Evolutionary Relationships of Vertebrate Lactate Dehydrogenase Isozymes A₄ (Muscle), B₄ (Heart), and C₄ (Testis)*

(Received for publication, May 19, 1982)

Steven S.-L. Li‡§, Walter M. Fitch¶, Yu-Ching E. Pan†, and Farida S. Sharief‡

From the ‡Laboratory of Genetics, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709 and ¶Department of Physiological Chemistry, University of Wisconsin, Madison, Wisconsin 53706

The amino acid sequence variation among seven lactate dehydrogenase isozymes from dogfish muscle, chicken muscle and heart, pig muscle and heart, and mouse and rat testes were compared with respect to the whole lactate dehydrogenase polypeptide chain as well as their four functional domains. The coenzyme-binding domain is more conserved than the substrate-binding domain. The sequence of the loop and helix aD region of testicular LDH-C₄ isozymes is very different from those of somatic LDH-A₄ and LDH-B₄ isozymes, while the NH₂-terminal arm is extremely variable. The most parsimonious phylogenetic tree among these seven vertebrate lactate dehydrogenase isozymes clearly indicates that the LDH-A₄ and LDH-B₄ isozymes are more closely related to each other than either to the LDH-C₄ isozymes.

In mammals and birds the LDH-A₄, LDH-B₄, and LDH-C₄ subunits are encoded by Ldh-a, Ldh-b, and Ldh-c genes, respectively (1). The three homotetrameric isozymes LDH-A₄, LDH-B₄, and LDH-C₄ possess distinct physical, catalytic, and immunological properties (2–8). It has been proposed that the Ldh-c gene was derived from the Ldh-b gene by duplication and subsequent mutations (1). In this study, the LDH-C₄ sequences failed to show a closer evolutionary relationship to the LDH-B₄ than to the LDH-A₄ sequences. In fact, the LDH-A₄ and LDH-B₄ subunits are more closely related to each other than either to the LDH-C₄ subunit.

EXPERIMENTAL PROCEDURES

Protein Sequence Comparison—The amino acid sequences of dogfish LDH-A₄, pig LDH-A₄, and LDH-B₄, and chicken LDH-A₄, and LDH-B₄ isozymes were determined previously (5, 9, 10), although the sequences for the chicken lactate dehydrogenase isozymes are incomplete. The primary structures of the LDH-C₄ isozymes from mouse and rat were established in this investigation (11). The lactate dehydrogenase polypeptide chains were divided into four functional domains: the NH₂-terminal arm (residues 1–20), coenzyme-binding domain (residues 21–95 and 118–163), loop and helix aD region (residues 96–117), and substrate-binding domain (residues 164–333), as described in the preceding paper (12). The amino acid sequences of these four regions as well as the total 333 residues are compared among the 7 lactate dehydrogenase isozymes.

Phylogenetic Tree Construction—The evolutionary relationship among these 7 vertebrate LDH-A₄, LDH-B₄, and LDH-C₄ isozymes was analyzed by the methods of Fitch et al. (13–16). The computer program used to construct the phylogenetic tree is based on the minimum number of assumptions necessary to explain the observed amino acid differences. The tree that requires the fewest nucleotide substitutions is called the most parsimonious phylogenetic tree. There are 945 possible unrooted tree topologies and all of them were examined.

RESULTS

Comparison of Lactate Dehydrogenase Sequences—The amino acid sequences of dogfish LDH-A₄, pig LDH-A₄, and LDH-B₄, and chicken LDH-A₄, and LDH-B₄, and mouse and rat LDH-C₄ isozymes are aligned in Fig. 1. The results of pairwise comparison for the four domains as well as for the whole lactate dehydrogenase polypeptide chain among 7 lactate dehydrogenase sequences are shown in Table I. The NH₂-terminal arm (residues 1–20) appears to be most variable; more than 60% of 26 amino acids are different among muscle, heart, and testis types of lactate dehydrogenase isozymes. The coenzyme-binding domain (residues 21–95 and 118–163) is more conserved than the substrate-binding domain (residues 164–333). The loop region (residues 96–117) is most conserved among somatic lactate dehydrogenase isozymes, while the loop sequence of testicular LDH-C₄ is much more variable than those of the coenzyme- and substrate-binding domains. As to the total 333 residues of lactate dehydrogenase subunits, there are fewer differences between the same type of isozymes in different species than between different types of isozymes within a given species. This result is consistent with the immunological properties of lactate dehydrogenase isozymes, that is, antibodies prepared against a given type of isozyme will cross-react with the same type of isozyme from different species but not with different types of isozymes from the same species (5–8).

Phylogenetic Relationship of Lactate Dehydrogenase Isozymes—The most parsimonious phylogenetic tree obtained from each of the four different structure-function domains of the lactate dehydrogenase polypeptides is congruent with that obtained from all 333 residues combined and thus only the most parsimonious tree of whole lactate dehydrogenase sequences is presented in Fig. 2. A minimum of 566 nucleotide substitutions was required to account for the observed differences among these 7 lactate dehydrogenase sequences. All 945 possible unrooted trees for 7 sequences were examined. The two next best trees required

*The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom all correspondence should be addressed.
¶ Recipient of Grant DEB84-14197 from the National Science Foundation.

The abbreviations used are: LDH-A₄, tetrameric lactate dehydrogenase A₄, isozyme from skeletal muscle; LDH-B₄, tetrameric lactate dehydrogenase B₄, isozyme from heart muscle; LDH-C₄, tetrameric lactate dehydrogenase C₄, isozyme from testis.
Evolution of Lactate Dehydrogenase Isozymes

Fig. 1. Amino acid sequences of dogfish LDH-A, chicken LDH-A, and LDH-B, pig LDH-A, and LDH-B, and mouse and rat LDH-C isozymes. The scientific names of animal species are as follow: dogfish, *Squalus acanthias*; chicken, *Gallus domesticus*; pig, *Sus scrofa*; mouse, *Mus musculus*; rat, *Rattus norvegicus*. The residues are numbered from 1 to 333, and these residue numbers are different from those of x-ray numbers for dogfish LDH-A and the chemical numbers of LDH-C subunits (5, 11). * signifies the absence (deletion) of amino acid(s), whereas X means the unknown residues. The identical residues of all 7 sequences at 159 positions are indicated by boxes.
Table 1
Sequence comparison of seven vertebrate lactate dehydrogenase isozymes

DA, CA, PA, CB, PB, MC, and RC signify the lactate dehydrogenase sequences from dogfish A, (muscle), chicken A, (muscle), pig A, (muscle), chicken B, (heart), pig B, (heart), mouse C, (testis), and rat C, (testis), respectively. The absence (deletion) of amino acids was counted as residue difference. The numbers in the upper right of the matrix are the number of amino acid differences along with the total number of residues compared; the numbers in the lower left of the matrix are the percentage of residue differences.

<table>
<thead>
<tr>
<th></th>
<th>DA</th>
<th>CA</th>
<th>PA</th>
<th>CB</th>
<th>PB</th>
<th>MC</th>
<th>RC</th>
<th>Residues compared</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>NH₂-terminal arm (residues 1-20)</td>
<td>15</td>
<td>12</td>
<td>12</td>
<td>14</td>
<td>13</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17</td>
<td>19</td>
<td>22</td>
<td>27</td>
<td>33</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
<td>22</td>
<td>27</td>
<td>33</td>
<td>121</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
<td>22</td>
<td>27</td>
<td>33</td>
<td>121</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>Coenzyme-binding (residues 21-95 and 118-163)</td>
<td>34</td>
<td>34</td>
<td>35</td>
<td>35</td>
<td>34</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>Loop and helix aD region (residues 96-117)</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>Substrate-binding (residues 164-333)</td>
<td>34</td>
<td>34</td>
<td>35</td>
<td>35</td>
<td>34</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.</td>
<td>Lactate dehydrogenase polypeptide (residues 1-333)</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Phylogenetic tree of seven vertebrate lactate dehydrogenase isozymes. The phylogenetic relationship among seven lactate dehydrogenase isozymes, namely, dogfish A, (muscle), chicken A, (muscle), pig A, (muscle), chicken B, (heart), pig B, (heart), mouse C, (testis), and rat C, (testis) is indicated by the most parsimonious tree. The number on each leg is the number of nucleotide substitutions required to account for the descent from the ancestor, and the number underlined is the average of its immediate descendants. The number of substitutions in any one interval is not an integer because of averaging over all equally most parsimonious solutions for that topology. The total number of nucleotide substitutions is 396. Count does not include insertions or deletions. The root is arbitrarily placed half-way between the two most distantly related groups.

Fig. 3. Comparison of NH₂-terminal sequences of dehydrogenases. A signifies the deletion of amino acid. Both soluble malate dehydrogenase isozyme (s-MDH) and malate dehydrogenase isozyme from mitochondria (m-MDH) from pig have two residues instead of one residue between the two most conserved glycines at positions 27 and 29. The first three residues of pig soluble malate dehydrogenase isozyme have not yet been sequenced. The arrangement of structure-function domains of horse liver alcohol dehydrogenase (ADH) is very different from those of other dehydrogenases, and its coenzyme-binding domain is reported to begin at residue 176. GADPH, glyceraldehyde-3-phosphate dehydrogenase.

Discussion
The extent of amino acid sequence variation analyzed in this study appears to be quite different for the four domains of lactate dehydrogenase: NH₂-terminal arm, coenzyme-binding domain, loop and helix aD region, and substrate-binding domain. The amino acid sequence of the coenzyme-binding domain is the most conserved of these 4 domains, and the structural similarity of this domain has been shown for lactate dehydrogenase, soluble malate dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase, and liver alcohol dehydrogenase. The catalytic domain of lactate dehydrogenase is more variable than the coenzyme-binding domain, and different dehydrogenases possess different catalytic domains.

The amino acid sequences of the loop and helix aD region of sonic LDH-A, and LDH-B, isozymes are very conserved, while a very different sequence of the loop and helix aD region was found in the testicular LDH-C, isozymes. The loop
and helix dD region of lactate dehydrogenase which is also present in malate dehydrogenase but absent in glyceraldehyde-3-phosphate dehydrogenase and alcohol dehydrogenase undergoes the largest structural alteration from apoenzyme to ternary complex with coenzyme and substrate, and this region had been implicated in the mechanism of action of lactate dehydrogenase catalysis (5, 19). Therefore, the very different amino acid sequence of the loop and helix CD region of testicular LDH-C4 isozymes might be one of the causes of their unique catalytic properties. The NH2-terminal 20 residues of lactate dehydrogenase isozymes are extremely variable, and their primary function apparently is to stabilize the quaternary structure of tetrameric lactate dehydrogenase through their interaction with the COOH-terminal region of the other subunits (19-21). However, no cooperative interaction between subunits of tetrameric lactate dehydrogenase has been observed (3, 4). Furthermore, the NH2-terminal arm of vertebrate lactate dehydrogenase is absent in malate dehydrogenase (24), glyceraldehyde-3-phosphate dehydrogenase (25), Drosophila alcohol dehydrogenase (29), and even microbial lactate dehydrogenase (27), as shown in Fig. 3. Thus, the functional significance of this NH2-terminal arm of tetrameric lactate dehydrogenase in vertebrates is not clear. In short, the complete structure of lactate dehydrogenase isozymes is composed of four domains and it has been suggested that the structure of dehydrogenases including lactate dehydrogenase has resulted from the combination of a variety of domains into a single polypeptide chain as a consequence of gene fusion (20, 21). In view of recent findings of eukaryotic gene organization, it would be of considerable interest if different domains of the dehydrogenases correspond to different exons, the coding regions of the dehydrogenase genes, as is the case of immunoglobulin genes (28).

There are three paralogous sets of Ldh gene sequences in these data and it is clear that the analysis separates them into their respective groups of Ldh-A, Ldh-B, and Ldh-C. Moreover, the orthologous set of LDH-A sequences yields a phylogeny congruent with the known evolutionary relationships of the three taxa. The rates of evolution shown in Fig. 2 are quite uniform although this may be at least partly a consequence of averaging over sufficiently long time intervals that larger short term changes in rate are masked. It does, however, add confidence to the treatment of peak heights as reasonably proportional to time except for the systematic bias inherent in a failure to correct for the possibility that one nucleotide difference is more likely to represent more than one nucleotide substitution the more remote their sequences are. This means that higher peaks are placed at heights that increasingly underestimate the total change. This will make the following statements conservative. The gene duplication leading to the Ldh-a and Ldh-b genes occurred at least one-third earlier than the divergence of the shark and higher vertebrates. That would place the duplication in the early Cambrian or even earlier. Thus, this separation of function would appear to go back at least to the earlier chordates. The gene duplication leading to the Ldh-c gene is even older, certainly being older than 750 million years, i.e., late Precambrian, or more closely, late Proterozoic. This would suggest the possibility that the first Ldh gene duplication originated in an invertebrate organism. An LDH-C4 subunit had been reported to be present in the eye and liver of advanced fishes and many tissues of primitive fishes, although the LDH-C4 isozyme has not been found in amphibians and reptiles (1). Further, the tryptic peptide maps, peptide compositions, and partial sequences of mouse LDH-A and LDH-B subunits were found to be more closely related to each other than the LDH-C subunit of the same species (2). Therefore, the order of the two duplication events leading to Ldh-a, Ldh-b, and Ldh-c genes is clear and rules out an earlier hypothesis that the Ldh-c gene derived from Ldh-b gene (1).

Acknowledgments—We thank Drs. C. Aquadro, F. B. Armstrong, J. H. Harrison, J. L. Irvin, G. B. Kitto, A. F. Ringe, and S. Slaughter for critical reading of the manuscript. Drs. B. H. Judd and H. V. Malling for their interest and support in this project, Dr. R. L. niece for computer analysis, and Leslie Gardner for her excellent help in preparing the manuscript.

REFERENCES


Downloaded from http://www.jbc.org/ by guest on November 12, 2017
Evolutionary relationships of vertebrate lactate dehydrogenase isozymes A4 (muscle), B4 (heart), and C4 (testis).

S S Li, W M Fitch, Y C Pan and F S Sharief


Access the most updated version of this article at http://www.jbc.org/content/258/11/7029

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/258/11/7029.full.html#ref-list-1