Vol. 177, No. 1, 1991
May 31, 1991

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 218-223

DIFFERENT SEGMENT SIMILARITIES IN LONG-CHAIN DEHYDROGENASES

Bengt Persson[1], Jonathan Jeffery[2] and Hans Jörnvall[1]

[1]Department of Chemistry I, Karolinska Institutet, S-104 01 Stockholm, Sweden

[2]Department of Molecular and Cell Biology, University of Aberdeen, Marischal College, Aberdeen AB9 1AS, United Kingdom

Received April 19, 1991

SUMMARY: Long-chain dehydrogenases were scrutinized for common patterns. Overall molecular similarities are not discerned, in contrast to the situation for several short-chain and medium-chain dehydrogenases, but coenzyme-binding segments are discernible. Species variants of glucose-6-phosphate dehydrogenase reveal about 20% strictly conserved residues, grouped into three segments and supporting assignments of sites for coenzyme-binding and catalysis. Glycine is overrepresented among the residues conserved, typical of distantly related proteins. Two of the enzymes within the pentose phosphate pathway reveal a distant similarity of interest for further evaluation, between a C-terminal 178-residue segment of glucose-6-phosphate dehydrogenase and the N-terminal part of 6-phosphogluconate dehydrogenase.

Dehydrogenase families have recently been shown to cover many different enzyme activities, such as diverse substrate specificities among some 20 short-chain dehydrogenases [1] and 7 different medium-chain dehydrogenases [2–4]. We have have therefore now compared "long-chain" glucose-6-phosphate dehydrogenase [5] and 6-phosphogluconate dehydrogenase [6] of the pentose phosphate pathway, with several similarly sized dehydrogenases in search of common patterns and possibly further family relationships. Some segments of similarity, probably of importance in coenzyme binding, are found between different long-chain dehydrogenases, but no entire-chain homologies are detected at this stage. Consequently, long-chain dehydrogenases thus far largely appear structurally unique in each case. However, some segment similarities are noticed. Apart from those known before, they reflect a segment of distant, possible similiarity between glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase, suggesting that NADP-dependent enzymes in the pentose phosphate pathway may contribute to tracings of further dehydrogenase relationships.

MATERIALS AND METHODS

Screenings were performed with the primary structure of *S. cerevisiae* glucose-6-phosphate dehydrogenase [5] and *E. coli* 6-phosphogluconate dehydrogenase [6] towards the proteins

0006-291X/91 $1.50

in the SwissProt data bank [7] utilizing the FASTA program [8]. The two dehydrogenases were further compared with more than 20 different dehydrogenases of 450–600 amino acid residues, utilizing both a program allowing spans of 15–45 residues in order to detect gap-derived frame-shifts, and one to construct hydrophilicity plots [9] to detect pattern similarities. Glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase were also compared with medium-chain [3] and short-chain [1] dehydrogenases.

RESULTS

The comparisons reveal one short segment of similarity between 6-phosphogluconate dehydrogenase and lipoamide dehydrogenase (Fig. 1A). This segment has been concluded to be NAD-binding in lipoamide dehydrogenase [10], and could therefore constitute part of the NADP-binding site in 6-phosphogluconate dehydrogenase. Furthermore, several FAD-dependent enzymes showed similarities to the FAD-binding site in lipoamide dehydrogenase (Fig. 1B), probably reflecting an FAD-binding region. A recently characterized flavine reductase [11] also bears similarities to this region (Fig. 1B). In addition, the comparisons support previously described similarities between aldehyde

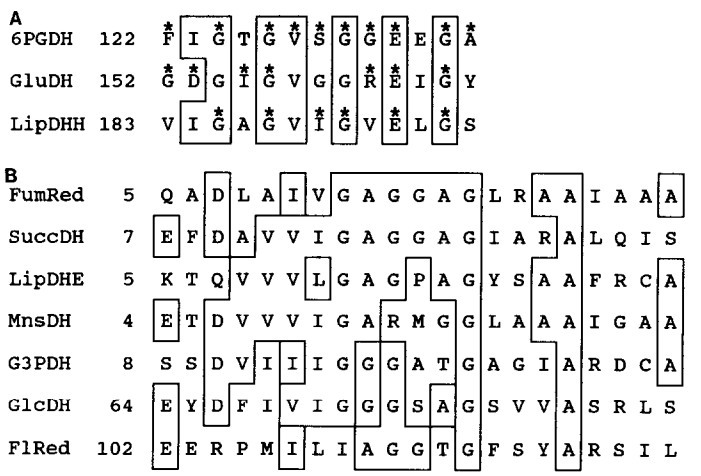

Figure 1. *Putative coenzyme-binding segments.* **A**, segment similarity between 6-phosphogluconate dehydrogenase (*E. coli*, 6PGDH), glutamate dehydrogenase type IV (*N. crassa*, GluDH), and lipoamide dehydrogenase (human, LipDHH). This segment is ascribed to NAD-binding in lipoamide dehydrogenase [10] and may constitute part of the NADP-binding site in the other dehydrogenases. Asterisks indicate positions which are conserved in other species [7]. **B**, segment similarity between different FAD-binding proteins, fumarate reductase (*E. coli*, FumRed), succinate dehydrogenase (*E. coli*, SuccDH), lipoamide dehydrogenase (*E. coli*, LipDHE), methoxyneurosporene dehydrogenase (*Rhodobacter capsulatus*, MnsDH), glycerol-3-phosphate dehydrogenase (*E. coli*, G3PDH), glucose dehydrogenase (*E. coli*, GlcDH), and flavine reductase (*E. coli*, FlRed). This segment is FAD-binding in lipoamide dehydrogenase [10] and is concluded to be FAD-binding also in the other enzymes. Identical residues are boxed (in B those identical in 3 or more structures). Numbers give starting position of each line. Sequences from [7,11].

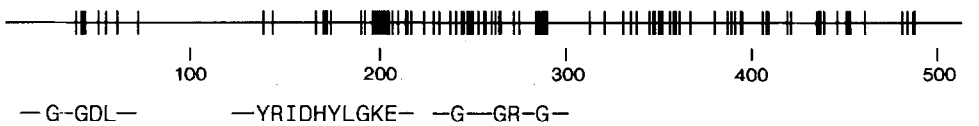

Figure 2. *Distribution of conserved residues in glucose-6-phosphate dehydrogenase.* Alignment of the human, rat, *D. melanogaster*, *S. cerevisiae*, *Z. mobilis* and *E. coli* enzymes [5,15,16] reveal 99 strictly conserved residues marked as vertical bars. Numbering refers to the rat enzyme. Expanded parts show putative coenzyme-binding and active sites.

dehydrogenase and delta-1-pyrroline-5-carboxylate dehydrogenase [12,13] on the one hand, and between fumarate reductase and succinate dehydrogenase on the other [14].

The recently characterized prokaryotic forms of glucose-6-phosphate dehydrogenase from *Z. mobilis* [15] and *E. coli* [16] together with mammalian, insect and yeast forms (cf. [5]) allow extended comparisons of species variants within a long-chain dehydrogenase. The results reveal that only 21% of the residues are strictly conserved through all lines (Fig. 2), affecting mainly Gly, Leu and Arg. The Gly overrepresentation indicates that the structures are similarly folded, in the same manner as a Gly conservation of other enzymes [1–3] has been found to be typical of distant relationships. Notably (Fig. 2), conserved residues are differently distributed in three separate regions. A central region covers a cluster of residue identities, while the whole C-terminal segment has residue identities fairly evenly distributed, and the N-terminal long segments with only few residues conserved. The residues flanking the putative active site residues His-200 and Lys-204 (rat enzyme positional numbers) are in the central region (Fig. 2). This region also includes one of the postulated [5] coenzyme-binding sites (241–247). However, the other site postulated for coenzyme-binding (positions 37–43) cannot be excluded since those residues are also conserved, with the exception of an Ala/Ser exchange in one species. Thus, with the addition also of prokaryotic forms of the enzyme, the number of strictly conserved residues is now considerably lower than in a previous study [1], emphasizing the importance of characterization of species variants in order properly to trace the functions of the protein.

Furthermore, a distant similarity (22%, 35 of 160 positions, excluding gaps) was found between the C-terminal 178-residue segment of *S. cerevisiae* glucose-6-phosphate dehydrogenase and the N-terminal part of *E. coli* 6-phosphogluconate dehydrogenase (Fig. 3). However, this residue conservation is not supported by inclusion of further species variants. On the other hand, hydrophilicity profiles [9] then show similarities for this segment (Fig. 3).

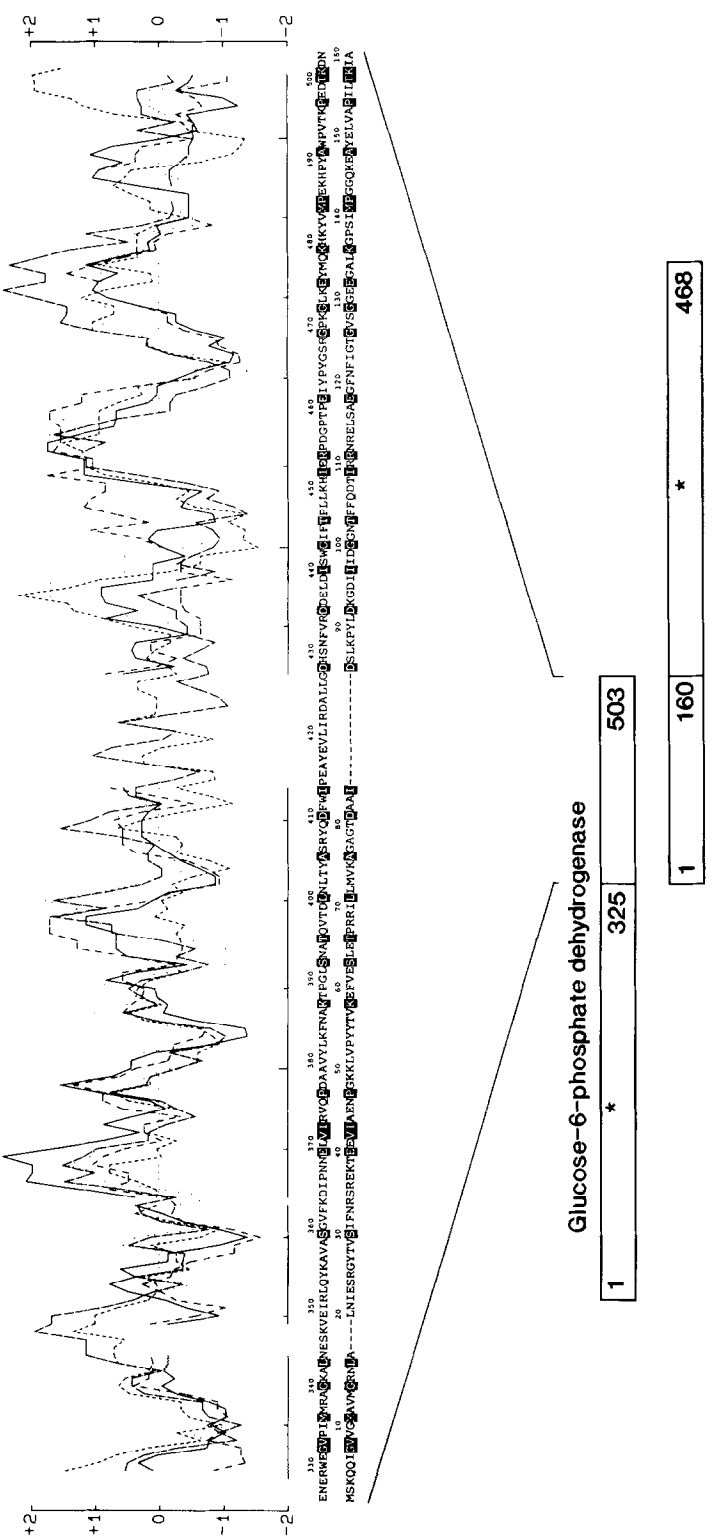

Figure 3. *Comparisons between segments of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase.* Hydrophilicity profiles (top) including species variants for glucose-6-phosphate dehydrogenase (*S. cerevisiae* [- - - - -], human [· · · · ·] and *D. melanogaster* [– · – · –]) and 6-phosphogluconate dehydrogenase (*E. coli* [———] and sheep [– – –]), with a short dash at every 10th and a long at every 50th dash. Alignment (below) refer to Bakers' yeast glucose-6-phosphate dehydrogenase and *E. coli* 6-phosphogluconate dehydrogenase. Conserved residues are outlined in black. Asterisks in bottom schemes indicate the chemically reactive lysine residues [17,18].

DISCUSSION

Glucose-6-phosphate dehydrogenase does not reveal entire-chain homology to other long-chain dehydrogenases, suggesting that it has a thus far largely unique structure. However, the comparisons reveal several short segments of similarity between potential coenzyme-binding sites (Fig. 1). Overall, the comparisons show that long-chain dehydrogenases do not appear to form easily traceable inter-related families, like those for the short-chain [1] or medium-chain [2] dehydrogenases. Some previously detected similarities were verified, concerning aldehyde dehydrogenase and delta-1-pyrroline-5-carboxylate dehydrogenase [12,13], and fumarate reductase and succinate dehydrogenase [14].

Two enzymes of the oxidative branch of the pentose phosphate pathway show distant similarities, between the C-terminal 178-residue segment of glucose-6-phosphate dehydrogenase and the N-terminal part of 6-phosphogluconate dehydrogenase (Fig. 3). They suggest the possibility of a common domain. Furthermore, both enzymes have a chemically reactive lysine residue, supposed to be important in catalysis [17,18]. This Lys residue is in both cases located in the middle of remaining parts, increasing the pattern symmetries (Fig. 3), and possibly indicating that the active sites, as for other enzymes [1], are located in an inter-domain region, dividing the large non-homologous segment in two parts roughly equal in size. However, the region homologous to other NAD(P)-binding enzymes in Fig. 1 is partly overlapping with this similarity (Fig. 3). Therefore, the significance of this distant similarity between the enzymes of the pentose phosphate pathway may be questioned, or may indicate very complicated ancestral relationships. Nevertheless, although the resemblance is not yet supported by other data, it constitutes a long-chain dehydrogenase similarity that may be of interest for further evaluation, especially since conclusions from segment distributions, lysine reactivity, and hydrophilicity patterns coincide.

ACKNOWLEDGMENTS

This work was supported by grants from the Swedish Medical Research Council (03X-3532, 03P-8864), the Swedish Society for Medical Research, and the Fund in Memory of Bengt Lundqvist.

REFERENCES

1. Persson, B., Krook, M., and Jörnvall, H. (1991) Eur. J. Biochem., in press.
2. Jörnvall, H., Persson, B., and Jeffery, J. (1987) Eur. J. Biochem. 167, 195–201.
3. Borrás, T., Persson, B., and Jörnvall, H. (1989) Biochemistry 28, 6133–6139.
4. Parés, X., Moreno, A., Cederlund, E., Höög, J.-O., and Jörnvall, H. (1990) FEBS Lett. 277, 115–118.

5. Persson, B., Wood, I., Jörnvall, H., and Jeffery, J. (1991) Eur. J. Biochem., in press.
6. Nasoff, M. S., Baker, II, H. V., and Wolf, Jr., R. E. (1984) Gene 27, 253–264.
7. Swiss-Prot Databank (1990), Release 16.0.
8. Pearson, W. R., and Lipman, D. J. (1988) Proc. Natl Acad. Sci. USA 85, 2444–2448.
9. Hopp, T. P., and Woods, K. R. (1981) Proc. Natl Acad. Sci. USA 78, 3824–3838.
10. McKie, J. H., and Douglas, K. T. (1991) FEBS Lett. 279, 5–8.
11. Spyrou, G., Haggård-Ljungquist, E., Krook, M., Jörnvall, H., Nilsson, E., and Reichard, P. (1991) J. Bacteriol., in press.
12. Krzywicki, K. A., and Brandriss, M. C. (1984) Mol. Cell. Biol. 4, 2837–2842.
13. Forte-McRobbie, C. M., and Pietruszko, R. (1986) J. Biol. Chem. 261, 2154–2163.
14. Wood, D., Darlison, M. G., Wilde, R. J., and Guest, J. R. (1984) Biochem. J. 222, 519–534.
15. Barnell, W. O., Yi, K. C., and Conway, T. (1990) J. Bacteriol. 172, 7227–7240.
16. Rowley, D. L., and Wolf, Jr., R. E. (1991) J. Bacteriol. 173, 968–977.
17. Carne, A., and Walker, J. E. (1983) J. Biol. Chem. 258, 12895–12906.
18. Jeffery, J., Hobbs, L., and Jörnvall, H. (1985) Biochemistry 24, 666–671.