Conserved supersecondary structural motif in NAD-dependent dehydrogenases

Alexey S. Kutzenko[a], Victor S. Lamzin[b], Vladimir O. Popov[c,*]

[a] *Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov str. 32, 117984 Moscow, Russia*
[b] *EMBL Outstation, DESY, Notkestraße 85, 22603 Hamburg, Germany*
[c] *Laboratory of Enzyme Engineering, A.N. Bakh Institute of Biochemistry, Russian Academy of Sciences, Leninsky pr., 33, 117071 Moscow, Russia*

Received 2 December 1997; revised version received 16 January 1998

Abstract L- and D-specific nicotinamide adenine dinucleotide (NAD)-dependent dehydrogenases map to the same structural protein superfamily as defined by the Structural Classification of Proteins (SCOP) and are based on the Rossmann fold type domains. A detailed classification of these domains is proposed using a novel diagnostic parameter of the rms per aligned pair. The catalytic domain in D-specific dehydrogenases shows a strong structural homology to the coenzyme binding domain. A topologically conserved part within the dehydrogenase superfamily reveals a supersecondary structural motif comprising the 5-stranded left-handedly twisted parallel β-sheet with one complete and one partial Rossmann fold units and two α-helices, the long helix, adjacent to and running roughly parallel with the β-sheet plane and the helix connecting two Rossmann folds.

Key words: Nicotinamide adenine dinucleotide-dependent dehydrogenase; D-specific Dehydrogenase; Conserved structural motif; Rossmann fold domain

1. Introduction

NAD(P)-dependent dehydrogenases comprise a numerous and well characterised group of enzymes. A number of high resolution crystal structures are available and a wealth of information on the catalytic and physico-chemical properties of dehydrogenases has been accumulated over the past decades [1–3].

Early X-ray studies have already revealed important similarities in the structures of NAD(P)-dependent enzymes [4]. These enzymes are typically composed of two or four identical polypeptide chains folded into tightly packed subunits. The subunits comprise two or more domains, which in some cases have a considerable freedom of movement via the interdomain hinge regions.

Each domain in dehydrogenases has a well defined function. The so-called coenzyme binding domain is located in the core of the molecule and has an evolutionarily conserved structure [4]. It is composed of a contiguous polypeptide chain of about 140–260 amino acids and serves a specific role of coenzyme recognition and binding. The core of the coenzyme binding domain is made up of a parallel left-handedly twisted β-sheet. The β-sheet includes two supersecondary structural elements each comprising three parallel β-strands connected by α-helices, β-strands or irregular loops, the so-called Rossmann fold (Figs. 1 and 2). Each of the Rossmann fold units binds a mononucleotide moiety (adenine or nicotinamide) of the coenzyme. Rossmann fold units provide a basis for structural motifs in a number of protein structures. These were grouped originally by Richardson [5] into several layers of loosening topological similarity. Coenzyme binding domains of NAD-dependent dehydrogenases present a classical example and top the list. This type of folding is prevalent in dehydrogenases but other topological motifs have been found to provide NAD(P) recognition in a number of proteins ([3] and references therein).

The role of catalytic domains in dehydrogenases is to provide residues essential for catalysis and substrate co-ordination. The active site of NAD-dependent dehydrogenases is formed by amino acid residues donated by both domains and is usually located at the domain interface. The catalytic domains of dehydrogenases were considered to have a variety of folds with no topological relation to the structure of the coenzyme binding domains. Recently determined crystal structures of the D-specific dehydrogenases of 2-hydroxy acids (acting on D-stereoisomers of the respective substrates as opposed to the L-specific enzymes) FDH [6], DGDH [7], PGDH[8] and DLDH [9] revealed that the fold of catalytic domains resembles the topology of the classic coenzyme binding domain but differs significantly from that of the catalytic domains in L-specific dehydrogenases. A high degree of internal symmetry for D-specific dehydrogenases has been mentioned [3,7,8,10], thus suggesting some general regularities for this phenomenon.

An additional structurally conserved unit in both D- and L-dehydrogenases is a long helix αA, adjacent to the β-sheet of the coenzyme binding domain (a general nomenclature of the secondary structural elements in dehydrogenases introduced by Adams and co-workers [11] as applied to FDH [6], Fig. 2, is used throughout the paper). It is located on the same side of the β-sheet as helices αB and αC from the Rossmann fold unit involved in binding of the adenine mononucleotide part of the coenzyme. Although the helix is conserved structurally, its position in the amino acid sequence differs from protein to protein [4,6–9].

As more structures of NAD(P)-dependent dehydrogenases become available, the comparison of common topological motifs is gradually evolving [12–15]. The Structural Classification of Proteins (SCOP) currently differentiates several levels of hierarchy: class-fold-superfamily-family-domain. It is anticipated that additional levels of classification may be intro-

*Corresponding author. Fax: (7) (95) 9580877.
E-mail: vpopov@glas.apc.org

Abbreviations: ADH, alcohol dehydrogenase; DHFR, dihydrofolate reductase; FD, flavodoxin; FDH, formate dehydrogenase; DGDH, D-glycerate dehydrogenase; GRS, glutathione reductase; GPDH, glyceraldehyde-3-phosphate dehydrogenase; LDH, lactate dehydrogenase; DLDH, D-lactate dehydrogenase; MDH, malate dehydrogenase; PGDH, D-phosphoglycerate dehydrogenase

PII S0014-5793(98)00074-X

duced, thus allowing better defined positioning of the protein structures [14].

Here we present a new structural motif found in several D-specific dehydrogenases. The motif has been identified in both coenzyme binding and catalytic domains thus emphasising an internal symmetry in their structural organisation. A more detailed classification of these domains is proposed. The domains can be ranked according to progressively loosening structural similarity in the succession: (1) coenzyme binding domains of D-specific dehydrogenases; (2) catalytic domains of D-specific dehydrogenases; (3) coenzyme binding domains of L-specific dehydrogenases. Classification may be further extended to the family level of the SCOP hierarchy differentiating L- and D-specific dehydrogenases.

2. Materials and methods

The following structures from the Protein Data Bank [16] were analysed: FDH (2nac, 2nad) [6], DGDH (1gdh) [7], PGDH (1psd) [8], DLDH (2dld) [9], GRS (3grs) [17], ADH (8adh) [18], DHFR (8dfr) [19], FD (2fcr) [20], GPDH (2gd1) [21], LDH (1ldm) [22], MDH (4mdh) [23].

A search for the homologous regions was performed by structure superposition using the algorithm of Rossmann and Argos [24] implemented in OVERLAP. Initial superposition of the structural elements was performed using EXIFIT [25]. Tek_FRODO [26], MOLSCRIPT [27] and ICM [28] graphics packages were used for structure visualisation and drawing.

3. Results and discussion

For quantitative comparisons the following set of structures was used: L-specific dehydrogenases (LDH, MDH, ADH, GPDH), D-specific dehydrogenases (DGDH, DLDH, PGDH), reductases (DHFR, GRS) and flavodoxin (FD). FDH (Fig. 2) from the methylotrophic bacterium *Pseudomonas* sp. 101, a member of the D-specific dehydrogenase family [29], was used as a reference protein.

Fig. 3 presents a stereo view of the two domains of FDH (coenzyme binding and catalytic) superimposed on one another. The topology (connectivity) of these domains (or parts thereof) is schematically shown in Fig. 1. Both domains of FDH have a Rossmann fold as a basic structural unit. The topology of the coenzyme binding domain of FDH, **αA**-α5-α6-**βA**-αB-**βB**-αC-βC-**α7-βD-αD-αE-βE**-αF-**βF**-αG-βG (equivalent fragments are marked in bold underlined), is close to the classical ones found in LDH and MDH, while the core of the catalytic domain, **β1**-α1-**β4-α2-β5-α3-β7**-α4-**β8**-(insert of coenzyme binding domain)-**α8**, may be regarded as a truncated copy of the coenzyme binding one. Superposition reveals 52 structurally equivalent pairs of Cα atoms with a rms deviation of about 1.1 Å. The alignment comprises vast stretches of the amino acid sequence and includes the entire β-sheet of the FDH catalytic domain as well as flanking (α8) and internal (α2, α3) helices. A notable feature of the alignment is the topological equivalence of two pairs of α-helices, α8-αA and α2-α7. The former two long helices link two domains, while the latter facilitate the advance of the main chain from one side of the β-sheet of the respective domains to the other.

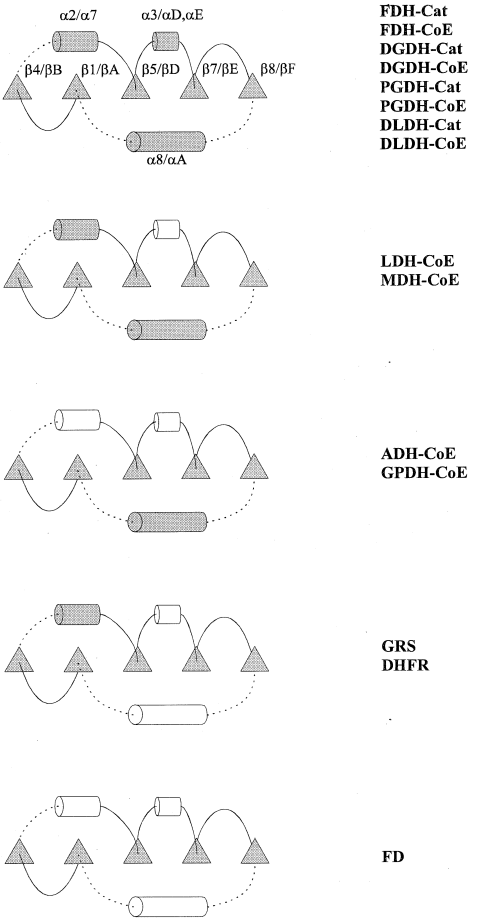

Fig. 1. Diagrammatic presentation of the structures of coenzyme binding and catalytic domains of NAD-dependent dehydrogenases and related proteins. Cylinders indicate α-helices, triangles β-sheets. Topologically equivalent parts of the structures are shaded. Cat: catalytic domain; CoE: coenzyme binding domain. Only conserved parts of the structures are shown (notation refers to FDH): catalytic domain (β1-β4-α2-β5-α3-β7-β8-α8); coenzyme binding domain (βA-βB-α7-βD-αD-βE-βF-αA).

No significant similarity in the primary structures of the two parts of the FDH polypeptide chain has been observed (the similarity between the two domains does not exceed 20%).

Comparison of the domain structures of NAD-dependent dehydrogenases and related enzymes (Fig. 1) reveals a com-

Table 1
Superposition of NAD(P)-binding Rossmann fold domains

Structural domain/comparison	Number of equivalent pairs of Cα atoms (N)	rms (Å)	R/N (Å)
Coenzyme binding domains of D-specific dehydrogenases (CoE-D)	138 ± 10	0.79 ± 0.06	0.006
Catalytic domains of D-specific dehydrogenases (Cat-D)	70 ± 13	1.3 ± 0.3	0.019
Coenzyme binding domains of L-specific dehydrogenases (CoE-L)	53 ± 9	1.3 ± 0.5	0.025
Catalytic domain of FDH vs. CoE-L	38 ± 6	1.2 ± 0.2	0.032
Coenzyme binding or catalytic domains of FDH vs. reductases and flavodoxin	28 ± 6	1.4 ± 0.3	0.051

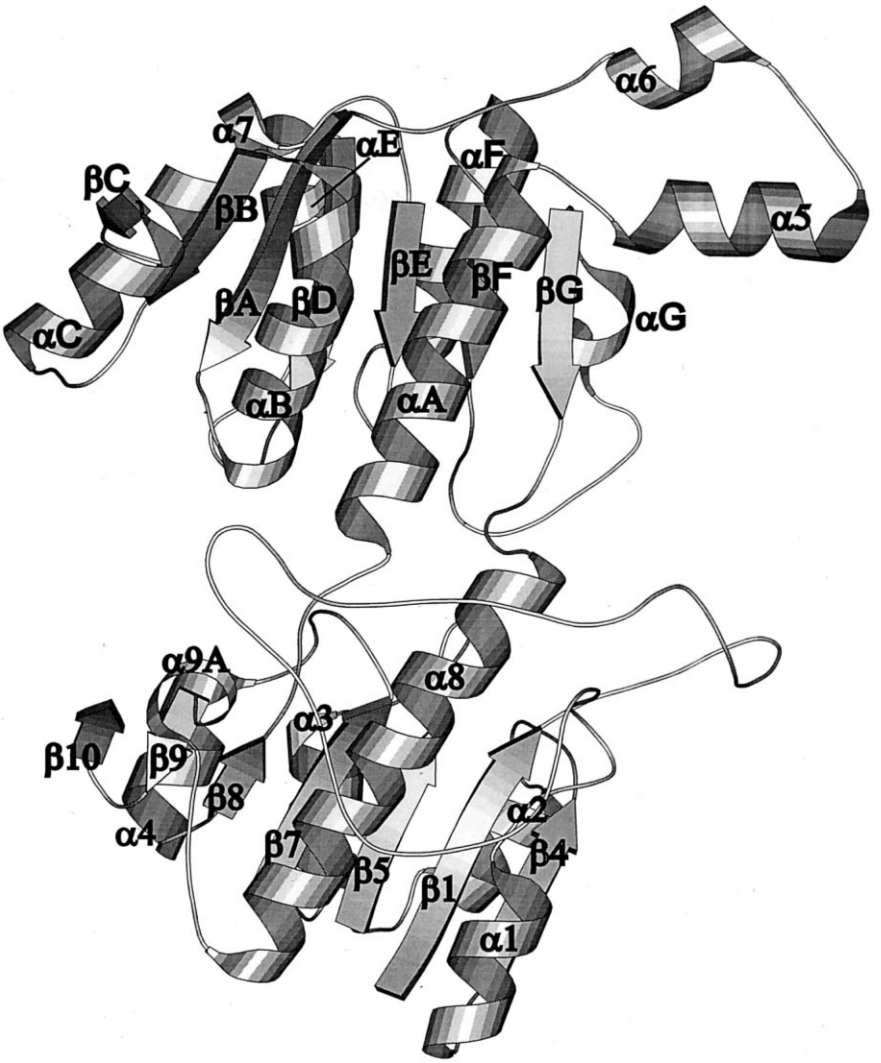

Fig. 2. MOLSCRIPT representation of the FDH structure.

mon structural motif that may be traced within the coenzyme binding domains of these proteins as well as in the catalytic domains of D-specific dehydrogenases. The structural motif, common to these structures, includes the 5-stranded β-sheet comprising one complete and one partial Rossmann fold and two α-helices. The helices are the long 'flanking' helix (αA for the coenzyme binding domain of FDH and α8 for the catalytic domain) running roughly parallel to the plane of the β-sheet and an α-helix which connects two Rossmann folds, 'interconnecting' (α7 and α2 respectively).

Several levels of progressively loosening structural similarity within the Rossmann fold domains can be traced (Table 1). As a measure of structural similarity the parameter [rms/number of equivalent pairs] or the rms per aligned pair (R/N) has been used. The lower the value of this parameter, i.e. the lower the rms or the higher the number of equivalent pairs, the closer is the structural relation between the proteins being compared. The R/N parameter has a clear compensational origin. The higher the value of rms, the higher should be the number of residues aligned for the structures to reveal topological similarity. When the complete subunits of D-specific dehydrogenases are compared (Table 2), they show average values of N, rms and R/N of 200 ± 27, 1.58 ± 051 Å, and 0.008 ± 0.002 Å respectively. For very closely related proteins, e.g. for the same protein acquiring a different conformation (e.g. apo versus holo), the parameter is close to zero (R/

Table 2
Superposition of the subunits of D-specific dehydrogenases

	FDH	DGDH	PGDH	DLDH
FDH		223/2.3	192/2.2	/∼6.0
DGDH	0.010		232/1.4	216/1.7 (186/1.3)
PGDH	0.011	0.006		201/1.5 (151/0.9)
DLDH	–	0.008 (0.007)	0.007 (0.006)	

Right side of the table: number of equivalent pairs of Cα atoms/rms (Å); left side: rms per aligned pair (R/N, Å). Figures in parentheses present the alignment with the least rms.

Fig. 3. Stereo view of the superposition of the catalytic (thin line) and the coenzyme binding (thick line) domains of FDH.

$N < 0.005$ Å for the alignment of the two conformations of the coenzyme binding domains in apo and holo forms of FDH [6]), while for distantly related structures (rms > 2–3 Å, $N < 20$–30) the parameter R/N exceeds 0.1 Å. The value of 0.1 may be considered a cut-off above which topological similarity becomes questionable.

Based on the values of R/N the Rossmann fold domains comprising NAD-dependent dehydrogenases can be subdivided into at least three structural groups or subclasses of loosening structural similarity within the group. The most extensive similarity is observed between the coenzyme binding domains of D-specific dehydrogenases. Catalytic domains of D-specific dehydrogenases comprise the second level. All five β-strands as well as 'flanking', 'interconnecting' and other helices found within Rossmann folds can be aligned (Fig. 1). The third level of structural homology is represented by the alignment of the coenzyme binding domains of L-specific dehydrogenases, some of which may already lack certain elements of the common structural motif. The alignment of the coenzyme binding domain of the FDH reference protein (D-specific enzyme) with the coenzyme binding domains of L-specific dehydrogenases falls within the same subclass ($N = 61 \pm 9$, rms = 1.50 ± 0.45 Å, R/N = 0.0246 Å) emphasising the close relation between the structures while the alignment of the catalytic domain of FDH with the coenzyme binding domains of L-specific dehydrogenases reveals a much higher R/N (0.0317 Å) pointing to a more distant relationship. The last level with the least similarity comprises alignments of catalytic or coenzyme binding domains of FDH with the structures of reductases and flavodoxin. Only the β-sheet and short segments (2–3 amino acid residues) of the 'flanking' or 'interconnecting' helices are preserved. Thus the parameter R/N provides a convenient way for comparison between the domains comprising NAD-dependent dehydrogenases and may possibly be used to compare structures from other families.

Classification may be further extended to the family level of the SCOP hierarchy differentiating L- and D-specific dehydrogenases of 2-hydroxy acids. The topological similarities inside the group are much more profound than between the groups.

The structural alignment of D-specific dehydrogenases (Table 2) is as good as superposition of L-LDH and L-MDH, whose substrates differ by only one carboxylic substituent and which are considered very close homologues giving 209 equivalent pairs of atoms with an rms of 1.42 Å [30]. Both subfamilies contain topologically related internal coenzyme binding domains, while the catalytic domain is different for L- and D-specific enzymes. Unlike L-specific enzymes D-specific dehydrogenases of 2-hydroxy acids appear to have a highly symmetrical organisation based on a single structural motif. They are composed of an even number of similar subunits which are organised into two domain structures. The domains (CoE-D and Cat-D) have a rather similar folds and can be represented as bodies related by a rotation axis. Moreover, each of the domains is composed of two Rossmann fold units or parts thereof related by another pseudosymmetry axis [10]. As pointed out earlier [14] proteins with similar folds in more than one domain are likely to be more closely related than those sharing common folds in the coenzyme domain only.

Acknowledgements: The work has been supported by grants from The International Science Foundation (MJS300) and INTAS (International Association for the Promotion of Co-operation with Scientists from the New Independent States of the Former Soviet Union) (94-1309). The authors would like to thank Prof. K. Wilson from York University for valuable comments.

References

[1] Grau, U.M. (1982) in: Pyridine Nucleotide Coenzymes (Everse, J., Anderson, B. and You, K.-S., Eds.), pp. 135–187, Academic Press, New York.
[2] Eklund, H. and Branden, C.-I. (1987) in: Pyridine Nucleotide Coenzymes (Dolphin, D., Avramovich, O. and Poulson, R., Eds.), pp. 51–98, John Wiley and Sons, New York.
[3] Popov, V.O. and Lamzin, V.S. (1994) Biochem. J. 301, 625–643.
[4] Rossman, M.G., Liljas, A., Branden, C.-I. and Banaszak, L.J. (1975) in: The Enzymes, 3rd edn. (Boyer, P.D., Ed.), Vol. 11, Part A, pp. 61–102, Academic Press, New York.
[5] Richardson, J.S. (1981) in: Advances in Protein Chemistry (Anfinsen, C.B., Edsall, J.T. and Richards, F.M., Eds.), Vol. 34, pp. 167–339, Academic Press, New York.
[6] Lamzin, V.S., Dauter, Z., Popov, V.O., Harutyunyan, E.H. and Wilson, K.S. (1994) J. Mol. Biol. 236, 759.
[7] Goldberg, J.D., Yoshida, T. and Brick, P. (1994) J. Mol. Biol. 236, 1123–1140.
[8] Schuller, D.J., Grant, G.A. and Banaszak, L.J. (1995) Nature Struct. Biol. 2, 69–75.
[9] Dunn, C.R. and Holbrook, J.J. (1996) Protein Data Bank entry 2DLD.

[10] Kutzenko, A.V., Korolev, S.V., Lamzin, V.S. and Popov, V.O. (1994) J. Mol. Biol. (Russia) 28, 626–632.
[11] Adams, M.J., Ford, G.C., Koekoek, R., Lentz, P.J., McPherson, A., Rossmann, M.G., Smiley, I.E., Shewitz, R.W. and Wonacott, A.J. (1970) Nature 227, 1098–1103.
[12] Lesk, A.M. (1995) Curr. Opin. Struct. Biol. 5, 775–783.
[13] Orengo, C.A., Jones, D.Y. and Thornton, J.M. (1994) Nature 372, 631–634.
[14] Brenner, S.E., Chothia, C., Hubbard, T.J.P. and Murzin, A. (1996) Methods Enzymol. 266, 635–643.
[15] Bellamachina, C.R. (1996) FASEB J. 10, 1257–1269.
[16] Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer Jr., E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) J. Mol. Biol. 112, 535–542.
[17] Karplus, P.A. and Schulz, G.E. (1987) J. Mol. Biol. 195, 701–729.
[18] Colonna-Cesari, F., Perahia, D., Karplus, M., Eklund, H., Branden, C.I. and Tapia, O. (1986) J. Biol. Chem. 261, 15273–15280.
[19] Davies, J.F., Matthews, D.A., Oatley, S.J., Burridge, J., Kaufman, B.T., Xuong, N.-H. and Kraut, J. (1989) Protein Data Bank entry 8DFR.
[20] Fukuyama, K., Matsubara, H. and Rogers, L.J. (1992) Protein Data Bank entry 2FCR.
[21] Skarzynski, T. and Wonacott, A.J. (1988) J. Mol. Biol. 203, 1097–1118.
[22] Abad-Zapatero, C., Griffith, J.P., Sussman, J.L. and Rossmann, M.G. (1987) J. Mol. Biol. 198, 445–467.
[23] Birktoft, J.J., Rhodes, G. and Banaszak, L.J. (1989) Biochemistry 28, 6065–6081.
[24] Rossmann, M.G. and Argos, P.J. (1975) J. Biol. Chem. 250, 7525–7532.
[25] McLachlan, A.D. (1985) EXIFIT protein superposition program.
[26] Vassylyev, D.G. and Adzhubei, A.A. (1992) J. Mol. Graphics 10, 101–104.
[27] Kraulis, P.J. (1991) J. Appl. Crystallogr. 24, 946–950.
[28] Abagyan, R., Totrov, M. and Kuznetsov, D. (1994) J. Comp. Chem. 15, 488–506.
[29] Vinals, K., Depiereux, E. and Feytmans, E. (1993) Biochem. Biophys. Res. Commun. 192, 182–188.
[30] Birktoft, J.J. and Banaszak, L.J. (1983) J. Biol. Chem. 258, 472–482.