Proc. Nat. Acad. Sci. USA
Vol. 70, No. 8, pp. 2295–2298, August 1973

Partial Similarities Between Yeast and Liver Alcohol Dehydrogenases

(primary structure/diagram comparison/conserved region/amino-acid exchanges)

HANS JÖRNVALL

Department of Chemistry I, Karolinska Institutet, S-104 01 Stockholm 60, Sweden

Communicated by Hugo Theorell, May 11, 1973

ABSTRACT **The primary structure of about half of the protein chain of yeast alcohol dehydrogenase has been determined and compared with the amino-acid sequences of other dehydrogenases. The enzyme is found to be distantly related to horse-liver alcohol dehydrogenase, although these two proteins have different quaternary structures and subunit sizes. Some regions show no significant similarities, but long segments within the N-terminal parts of the molecules are homologous, suggesting a common and important function for these segments. Ancestral connections between some different dehydrogenases can be concluded and the degree of evolutionary changes may be estimated.**

Many attempts to trace evolutionary connections between different dehydrogenases have been made. Similarities are indicated from comparisons without knowledge of enzyme structures (1–3) and from comparisons of short regions of glyceraldehyde phosphate dehydrogenase (GPDH), liver alcohol dehydrogenase (LADH), yeast alcohol dehydrogenase (YADH), lactate dehydrogenase (LDH), or glutamate dehydrogenase (GDH) (4–10). Long segments of possible but distant homology are found between GPDH and LADH (11) or GDH (12). Similarities in the tertiary structures of LDH (13), soluble malate dehydrogenase (14), LADH (15), and GPDH (16) are also evident. Ancestral connections between different dehydrogenases are therefore likely, although evolutionary relationships are not clear.

In this respect, the structure of YADH is of particular interest since some of its properties are identical to those of LADH, whereas others are identical to those of LDH and GPDH. Thus, both YADH and LADH have alcohol as substrate and contain zinc (17, 18), in contrast to LDH (13) and GPDH (19). On the other hand, YADH, LDH, and GPDH are all tetrameric with subunits of molecular weight about 36,000 (7, 13, 20), whereas LADH is dimeric with subunits of molecular weight 40,000 (21).

The primary structure of YADH was therefore studied. The entire amino-acid sequence has not been fully established, but long segments of the molecule could be deduced and compared with other dehydrogenases. YADH was then found to be distantly related to LADH, indicating ancestral connections, although similarities are not evenly distributed over the whole molecules. A region within the N-terminal third of LADH has a more conserved homology towards YADH and may serve a common and important function, such as coenzyme binding. These results and those in the accompanying paper (15) therefore support each other and it is possible to estimate the degree of evolutionary change between some dehydrogenases.

Abbreviations: GPDH, glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12); LADH, horse-liver alcohol dehydrogenase (EC 1.1.1.1), isoenzyme EE; YADH, yeast alcohol dehydrogenase (EC 1.1.1.1); LDH, lactate dehydrogenase (EC 1.1.1.27); GDH, glutamate dehydrogenase (EC 1.4.1.3).

Work on the primary structure of YADH

Peptide mixtures obtained after treatment of [^{14}C]carboxymethylated YADH with trypsin, chymotrypsin, pepsin, thermolysin, or cyanogen bromide, as well as tryptic digests of the maleylated (22) protein, were fractionated by exclusion chromatography on Sephadex G-50 (Pharmacia, Sweden) followed by high-voltage electrophoresis and chromatography on paper. Sequences of pure peptides were determined by the dansyl-Edman procedure (23, 24). The methods used for modification, digestion, purification, and analysis have been described (21, 11).

Peptide overlapping in YADH is complicated by the presence of many adjacent peptide bonds susceptible to hydrolysis by all the enzymes used. For example, leucine or an aromatic residue occur adjacent to an arginine or lysine residue in at least 12 places and, consequently, overlapping fragments are difficult to obtain. Nevertheless, 333 residues in the protein have been characterized in peptides. This is close to the total number (7, 25) of residues, and long segments may be ordered into a tentative amino-acid sequence for the subunit of YADH.

N- and C-terminal parts, which account for about half the subunit size of YADH, are shown in Fig. 1. These parts are deduced from the structures of peptides formed after the different proteolytic treatments mentioned above. At four places (indicated in Fig. 1) overlapping segments are short, but the sequence is supported by the fragmentation patterns of long maleylated tryptic fragments. This structure is also compatible with previous reports concerning partially characterized tryptic peptides (7, 25, 26) but will be considered tentative until the complete sequence of YADH has been established. The general validity of conclusions based on the comparisons below is not influenced by minor errors, if any, in the proposed sequence of YADH.

The N-terminus of YADH is blocked by an acyl group. The reactive cysteine residue (7) is present in the region shown in Fig. 1 (number 43), and four cysteine residues, or half the total number of cysteine residues in the protein (7, 25, 26), occur close together (positions 96, 100, 103, and 111 in Fig. 1).

Structure of YADH compared with other dehydrogenases

A direct comparison between YADH and other dehydrogenases of known structure does not reveal any obvious

Acyl-Ser-Ile-Pro-Glu-Thr-Gln-Lys-Gly-Val-Ile-Phe/Tyr-Glu-Ser-His-Gly-Lys-
-Leu-Glu-Tyr-Lys-Asp-Ile-Pro-Val-Pro-Lys-Pro-Lys-Ala-Asn-Glu-Leu-Leu/Ile-
-Asn-Val-Lys-Tyr/Ser-Gly-Val-Cys-His-Thr-Asp-Leu-His-Ala-Trp-His-Gly-Asp-
-Leu-Pro-Trp-Pro-Thr/Lys-Leu-Pro-Leu-Val-Gly-Gly-His-Glu-Gly-Ala-Gly-Val-
-Val-Val-Gly-Met-Gly-Glu-Asn-Val-Lys-Gly-Trp-Lys-Ile-Gly-Asp-Tyr-Ala-Gly-
-Ile-Lys-Trp-Leu-Asp-(Ser, Cys,Gly)-Met-Ala-Cys-Glu-Tyr-Cys-Glu-Leu-Gly-
-Asn-Glu-Ser-Asn-Cys-Pro-His-Ala-Asp-Leu-... | ...-Glu-Ala-Leu-Asp-Phe-
-Phe-Ala-Arg-Gly-Leu-Ile-Lys-Ser-Pro-Ile-Lys-Val-Val-Gly-Leu-Ser-Thr-Leu-
-Pro-Glu-Ile-Tyr-Glu-Lys-Met-Glu-Lys-Gly-Gln-Val-Val-Gly-Arg-Tyr-Val-Val-
-Asp-Thr-Ser-Lys-COOH.

FIG. 1. N- and C-terminal segments of YADH. Numbering after residue 11 is tentative. *Vertical bar* indicates positions where peptide overlapping is short.

similarities in the terminal regions except that YADH has an acylated N-terminal serine residue like lobster GPDH (27) and LADH (21). The available structure of YADH was therefore systematically searched for identities with LADH (28), GPDH (10), and known segments of LDH (13, 29), by the diagram method (30), in which "diagonal matching" indicates similarities.

Comparisons with GPDH and LDH yield no conclusive results. Only similarities that are short, require gaps, or fit better when one of the sequences is reversed are seen. If these reflect ancestral connections, exact relationships are impossible to discern without knowledge of interconnecting molecules. Comparison with another enzyme from the same species, yeast GPDH (10), does not yield any more definitive results.

Comparison of YADH with LADH, however, reveals distant similarities. A diagonal comparison between LADH and the N-terminal segment of YADH is shown in Fig. 2. A diagonal line is evident, indicating homology, from about residue 30 to about residue 85 in LADH, and to a lesser extent to the end of this part of YADH. A two-step upward displacement of the line is visible in the middle, indicating insertions in YADH or deletions in LADH. No clear similarities are seen between the N-terminal 25-residue regions (Fig. 2). Neither is any similarity found between the C-terminal 45-residue fragment of YADH and any part of the whole LADH chain when they are compared in the same way. The homology between YADH and LADH is, therefore, not uniformly distributed over the molecules. This conclusion is independent of the tentative nature of the structure of YADH, since minor alterations would neither change the general similarity in homologous regions nor the dissimilarity in nonhomologous regions.

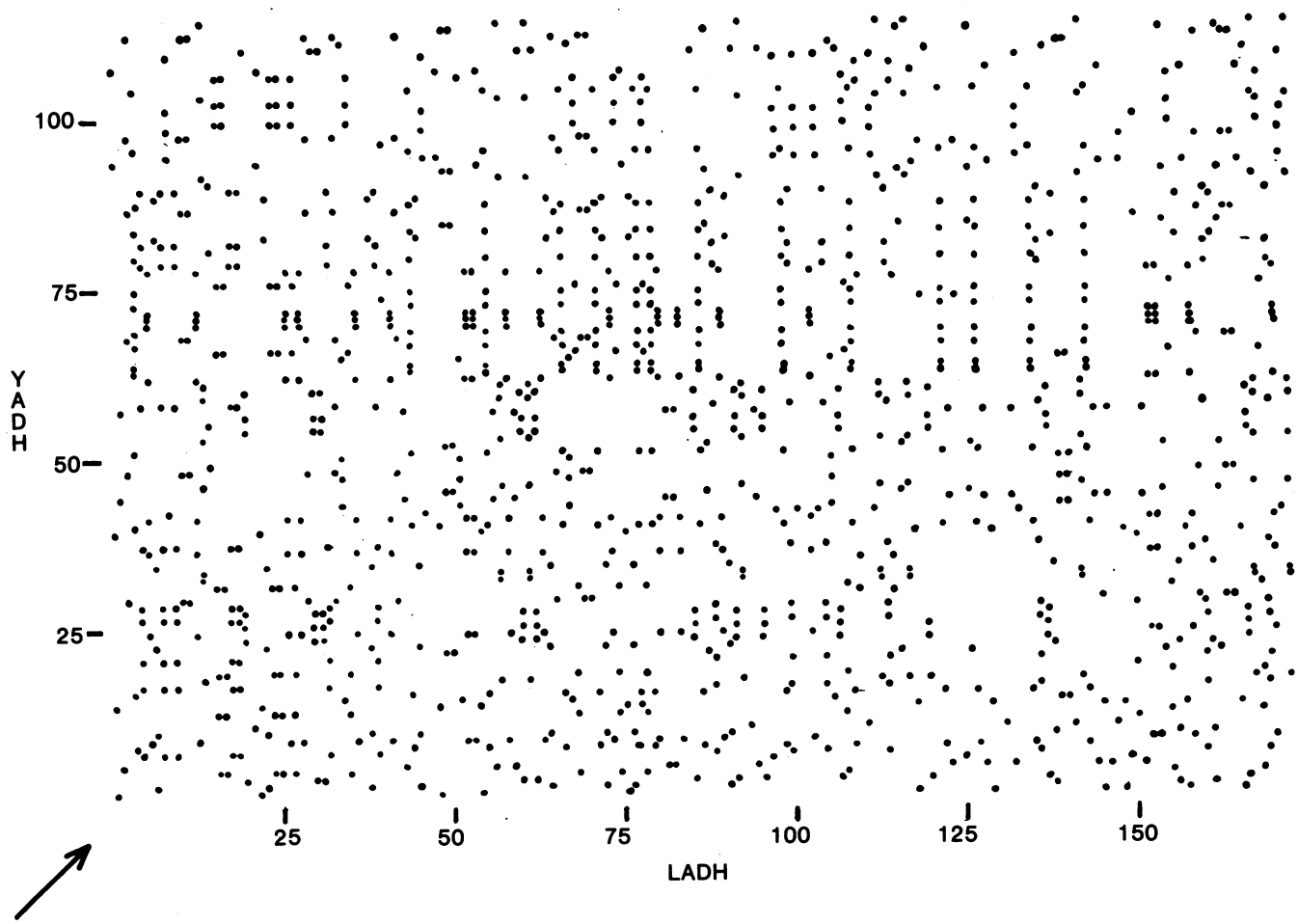

FIG. 2. Diagonal comparison between YADH (first 116-residue fragment in Fig. 1) and LADH [N-terminal 170-residue fragment (28)]. *Arrow* indicates diagonal line, most clearly seen in the middle of the figure and with line of sight closely parallel to the diagonal.

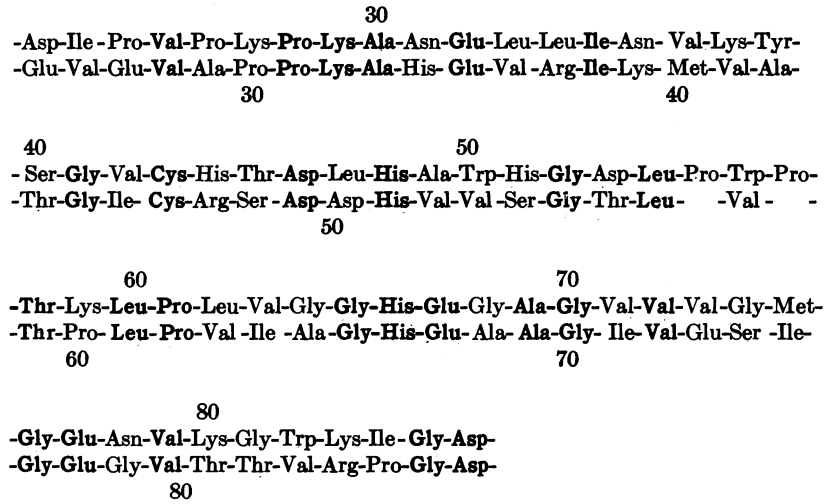

FIG. 3. Comparison of sequences of YADH and LADH at positions of greatest similarity in Fig. 2. YADH *above* (residues 22–86 in Fig. 1) and LADH *below* [residues 25–87 (28)]. 65 positions were compared. Identical residues at 26 positions (40%); with conservative exchanges (Asp/Glu, Lys/Arg,Thr/Ser, Val/Leu/Ile) included: similar residues at 36 positions (55%). Differences (except for two gaps) compatible with one-base exchanges in the genetic code at 23 positions (35%), with two-base exchanges at 14 positions (22%), with three-base exchanges at no positions.

The sequences corresponding to the diagonal seen in Fig. 2 are shown in Fig. 3. The previously noticed similarities (7) around the essential cysteine residues (residue 46 in LADH) are within this region. The identity of 40% is supplemented by several conservative amono-acid exchanges and an excess of exchanges compatible with one-base differences in the genetic code. No exchanges require three-base differences (Fig. 3). It is thus clear that these regions of YADH and LADH are both functionally and genetically related, and it can be concluded that they have a common evolutionary origin. Subsequent divergence, however, has been great. YADH and LADH differ in quaternary structure and subunit size and are identical only to about 40% in the area of similarity. This may be compared with the identity between yeast and pig GPDH, which is 68% (10) over the total subunits. Comparatively rapid evolutionary changes in alcohol dehydrogenases are thus indicated, in agreement with other results (31, 32).

Individual residues are also worth considering. Glycine is the most conserved one (12 of 17 are identical in Fig. 3), supporting conclusions (33) that glycine is of particular importance in the tertiary structure of proteins. All tryptophans in Fig. 3 (residues 50, 56, and 82 in YADH) are replaced by valines in the homologous enzyme. The two remaining tryptophans in the structures compared in Fig. 2 are replaced by tyrosine and phenylalanine. A specific role of tryptophan (34) in these dehydrogenases, except as a large hydrophobic residue, is therefore unlikely, in agreement with other results (28). Lysine and arginine are not conserved, and tryptic fragmentations of the two proteins, therefore, yield completely different pictures. Of the 12 tryptic peptides covering the regions of the proteins shown in Fig. 3, not a single corresponding pair is similar in size. Peptide mapping does not, therefore, reveal any similarities, not even in the regions of detectable homology. Due to the distant relationship, homology is of no direct help in overlapping remaining segments of YADH. Three of the short "overlaps" in Fig. 1 are, however, supported (Fig. 3).

The existence of a region in alcohol dehydrogenases with a homology more conserved than in the remainder of the molecules indicates restricted evolutionary changes in this segment. It may, therefore, serve an important and common function, such as forming part of the active site and coenzyme binding structures. Such an interpretation is favored by the fact that the reactive SH-group, suggested to be at the active site (7, 8) in YADH and LADH, is inside the region of distinguishable homology. The accompanying paper (15) also demonstrates that the repeating β-stretches, forming the coenzyme binding site and reminiscent of similar regions in other dehydrogenases, are indeed formed by those residues in LADH that yield the visible line in Fig. 2. Conclusions from work on primary and tertiary structures thus support each other.

The conserved region in LADH contains the "functional" (35, 36) zinc atom (15), which may therefore also occur in YADH. Since there is only one zinc atom per subunit of YADH (17), it would appear that the "structural" (35, 36) zinc atom is missing. Its function in the dimeric LADH may, presumably, therefore be replaced by the tetrameric structure of YADH.

From the comparison between YADH and LADH, the degree of possible similarity with other dehydrogenases may also be estimated. Evolutionary changes are not equal in different classes of dehydrogenases, as indicated by the species variations in GPDH (10) compared to YADH/LADH. Even within the conserved region of these alcohol dehydrogenases, identities are below 50% and deletions/insertions occur. Hence, the relationship between different groups of dehydrogenases may be difficult to discern in present structures and the distant similarity between GPDH and LADH in the N-terminal regions (11) seems reasonable. The tracing of evolutionary lines in greater detail needs further work. At present, however, primary and tertiary structures establish ancestral connections between certain dehydrogenases and show restrictions on evolutionary changes by the coenzyme binding.

The skilful technical assistance of Miss Ella Johnson-Finne is gratefully acknowledged. This work was supported by a grant from the Swedish Medical Research Council (Project no. 13X-3532).

1. Fondy, T. P., Ross, C. R. & Sollohub, S. J. (1969) *J. Biol. Chem.* **244**, 1631–1644.
2. Holland, M. J. & Westhead, E. W. (1969) *Fed. Proc.* **28**, 728.
3. Fondy, T. P. & Holohan, P. D. (1971) *J. Theor. Biol.* **31**, 229–244.
4. Fondy, T. P., Everse, J., Driscoll, G. A., Castillo, F., Stolzenbach, F. E. & Kaplan, N. O. (1965) *J. Biol. Chem.* **240**, 4219–4234.
5. Holbrook, J. J., Pfleiderer, G., Mella, K., Volz, M., Leskowac, W. & Jeckel, R. (1967) *Eur. J. Biochem.* **1**, 476–481.
6. Harris, I., Meriwether, B. P. & Park, J. H. (1963) *Nature* **198**, 154–157.
7. Harris, I. (1964) *Nature* **203**, 30–34.
8. Li, T.-K. & Vallee, B. L. (1964) *Biochemistry* **3**, 869–873.
9. Smith, E. L., Landon, M., Piszkiewicz, D., Brattin, W. J., Langley, T. J. & Melamed, M. D. (1970) *Proc. Nat. Acad. Sci. USA* **67**, 724–730.
10. Jones, G. M. T. & Harris, J. I. (1972) *FEBS Lett.* **22**, 185–189.
11. Jörnvall, H. (1970) *Eur. J. Biochem.* **14**, 521–534.
12. Engel, P. C., (1973) *Nature* **241**, 118–120.
13. Rossmann, M. G., Adams, M. J., Buehner, M., Ford, G. C., Hackert, M. L., Lentz, P. J., Jr., McPherson, A., Jr., Schevitz, R. W. & Smiley, I. E. (1971) *Cold Spring Harbor Symp. Quant. Biol.* **36**, 179–191.
14. Hill, E., Tsernoglou, D., Webb, L. & Banaszak, L. J. (1972) *J. Mol. Biol.* **72**, 577–591.
15. Brändén, C.-I., Eklund, H., Nordström, B., Boiwe, T., Söderlund, G., Zeppezauer, E., Olsson, I. & Åkeson, Å. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 2439–2442.
16. Buehner, M., Ford, G. C., Moras, D., Olsen, K. W. & Rossmann, M. G. (1973) *Proc. Nat. Acad. Sci. USA* **70**, in press.
17. Vallee, B. L. & Hoch, F. L. (1955) *Proc. Nat. Acad. Sci. USA* **41**, 327–338.
18. Vallee, B. L. & Hoch, F. L. (1957) *J. Biol. Chem.* **225**, 185–195.
19. Barkman, D. E., Sandstead, H. H. & Park, J. H. (1970) *J. Biol. Chem.* **245**, 1036–1040.
20. Harrington, W. F. & Karr, G. M. (1965) *J. Mol. Biol.* **13**, 885–893.
21. Jörnvall, H. & Harris, J. I. (1970) *Eur. J. Biochem.* **13**, 565–576.
22. Butler, P. J. G., Harris, J. I., Hartley, B. S. & Leberman, R. (1967) *Biochem. J.* **103**, 78P–79P; (1969) *Biochem. J.* **112**, 679–689.
23. Gray, W. R. & Hartley, B. S. (1963) *Biochem. J.* **89**, 379–380.
24. Gray, W. R. (1967) *Methods Enzymol.* **11**, 469–475.
25. Butler, P. J. G. (1967) *Studies on the Structure of Yeast Alcohol Dehydrogenase*, (PhD Dissertation), Cambridge, England.
26. Butler, P. J. G. & Harris, J. I. (1968) *FEBS Abstr.* **5**, 186.
27. Davidson, B. E., Sajgò, M., Noller, H. F. & Harris, J. I. (1967) *Nature* **216**, 1181–1185.
28. Jörnvall, H. (1970) *Eur. J. Biochem.* **16**, 25–40.
29. Allison, W. S., Admiraal, J. & Kaplan, N. O. (1969) *J. Biol. Chem.* **244**, 4743–4749.
30. Gibbs, A. J. & McIntyre, G. A. (1970) *Eur. J. Biochem.* **16**, 1–11.
31. Jörnvall, H. (1970) *Eur. J. Biochem.* **16**, 41–49.
32. Jörnvall, H. & Pietruszko, R. (1972) *Eur. J. Biochem.* **25**, 283–290.
33. Smith, E. L. & Margoliash, E. (1964) *Fed. Proc.* **23**, 1243–1247.
34. Schellenberg, K. A. (1965) *J. Biol. Chem.* **240**, 1165–1169.
35. Drum, D. E., Harrison, J. H., IV, Li, T.-K., Bethune, J. L. & Vallee, B. L. (1967) *Proc. Nat. Acad. Sci. USA* **57**, 1434–1440.
36. Drum, D. E., Li, T.-K. & Vallee, B. L. (1969) *Biochemistry* **8**, 3792–3797.