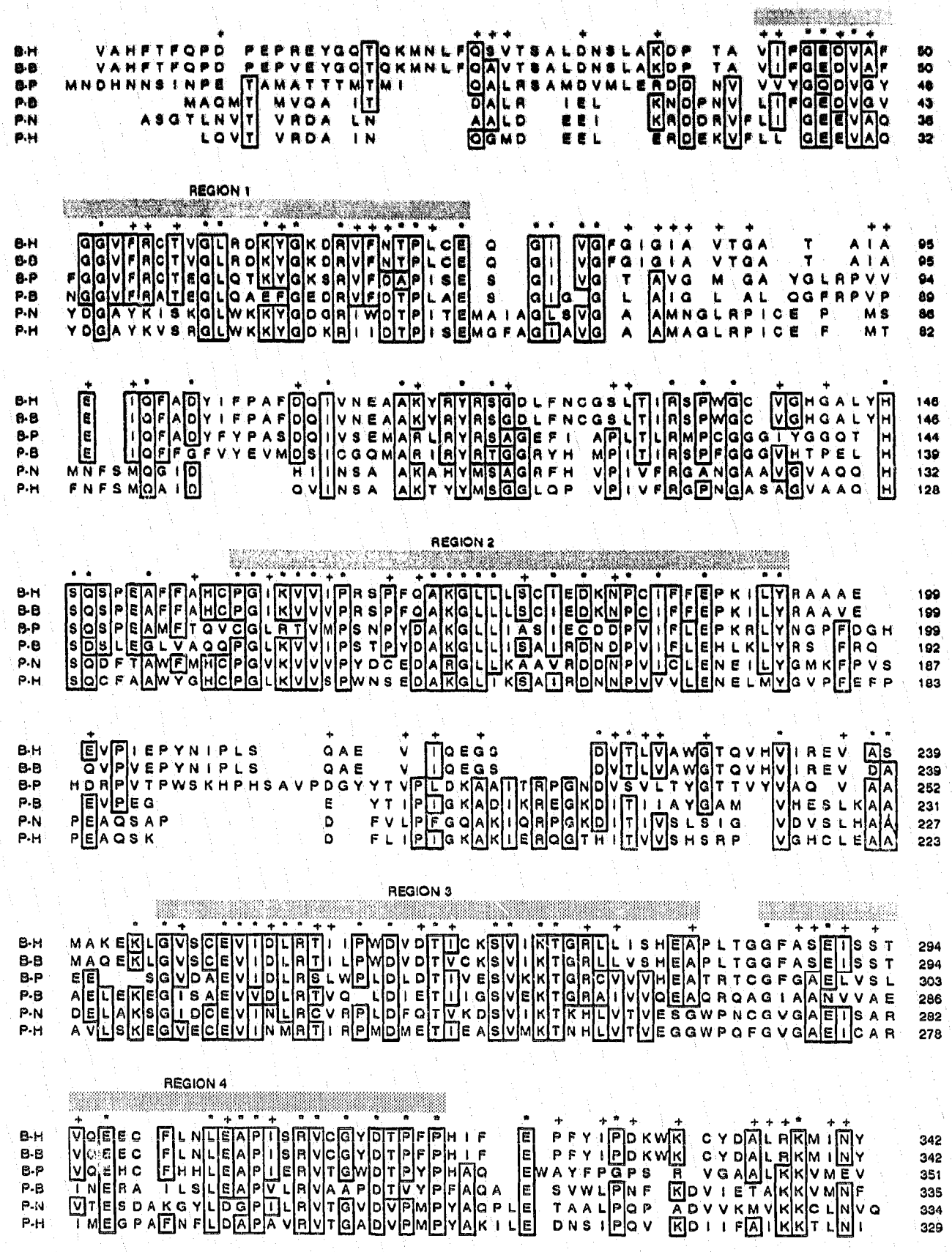
```
                    *  *     +     *  *  +  *
E. coli E1          D  G  L  E  R  M  Y  G  E  K
Human E1β           R  G  L  W  K  K  Y  G  D  K
```

This is the only region besides the thiamin pyrophosphate binding domain that had areas of homology with *E. coli* E_1 [14]. It is interesting that this homology exists since the E_1 from *E. coli* appears to be the result of evolutionary divergence [24]. It is tempting to speculate that just as the thiamin pyrophosphate binding motif is similar for all enzymes requiring thiamin pyrophosphate as a cofactor, presumably because binding to thiamin dictates a certain structure, this region of homology might also serve in a similar capacity for both *E. coli* E_1 and PDC $E_1\beta$. Since both the E_1 polypeptide of *E. coli* and the $E_1\beta$ of PDC bind to PDC E_2 (which has extensive homology with E_2s from other α-keto acid dehydrogenase complexes [24]), it is possible that this conserved stretch of amino acids in region 1 is one of the sites at which $E_1\beta$ binds to E_2. Region 2, spanning amino acids 140–176, is also very well conserved with 49% identity and 65% homology among at least 5 of the 6 sequences compared. There is very strong homology between Region 2 and 3-isopropyl malate dehydrogenase from *Thermus aquaticus* (amino acids 108–145) [25]. When compared to human $E_1\beta$, there is 30% identity and 49% homology as shown below.

```
3-Isopropylmalate dehydrogenase
***  +**   * *+*+    +*+  +  +   *   *
PGLERLSPLKEEIARGVDVLIVRELTGGIYFGEPRGMS
PGLKVVSPWNSEDAKGLIKSAIRDNNPVV-VLENELMY
Human E1β
```

Fig. 1. Alignment of 8 $E_1\alpha$ sequences from PDC and BCKDC. An asterisk (*) over a column denotes identity for at least 7 of the 8 amino acids in that column, and a plus sign (+) signifies homology among at least 7 of the amino acid residues in a column. Amino acids are boxed in if 4 or more residues are identical in a column. Shaded serine residues are sites of phosphorylation. Shaded lines indicate regions of increased homology. *Key:* B-H, BCKDC human $E_1\alpha$ [20]; B-B, BCKDC bovine $E_1\alpha$ [26]; B-R BCKDC rat $E_1\alpha$ [27]; B-P, BCKDC *Pseudomonas putida* $E_1\alpha$ [23]; P-B, PDC *Bacillus stearothermophilus* $E_1\alpha$ [4]; P-Y, PDC yeast $E_1\alpha$ [10]; P-H4, PDC human processed gene $E_1\alpha$ located on chromosome 4 [11]; and P-HX, PDC human $E_1\alpha$ located on chromosome X [9].


THIAMIN PYROPHOSPHATE BINDING MOTIF
PUTATIVE SUBUNIT INTERACTION SITE
PHOSPHORYLATION SITES
SITE 1
SITE 2
SITE 3


REGION 1
REGION 2
REGION 3
REGION 4

The similarity of region 2 to 3-isopropylmalate dehydrogenase, a non-thiamin pyrophosphate binding enzyme, catalyzing the oxidative decarboxylation of 3-isopropylmalate to α-ketoisocaproate, suggests that this region may represent a site that is involved in the oxidative decarboxylation of pyruvate and the branched-chain α-keto acids. This hypothesis is also supported by the fact that consensus sequences generated from regions 2 and 3 were shown by a computer search of the NBRF protein sequence data bank to have moderate degrees of homology with other dehydrogenases including *E. coli* NADH dehydrogenase (30% identity and 57% homology with 37 amino acids of region 2), *Xenopus laevis* NADH-ubiquinone dehydrogenase (24% identity and 49% homology with 37 amino acids of region 2), porcine malate dehydrogenase (28% identity and 48% homology with 29 amino acids in region 3), *E. coli* glutathione (NADPH) reductase (40% identity and 60% homology with 20 amino acids in region 3), and *Corynebacterium glutamicum* homoserine dehydrogenase (29% identity and 48% homology with 21 amino acids of region 3). Region 4, spanning amino acids 270–303 near the C-terminus, contains 12 identical amino acid residues and 7 conserved amino acid residues in at least 5 of the 6 sequences. A possible function has not been suggested for this region.

The above analysis shows that certain regions of both $E_1\alpha$ and $E_1\beta$ are conserved among the α-keto acid dehydrogenases. With the exception of the thiamin pyrophosphate binding motif and the phoshorylation sites, no definitive functions can be assigned to these regions at this time. However, the fact that these regions are so conserved among species suggests that they do serve some critical function that is shared among all the α-ketoacid dehydrogenase complexes that have α and β subunits for the E_1 component.

Acknowledgements: This work was supported in part by NIH grants DK 20478 and DK 42885. This work was aided by a Basil O'Connor Starter Scholar Research Award No. 5-759. I.D.W. is a recipient of a NIH physician scientist award HD 00878. The authors wish to thank Dr Keith Johnson (University of Toledo, Toledo, OH) for providing the β sequence of PDC for nematode prior to its publication, and Drs Douglas S. Kerr, Joyce E. Jentoft, William C. Merrick (all of Case Western Reserve University), and Thomas E. Roche (Kansas State University) for their critical reading of the manuscript. We would also like to thank Drs Paul and Linda Brady (University of Minnesota) for their assistance with the computer analysis.

REFERENCES

[1] Reed, L.J. (1974) Acc. Chem. Res. 7, 40–46.
[2] Patel, M.S. and Roche, T.E. (1990) FASEB J. 4, 3224–3233.
[3] Camp, P.J. and Randall, D.D. (1985) Plant Physiol. 77, 571–577.
[4] Hawkins, C.F., Borges, A. and Perham, R.N. (1990) Eur. J. Biochem. 191, 337–346.
[5] Roche, T.E. and Reed, L.J. (1972) Biochem. Biophys. Res. Commun. 48, 840–846.
[6] Hawkins, C.F., Borges, A. and Perham, R.N. (1989) FEBS Lett. 255, 77–82.
[7] Khailova, L.S., Korochkina, L.G. and Severin, S.E. (1989) Ann. N.Y. Acad. Sci. 573, 36–54.
[8] Rahmatullah, M., Gopalakrishnan, S., Andrews, P.C., Chang, C.L., Radke, G.A. and Roche, T.E. (1989) J. Biol. Chem. 264, 2221–2227.
[9] Ho, L., Wexler, I.D., Liu, T-C., Thekkumkara, T.J. and Patel, M.S. (1989) Proc. Natl. Acad. Sci. USA 86, 5330–5334.
[10] Behal, R.H., Browning, K.S. and Reed, L.J. (1989) Biochem. Biophys. Res. Commun. 164, 941–946.
[11] Dahl, H.H., Brown, R.M., Hutchison, W.M., Maragos, C. and Brown, G.K. (1990) Genomics 8, 225–232.
[12] Ho, L. and Patel, M.S. (1990) Gene 86, 297–302.
[13] Wheelock, M.J., Komuniecki, R., Duran, E. and Johnson, K.R. (1991) Mol. Biochem. Parasit. (in press).
[14] Stephens, P.E., Darlison, M.G., Lewis, H.M. and Guest, J.R. (1983) Eur. J. Biochem. 133, 155–162.
[15] Darlison, M.G., Spencer, M.E. and Guest, J.R. (1984) Eur. J. Biochem. 141, 351–359.
[16] Repetto, B. and Tzagoloff, A. (1989) Mol. Cell. Biol. 9, 2695–2705.
[17] Wexler, I.D., Hemalatha, S.G., Liu, T-C., Berry, S.A., Kerr, D.S. and Patel, M.S. (1990) Pediatr. Res., 27, 194A.
[18] Kerr, D.S., Berry, S.A., Lusk, M.M., Ho, L. and Patel, M.S. (1988) Pediatr. Res. 24, 95–100.
[19] Ho, L., Wexler, I.D., Kerr, D.S. and Patel, M.S. (1989) Ann. N.Y. Acad Sci. 573, 347–359.
[20] Fisher, C.W., Chuang, J.L., Griffin, T.A., Lau, K.S., Cox, R.P. and Chuang, D.T. (1989) J. Biol. Chem. 264, 3448–3453.
[21] Nobukuni, Y., Mitsubuchi, H., Endo, F., Akaboshi, I., Asaka, J. and Matsuda, I. (1990) J. Clin. Invest. 86, 242–247.
[22] Nobukuni, Y., Mitsubuchi, H., Endo, F., Asaka, J., Oyama, R., Titani, K. and Matsuda, I. (1990) Biochemistry 29, 1154–1160.
[23] Burns, G., Brown, T., Hatter, K., Idriss, J.M. and Sokatch, J.R. (1988) Eur. J. Biochem. 176, 311–317.
[24] Guest, J.R., Angier, S.J. and Russell, G.C. (1989) Ann. N.Y. Acad. Sci. 573, 76–99.
[25] Kagawa, R., Nojima, H., Nukiwa, N., Ishizuka, M., Yasuhara, T., Tanaka, T. and Oshima, T. (1984) J. Biol. Chem. 259, 2956–2960.
[26] Hu, C-W.C., Lau, K.S., Griffin, T.A., Chuang, J.L., Fisher, C.W., Cox, R.P. and Chuang, D.T. (1988) J. Biol. Chem. 263, 9007–9014.
[27] Zhang, B., Kuntz, M.J., Goodwin, G.W., Harris, R.A. and Crabb, D.W. (1987) J. Biol. Chem. 262, 15220–15224.

←

Fig. 2. Alignment of 6 $E_1\beta$ sequences from PDC and BCKDC. An asterisk (*) over a column denotes identity among at least 5 of the 6 amino acids in that column, and a plus sign (+) signifies homology of at least 5 of the amino acid residues in a column. Amino acids are boxed in if 4 or more residues in a column are identical. Shaded lines indicate regions of increased homology. *Key:* B-H, BCKDC human $E_1\beta$ [21]; B-B, BCKDC bovine $E_1\beta$ [22], B-P, BCKDC *Pseudomonas putida* $E_1\beta$ [23]; P-B, PDC *Bacillus stearothermophilus* $E_1\beta$ [4]; P-N, PDC nematode *Ascaris suum* $E_1\beta$ [13]; and P-H, PDC human $E_1\beta$ [12].