Biochemical and Biophysical Research Communications 396 (2010) 125–130

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Superfamilies SDR and MDR: From early ancestry to present forms. Emergence of three lines, a Zn-metalloenzyme, and distinct variabilities

Hans Jörnvall [a,*], Joel Hedlund [b], Tomas Bergman [a], Udo Oppermann [c,1], Bengt Persson [b,2]

[a] Department of Medical Biochemistry and Biophysics (MBB), Karolinska Institutet, SE-171 77 Stockholm, Sweden
[b] IFM Bioinformatics, Linköping University, SE-581 83 Linköping, Sweden
[c] Structural Genomics Consortium, University of Oxford, OX3 7LD, UK

ARTICLE INFO

Article history:
Received 11 March 2010

Keywords:
Dehydrogenases
Reductases
Molecular architecture
Molecular evolution
Zinc metalloproteins
Domain variabilities

ABSTRACT

Two large gene and protein superfamilies, SDR and MDR (short- and medium-chain dehydrogenases/reductases), were originally defined from analysis of alcohol and polyol dehydrogenases. The superfamilies contain minimally 82 and 25 genes, respectively, in humans, minimally 324 and 86 enzyme families when known lines in other organisms are also included, and over 47,000 and 15,000 variants in existing sequence data bank entries. SDR enzymes have one-domain subunits without metal and MDR two-domain subunits without or with zinc, and these three lines appear to have emerged in that order from the universal cellular ancestor. This is compatible with their molecular architectures, present multiplicity, and overall distribution in the kingdoms of life, with SDR also of viral occurrence. An MDR-zinc, when present, is often, but not always, catalytic. It appears also to have a structural role in inter-domain interactions, coenzyme binding and substrate pocket formation, as supported by domain variability ratios and ligand positions. Differences among structural and catalytic zinc ions may be relative and involve several states. Combined, the comparisons trace evolutionary properties of huge superfamilies, with partially redundant enzymes in cellular redox functions.

1. Background

Alcohol dehydrogenase (ADH) has long been the subject of molecular studies and now illustrates several evolutionary properties common to proteins in general. Major steps in this chain of results are reviewed in this chapter. They cover over 60 years of research at Karolinska Institutet since the mammalian protein was crystallized [1], and close to 45 years since it was found to be N-terminally acetylated [2]. ADH was initially selected for studies because of its dehydrogenase function requiring cofactors with resulting versatile kinetics [3]. Its zinc metalloprotein nature and medically important role in alcohol elimination also motivated many studies. It was one of the dehydrogenases early defined in primary and tertiary structure, contributing to the then growing knowledge of a coenzyme-binding domain, the now famous Rossmann-fold [4]. It occurs in many dehydrogenases, kinases and other [5,6] proteins. Two superfamilies were found and defined around ADH activities [7] and are now known as MDR and SDR, for Medium-chain and Short-chain Dehydrogenases/Reductases, respectively. Many years of work then established the family members, showed a series of MDR gene duplications during vertebrate radiation, enzymogenesis of novel functions, an origin of MDR-ADH from GSH-linked formaldehyde dehydrogenase, and distinctive evolutionary patterns for MDR [8,9] and SDR [10] relationships.

From the "explosion" of organisms and forms since characterized in many genome projects, the SDR and MDR superfamilies are now known in a multitude of forms. They include about 82 and 25 genes, respectively, in the human, manifold these numbers when also other lines of life are included, and several tens of thousands of variant forms in the sequence data banks [11,12]. From this multitude of data and from observations on occurrences, multiplicities and variabilities, it is now possible to link previous observations with novel conclusions. Combined, they illustrate wide ranges of molecular evolution, including questions on ancestral origins, zinc metalloenzyme properties, and complex functional aspects, as now outlined.

2. The MDR and SDR superfamilies

The huge group of proteins (and their genes) with related redox coenzyme-binding domains as parts of their structures has a dis-

* Corresponding author. Fax: +46 8 337462.
E-mail address: Hans.Jornvall@ki.se (H. Jörnvall).
[1] Present address: Nuffield Department of Orthopedics, Rheumatology and Musculoskeletal Sciences, NIHR Biomedical Research Unit, University of Oxford, OX3 7LD, UK.
[2] Additional address: Department of Cell and Molecular Biology, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

0006-291X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2010.03.094

tant origin. The group is subdivided into superfamilies, also of ancient origin, which represent recombinatorial events putting together different building units into separate orders, numbers, sizes and arrangements, forming the basis for further evolution. In the case of ADHs, at least four superfamilies of different basic solutions are known. They share similarities in the Rossmann-fold of the coenzyme-binding domains, and also share some end substrates ("alcohol" and other molecules) but have different overall molecular architectures and different catalytic mechanisms. Two of these superfamilies containing ADHs are well-characterized, MDR and SDR, now representing many enzyme activities.

The general architectures, relationships, structures and family properties of the MDR and SDR superfamilies were recently reviewed [11–14]. In short, the MDR proteins typically have two-domain subunits, with the coenzyme-binding domain and its building units in a consecutive fold covering much of the C-terminal half. The other domain, the "catalytic" domain, presumably derived from GroES chaperon building elements [15], is typically co-folded from different parts, including much of the N-terminal half and a C-terminal segment. In contrast, the SDR proteins have a simpler molecular architecture, with largely one-domain subunits, derived from just the coenzyme-binding fold, with more or less extensions.

Both these superfamilies occur in all life forms. This wide-spread, homogeneous nature of the families in all kingdoms of life and the recombinatorial formation of both subunit types from building blocks, suggest that they are remnants from horizontal gene transfers. This would apply to the communal pool of the last universal cellular ancestor and would allow the genes then to emerge with the kingdoms [16] for further Darwinian evolution. This conclusion is supported by recent data from random genome screenings of microorganisms and viruses from sea water, where SDR genes are among the genes most abundant [17]. Hence, it may be concluded that SDR genes are of extreme age, and emerged early from the communal pool during cellular evolution [16]. This conclusion is compatible also with the fact that SDR, and not MDR, has the simplest structure, with just one domain and building block, and the most wide-spread occurrence in nature [11], being one of the early characteristics of life [17] at the stage when viral gene transfer appears to have been a common feature [18]. The viral gene transfers and the fact that these enzymes have ribonucleotide cofactors, can also be taken to reflect even earlier origins, from the stages of ancestral RNA worlds [19].

MDR, too, fits into an early formation pattern, but appears to be of a later emergence from spillover at the communal pool stage, compatible with its more complex architecture with different building units, and its less abundant spread in life forms today. Furthermore, the MDR superfamily has a special pattern of two types, many being either metalloenzymes containing zinc, or just non-metalloproteins without zinc [20]. Interestingly, the zinc-free MDR families are more common in prokaryotes and with structural features favouring NADPH binding, and hence reductase functions, while the zinc metalloenzyme MDR families are more common in eukaryotes and with structural features favouring NAD binding and hence primarily dehydrogenase functions [20]. In conclusion, it appears as if the three protein subunit types, SDR, non-Zn-MDR and Zn-MDR, represent consecutive emergencies from the communal pool, first SDR, then metal-free MDR, and finally zinc metalloprotein-MDR (Fig. 1). All three have since diverged and are now of wide-spread occurrence, with many enzyme types. However, the family patterns still differ, with more early branchings in the SDR than MDR superfamily, as visible by the evolutionary trees of present SDR and MDR forms in mammals (Fig. 2), compatible with the earlier emergence of SDR (Fig. 1).

3. Zinc metalloenzyme nature

The 2-fold pattern of MDR enzymes, with and without zinc, forms a link with function (above) and appears also to illustrate characteristics of the metalloenzyme nature (Fig. 3). Thus, the zinc-free MDR forms appear to be older than the zinc-containing forms, are of mainly prokaryotic abundance and reductive. It even appears possible that the early recruitment of the metal coincided with the emergence of oxygen in the Earth's atmosphere, and that just zinc was then beneficial because of its inertness to redox-mediated valence changes, providing a basis for its function as a Lewis acid catalyst in the enzyme reaction mechanism. The zinc-free forms also appear, in the large family evaluations now possible from the data banks, to be associated with a domain variability different from that of the metalloprotein subunit type [20]. Thus, all MDR forms with a conservation of the coenzyme-binding domain more than 2-fold that of the catalytic domain are of the non-Zn type, and of the 10 families with the highest conservation in the coenzyme-binding domain, eight are non-Zn-MDR forms. Hence, forms with more variability in this domain (or either domain) include Zn-MDRs, almost as if zinc might replace amino acid conservation in maintaining the domain functional. All comparisons of domains, with and without zinc, with and without extreme residue variability, or with reductive or oxidative enzymatic direction, show surprisingly large numerical differences, suggesting that the presence of zinc has extensive effects [20]. Notably, the catalytic zinc, when present, is bound with three protein ligands to the catalytic domain but has a function, apart from catalysis, to also contribute to maintain the architecture for binding the coenzyme and for a conformational movement forming the substrate binding site [13], giving an ordered reaction mechanism. Thus, the catalytic zinc appears to allow for domain communications, compatible with the observation that catalytic zinc atoms frequently have one large, inter-ligand distance along the protein chain [21].

Of additional interest is that Zn-MDR proteins, like mammalian ADH, frequently have also a second zinc ion bound per subunit. This is then a "structural" zinc, differently coordinated [21] with four closely spaced protein ligands [13]. This ion center, when present, is in a variable region, absent or different in other related forms (like sorbitol DH) and appears to contribute to subunit interactions (cf. Fig. 3), influencing the quaternary structure of the protein [22]. Sequence alterations in this region allow ligand shifts, suggesting that the structural zinc atom [23], like the catalytic one, may be in a strained conformation, typical of an entatic state of the zinc [24]. Furthermore, also the active-site zinc in some MDR-ADH forms may acquire ligand shifts [25] and this is then associated with considerable zinc movement and conformational changes that affect coenzyme binding [26] and may participate in product release during the reaction cycle [27].

As stated, the major division among the MDR-enzyme types is between the presence or absence of the catalytic zinc ion. The structural zinc, when present, usually occurs in addition to the catalytic zinc, hence appears to form a secondary variable, of later origin (or loss) than the catalytic ion. Nevertheless, the data bank comparisons show that two MDR families (numbered 58 and 83 [20]) exist in which the protein has just the structural zinc ion. Hence, in special cases, either zinc ion may emerge (or be lost) during evolution. In conclusion, the combined data suggest that catalytic and structural zinc may both contribute profound effects on the protein and supply domain intercommunications, subunit interactions, or substrate bindings. In short, the zinc appears to give the protein an additional strength, as if compensating for domain variability and for evolutionary changes in the protein scaffold.

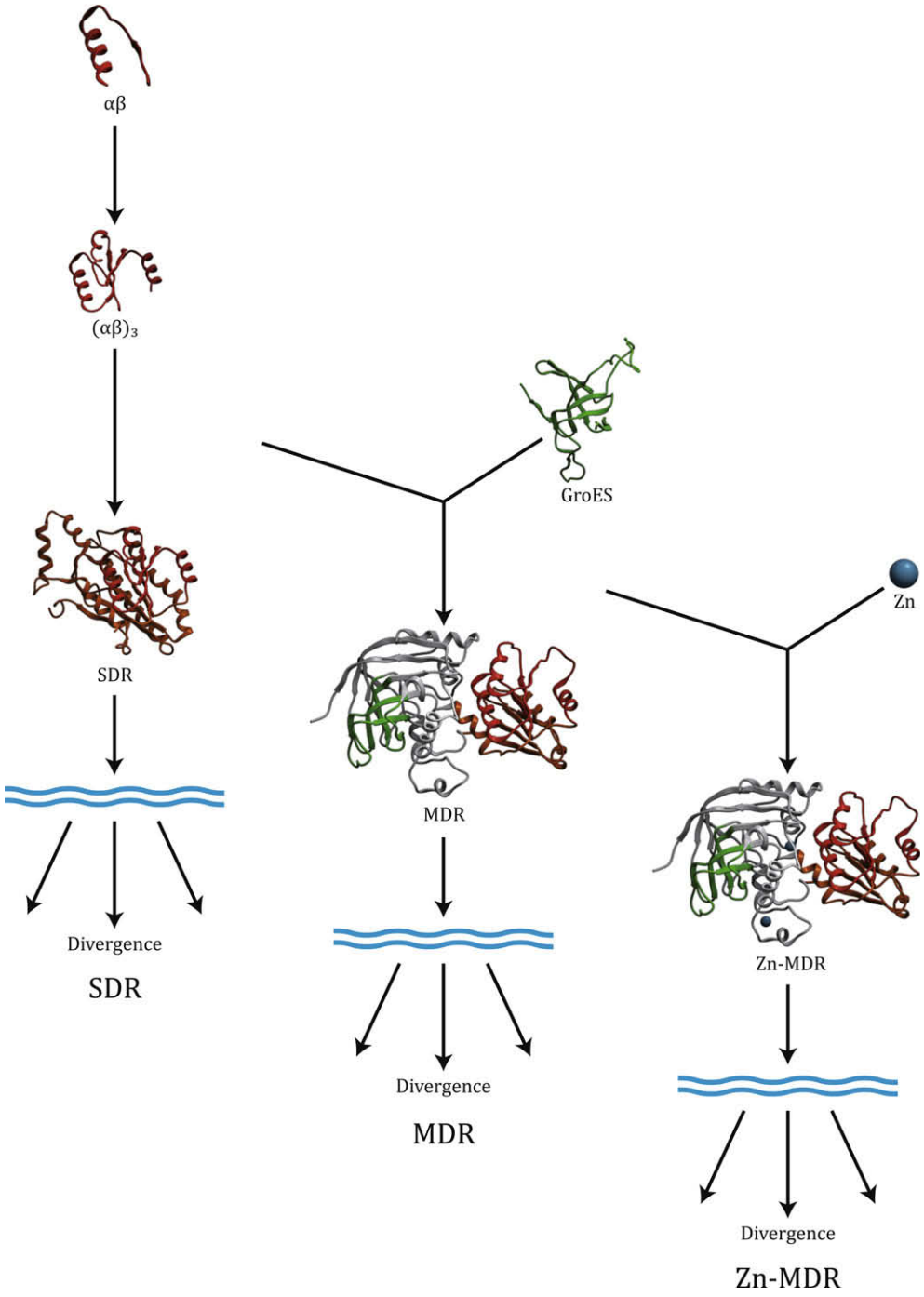

Fig. 1. Ancestral origins of SDR and MDR forms at separate stages. As given in the text, SDR has the simplest build-up and widest spread in nature, suggesting a very early origin from αβ elements to a Rossmann-fold domain in the universal cellular ancestor (wavy blue lines) for subsequent Darwinian evolution in the cells of all three kingdoms of life. MDR has a more complex build-up, is less wide-spread, and compatible with a later exit into Darwinian evolution, and still further so for Zn-MDR. Timings of the origins are still unclear, but here given as being before the last cellular ancestor although especially the third line could have later origin(s), possibly correlated with additional events (oxidative atmosphere, cf. text).

Finally, a note on the feature of the similar SDR and MDR catalytic mechanisms. When present, a catalytic zinc with a hydroxyl from dissociated water provides a ligand for deprotonation of the alcohol substrate, to allow for hydride transfer to the adjacent nicotinamide ring of the coenzyme [13]. In the metal-free SDR enzymes, a hydroxyl-tyrosinate ion stabilized by an adjacent Lys residue gives a similar acid–base catalyzed mechanism and hydride transfer to the coenzyme [11]. However, in both cases, detailed reaction mechanisms are complex and still the subject of study at increased resolution [13,14,28].

4. Family divergence

In addition to the superfamily relationships, reflecting early formations of the SDR and MDR systems, and the roles of zinc as outlined above, both superfamilies also exhibit extensive divergence,

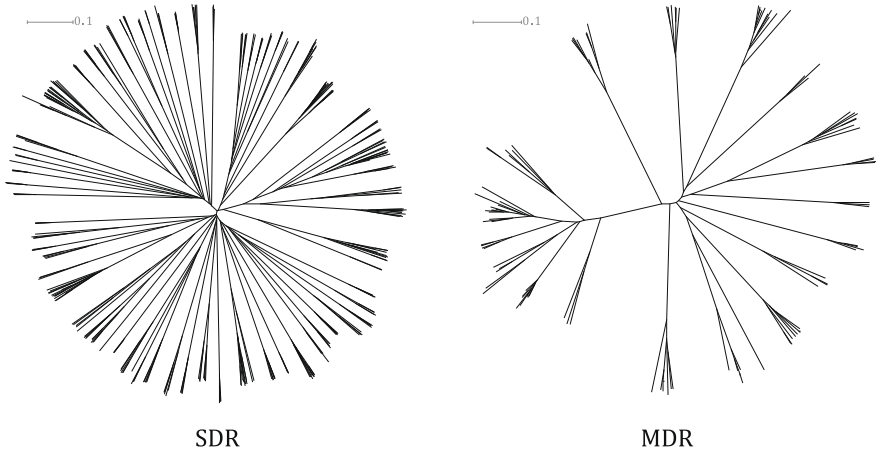

Fig. 2. Separate molecular patterns of the constituent enzyme family branches in the mammalian SDR and MDR forms. SDR has more family branches overall, much early branching, and less frequent later branching than MDR, compatible with an earlier origin of SDR and earlier lockings into defined functions. Scale bars show the cumulative horizontal branch length corresponding to 10% sequence differences.

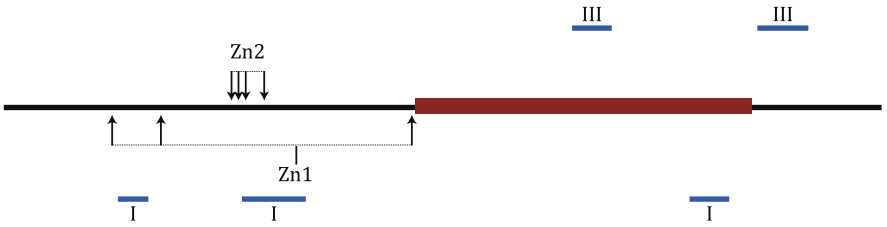

Fig. 3. Zn-ADH variability patterns. The middle, continuous line represents the subunit architecture, with a catalytic (black) and a coenzyme-binding (red) domain, and two zinc ions, one "catalytic" (Zn1) protein–liganded by three well-spread residues covering much of the catalytic domain and bordering to the second domain, and one "structural" (Zn2) liganded by four closely spaced residues well inside one domain. Blue regions indicate the most variable segments in the "variable" classes of the protein (marked I) and in the "constant" classes (marked III), clearly showing the separate variability patterns between the enzymes of classes I and III.

producing separate enzyme activities in family relationships. Two trends are relevant to consider in relation to this divergence, one regarding SDR, the other regarding MDR.

5. EC enzyme class divergence in SDR

Considering the early origin of SDR (Fig. 1), the subsequent divergence has had time to progress far. Hundreds of SDR enzyme activities and corresponding families have been detected. Keeping similar coenzyme-binding structures and active-site relationships, other variability patterns differ greatly, and at least five sub-superfamily types of protein chain have been discerned in the data bank comparisons: "classical", "extended", "intermediate", "divergent" and "complex" SDR enzymes [29]. This divergence has also evolved novel enzyme types regarding the EC classification of enzymes into six classes. Thus, most of the SDR families are still dehydrogenases or reductases of the EC 1 class (219 families recognized last year), but some lyases (EC 4 enzymes) have also evolved (26 families recognized last year), as well as some isomerases (EC 5 enzymes). The active-site Tyr residue, assisted by adjacent Lys, Asn, and Ser residues, has been found to adjust to the basic reaction mechanisms in these cases, too, still with acid–base catalysis and proton transfers [11]. Thus, SDR proteins do not only represent very distant origin, including viral representation [11,17] presumably from a time when virally-mediated lateral gene transfers commonly occurred [18], but also a wide activity type, involving half of all enzyme activity types (three of six EC classes, thus far). Few gene/protein superfamilies exhibit this great divergence.

6. Enzymogenesis and recent isozyme divergence in MDR

A more recent divergence pattern, but still of old age (~500 million years [30]), is apparent in the ADH line of the MDR superfamily and has paralleled the vertebrate radiation by a series of gene duplications (Fig. 4). The parent form has been shown to be the GSH-dependent ADH 3 type [31], which is a formaldehyde dehydrogenase [32] with the GSH adduct hydroxymethylglutathione as substrate. Notably, handling of formaldehyde is an old prerequisite, since formaldehyde is a compound in interconversions within the one-carbon pool, involving both serine and glycine relationships as well as methylation reactions. The ADH 3 type also has an inherent classical ADH activity [33]. This activity is apparently sufficient for aldehyde and alcohol elimination in sub-vertebrate ocean life. However, starting at the origin of osseous fish [8], a series of gene duplications has occurred during vertebrate radiation, allowing for creation of additional enzyme activities and leading to the classical liver ADH (the ADH 1 form) in higher vertebrates. These various ADH types are also called classes but are distinct from the EC classes above, and contain only EC 1 enzymes. In the human, we have six such ADH genes, while other species have additional genes of this type from further duplications. In addition, we (and other vertebrates) have recent sets of duplications, giving rise to traditional isozymes, also line-specific. In this manner, the human now has seven genes for ADH, all clustered in 100.0–100.4 Mb (q23) of chromosome 4 [12]. The duplications are of separate age, and reflect functional evolution, with enzymogenesis of new activities from the ancestral form of GSH-coupled formalde-

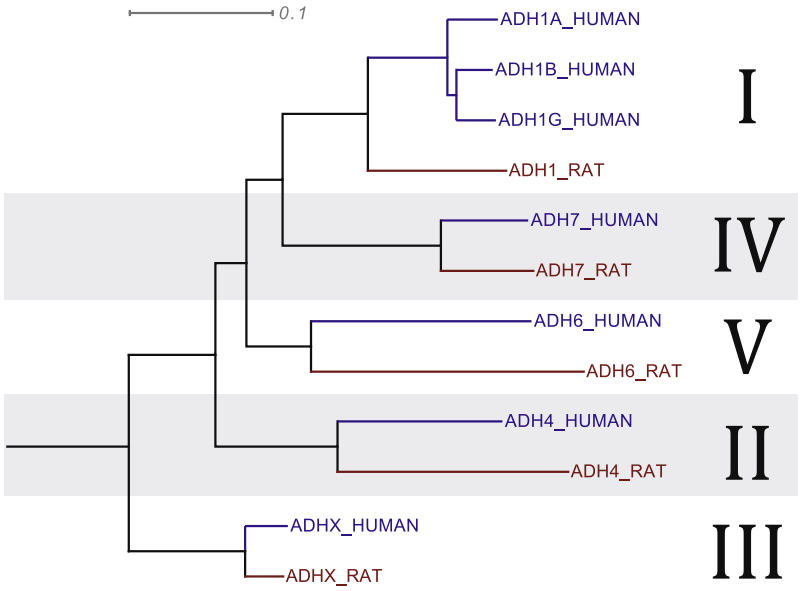

Fig. 4. Phylogenetic tree linking the gene relationships among all human ADH classes and their counterparts in the rat, showing the different origins and evolutionary speeds of the separate forms. Scale bar shows the cumulative horizontal branch length corresponding to 10% sequence differences.

hyde dehydrogenase. Expression patterns, organ distributions and molecular evolutionary properties differ among these genes [33]. In total, the duplication pattern during vertebrate evolution (Fig. 4) very much resembles the globin tree in vertebrates, where each duplication has generated a new functional variant. This has allowed for the evolution of allosteric hemoglobin to permit oxygen transport, and additional duplications to give embryonic and fetal variants for placental life at different oxygen tensions, much like we have a series of differently adjusted ADHs. Notably, one human ADH form is also of special fetal expression [33].

The ADH classes and more recent isozymes have been studied in detail and have revealed two interesting observations, one on segment variability along the protein chain [8], the other on parallel isozyme variability in many other dehydrogenases [34]. Both variations are further treated below.

7. Segment variability, with one "constant" form and one "evolving" form

Upon discovery of these multiple ADH structures, their species variabilities were traced, initially to get functional information on which residues that are important. But with results on the residue variabilities, interesting patterns were observed tracing separate lines. Thus, our seven ADH genes exhibit two types of evolutionary pattern. One, the ADH 3 gene (and protein), is comparatively constant among species and varies like for many basic metabolic enzymes (with ~6% divergence between the human and rat proteins), while the other ADHs (ADH non-3 type) overall vary about 2- to 5-fold more. This was originally unexpected but has now been shown in also other dehydrogenases (below). Thus, ADH (and several dehydrogenases) have multiple genes, where one type is fairly constant, like for fundamental enzymes in general, and another is considerably more variable [34]. The different pressure on their evolutionary conservation must reflect separate functional roles. Notably, the "constant" ADH (ADH 3) is the parent form and has a strict major enzyme activity (GSH-coupled formaldehyde DH activity), while the secondary duplications have given rise to more variable forms. Apparently, the parent form keeps the original, basic enzyme activity, while the secondary forms evolve faster and adjust to novel functions, giving rise to enzymogenesis, including of the classical liver ADH activity. Again, this is to some extent paralleled with the globin family, where one functional gene (for α-chains) keeps the original activity, while the other (β) evolves separate functions (for γ and δ type of activities).

Furthermore, segment variability within these ADH subunits differ. One form (the "constant" ADH 3 form) appears to have its most variable segments at surface positions or else at "non-functional" positions (Fig. 3), whereas the other forms (the "evolving" non-3 ADH forms) have their most variable segments at functional regions of the active-site, the subunit interactions, and the structural zinc loop [8] (Fig. 3). Thus, not only do the two structural types differ overall in evolutionary rate, but also in segment positions for the variabilities that do occur. The "constant" form is, both overall and in segments, the parent form, while the "variant" form in both senses is the secondary, duplicatory product. Finally, although with less known support, the two evolutionary types appear to differ in duplicatory tendency, the "constant" type not forming many isozymes by further duplications. In contrast, the "variable" form giving rise to more isozyme duplications and more allelic variability, at least in the human. Hence the differences appear fairly extensive between the "constant" and the "variant" type, affecting both local segments and the entire conservation, even possibly correlating with duplicatory tendencies. The two forms keep their considerably different evolutionary rates over extensive time periods.

8. Parallel evolutionary patterns in many dehydrogenases

With the above patterns noticed, we also find similar two-type constant and variable enzyme forms in several other dehydrogenases, including aldehyde dehydrogenase [33] (of another structural family). In fact, some dehydrogenases are now known to be highly multiple, especially ADH and aldehyde DH. The latter with no less than 18 genes in the human, and with a duplicatory variability apparently larger than for ADH in the sense that aldehyde dehydrogenase genes now are spread over many chromosomes in the human, while all human ADH genes are in one locus. Combined, these observations suggest that ADH and several other DHs have important functional roles. They appear to be essential both to maintain their conserved and, in parallel, to evolve further. This

multiplicity explains why ADH single-gene knock-out modified animals show few phenotypic differences, because of the many functionally redundant forms. Overall, though, such redundancy suggests vital functions for ADH and several DHs.

Finally, some functional conclusions as to the physiological roles of these enzymes can be drawn. The overall conclusion is that we need ADHs, aldehyde dehydrogenases and other dehydrogenases for protection against a variety of toxic alcohols, aldehydes and other compounds, more or less like we need (and have) multiple cytochrome P450s in detoxification reactions [35]. Functional mutants, by absence or knock-outs, in both the P450 case and the DH cases may have limited immediate effects but give long-term reduced capacity, eventually leading to toxicity and presumably loss of cell functions, and hence cancer. Notably, although P450 and ADH can convert similar substrates, they differ in long-term, with P450 forms having high capacity but DH forms giving less dangerous products, with no reactive oxygen species, and only mild DH reactions. At the same time, the constant evolution of variable forms opens for generation of novel activities and thus novel functions. Two such "novel" and potentially interesting functions regarding ADH are, on the one hand, the role of both SDR and MDR enzymes in retinoid metabolism [33,36], with much influence on regulatory functions [33], and on the other hand, the fairly recently found ADH function in chemical elimination of the NO molecule, like ADH 3 does, as recently reviewed [36]. Hence, continued studies of the "old" ADH enzyme functions are still of importance and give unexpected results. In conclusion, the ADH activity represents a "living" and evolving gene "battery", can now be traced back to an early vertebrate origin, and further back to ancestral relationships at the build-up of cellular life and the start of divergent evolution.

Acknowledgments

We are grateful to many students passing through the MDR/SDR field in the parent laboratory at MBB, Karolinska Institutet, and to several technical assistants for much help during many years, in particular Ella Cederlund and Carina Palmberg. We are also grateful to Karolinska Institutet, Linköping University, the Swedish Research Council, the Knut and Alice Wallenberg Foundation for much research support. The Structural Genomics Consortium is a registered charity (No. 1097737). Its funding details can be found at http://www.thesgc.org/.

References

[1] R.K. Bonnichsen, A.M. Wassén, Crystalline alcohol dehydrogenase from horse liver, Arch. Biochem. Biophys. 18 (1948) 361–363.
[2] H. Jörnvall, Acetyl group determinations in horse liver alcohol dehydrogenase and in horse myoglobin, Acta Chem. Scand. 21 (1967) 1805–1811.
[3] H. Theorell, B. Chance, Studies of alcohol dehydrogenase. II. The kinetics of the compound of horse liver alcohol dehydrogenase and reduced diphosphopyridine nucleotide, Acta Chem. Scand. 5 (1951) 1127–1144.
[4] M.G. Rossmann, D. Moras, K.W. Olsen, Chemical and biological evolution of nucleotide-binding protein, Nature 250 (1974) 194–199.
[5] C.L. Drennan, S. Huang, J.T. Drummond, R.G. Matthews, M.L. Lidwig, How a protein binds B12: a X-ray structure of B12-binding domains of methionine synthase, Science 266 (1994) 1669–1674.
[6] L. Aravind, L.M. Iyer, Provenance of SET-domain histone methyltransferases through duplication of a simple structural unit, Cell Cycle 2 (2003) 369–376.
[7] H. Jörnvall, M. Persson, J. Jeffery, Alcohol and polyol dehydrogenases are both divided into two protein types, and structural properties cross-relate the different enzyme activities within each type, Proc. Natl. Acad. Sci. USA 78 (1981) 4226–4230.
[8] O. Danielsson, S. Atrian, T. Luque, L. Hjelmquist, R. Gonzalez-Duarte, H. Jörnvall, Fundamental molecular differences between alcohol dehydrogenase classes, Proc. Natl. Acad. Sci. USA 91 (1994) 4980–4984.
[9] B. Persson, J.S. Zigler Jr., H. Jörnvall, A super-family of medium-chain dehydrogenases/reductases (MDR). Sub-lines including ζ-crystalline, alcohol and polyol dehydrogenases, quinine oxidoreductases, enoyl reductases, VAT-1 and other proteins, Eur. J. Biochem. 226 (1994) 15–22.
[10] H. Jörnvall, B. Persson, M. Krook, S. Atrian, R. Gonzalez-Duarte, J. Jeffery, D. Ghosh, Short-chain dehydrogenases/reductases (SDR), Biochemistry 34 (1995) 6003–6013.
[11] K.L. Kavanagh, H. Jörnvall, B. Persson, U. Oppermann, The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes, Cell. Mol. Life Sci. 65 (2008) 3895–3906.
[12] B. Persson, J. Hedlund, H. Jörnvall, The MDR superfamily, Cell. Mol. Life Sci. 65 (2008) 3879–3894.
[13] H. Eklund, S. Ramaswamy, Three-dimensional structures of MDR alcohol dehydrogenases, Cell. Mol. Life Sci. 65 (2008) 3907–3917.
[14] R. Ladenstein, J.-O. Winberg, J. Benach, Structure–function relationships in short-chain alcohol dehydrogenases, Cell. Mol. Life Sci. 65 (2008) 3918–3935.
[15] B. Taneja, S.C. Mande, Conserved structural features and sequence patterns in the GroES fold family, Protein Eng. 12 (1999) 815–818.
[16] C.R. Woese, The universal ancestor, Proc. Natl. Acad. Sci. USA 95 (1998) 6854–6859.
[17] J.C. Venter, K. Remington, J.F. Heidelberg, A.L. Halpern, D. Rusch, J.A. Eisen, D. Wu, I. Paulsen, K.E. Nelson, W. Nelson, D.E. Fouts, S. Levy, A.H. Knap, M.W. Lomas, K. Nelson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y.-H. Rogers, H.O. Smith, Environmental genome shotgun sequencing of the Sargasso Sea, Science 304 (2004) 66–74.
[18] J. Filée, P. Forterre, J. Laurent, The role played by viruses in the evolution of their hosts: a view based on informational protein phylogenies, Res. Microbiol. 154 (2003) 237–243.
[19] P. Forterre, The two ages of the RNA world, and the transition of the DNA world, a story of viruses and cells, Biochimie 87 (2005) 793–803.
[20] J. Hedlund, H. Jörnvall, B. Persson, Subdivision of the MDR superfamily of medium-chain dehydrogenases/reductases through iterative hidden Markov model refinement, BMC Bioinformatics, in press.
[21] B.L. Vallee, D.S. Auld, Zinc coordination, function, and structure of zinc enzymes and other proteins, Biochemistry 29 (1990) 5647–5659.
[22] H. Jörnvall, H. von Bahr-Lindström, J. Jeffery, Extensive variations and basic features in the alcohol dehydrogenase–sorbitol dehydrogenase family, Eur. J. Biochem. 140 (1984) 17–23.
[23] T. Bergman, K. Zhang, C. Palmberg, H. Jörnvall, D.S. Auld, Zinc binding to peptide analogs of the structural zinc site in alcohol dehydrogenase: implications for an entatic state, Cell. Mol. Life Sci. 65 (2008) 4019–4027.
[24] B.L. Vallee, R.J.P. Williams, Metalloenzymes: the entatic nature of their active sites, Proc. Natl. Acad. Sci USA 59 (1968) 498–505.
[25] U. Ryde, The coordination of the catalytic zinc in alcohol dehydrogenase studied by combined quantum-chemical and molecular mechanics calculations, J. Comput. Aided Mol. Des. 10 (1996) 153–164.
[26] P.C. Sanghani, H. Robinson, W.F. Bosron, T.D. Hurley, Human glutathione-dependent formaldehyde dehydrogenase. Structures of apo, binary, and inhibitory ternary complexes, Biochemistry 41 (2002) 10778–10786.
[27] U. Ryde, Molecular dynamics simulations of alcohol dehydrogenase with a four- or five-coordinate catalytic zinc ion, Proteins 21 (1995) 40–56.
[28] R. Meijers, H.W. Adolph, Z. Dauter, K.S. Wilson, V.S. Lamzin, E.S. Cedergren-Zeppezauer, Structural evidence for a ligand coordination switch in liver alcohol dehydrogenase, Biochemistry 46 (2007) 5446–5454.
[29] Y. Kallberg, U. Oppermann, H. Jörnvall, B. Persson, Short-chain dehydrogenases/reductases (SDRs), Eur. J. Biochem. 269 (2002) 4409–4417.
[30] C. Canestro, R. Albalat, L. Hjelmqvist, L. Godoy, H. Jörnvall, R. Gonzalez-Duarte, Ascidian and amphioxus Adh genes correlate functional and molecular features of the ADH family expansion during vertebrate evolution, J. Mol. Evol. 54 (2002) 81–89.
[31] O. Danielsson, H. Jörnvall, "Enzymogenesis": classical liver alcohol dehydrogenase origin from the glutathione-dependent formaldehyde dehydrogenase line, Proc. Natl. Acad. Sci. USA 89 (1992) 9247–9251.
[32] M. Koivusalo, M. Baumann, L. Uotila, Evidence for the identity of glutathione-dependent formaldehyde dehydrogenase and class III alcohol dehydrogenase, FEBS Lett. 257 (1989) 105–109.
[33] X. Parés, J. Farrés, N. Kedishvili, G. Duester, Medium-chain and short-chain dehydrogenases/reductases in retinoid metabolism, Cell. Mol. Life Sci. 65 (2008) 3936–3949.
[34] L. Hjelmqvist, A. Norin, M. El-Ahmad, W. Griffiths, H. Jörnvall, Distinct but parallel evolutionary patterns between alcohol and aldehyde dehydrogenases: addition of fish/human betaine aldehyde dehydrogenase divergence, Cell. Mol. Life Sci. 60 (2003) 2009–2016.
[35] H. Jörnvall, B. Persson, M. Krook, J. Hempel, Alcohol and aldehyde dehydrogenases, in: N. Palmer (Ed.), The Molecular Pathology of Alcoholism, Oxford Univ. Press, 1991, pp. 130–156.
[36] C.A. Staab, M. Hellgren, J.-O. Höög, Dual functions of alcohol dehydrogenase 3: implications with focus on formaldehyde dehydrogenase and *S*-nitrosoglutathione reductase activities, Cell. Mol. Life Sci. 65 (2008) 3950–3960.