

Amino Acid Residues that Determine Functional Specificity of NADP- and NAD-Dependent Isocitrate and Isopropylmalate Dehydrogenases

Olga V. Kalinina¹ and Mikhail S. Gelfand^{1,2*}

¹Department of Bioengineering and Bioinformatics, Moscow State University, Moscow, Russia

²Institute for Information Transmission Problems, Moscow, Russia

ABSTRACT Isocitrate and isopropylmalate dehydrogenases are homologous enzymes important for the cell metabolism. They oxidize their substrates using NAD or NADP as cofactors. Thus, they have two specificities, towards the substrate and the cofactor, appearing in three combinations. Although many three-dimensional (3D) structures are resolved, identification of amino acids determining these specificities remains a challenge. We present computational identification and analysis of specificity-determining positions (SDPs). Besides many experimentally proven SDPs, we predict new SDPs, for example, four substrate-specific positions (103Leu, 105Thr, 337Ala, and 341Thr in IDH from *E. coli*) that contact the cofactor and may play a role in the recognition process. *Proteins* 2006;64:1001–1009. © 2006 Wiley-Liss, Inc.

Key words: specificity determinants; specificity-determining positions (SDP); isocitrate dehydrogenase; protein function

INTRODUCTION

Isocitrate dehydrogenase (IDH) is one of central enzymes of the carbon metabolism. It participates in the tricarboxylic acid cycle in a variety of organisms.^{1–3} IDH catalyzes oxidation of isocitrate to α -ketoglutarate and CO₂ using either NAD (EC: 1.1.1.41) or NADP (EC: 1.1.1.42) as a cofactor. Eukaryotes have three IDH isoenzymes, two mitochondrial (NAD dependent and NADP dependent) and one cytosolic (NADP dependent). Prokaryotes have only one IDH, whose dependence on NADP or NAD perfectly correlates with the presence or absence of the glyoxylate bypass in the organism.⁴

Isopropylmalate dehydrogenase (IMDH) catalyzes oxidative decarboxylation of 3-isopropylmalate into 2-oxo-4-methylvalerate, the third step in the leucine biosynthesis in bacteria and fungi.^{5,6} All known IMDHs are NAD dependent (EC: 1.1.1.85).

IDH and IMDH are evolutionary related.^{2–5} They are about 25% identical, have a common protein fold, and the same catalytic mechanism.^{2,3,7} In most species they function as homodimers or, as some eukaryotic IDHs, more complicated complexes.

Numerous specificity determinants were identified in bacterial, archaeal, and eukaryotic IDHs and IMDHs.^{8–13} Mutants with inverted substrate and cofactor specificities

were created.^{7,9,14–17} However, the activity of the specificity-reversal mutants was usually lower than that of wild-type proteins, and sometimes residues far from the functional sites had to be mutated in order to increase it.^{7,17} This implies that the specific recognition of a substrate does not simply depend on a few amino acids in the binding pocket.

The aim of this study was computational identification of positions determining specificity of IDHs and IMDHs towards their substrates and cofactors.

METHODS

All isocitrate or isopropylmalate dehydrogenases, whose annotated function relies on experimental support, were taken from the Swiss-Prot database (<http://www.expasy.ch/sprot/sprot-top.html>). They were split into four specificity groups: NAD-dependent IDHs from mitochondria and *P. furiosus*; type I NADP-dependent IDHs from bacteria and archaea; IMDHs (all NAD-dependent); and type II IDHs from eukaryota, both mitochondrial and cytosolic. The annotation of specificity was taken from Swiss-Prot as well.

Proteins within groups were aligned using CLUSTALW,¹⁸ then the obtained profiles were aligned, and the result was manually edited to satisfy available experimental and structural data. The final alignment is available in the Supplementary Material. The maximum likelihood phylogenetic tree was built using the PROML utility from PHYLIP.¹⁹

Abbreviations: IDH, isocitrate dehydrogenase; IMDH, isopropylmalate dehydrogenase; NAD, nicotinamide adenine dinucleotide; NADP, nicotinamide adenine dinucleotide phosphate; MSA, multiple sequence alignment; SDP, specificity-determining position.

The Supplementary Material referred to in this article can be found at <http://www.interscience.wiley.com/jpages/0887-3585/suppmat/>

Grant sponsor: Howard Hughes Medical Institute; Grant number: 55000309; Grant sponsor: Ludwig Institute for Cancer Research; Grant number: CRDF RBO-1268; Grant sponsor: Russian Academy of Sciences; Grant sponsor: Russian Science Support Fund; Grant sponsor: INTAS Fellowship Grant for Young Scientists; Grant number: 04-83-3704.

*Correspondence to: Mikhail S. Gelfand, Institute for Information Transmission Problems, Bolshoi Karetny per., 19, Moscow, 127994, Russia. E-mail: gelfand@iitp.ru

Received 5 October 2005; Revised 7 February 2006; Accepted 27 February 2006

Published online 9 June 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21027

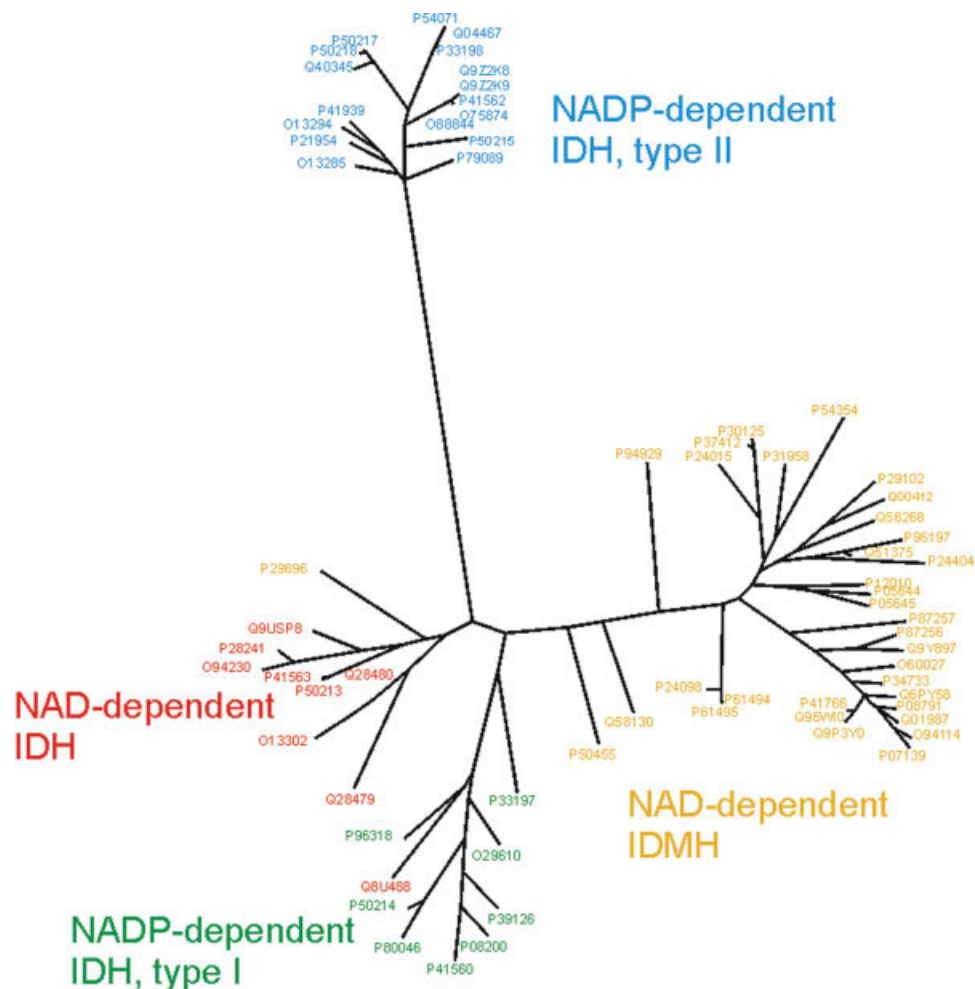


Fig. 1. The phylogenetic tree of the NAD- and NADP-dependent IDHs and IMDHs. Proteins are identified by their Swiss-Prot accession numbers. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Specificity determining positions (SDPs) were identified using SDPpred²⁰ (<http://monkey.belozersky.msu.ru/~psn>). SDPpred searches for positions in a multiple alignment that are well conserved within specificity groups but differ between these groups. For each alignment column, it computes a measure of its statistical significance, *z*-score. The higher *z*-score, the more strictly the distribution of amino acids in this position is associated with specificity, and thus, by assumption, the more likely the position is a true specificity determinant. Recently, a number of techniques dealing with similar problems has emerged.^{21–26} The statistical rationale of SDPpred is most similar to those of Hannenhalli and Russell,²³ Mirny and Gelfand,²⁴ and Donald and Shakhovich.²⁶ Still, it differs from most of them in a number of technical details, which make SDPpred procedure more reliable and fully automated: SDPpred takes into account evolutionary and physicochemical similarities of amino acids enclosed in amino acid substitution matrices; it corrects the statistics basing on real evolutionary distances within and between specificity groups; and it has an automatic procedure of choosing the

number of positions to be SDPs, which is least probable to appear by chance for the given alignment.

Predicted SDPs were mapped to three-dimensional (3D) structures of IDHs and IMDHs listed in the Supplementary Material. These structures were selected using two criteria: (1) no mutations and (2) presence of either substrate or its analog, or cofactor, or both. The only exception was the cofactor-specificity inverted mutant *liso*. Structural analysis was performed using RASMOL.²⁷

RESULTS AND DISCUSSION

The phylogenetic tree of the family is shown in Figure 1. Eukaryotic type II NADP-IDHs are very distant members of the family. The phylogenetic relationships among the remaining three groups can be hardly resolved. We considered three groupings by specificity: (i) by cofactor: all NAD-dependent enzymes (NAD-IDHs and IMDHs) versus all NADP-dependent enzymes (type I and II NADP-IDHs); (ii) by substrate: all IDHs versus all IMDHs; and (iii) all four groups. The predicted SDPs are listed in the Supplementary Material (alignment positions numbered as in

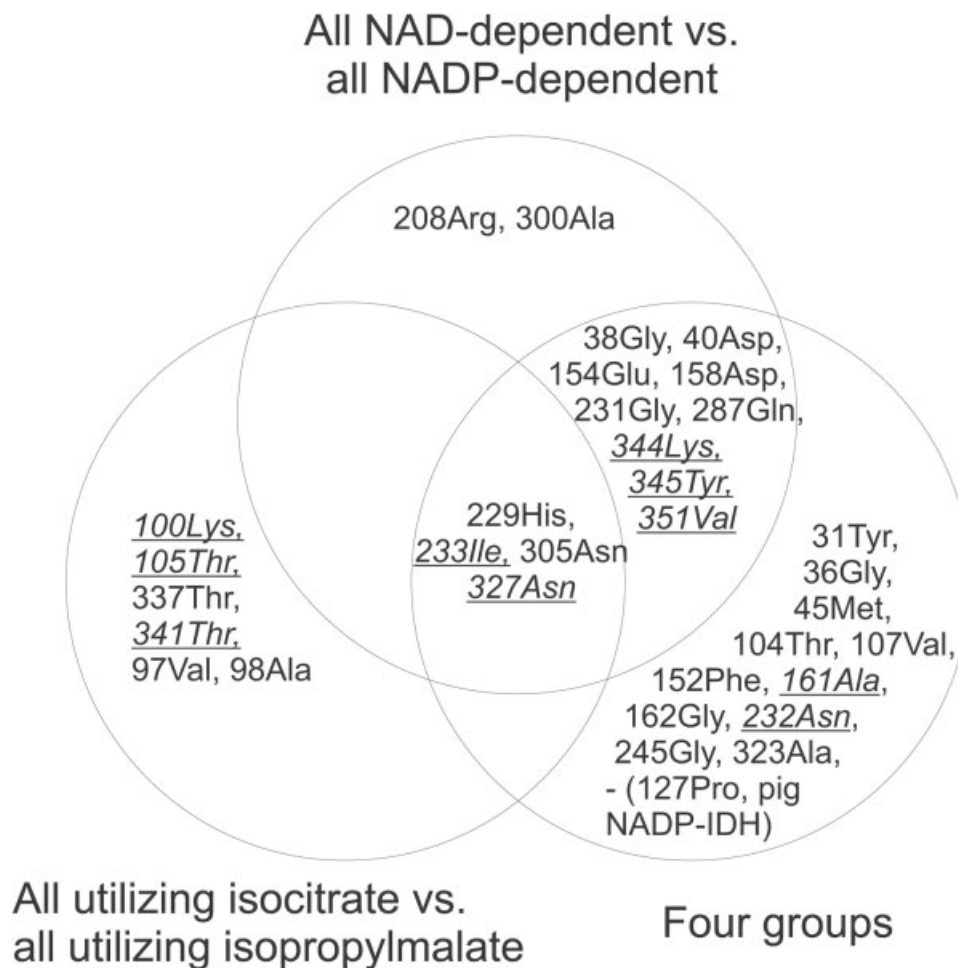


Fig. 2. Overlap of cofactor-specific, substrate-specific, and four group SDPs. Experimentally supported functionally important positions italicized and underlined.

NADP-IDH from *E.coli*, Swiss-Prot P08200). The overlap of the three sets of predicted SDPs is shown in Figure 2. The third comparison produces a larger number of SDPs than the first two, although z -scores for them are not so high. One can describe four-group SDPs roughly as a union of cofactor-specific and substrate-specific SDPs, together with some new positions. This can be explained by the fact that dividing the family into four groups account for both types of specificity, whereas the other two divisions ignore one of them. There are no SDPs common for the first and second comparisons that do not appear in the third one. One can notice that experimentally verified SDPs have usually high z -scores, although there are positions with equally high scores, but no experimental support.

SDPs were mapped onto 3D structures of IDHs and IMDHs (e.g., Fig. 3). For cofactor-specificity groups, many predicted SDPs lie in the cofactor-binding part of the molecule [Fig. 3(A)]. For substrate-specificity groups, all SDPs are located in the substrate-binding part [Fig. 3(B)]. In all three comparisons, a substantial fraction of the predicted SDPs lie on the intersubunit interface (see Table I and below for detail).

The distances between SDPs and the cofactor, the substrate, or the other subunit in different structures for each comparison are presented in the Supplementary Material, whereas the averages and standard deviations are shown in Figure 4. Fourteen out of 16 SDPs identified in the cofactor-specificity comparison, 13 out of 15 substrate-specificity SDPs, and 27 of 31 four-group SDPs lie in the vicinity of the substrate, the cofactor, or the other subunit in at least one of the considered structures (see Supplementary Material). For many SDPs there exist experimental data showing that they are indeed involved in the specific binding of the cofactor or the substrate. Despite the large number of predicted SDPs, they are, on average, more closely located to the important protein regions, and the fraction of experimentally proven positions important for function is also higher among SDPs in all three predictions than in the protein on average [Fig. 5(A)].

Nevertheless, a number of functionally proven important amino acids was not predicted as SDPs. The most common reason is buried in the definition of SDP. SDP is a position, which is important for protein specificity, not for protein

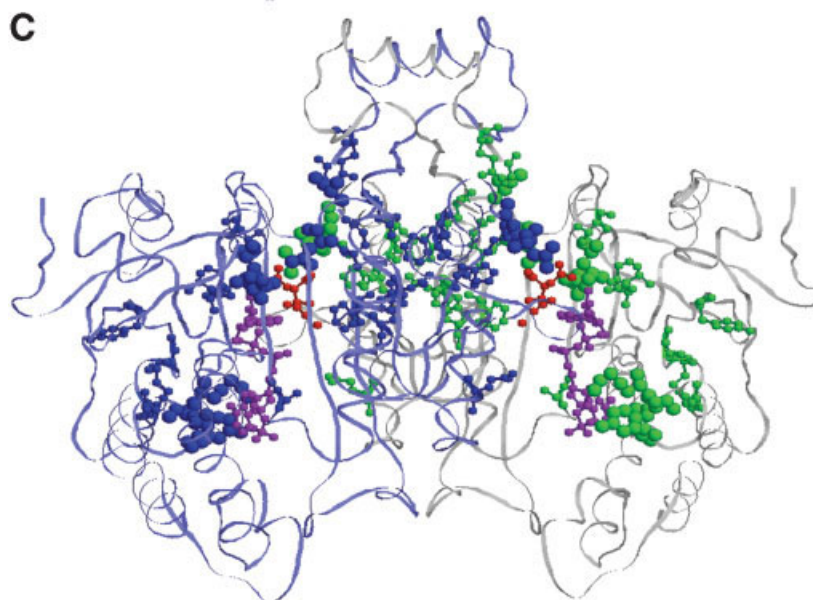
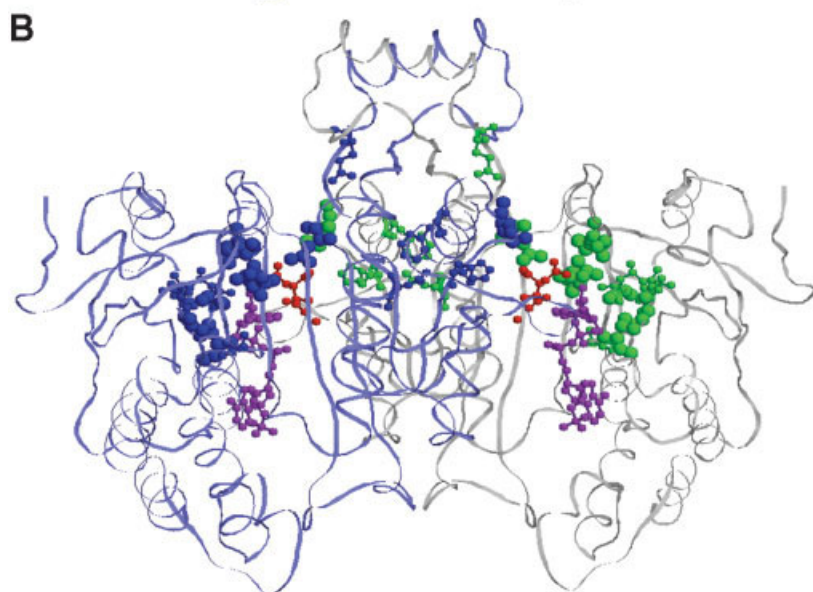
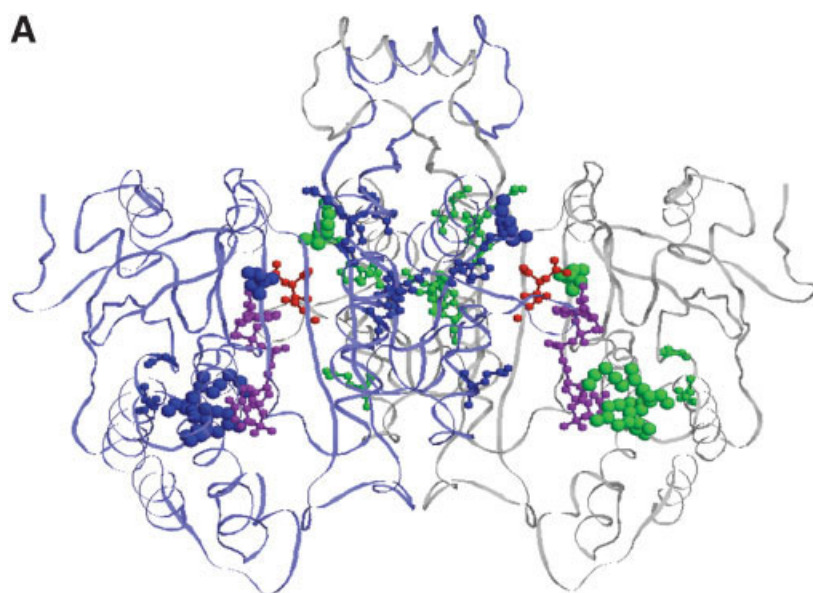


TABLE I. Predicted SDPs. (1) Amino acid in NADP-IDH from *E.coli* (P08200). (2) z-score in cofactor-specific comparison (blank if not identified). (3) z-score in substrate-specific comparison (blank if not identified). (4) z-score in four group comparison (blank if not identified). (5) Contacts in 3D structure 1ai2: C – with cofactor, S – with substrate, O – with the other subunit; capital letter: <5Å, lower case letter: <7Å. (6) Comments: experimentally shown determinants of specificity towards (arrow) cofactors (NAD, NADP) and substrates (IC, isocitrate; IPM, isopropylmalate) in proteins from *Escherichia coli* (Ec), *Thermus thermophilus* (Tt), *Thermoplasma ferrooxidans* (Tf), human cytosole (Hcyt), pig mitochondria (Pmit).

(1)	(2)	(3)	(4)	(5)	(6)
31Tyr			4.93		
36Gly			4.91	c	
38Gly	6.06		6.16		
40Asp	6.15		5.55		
45Met			5.66		
97Val		7.51			
98Ala		6.25			
100Lys		9.04		Cs	NADP-IDH_Hcyt (72Lys) → NADP [29]
103Leu		6.47	5.88	CS	
104Thr			7.1	Cs	
105Thr		9.82		Cso	NADP-IDH_Hcyt (77The) → IC, NADP; NADP-IDH_Pmit (86Thr) → IC [29,34]
107Val			3.95	O	
115Asn		7.35	6.03	CS	Key specificity determinant [1,8,9,11,14,16,31,31,32]
143(E)			5.34		Fifth gap after 143His in the alignment
152Phe			7.73		
154Glu	5.94		5.95	o	
155Asn		8.47	7.94	s	IMDH_Tt (134Leu) → IPM [5]
158Asp	12.55		5.47	o	
161Ala			6.05	O	IMDH_Tt (140Phe) → NAD [5]
162Gly			4.71	O	
164Glu		6.08	4.67	O	
208Arg	5.7			O	
229His	5.9	8.98	7.96	o	
231Gly	6.43		6.13	o	
232Asn			4.8	csO	NADP-IDH_Hcyt (72Lys) → IC [34]
233Ile	6.56	15.17	6.42	SO	IMDH_Tf (193Val) → IPM [30]
241Phe		12.02	5.76		
245Gly			4.6		
287Gln	5.95		6.04	csO	
300Ala	6.64				
305Asn	7.96	8.8	10.03	o	
308Tyr	6.5		6.09	sO	
323Ala			6.69		
327Asn	5.83	7.61	7.88	s	IMDH_Tt (261Ser) → IPM [5]
337Ala		6.3		Cs	
341Thr		8.77		C	NADP-IDH_Hcyt (311Thr) → NADP [34]
344Lys	8.03		6.73	C	IMDH_Tt (278Asp) → NAD [4,7,14]
345Tyr	7.62		5.44	C	NADP-IDH_Ec → NADP; IMDH_Tt (279Ile) → NAD [4,7,14]
351Val	5.65		5.37	C	NADP-IDH_Ec → NADP; IMDH_Tt (285Ala) → NAD [4,7,14]

function in general. Thus, some essential amino acids, which serve for the general biochemical function and are conserved in the whole protein family, are not identified by SDPpred. The literature mining revealed the total of 52 positions, which could be implicated in the function. We consider these

positions as experimentally proven to be biologically important. Among them, 11 are predicted as SDPs, 9 are at least 91% conserved, and 29 are within 5 Å distance from either a SDP or a conserved residue. This leaves only three positions, which can be considered as true false negatives. When we restrict the definition of important positions to those, which have been specifically described as influencing the enzyme specificity either towards substrate or cofactor, we obtain 14 positions, 5 of which are predicted as SDPs and 8 lie within 5 Å of SDPs. Naturally, there are no conserved positions in this set, and this leaves only one position as a false negative. In Figure 5(C), one can see that the fraction of false negatives

Fig. 3. Cofactor-specific (A), substrate-specific (B), and four group (C) SDPs mapped onto the structure of IDH from *E. coli* (PDB code: 1ai2). Substrate (isocitrate calcium complex) is shown in red; cofactor (NADP) is shown in magenta; predicted SDPs are shown in green, ball-and-stick model. The second subunit of the dimer is in blue. Experimentally supported SDPs are shown in spacefill model.

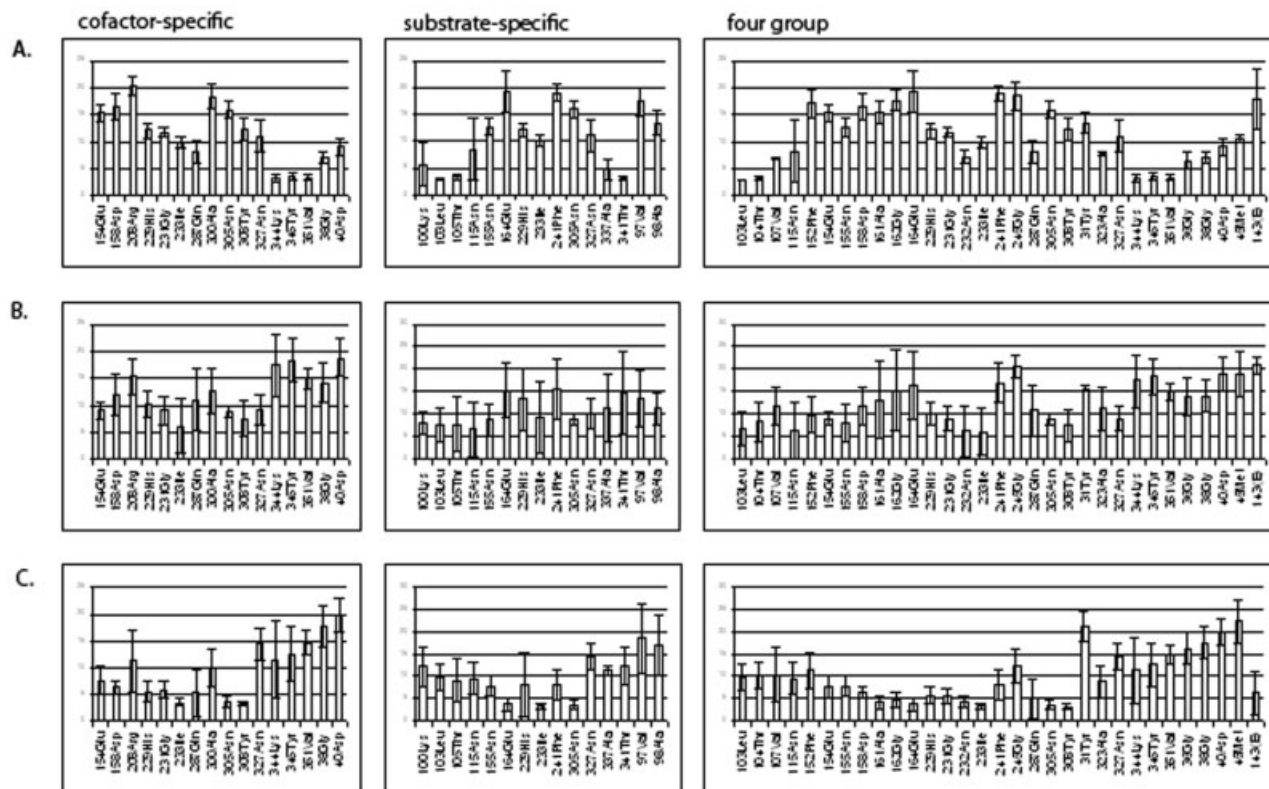


Fig. 4. Distances to the cofactor (A), the substrate (B), and the other subunit (C) in different IDH and IMDH structures. Vertical axis: averages and standard deviations.

(Other) and positions within 5 Å (Neighbors) are similar in the initial and restricted sets of experimentally proven positions, whereas among the positions not described as important, the fraction of neighbors is slightly lower and the fraction of distal positions is much higher. These observations indicate that not only positions with peculiar statistical properties (SDPs or conserved) can be implicated in function or specificity, but residues in the vicinity of such amino acids can be of interest for an experimentalist.

Among the three false negatives, 394Arg was described as one of amino acids most important for the cofactor binding.^{4,8} From the six amino acids, 292Arg, 344Lys, 345Tyr, 351Val, 391Tyr, and 394Arg, known to be crucial for the specificity towards the cofactor, five are shown to contribute additively to enzyme preference for NAD or NADP.²⁸ Three of them (344Lys, 345Tyr, and 351Val) were predicted as SDPs, 292Arg and 391Tyr are well conserved in NADP-dependent enzymes but not conserved in NAD-dependent enzymes. 394Arg, not studied by Lunzer and colleagues,²⁸ but known to contribute to the cofactor-binding site,^{4,8} lies in a poorly aligned region, which has a different 3D structure in NAD- and NADP-dependent enzymes. The remaining two false negatives are 110Gly and 135Gln, which correspond to 82Arg and 109Arg in human cytosolic NADP-dependent IDH and contact with either NADP (82Arg) or isocitrate (109Arg) in its crystal structure.²⁹

Apart from 344Lys, 345Tyr, and 351Val, several other amino acids contact cofactor in most considered structures: substrate-specific residues 103Leu, 105Thr, 337Ala, and 341Thr contact the nicotinamide nucleotide and thus spatially lie between the cofactor-binding and the substrate-binding pockets. They may act as primary recognition site for the substrate. There are no experimental data on the physiological role of the latter residues, which makes them a promising target for mutagenesis studies. The distance to the substrate varies greatly for cofactor-specific SDPs, but among the substrate-specific and the four group SDPs, there are several consistently contacting positions, some of which are known specificity determinants (115Asn, 232Asn, 233Ile) and some are new ones (103Leu). Analysis of the intersubunit contacts leads to identification of two spatial clusters of contacting SDPs, one formed by 305Asn and 308Tyr from both subunits (305Asn and 308Tyr from different subunits may possibly make a hydrogen bond between the asparagine oxygen and the tyrosine hydroxyl group), and the other by 107Val and 164Glu from one subunit and 232Asn and 233Ile from the other (Fig. 6).

By definition, the predicted SDPs are positions, whose evolution rate is unusual given the phylogenetic tree. We assume that such abnormalities reflect adaptation to different interacting molecules and thus are true specificity determinants. They are highly influenced by the adaptation to changes in specificity that occur after gene

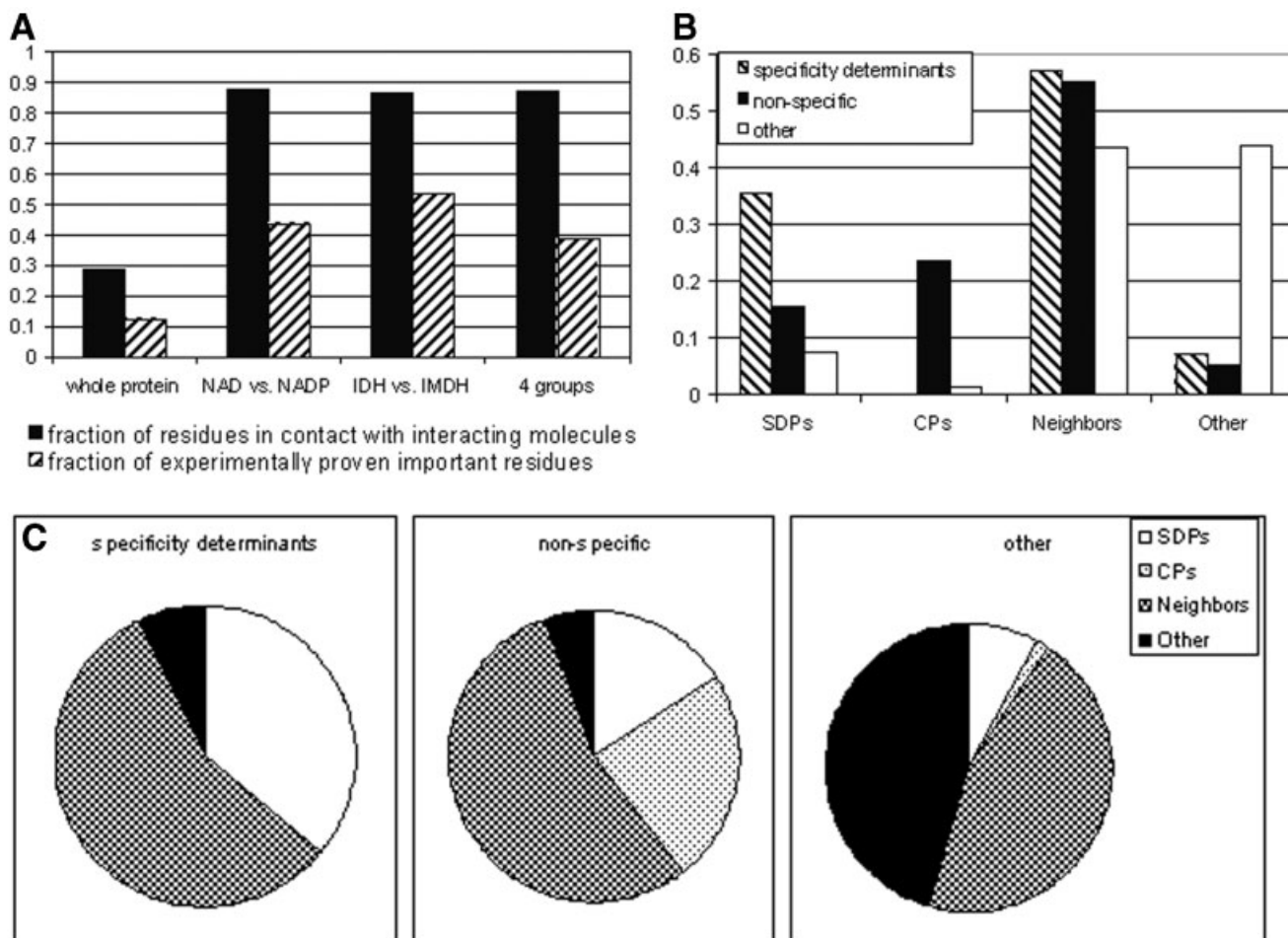


Fig. 5. (A) Fraction of amino acid residues in close contact ($< 7 \text{ \AA}$) with an interaction partner and experimentally proven to be important for function in protein on average and in each comparison. Experimental data collected for all proteins of the family, contacts calculated for IDH from *E. coli*, 1ai2. (B) Fraction of amino acid residues of different functional types: *specificity determinants* are described in the literature as essential for the enzyme specificity; *nonspecific* are mentioned as related to function or contacting substrate or cofactor (specificity determinants excluded); *other* are the remaining residues. Prediction classes: SDPs, predicted as SDPs; CPs, conserved positions (at least 91% conservation in the family alignment); Neighbors, amino acids in close contact ($< 5 \text{ \AA}$) with SDPs or CPs; Other, remaining residues. Contacts were calculated for IDH from *E. coli*, 1ai2. (C) Fraction of amino acids belonging to different prediction classes, for different functional types of residues (notation as in B).

duplication. Indeed, almost all SDPs in each comparison are located in the zone of possible interaction with the cofactor, the substrate, or the other subunit, and some have been experimentally identified as specificity determinants.

Still, in each case there are SDPs, to which no function can be attributed, because in all considered structures they are far from the sites of contact and they have never been reported to be important for the function in literature. There are two such residues in the cofactor-specificity comparison (33Asp and 304Ala), two in the substrate-specificity comparison (97Val and 98Ala), and five in the four-group comparison (31Tyr, 40Asp, 45Met, 152Phe, and 245Gly). We formally consider them as overprediction, or false positives.

However, even in cases when the specificity of an enzyme to the substrate or to the cofactor has been successfully inverted,^{4,7,9,14,15,17} the kinetic parameters of the mutants are inferior compared to the wild type, while

substitutions of residues distant from any compound may improve the kinetics.^{7,17} Such substituted residues are often close to some of the predicted SDPs. For example, 154Glu, 158Asp, and 327Asn are located within 7 \AA from 201Cys and 332Cys, which enhance the NAD-specificity in the coenzyme-specificity inverted mutant.^{7,17} These and other predicted SDPs may thus play a role in the mechanism of cofactor recognition, and so we propose that some apparently false-positive SDPs may in fact be important for the specific recognition.

It is impossible to separate to separate NAD-dependent from NADP-dependent enzymes by cutting the family phylogenetic tree (Fig. 1) at a single point, and thus the specificity to NADP as the cofactor may have arisen twice.⁴ Closer inspection of amino acid residues occupying SDPs in the comparison by the cofactor specificity indicates that indeed the adaptation to NADP may have taken different routes. There are very few SDPs, in which amino acids in type I and type II NADP-IDHs coincide.

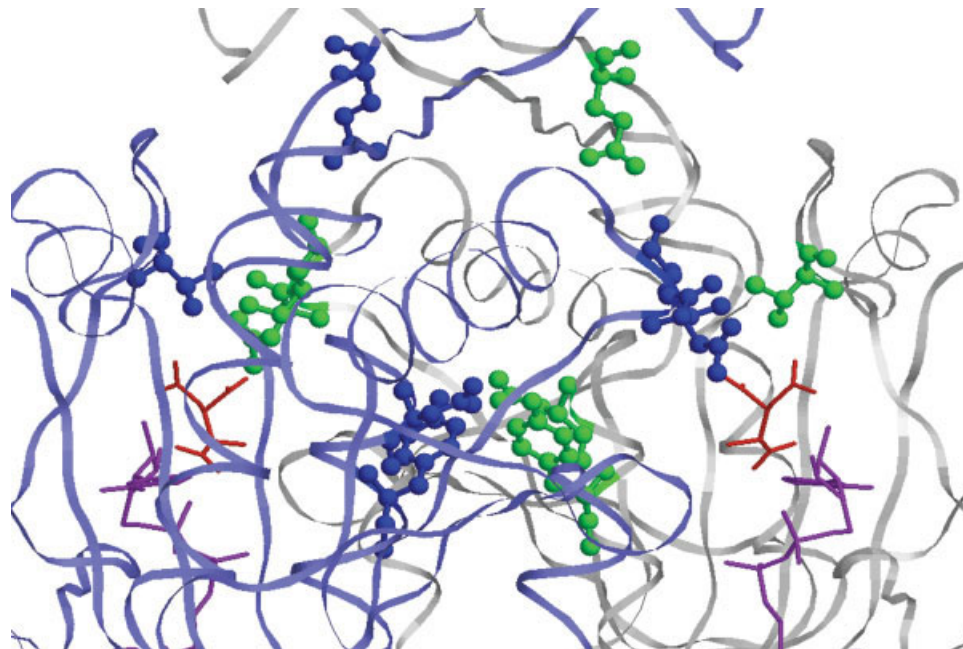


Fig. 6. Clusters of SDPs on the intersubunit contact interface. Notation as in Figure 3. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

CONCLUSIONS

We present a computational analysis of the specificity-determining positions in the IDH/IMDH family. Because the enzymes of the family have two specificities, towards the substrate and towards the cofactor, which appear in different combinations, we performed three comparisons (substrate-specific, cofactor-specific, and combined). The predicted in all three comparisons SDPs agree well with the experimental and structural data: 40% to 50% of the predicted SDPs were previously shown to contribute to protein function in experimental assays, and up to 90% lie in regions of contact with either cofactor, or substrate, or the other subunit, a fraction significantly higher than in the whole protein.

Still, a number of experimentally proven important positions were not identified in our study. Some of them are totally or nearly totally conserved, and most of the rest lies in the vicinity of either SDPs or conserved positions, which suggests their possible indirect role in the enzyme function. In the cofactor-binding pocket, two experimentally known important positions are conserved among NADP-dependent, but variable among NAD-dependent enzymes, which implies that they are crucial for recognition of NADP, but not of NAD.

On the other hand, we identify a number of new interesting mutation targets. Experimental assays may test the role of 103Leu, 105Thr, 337Ala, and 341Thr in substrate recognition and of 305Asn, 308Tyr or 107Val, 164Glu in dimerization and active site formation.

Interestingly, in all three comparisons, about half of the SDPs concentrate on the intersubunit interface. Sometimes SDPs are arranged in spatial clusters. We have seen similar clusters in other protein families.^{30,31} We propose

that these positions provide for the correct recognition of the partner subunit to exclude the possibility of formation of chimeric complexes. Another possible explanation for the studied enzyme family is that the active site is formed by residues from both subunits, and so some of the intersubunit contacts are actually the site residues. Anyway, a number of intersubunit contact SDPs lie at large distance from the binding sites. The function of these and other predicted SDPs awaits experimental elucidation.

ACKNOWLEDGMENTS

We are grateful to Aleksandra Rakhmaninova and Andrei Mironov for many useful discussions and to anonymous reviewers for valuable comments.

REFERENCES

1. Cozzone AJ. Regulation of acetate metabolism by protein phosphorylation in enteric bacteria. *Annu Rev Microbiol* 1998;52:127–164.
2. Hurley JH, Thorness PE, Ramalingam V, Helmers NH, Koshland DE Jr, Stroud RM. Structure of a bacterial enzyme regulated by phosphorylation, isocitrate dehydrogenase. *Proc Natl Acad Sci U S A* 1989;86:8635–8639.
3. Cupp JR, McAlister-Henn L. NAD(+)-dependent isocitrate dehydrogenase. Cloning, nucleotide sequence, and disruption of the IDH2 gene from *Saccharomyces cerevisiae*. *J Biol Chem* 1991;266:22199–22205.
4. Zhu G, Golding GB, Dean AM. The selective cause of an ancient adaptation. *Science* 2005;307:1279–1282.
5. Zhang T, Koshland DE Jr. Modeling substrate binding in *Thermus thermophilus* isopropylmalate dehydrogenase. *Protein Sci* 1995;4:84–92.
6. Imada K, Sato M, Katsube Y, Matsuura Y, Oshima T. Three-dimensional structure of a highly thermostable enzyme, 3-isopropylmalate dehydrogenase of *Thermus thermophilus* at 2.2 Å resolution. *J Mol Biol* 1991;222:725–738.
7. Chen R, Greer A, Dean AM. A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. *Proc Natl Acad Sci U S A* 1995;92:11666–11670.

8. Dean AM, Shiau AK, Koshland DE Jr. Determinants of performance in the isocitrate dehydrogenase of *Escherichia coli*. *Protein Sci* 1996;5:341–347.
9. Doyle SA, Fung S-YF, Koshland DE Jr. Redesigning the substrate specificity of an enzyme: isocitrate dehydrogenase. *Biochemistry* 2000;39:14348–14355.
10. Steen IH, Lien T, Madsen MS, Birkeland N-K. Identification of the cofactor discrimination sites in NAD-isocitrate dehydrogenase from *Pyrococcus furiosus*. *Arch Microbiol* 2002;178:297–300.
11. Fujita M, Tamegai H, Eguchi T, Kakanuma K. Novel substrate specificity of designer 3-isopropylmalate dehydrogenase from *Thermus thermophilus* HB8. *Biosci Biotechnol Biochem* 2001;65:2695–2700.
12. Lee P, Colman R. Implication by site-directed mutagenesis of Arg314 and Tyr316 in the coenzyme site of pig mitochondrial NADP-dependent isocitrate dehydrogenase. *Arch Biochem Biophys* 2002;401:81–90.
13. Soundar S, Park J-H, Huh T-L, Colman RF. Evaluation by mutagenesis of the importance of 3 arginines in alpha, beta and gamma subunits of human NAD-dependent isocitrate dehydrogenase. *J Biol Chem* 2003;278:52146–52153.
14. Dean AM, Golding GB. Protein engineering reveals ancient adaptive replacements in isocitrate dehydrogenase. *Proc Natl Acad Sci U S A* 1997;94:3104–3109.
15. Chen R, Greer A, Dean AM. Redesigning secondary structure to invert coenzyme specificity in isopropylmalate dehydrogenase. *Proc Natl Acad Sci U S A* 1996;93:12171–12176.
16. Doyle SA, Beernink PT, Koshland DE Jr. Structural basis for a change in substrate specificity: crystal structure of S113E isocitrate dehydrogenase in a complex with isopropylmalate, Mg²⁺ and NADP. *Biochemistry* 2001;40:4234–4241.
17. Hurley JH, Chen R, Dean AM. Determinants of cofactor specificity in isocitrate dehydrogenase: structure of an engineered NADP⁺→NAD⁺ specificity-reversal mutant. *Biochemistry* 1996;35:5670–5678.
18. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997;25:4876–4882.
19. Felsenstein J. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol* 1996;266:418–427.
20. Kalinina OV, Novichkov PS, Mironov AA, Gelfand MS, Rakhmaninova AB. SDPpred: a tool for prediction of amino acid residues that determine differences in functional specificity of homologous proteins. *Nucleic Acids Res* 2004;32:W424–W428.
21. Lichtarge O, Bourne HR, Cohen FE. An evolutionary trace method defines binding surfaces common to protein families. *J Mol Biol* 1996;257:342–358.
22. Lichtarge O, Yamamoto KR, Cohen FE. Identification of functional surfaces of the zinc binding domains of intracellular receptors. *J Mol Biol* 1997;274:325–337.
23. Hannenhalli SS, Russell RB. Analysis and prediction of functional sub-types from protein sequence alignments. *J Mol Biol* 2000;303:61–76.
24. Mirny LA, Gelfand MS. Using orthologous and paralogous proteins to identify specificity-determining residues in bacterial transcription factors. *J Mol Biol* 2002;321:7–20.
25. Valencia A, Pazos F. Computational methods for the prediction of protein interactions. *Curr Opin Struct Biol* 2002;12:368–373.
26. Donald JE, Shakhnovich EI. Predicting specificity-determining residues in two large eukaryotic transcription factor families. *Nucleic Acids Res* 2005;33:4455–4465.
27. Bernstein HJ. Recent changes to RasMol, recombining the variants. *Trends Biochem Sci* 2000;25:353–355.
28. Lunzer M, Miller SP, Felsheim R, Dean AM. The biochemical architecture of an ancient adaptive landscape. *Science* 2005;310:499–501.
29. Xu X, Zhao J, Peng B, Huang Q, Arnold E, Ding J. Structure of human cytosolic NADP-dependent dehydrogenase reveals a novel self-regulatory mechanism of activity. *J Biol Chem* 2004;279:22946–33957.
30. Kalinina OV, Mironov AA, Gelfand MS, Rakhmaninova AB. Automated selection of positions determining functional specificity of proteins by comparative analysis of orthologous groups in protein families. *Protein Sci* 2004;13:443–456.
31. Kalinina OV, Gelfand MS, Mironov AA, Rakhmaninova AB. Amino acid residues forming specific contacts between subunits in tetramers of the membrane channel GlpF. *Biophysics (Moscow)* 2003;48:S141–S145.
32. Imada K, Inagaki K, Matsunami H, Kawaguchi H, Tanaka H, Tanaka N, Namba K. Structure of 3-isopropylmalate dehydrogenase at 2.0 Å resolution: the role of Glu88 in the unique substrate-recognition mechanism. *Structure* 1998;6:971–982.
33. Dean AM, Dvorak L. The role of glutamate 87 in the kinetic mechanism of *Thermus thermophilus* isopropylmalate dehydrogenase. *Protein Sci* 1995;4:2156–2167.
34. Miyazaki K, Yaoi T, Oshima T. Expression, purification, and substrate specificity of isocitrate dehydrogenase from *Thermus thermophilus* HB8. *Eur J Biochem* 1994;221:899–903.
35. Ceccarelli C, Grodsky NB, Ariyarante N, Colman RF, Bahson BJ. Crystal structure of the porcine mitochondrial NADP-dependent isocitrate dehydrogenase complexed with Mn²⁺ and isocitrate. *J Biol Chem* 2002;277:43454–434362.