Volume 324, number 1, 9–14 FEBS 12516 June 1993

Dual relationships of xylitol and alcohol dehydrogenases in families of two protein types

Bengt Persson[c], Johan Hallborn[b], Mats Walfridsson[b], Bärbel Hahn-Hägerdal[b], Sirkka Keränen[a], Merja Penttilä[a] and Hans Jörnvall[c]

[a]*VTT, Biotechnical Laboratory, PO Box 202, SF-02151 Espoo, Finland,* [b]*Department of Applied Microbiology, University of Lund, PO Box 124, S-221 00 Lund, Sweden* and [c]*Department of Chemistry I, Karolinska Institutet, S-104 01 Stockholm, Sweden*

Received 5 April 1993; revised version received 14 April 1993

Xylitol dehydrogenase encoded by gene *XYL2* from *Pichia stipitis* is a member of the medium-chain alcohol dehydrogenase family, as evidenced by the domain organization and a distant homology (24% residue identity with the human class I_{γ_1} alcohol dehydrogenase). Much of a loop structure is missing, like in mammalian sorbitol and prokaryotic threonine dehydrogenases, many additional differences occur, and relationships are closest with the sorbitol dehydrogenase, the equivalence of which in *P. stipitis* may actually be the xylitol dehydrogenase. A second *P. stipitis* gene, also cloned and corresponding to a xylitol dehydrogenase, is highly different from *XYL2*, but encodes an enzyme with structural properties typical of the short-chain dehydrogenase family, which also contains an alcohol dehydrogenase (from *Drosophila*). Thus, yeast xylitol dehydrogenases, like alcohol and polyol dehydrogenases from other sources, have dual derivations, combining similar enzyme activities in separate protein families. In contrast to the situation with the other enzymes, both forms of xylitol dehydrogenase are present in one organism.

Enzyme relationship; Protein family; Homology; Short-chain dehydrogenase; Medium-chain alcohol dehydrogenase; Active site

1. INTRODUCTION

One form of xylitol dehydrogenase from the yeast *Pichia stipitis* has been structurally characterized from the corresponding gene, *XYL2*, and was then found to have a segment typical of NAD-binding dehydrogenases [1], but little is known about the structure–function relationships of the protein. However, the NAD-binding part has a position in the polypeptide chain [1] similar to that in classical liver alcohol dehydrogenase and other enzymes within the medium-chain alcohol dehydrogenase family [2]. Consequently, a more complete evaluation of the xylitol dehydrogenase structure versus that of the medium-chain enzymes is of interest to elucidate the functional characteristics. This is important since the protein family exhibits different types of evolutionary change [3], and has variable segments influencing also active site relationships [4–6]. We now examined this *P. stipitis* xylitol dehydrogenase structure in relation to the structures of known forms of alcohol, sorbitol, and threonine dehydrogenases, and ζ-crystallin. The results reveal extensive differences within an overall basic relationship, linking the *P. stipitis* xylitol dehydrogenase to the medium-chain enzyme family.

In addition, recent analysis [7] suggests the presence of a second xylitol dehydrogenase in *P. stipitis*. DNA analysis of a corresponding clone shows a deduced protein structure that is related to that of another dehydrogenase family, short-chain dehydrogenases, which includes insect alcohol dehydrogenase [8]. This allows the delineation of the functional organization and active site relationships also of the second form of xylitol dehydrogenase. The results further show that several dehydrogenases are represented by enzymes within two different protein families, which in the case of xylitol dehydrogenase apparently both are present in the same organism.

Correspondence address. H. Jörnvall, Department of Chemistry I, Karolinska Institutet, S-104 01 Stockholm, Sweden. Fax: (46) (8) 33 74 62.

2. MATERIALS AND METHODS

Structures of xylitol dehydrogenases from *Pichia stipitis* [1,7] and of the short-chain enzymes [8] were those reported. For the medium-chain alcohol dehydrogenases [2], the enzyme structures include the recently discovered novel classes of the mammalian enzyme [9,10]. All were compared in alignment programs which allowed on-screen evaluation of residue matches, graphic representations and estimates of secondary structures and hydropathy profiles [8].

3. RESULTS AND DISCUSSION

3.1. *Alignment of* Pichia stipitis XYL2 *gene encoded xylitol dehydrogenase with medium-chain alcohol dehydrogenases*

The primary structure of the xylitol dehydrogenase encoded by the *XYL2* gene is homologous with enzymes of the medium-chain alcohol dehydrogenase family, i.e. mammalian sorbitol dehydrogenases, prokaryotic thre-

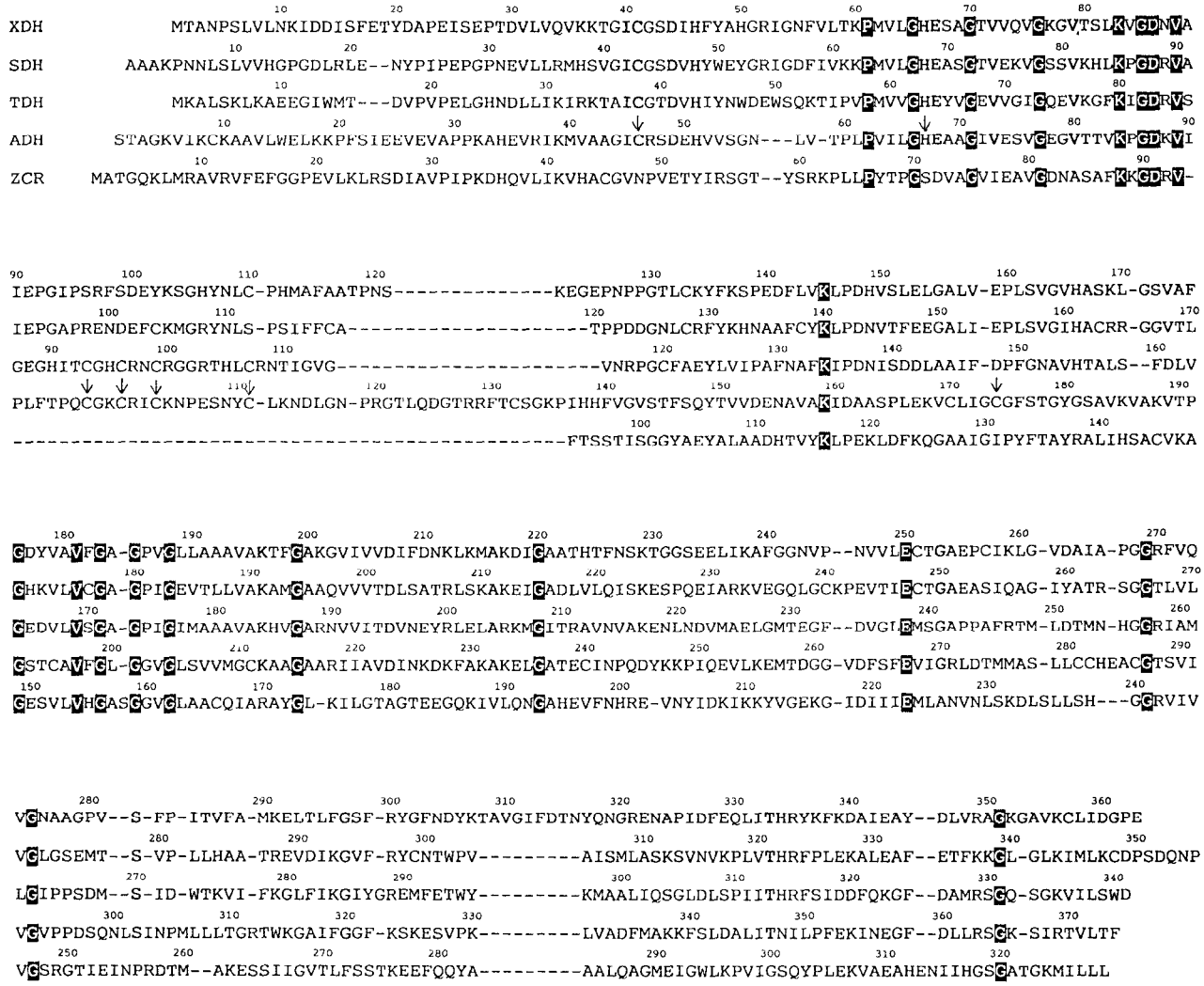

Fig. 1. Alignment of *Pichia stipitis* xylitol dehydrogenase encoded by the *XYL2* gene with four different members of the medium-chain alcohol dehydrogenase family. In each case, the human variant is shown, when characterized, i.e. for sorbitol dehydrogenase (SDH), alcohol dehydrogenase (ADH), and ζ-crystallin (ZCR). For ADH, the class I_{γ_1} subunit is shown, since that is the one most closely related to the parent form [2]. For the threonine dehydrogenase (TDH) the form characterized is from *Escherichia coli* [5]. Strictly conserved residues in inverted colours, those conserved in the dehydrogenases shaded. Positions of zinc ligands in ADH [13] are indicated by arrows.

onine dehydrogenase, zinc-containing alcohol dehydrogenases of five classes, and ζ-crystallin (Fig. 1). Considering the wide spread of these enzymes in activity and origins, the relationships are impressive. Overall residue identities range down to 20–25%, as for the other members of this family [2], but are higher for the xylitol dehydrogenase/sorbitol dehydrogenase pair (Table IA). This structural similarity parallels the fact that sorbitol dehydrogenases functionally have xylitol dehydrogenase activity [11]. In particular, the present yeast enzyme has both activities, the activity with sorbitol being about half that with xylitol; in addition, there is also an activity with ribitol (slightly less than with xylitol). Furthermore, the structural features typical of characterized sorbitol dehydrogenases, like a missing segment versus alcohol dehydrogenase (section 3.1.2 below) and Glu rather than Cys as the third ligand to the active site zinc atom (section 3.1.3 below), are also present in the xylitol dehydrogenase. These coincidences may even suggest that the xylitol dehydrogenase constitutes the equivalence in yeast of mammalian sorbitol dehydrogenase. Residues strictly conserved (below) are characteristic of the family and distinguish the dehydrogenases from ζ-crystallin. Thus, the *P. stipitis* xylitol dehydrogenase is a typical member of the medium-chain alcohol dehydrogenase family (Fig. 2), with closest relationships to the characterized sorbitol dehydrogenases, intermediate relationships to the other dehydrogenases, and most distant relationships to the crystallin, which enzymatically functions as a reductase [12] in the opposite direction. Apart from establishing the overall relationships, the alignment in Fig. 1 reveals aspects on residue

distributions, deletions and insertions, and zinc-ligands of xylitol dehydrogenase, as outlined below.

3.1.1. Residue distributions

A glycine conservation along much of the entire molecules (Fig. 1) constitutes an extraordinary pattern (Table IIA). Although the proteins are distantly related with single or zero other residues strictly conserved (Table IIA) over one-third of *all* glycine residues in xylitol dehydrogenase (13 of 37, i.e. 35%) are strictly conserved in all five enzyme types, and glycine accounts for 65% (13 of 20) of all residues conserved among the five proteins. It is the residue by far most strictly conserved in two segments, at positions 61–88 and 176–220 of xylitol dehydrogenase (Fig. 1). Both these segments correspond to special functional units of the enzymes, the former including the second ligand to the active site zinc atom, the latter forming a critical part of the coenzyme-binding site (cf. [13]). The conserved Gly residues correspond in most cases to reverse turns in the alcohol dehydrogenase conformation (cf. Fig. 2 in [6]). It is concluded that xylitol dehydrogenase has a largely conserved backbone conformation in relation to the other enzymes, and that this also relates to critical residues at the active site. For the whole family, the patterns in Fig. 1 and Table II extraordinarily well illustrate Gly conservation in a family of distantly related proteins.

3.1.2. Deletions/insertions

P. stipitis medium-chain xylitol dehydrogenase exhibits one missing segment toward alcohol dehydrogenase, between positions 121 and 122 (Fig. 1). This constitutes a large part of the segment missing also in liver sorbitol and prokaryotic threonine dehydrogenases, and corresponds to a separate loop in the subunit of alcohol dehydrogenase, the absence of which has been concluded to be associated with subunit interactions in the tetrameric enzymes [4,14]. Significantly, xylitol dehydrogenase is a tetramer (J.H., unpublished), thus strengthening the concept that absence of this loop is important in yielding dimer–dimer interactions in all enzymes of the family.

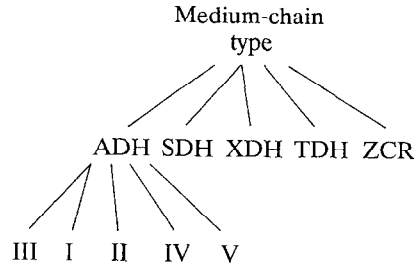

Fig. 2. Present pattern of the enzymes within the medium-chain alcohol dehydrogenase family. Abbreviations as in Fig. 1 and Table I. Regarding the enzymes (middle line), the xylitol dehydrogenase is the novel addition, although it may be the yeast equivalence of liver sorbitol dehydrogenase (cf. text); regarding the classes (bottom line) class IV and V constitute the recent additions [9]. Physical contacts between lines at the top of each branching (⋀) indicate a closer enzyme relationship than for those enzyme pairs represented by lines more widely separated at the top (/ \), but exact topography is still incompletely known.

Another segment difference is at positions 308–316, where xylitol dehydrogenase uniquely has extra residues (Fig. 1). This region corresponds to a domain border between the end of the coenzyme-binding domain and the second part of the catalytic domain in alcohol dehydrogenase [13]. It is generally less conserved and differs considerably also in the other enzymes (Fig. 1).

3.1.3. Zinc ligands and the active site

The yeast medium-chain xylitol dehydrogenase has three protein ligands typical of binding an active site zinc atom, Cys^{41}, His^{66}, and Glu^{159} (cf. arrows in Fig. 1). Of these, the Cys and His ligands are conserved in all the zinc-binding dehydrogenases, and the Glu position is occupied by either Glu, Asp or Cys in the other enzymes of the family [4,5], compatible with the rules for binding of catalytic zinc atoms in proteins in general [15]. It may be concluded that yeast xylitol dehydroge-

Table I

Pichia stipitis xylitol dehydrogenase relationships with members of the medium-chain alcohol dehydrogenase family (A), and the short-chain dehydrogenase family (B)

Enzymes compared (XDH_m versus)	Residue identities		Enzymes compared (XDH_s versus)	Residue identities	
	(Absolute numbers)	(%)		(Absolute numbers)	(%)
(A)			(B)		
SDH	134	39	DADH	41	19
TDH	97	29	3βStDH	48	22
ADH	83	24	20βStDH	57	26
ZCR	62	20	PGDH	35	16

(A) *P. stipitis XYL2* gene encoded xylitol dehydrogenase (XDH_m), liver sorbitol dehydrogenase (SDH), prokaryotic threonine dehydrogenase (TDH), human class I_{γ_1} alcohol dehydrogenase (ADH), human ζ-crystallin (ZCR). (B) *P stipitis* putative xylitol dehydrogenase of short-chain type (XDH_s), *Drosophila* alcohol dehydrogenase (DADH), prokaryotic 3β-hydroxysteroid (3βStDH) and 3α/20β-hydroxysteroid (20βStDH) dehydrogenases, and human NAD^+-dependent 15-hydroxysteroid dehydrogenase (PGDH). Gap regions are not included in the identities given.

nase is a zinc enzyme with the same catalytic mechanism as the other dehydrogenases, but that they all differ in this respect from the ζ-crystallin with its reductase activity [12].

The ligands of the second zinc atom of the medium-chain enzymes (Cys^{97}, Cys^{100}, Cys^{103}, Cys^{111} in alcohol dehydrogenase, arrows in Fig. 1) are not conserved in the yeast xylitol dehydrogenase, suggesting that this structural zinc atom is missing in xylitol dehydrogenase. This is also typical of sorbitol dehydrogenase [6], further emphasizing the link, or even equivalence, between the xylitol and sorbitol dehydrogenases. Within the protein family, the pattern is variable. The four cysteine residues involved, which in an isolated peptide segment also constitute a zinc-binding site [16], are conserved in threonine dehydrogenase, but not in the two sugar dehydrogenases except for single residues, in liver sorbitol dehydrogenase corresponding to the third Cys and in the *P. stipitis* xylitol dehydrogenase to the fourth (Fig. 1).

3.1.4. Conclusion

P. stipitis xylitol dehydrogenase is a member of the medium-chain alcohol dehydrogenase family, having similar subunit interactions as liver sorbitol dehydrogenase and prokaryotic threonine dehydrogenase, but lacking the second zinc site as does sorbitol dehydrogenase. The patterns illustrate both remarkable conservation in glycine distributions and considerable variations in metal binding and deletions/insertions.

3.2. *A second xylitol dehydrogenase gene from* P. stipitis *corresponding to a short-chain dehydrogenase*

Recently another *P. stipitis* gene was cloned from a cDNA expression library established in *E. coli* [7]. Based on screening with a zymogram technique [7], it encodes xylitol dehydrogenase activity. In addition, when cloned into a cosmid construction and expressed in *Saccharomyces cerevisiae*, the transformant exhibited xylitol dehydrogenase activity. The nucleotide-deduced amino acid sequence from this second dehydrogenase gene of *P. stipitis* is highly different from that of the *XYL2* product but shows typical relationships with the short-chain enzymes (Fig. 3). Overall values are given in Table IB and residues conserved in Table IIB. This short-chain enzyme family has recently been characterized also in tertiary structure [17,18] and utilized for model building [19,20]. In relation to these basic aspects, the novel short-chain *P. stipitis* enzyme is of interest for functional interpretations of features common to short-chain dehydrogenases.

3.2.1. Residue distribution

The alignment patterns are related to those of the medium-chain dehydrogenases in the sense that residue conservation is spread over the entire chains (Fig. 3) and include only a limited number of residues (Table IIB). However, glycine conservation is not equally pronounced as in the medium-chain family, although exclusion of one enzyme from the comparison results in a tendency of glycine conservation (left column, Table

Table II

Strictly conserved residues in alignments of different members of the medium-chain alcohol dehydrogenase family (A) and the short-chain dehydrogenase family (B)

Residue conserved	The four dehydrogenases	All five enzymes	Residue conserved	Four of the dehydrogenases	All five dehydrogenases
(A)			(B)		
Gly	15	13	Gly	10	4
Val	4	2	Leu	5	1
Lys	3	2	Asn	4	2
Asp	3	1	Asp	3	2
Glu	2	1	Ala	3	1
Ala	2	-	Val	3	-
His	2	-	Ser	2	1
Pro	1	1	Thr	2	1
Cys	1	-	Glu	1	1
Ile	1	-	Ile	1	1
Ser	1	-	Lys	1	1
Thr	1	-	Pro	1	1
			Tyr	1	1
Sum	36	20	Arg	1	-
			Sum	38	17

(A) The enzymes are the five aligned in Fig. 1, i.e. four dehydrogenases (xylitol dehydrogenase, sorbitol dehydrogenase, threonine dehydrogenase and alcohol dehydrogenase) and ζ-crystallin. (B) The enzymes are those in Fig. 3, i.e. the short-chain *Pichia* xylitol dehydrogenase, *Drosophila* alcohol dehydrogenase, *Pseudomonas* 3β-hydroxysteroid dehydrogenase, and *Streptomyces* 3α/20β-hydroxysteroid dehydrogenase (for the four enzymes), and human placental NAD^+-dependent 15-hydroxysteroid dehydrogenase.

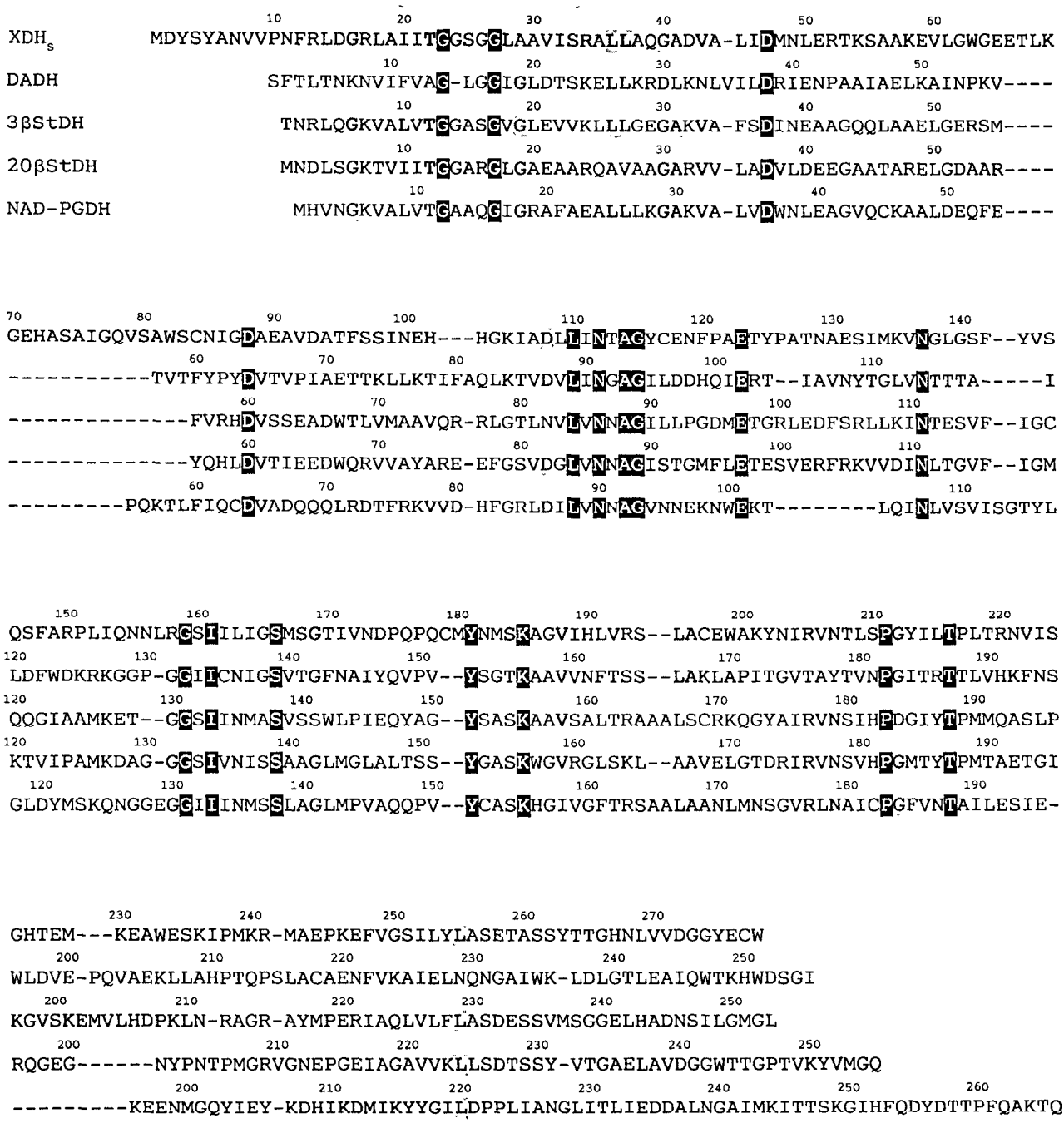

Fig. 3. Alignment of a second, putative *Pichia stipitis* xylitol dehydrogenase with short-chain dehydrogenases. The short-chain family has many different enzyme members [8]. In order to get equally many representatives as for the medium-chain family (Fig. 1), only four other members were included in the comparison, and were selected to include the four different types structurally analyzed in our laboratories, i.e. *Drosophila* alcohol dehydrogenase (DADH), *Pseudomonas* 3β-hydroxysteroid dehydrogenase (3βStDH), *Streptomyces* 3α/20β-hydroxysteroid dehydrogenase (20βStDH), and human placental NAD^+-dependent prostaglandin dehydrogenase (NAD-PGDH). Structures from [7,8]. Strictly conserved residues in all five enzymes in inverted colours, those conserved in any four of the sequences shaded The alignment is based upon multiple-sequence comparisons also involving other short-chain enzymes [8] and explaining the introductions of gaps, which for the C-terminal parts are not otherwise obvious from just the present structures.

IIB). Furthermore, the N-terminal part of the protein encoded by the novel short-chain dehydrogenase gene includes three glycine residues that are conserved in other short-chain dehydrogenases and characteristic of the coenzyme binding fold [8,21]. The short-chain family therefore follows the general pattern of glycine conservation in families of distantly related enzymes, but the pattern is less pronounced than in the medium-chain

family, suggesting that conformational differences are larger in the short-chain family. This conclusion is supported by the frequent presence of gaps and insertions in the alignment of the short-chain enzymes (Fig. 3), as also evident from model buildings [19,20].

3.2.2. Active site residues

The xylitol dehydrogenase has Tyr^{181} and Lys^{185}. They correspond to Tyr^{152} and Lys^{156} of 3α/20β-hydroxysteroid dehydrogenase (Fig. 3), residues which have been ascribed a functional importance from sequence comparisons [8], site-directed mutagenesis [22,23], chemical modification [24], and known conformations [17]. Consequently, the novel xylitol dehydrogenase is expected to function like the other short-chain enzymes, with similar catalytic mechanisms.

3.2.3. Conclusion

P. stipitis seems to contain a second xylitol dehydrogenase which belongs to the short-chain dehydrogenase family. The spread of variation in this family, including insertions and deletions, is wider than for the medium-chain enzymes. Nevertheless, a weak tendency to glycine conservation is observed, and a weak similarity is compatible with largely conserved folds, but with additions and removals of extra elements into the common pattern as shown for prostaglandin dehydrogenases [19,20]. Functionally important tyrosine and lysine residues are strictly conserved also in the xylitol dehydrogenase.

3.3. *Perspectives*

The two recently characterized *P. stipitis* enzymes of xylitol dehydrogenase derivation are members of two different protein families. Within each family, both exhibit discernible homologies, and apparently largely conserved active sites and reaction mechanisms. Nevertheless, overall residue distributions are markedly different, with the short-chain enzymes in general less well conserved. Since the two families have different molecular organizations and represent separate lineages, the two xylitol dehydrogenase activities constitute an example of convergence in function toward an activity, but divergence in structure within each family. Noticeably, this is observed also for other activities within these two lineages. Thus, alcohol dehydrogenase activity occurs in both families, as well as sorbitol/ribitol dehydrogenase activity [11], but in those cases the two lines have different organism distributions. In the present case, however, both the medium-chain and the short-chain enzyme are encoded by genes from the same organism.

With the addition of the present enzyme pair, most of the enzymes in the medium-chain alcohol dehydrogenase family have characterized counterparts with similar enzyme activity in the short-chain enzyme family of completely different domain organization and catalytic mechanism. The spread of all these dehydrogenases into two different families indicates a considerable divergence, and the common feature of dual enzymes with the same activity forms a distinct pattern within this system of oxidoreductases.

Acknowledgements: This work was supported by the Swedish Medical Research Council (project 03X-3532), the Swedish Alcohol Research Fund, the Swedish Cancer Society (1806), the Nordic Industrial Foundation (86199), the National Swedish Board for Technical Development (90-00978), the Swedish National Science Research Council (EEG04777-300), and the Swedish Ethanol Foundation

REFERENCES

[1] Kotter, P., Amore, R., Hollenberg, C.P. and Ciriacy, M. (1990) Curr. Genet. 18, 493–500
[2] Jörnvall, H., Persson, B. and Jeffery, J. (1987) Eur. J. Biochem. 167, 195–201.
[3] Yin, S.-J., Vagelopoulos, N., Wang, S.-L. and Jörnvall, H. (1991) FEBS Lett. 283, 85–88.
[4] Eklund, H., Horjales, E., Jörnvall, H., Brändén, C.-I. and Jeffery, J. (1985) Biochemistry 24, 8005–8012.
[5] Aronson, B D., Somerville, R.L., Epperly, B R. and Dekker, E.E. (1989) J. Biol. Chem. 264, 5226–5232.
[6] Borrás, T., Persson, B and Jörnvall, H. (1989) Biochemistry 28, 6133–6139.
[7] Hallborn, J., Penttilä, M., Ojamo, H., Walfridsson, M., Airaksinen, U., Keränen, S. and Hahn-Hägerdal, B. (1991) Int. Pat. Publ. WO 91/15588.
[8] Persson, B., Krook, M. and Jörnvall, H. (1991) Eur J Biochem. 200, 537–543.
[9] Parés, X., Cederlund, E., Moreno, A., Saubi, N., Hoög, J.-O. and Jörnvall, H. (1992) FEBS Lett. 303, 69–72.
[10] Yasunami, M., Chen, C.-S. and Yoshida, A. (1991) Proc. Natl. Acad. Sci. USA 88, 7610–7614.
[11] Jeffery, J. and Jornvall, H. (1988) in Adv. Enzymol (Meister, A. Ed.) 61, 47–106.
[12] Rao, P V., Krishna, C.M. and Zigler Jr., J.S. (1992) J. Biol. Chem. 267, 96-102.
[13] Eklund, H., Nordstrom, B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T., Söderberg, B.-O., Tapia, O., Brändén, C.-I. and Åkeson, Å. (1976) J Mol. Biol. 102, 27–59.
[14] Roumi, P., Loomes, K. and Jörnvall, H (1993) Eur. J. Biochem. 213, 487–492.
[15] Vallee, B.L. and Auld, D.S. (1990) Biochemistry 29, 5647–5659.
[16] Bergman, T., Jörnvall, H., Holmquist, B. and Vallee, B.L. (1992) Eur. J. Biochem. 205, 467–470.
[17] Ghosh, D., Weeks, C.M., Grochulski, P., Duax, W.L., Erman, M., Rimsay, R.L. and Orr, J.C. (1991) Proc. Natl. Acad. Sci. USA 88, 10064–10068.
[18] Varughese, K.I., Skinner, M.M., Whiteley, J.M., Matthews, D.A. and Xuong, N.H (1992) Proc. Natl. Acad. Sci. USA 89, 6080–6084.
[19] Krook, M., Ghosh, D., Strömberg, R., Carlquist, M. and Jörnvall, H. (1993) Proc. Natl. Acad. Sci. USA 90, 502–506.
[20] Krook, M., Ghosh, D., Duax, W and Jornvall, H. (1993) FEBS Lett. 322, 139–142.
[21] Wierenga, R.K., Maeyer, M.C. and Hol, W.G J. (1985) Biochemistry 24, 1346–1357.
[22] Ensor, C.M. and Tai, H.-H. (1991) Biochem. Biophys. Res. Commun. 176, 840–845.
[23] Albalat, R., Gonzàlez-Duarte, R. and Atrian, S. (1992) FEBS Lett. 308, 235–239.
[24] Krook, M., Prozorovski, V., Atrian, S., Gonzàlez-Duarte, R. and Jörnvall, H. (1992) Eur. J Biochem. 209, 233–239