Protein Science (1995), *4*:84–92. Cambridge University Press. Printed in the USA.

Modeling substrate binding in *Thermus thermophilus* isopropylmalate dehydrogenase

TAO ZHANG AND DANIEL E. KOSHLAND, JR.

Department of Molecular and Cell Biology, Division of Biochemistry, University of California, Berkeley, California 94720

(RECEIVED August 9, 1994; ACCEPTED October 25, 1994)

Abstract

The *Thermus thermophilus* 3-isopropylmalate dehydrogenase (IPMDH) and *Escherichia coli* isocitrate dehydrogenase (ICDH) are two functionally and evolutionarily related enzymes with distinct substrate specificities. To understand the determinants of substrate specificities of the two proteins, the substrate and coenzyme in IPMDH were docked into their respective binding sites based on the published structure for apo IPMDH and its sequence and structural homology to ICDH. This modeling study suggests that (1) the substrate and coenzyme (NAD) binding modes of IPMDH are significantly different from those of ICDH, (2) the interactions between the substrates and coenzymes help explain the differences in substrate specificities of IPMDH and ICDH, and (3) binding of the substrate and coenzyme should induce a conformational change in the structure of IPMDH.

Keywords: docking; isocitrate dehydrogenase; isopropylmalate dehydrogenase; substrate binding

Thermus thermophilus 3-isopropylmalate dehydrogenase (IPMDH) and *Escherichia coli* isocitrate dehydrogenase (ICDH) are two closely related enzymes that catalyze similar reactions in different metabolic pathways. IPMDH functions in the leucine biosynthesis pathway and catalyzes the oxidative decarboxylation of isopropylmalate to α-ketoisocaproate (Imada et al., 1991). ICDH catalyzes a similar decarboxylation reaction in the carbohydrate metabolism (Kornberg, 1966), and converts isocitrate to α-ketoglutarate. The two reactions are shown in Figure 1A. Both reactions are postulated to proceed in two steps, with dehydrogenation preceding decarboxylation (Fig. 1B) (Grissom & Cleland, 1985; Pirrung et al., 1994).

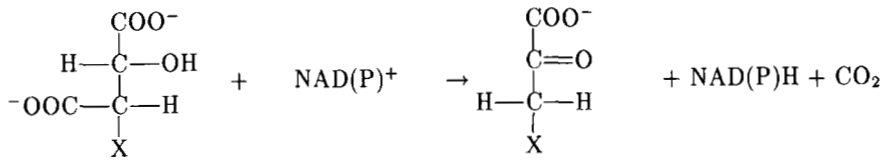
Most of the postulated key catalytic residues in *E. coli* ICDH (Hurley et al., 1991) are conserved in *T. thermophilus* IPMDH (Imada et al., 1991; Miyazaki et al., 1992). The two proteins have an overall 30% sequence identity (Fig. 2). The conservation of catalytic residues also extends to the IPMDHs and ICDHs of other species (Fig. 2), suggesting that these proteins may have a common evolutionary origin and structural, functional properties.

The substrate specificities of *T. thermophilus* IPMDH and *E. coli* ICDH are, however, very different. ICDH shows no detectable activities for isopropylmalate and IPMDH shows no activity for isocitrate (Miyazaki et al., 1993). A mammalian form of ICDH has also been extensively investigated (Colman, 1973; Ehrlich & Colman, 1978). Miyazaki et al. (1993) showed that *T. thermophilus* IPMDH has broad substrate specificities against alkyl-malates. Substituting the γ-moiety of isopropylmalate with methyl, ethyl, or propyl groups has little effect on the overall catalytic efficiency (k_{cat}/K_m), suggesting that the hydrophobic packing between the substrate and the enzyme-coenzyme complex is not very specific. On the other hand, IPMDH shows no detectable activity against isocitrate, suggesting that the polar or charge-charge interaction with the γ-carboxyl group destabilizes the IPMDH-substrate-coenzyme complex (Miyazaki et al., 1993). Studies on the roles of catalytic residues of IPMDH are limited. It was shown in a recent study that substituting Tyr 139 with phenylalanine resulted in a 10-fold decrease in the K_m of the coenzyme NAD but little effect on the K_m of the substrate (Miyazaki et al., 1993). The same mutation also resulted in a 10-fold decrease in the turnover number (k_{cat}) and therefore did not change the overall catalytic activity at low substrate concentration. However, a similar change in the corresponding residue in ICDH (Tyr 160) resulted in a different change in the k_{cat} and K_m of isocitrate, suggesting that the two residues may play different roles in substrate binding and catalysis. Therefore, despite the sequence homology between IPMDH and ICDH, the two enzymes may have differences in their substrate binding and catalytic groups.

To interpret these distinct mechanisms and specificities, it is desirable to know the binding site of isopropylmalate to the IPMDH. To do so, the apo structure of IPMDH was aligned to the structure of ICDH based on the sequence homology (Fig. 2). This structural alignment reveals a general structural homology between the two proteins and suggested the possible

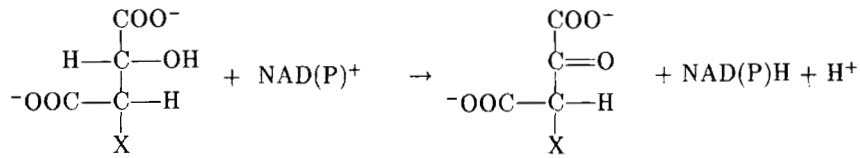
Reprint requests to: Daniel E. Koshland, Jr., Department of Molecular and Cell Biology, Division of Biochemistry, 229 Stanley Hall, University of California, Berkeley, California 94720; e-mail: daniel_koshland@maillink.berkeley.edu.

A



Where X= $C(CH3)_2$ (giving IPMDH and NAD^+)
CH_2COO^- (giving ICDH and $NADP^+$)

B



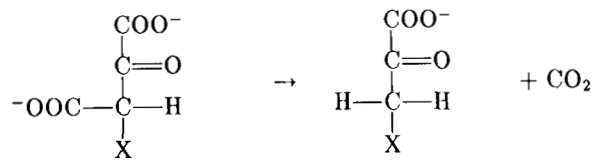


Fig. 1. A: Reactions catalyzed by *T. thermophilus* IPMDH (Imada et al., 1991) and *E. coli* ICDH (Kornberg, 1966). X = $C(CH_3)_2$ gives IPMDH and NAD^+, and X = CH_2COO^- gives ICDH and $NADP^+$. **B:** Both reactions are postulated to proceed in two steps, with dehydrogenation preceding decarboxylation (Grissom & Cleland, 1985; Pirrung et al., 1994).

substrate binding site of IPMDH. The isopropylmalate and NAD molecules were then docked to the protein surface of IPMDH by an automated docking method (Goodsell & Olson, 1990) to determine the possible substrate binding mode.

Results and discussion

Structural alignment of IPMDH and ICDH

To align structures of IPMDH and ICDH, we selected residues from each protein to define the transformation matrix that superimposes the two structures. The selection of these residues was based on (1) functional importance in the active sites and (2) evolutionary conservation among different species. The seven active site residues of *E. coli* ICDH, R129, R153, Y160, K230′, D283′, D307, and D311 were used (a prime, ′, indicates the residue of the second subunit). The structural and functional importance of these residues was supported by their conservation in ICDHs from *E. coli* to mammals (Fig. 2). Sequence alignment of the ICDH and IPMDH showed the corresponding residues in IPMDH were: R104, R132, Y139, K185′, D217′, D241, and D245. These residues are also conserved in IPMDHs from different species (Fig. 2).

The atoms of the selected IPMDH residues were rigidly fitted to those corresponding atoms of ICDH by rotation and translation to minimize the RMS deviations between the coordinates of the two sets of atoms. The resulting transformation matrix was then used to align the IPMDH structure to that of ICDH. Views of the ICDH active site and the corresponding region of the aligned IPMDH are shown in Figure 3. The RMS deviation among the selected active site residues in the aligned structures is about 0.8 Å.

In previous studies, the two proteins were shown to have similar sequence and topology of secondary structures (Imada et al., 1991). The structural alignment shows that the tertiary structural homology between the two proteins is even more striking. Both proteins form functional dimers with the two-fold symmetry axes at the center of the dimer interfaces. Each dimeric complex contains two outside domains, one central domain, and two interdomain pockets. It was shown in the alignment that the two central domains from the two proteins were virtually superimposable. For example, the four-helix-bundles of the two proteins align well with each other at the dimer interface. So do other helices and β-strands in the central domains of the two proteins. The central and outside domains of the two proteins are connected by a group of β-strands, which also form the bottom of the interdomain pockets of the two proteins. These β-strands are superimposable in the two aligned structures. Although the two outside domains of the two proteins have similar structures, they were displaced from each other in the structural alignment. This displacement seems a result of a collective interdomain movement, such as a hinge motion. A tentative interdomain hinge motion in IPMDH may reduce the size of its interdomain pocket to that of ICDH's and, at the same time align its outside domain better with that of ICDH (Fig. 3A,B).

The structural alignment and homology suggest that the enzyme active site of IPMDH is located in its interdomain pocket. This is similar to ICDH, where the catalytic residues are located on the α-helices and β-strands that form the sides and bottom of the interdomain pocket. The corresponding IPMDH residues are virtually superimposable on these ICDH catalytic residues (Fig. 3C). The side chains of R129 and 153 in ICDH, for example, took similar spatial positions as those of R104 and 132 in IPMDH. The aspartates in the active sites of IPMDH (D217′,

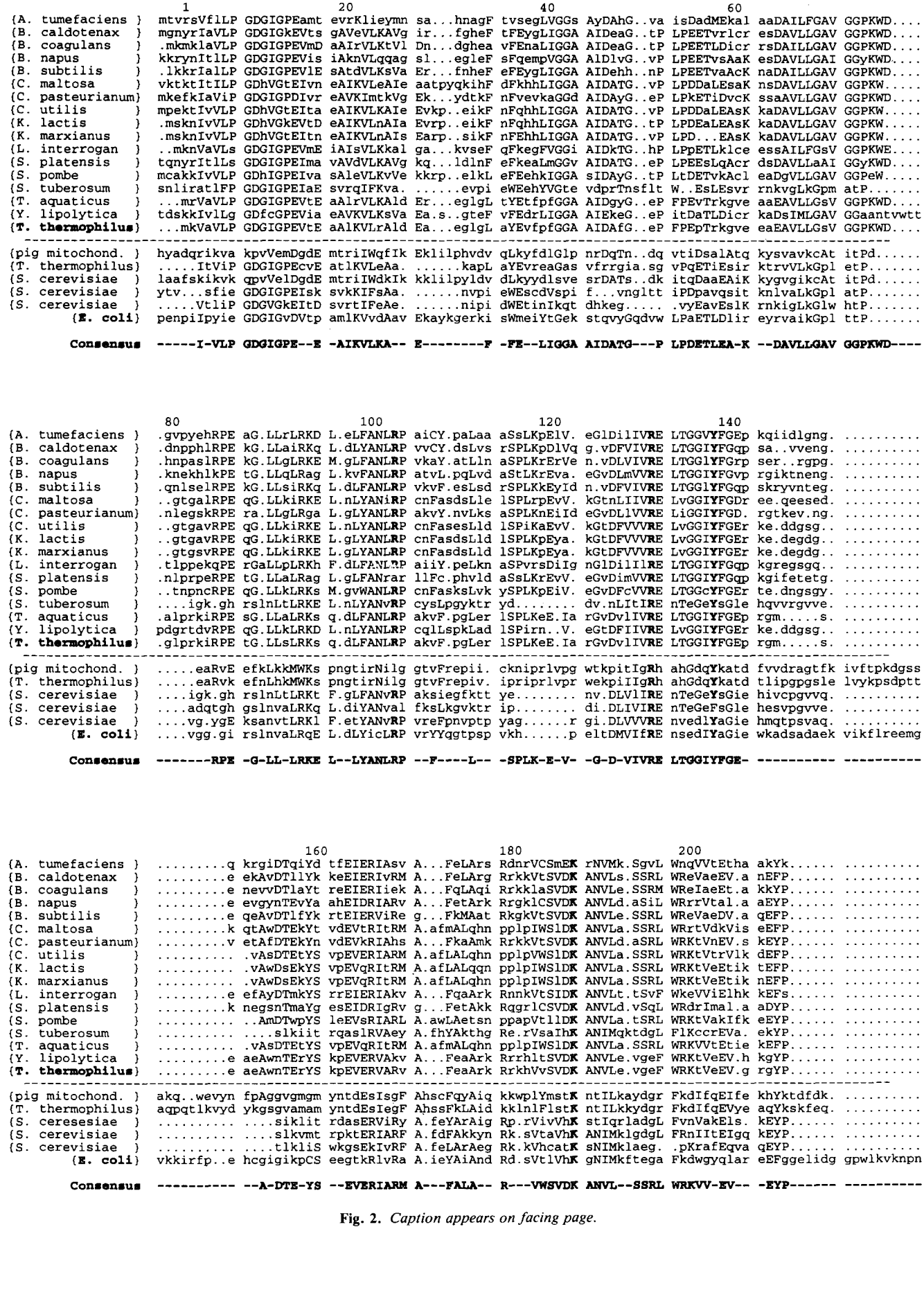


Fig. 2. *Caption appears on facing page.*

```
                          220                 240                 260                 280
{A. tumefaciens }  ...DVqLeHm LaDaggMQLV rkPkQF..DV IVTdNLFGDm LSDv.Aamlt GSLGMLPSAS Lga.pdaktg krkaMYEPVH GSAPDIAGKs
{B. caldotenax  }  ...DVtLeHm LVD.mrMQLI raPkQF..DV IVTeNMFGDI LSDE.ASmls GSLGMLPSAS Lsa.sgp... ...sLYEPVH GSAPDIAGmn
{B. coagulans   }  ...DVeLsHm LVDstsMQLI anPgQF..DV IVTeNMFGDI LSDE.ASvIt GSLGMLPSAS Lrs.drf... ...GMYEPVH GSAPDIAGqg
{B. napus       }  ...DVeLsHm yVDnAAMQLV rdPkQF..Dt IVTnNiFGDI LSDE.ASmIt GSiGMLPSAS Lsd.sgp... ...GLFEPIH GSAPDIAGqd
{B. subtilis    }  ...DVkLeHm LVDnAAMQLI yaPnQF..DV VVTeNMFGDI LSDE.ASmlt GSLGMLPSAS Lss.sgl... ...hLFEPVH GSAPDIAGKg
{C. maltosa     }  ...tlsvqHq LIDsAchdFn skP....nqI ewynyhF*.. .......... .......... .......... .......... ..........
{C. pasteurianum}  ...EVeLsHl yVDntAMQLV rkPsQF..DV IlTnNiFGDI LSDE.ASmIt GSiGMLaSsS ire.dsf... ...GMYEPIH GSAPDIAGld
{C. utilis      }  ...qltvqHq LIDsAAMiLI KyPtQL.ngI VITsNMFGDI iSDE.ASvIp GSLGLLPSAS Laslpds..n kafGLYEPcH GSAPDl.pan
{K. lactis      }  ...qltvqHq LIDsAAMiLV KsPtkL.ngV VITnNMFGDI iSDE.ASvIp GSLGLLPSAS Laslpdt..n kafGLYEPcH GSAPDl.pan
{K. marxianus   }  ...qltvqHq LIDsAAMiLV KsPtkL.ngI VITnNMFGDI iSDE.ASvIp GSLGLLPSAS Laslpdt..n kafGLYEPcH GSAPDl.pan
{L. interrogan  }  ...DVqLnHl yVDnAAMQLI vnPkQF..DV VlceNMFGDI LSDE.ASiIt GSiGMLPSAS Lse.sgf... ...GLYEPsg GSAPDIAGKg
{S platensis    }  ...EVeLsHl yVDnAAMQLV rwPkQF..Dt IVTgNLFGDI LSDa.Aamlt GSiGMLPSAS Lga.sgp... ...GvFEPVH GSAPDIAGqd
{S. pombe       }  ...hltLknq LIDsAAMlLV KsPrtL.ngV VlTdNLFGDI iSDE.ASvIp GSLGLLPSAS Lsgvvgksee kvhcLvEPIH GSAPDIAGKg
{S. tuberosum   }  ...EIkyeev vIDnccMmLV KnPalF..DV lVmpNLYGDI iSDlcAgl.i GgLGLtPScn ig.......e ggiaLaEaVH GSAPDIAGKn
{T. aquaticus   }  ...DVaLdHq yVDamAMhLV KnParF..DV VVTgNiFGDI LSDl.ASvlp GSLGLLPSAS L.g.rgt... ...pvFEPVH GSAPDIAGKg
{Y. lipolytica  }  ...qleLnHq LIDsAAMiLI KqPskm.ngI IITtNMFGDI iSDE.ASvIp GSLGLLPSAS Laslpdt..n eafGLYEPcH GSAPDl.GKq
{T. thermophilus}  ...DVaLeHq yVDamAMhLV rsParF..DV VVTgNiFGDI LSDl.ASvlp GSLGLLPSAS L.g.rgt... ...pvFEPVH GSAPDIAGKg
  ------------------------------------------------------------------------------------------------------------------
{pig mitochond. }  ..ykIwyeHr LIDdmvaQvl KssggF...V wackNydGDV qSDilAqgf. GSLGLMtSvl vcpdgktiea eaahgtvtrH yrehqkgrpt
{T. thermophilus}  ..lgIhyeHr LIDdmvaQMI KskggF...I malkNydGDV qSDivAqgf. GSLGLMtSil vtpdgktfes eaahgtvtrH yrkyqkgeet
{S. cerevisiae  }  ...DltLete LIDnsvLkvV tnPsaYtdaV sVcpNLYGDI LSDlnsglsa GSLGLtPSAn ig.......h .kisiFEaVH GSAPDIAGqd
{S. cerevisiae  }  ...DIdvssi iVDnAsMQaV akPhQF..DV lVTpsMYGtI LgnigAal.i GgpGLvagAn Fg.......r dyavFepgsr hvglDIkGqn
{S. cerevisiae  }  ...DIeavHi iVDnAAhQLV KrPeQF..EV IVTtNMnGDI LSDltsgl.i GgLGFaPSAn ig.......n e.vaiFEaVH GSAPkyAGKn
        {E. coli}  tgkEIvikdv iaDaflqQil lrPaeY..DV IacmNLnGDy iSDalAa.qv GgiGiaPgAn ig.......d e.caLFEatH GtAPkyAGqd

      Consensus    ---DV-L-H- L-D-AAMQLV K-P-QF--DV IVT-NMFGDI LSDE-AS-I- GSLGLLPSAS L--------- ---GLYEPVH GSAPDIAGK-
```

```
                          300                 320                 340
{A. tumefaciens }  iANPIAmIa. .SfAMcLRYS Fn....MvDE AtklEaAIan VLdk...... GirTaDimad gcrq....vg TsDMgDaVla eFkalsa*..
{B. caldotenax  }  KANPIAaIL. .SAAMMLRLS Fg....LtaE A.ggrarVwq aLal...... Gsgsrlgq.r rphl.....s TnEMvEeIka avldytaiaq
{B. coagulans   }  KANPlgtVL. .SAALMLRYS Fg....LekE AaaIEkAVdd VLqd...... GycTgDlqva ngkv....vs TiELtDrlie kLnnsaagpr
{B. napus       }  KANPlAtIL. .SAAMLLkYg Lg....eeka AkrIEdAVlg aLnk...... GfrTgDiysa gtkl....vg ckEMgEeVlk svdshvqasv
{B. subtilis    }  mANPfAaIL. .SAAMLLRtS Fg....LeEE AkaVEdAVnK VLas...... GkrTrDla.r seef....ss TqaitEeVka aimsentisn
{C. maltosa     }  .......... .......... .......... .......... .......... .......... .......... .......... ..........
{C. pasteurianum}  iANPlAqIL. .SAAMLMeYS Ld....MsEa ArdVEaAVeK VLne...... GyrTaDiyie gtkk....vg TsEMgnlVle rL*.......
{C. utilis      }  KvNPIAtIL. .SAAMMLkLS Ld....LyEE gvaVEtAVkq VLda...... GirTgDlkgt nstt....ev gdavaEaVkk iLa*......
{K. lactis      }  KvNPIAtIL. .SAAMMLkLS Ld....LvEE gralEeAVrn VLda...... GvrTgDlggs nstt....ev gdaiakaVke iLa*......
{K. marxianus   }  KvNPIAtIL. .SAAMMLkLS Ld....LvEE graVEeAVrK VLds...... GirTgDlggs nstt....ev gdavakaVke iLa*......
{L. interrogan  }  vANPIAqVL. .SAALMLRYS Fs....MeEE AnkIEtAVrK tias...... GkrTrDiaev gsti....vg TkEigqlIes fL*.......
{S. platensis   }  KANPlAqVL. .SAAMMLRYg Ld....epaa sdpVEkAVlK VLdw...... GypTgDimse gmka....vg crkwgicy.. ..........
{S. pombe       }  ivNPVgtIL. .SAsLLLRYg Ln....apkE AeaIEaAVrK VLddtsiggr GlyTrDlgge asta....di TkavvEelek iL*.......
{S. tuberosum   }  lANPtAllL. .SsvsMLRh. Le....LhDk AdrIqdAIlK tiaggkvpnw r..p...wrh cy.......n n*.......... ..........
{T. aquaticus   }  iANPtAaIL. .SAAMMLeha Fg....LvEl ArrVEaAVaK aLre...... tpppdlggsa .........g TqaFtEeVlr hL*.......
{Y. lipolitica  }  KvNPIAtIL. .SAAMMLkFS Ln....Mkpa gdaVEaAVke svea...... GitTaDiggs ssts....ev gdlLptrsrs csrrsksflr
{T. thermophilus}  iANPtAaIL. .SAAMMLeha Fg....LvEl ArkVEdAVaK aLle...... tpppdlggsa .........g TeaFtatVlr hLa.......
  ------------------------------------------------------------------------------------------------------------------
{pig mitochond. }  stNPIAsIFa wtrgLehRgk LdgnqdLirf AqtlEkvcve tv.esgamtk dlagcihgls nvklnehfln TsDFlDtIks nLdralgrq.
{T. thermophilus}  stNsIAsIFa wSrgLLkRge LdntpaLckf AnilEsAtln tvqqdgimtk dlala....c gnnersayvt TeEFlDaVek rLqkeiksie
{S. cerevisiae  }  KANPtAllL. .SsvMMLnh. Mg....Ltnh AdqIqnAVls tiasgpenrt G..d...lag ta.......t TssFtEaVik rL*.......
{S. cerevisiae  }  vANPtAmIL. .SstLMLnh. Lg....LnEy AtrIskAVhe tiaegk.htt r..d...igg ss.......s TtDFtneIin kLstm.....
{S. cerevisiae  }  viNPtAvlL. .SAvMMLRY. Le....efat AdlIEnAlly tLeegrvltg dvvg...ydr ga.......k TtEytEaIiq nLgktprktq
        {E. coli}  KvNPgsiIL. .SAeMMLRh. Mg....wtEa AdlIvkgmeg ai.naktvty dfer...lmd gakl....lk csEFgDaIie nM*.......

      Consensus    KANPIA-IL- -SAAMMLRYS L-----L-EE A--IE-AV-K VL-------- G--T-D---- ---------- T-EF-E-V-- -L--------
```

Fig. 2. Alignment of amino acid sequences of isopropylmalate dehydrogenases (IPMDH) and isocitrate dehydrogenases (ICDH). All sequences are retrieved from GenBank. Capital letters indicate that the residues are conserved in 10 of the 24 total sequences. The consensus sequence is also listed at the bottom of the aligned individual sequences. Sources of the IPMDH sequences and their corresponding GenBank codes are: *Amyloliquefaciens tumefaciens*, atumadh15; *Bacillus caldotenax*, bc3imd; *Bacillus coagulans*, bacipmd; *Bacillus napus*, bnipmdh; *Bacillus subtilis*, bsleuc; *Clostridium maltosa*, cmlebbid; *Clostridium pasteurianum*, cloleuilv; *Clostridium utilis*, ysaimdh; *Klebsiella lactis*, klklleu2g; *Klebsiella marxianus*, ekleu2g; *Leptospira interrogans*, lepleub; *Staphylococcus platensis*, spu3id; *Schizosaccharomyces pombe*, yspleu1a; *Staphylococcus tuberosum*, tbbisodeh; *Thermus aquaticus*, d10700; *Thermus thermophilus*, tthleub; *Yersinia lipolytica*, ysjleu2b. Sources of the ICDH sequences and their corresponding GenBank codes are: pig mitochond., pigmtnadp; *T. thermophilus*, tthisocitd; *Saccharomyces cerevisiae* (NADPH-specific), yscidp1; *S. cerevisiae*, yscidh2a; *S. cerevisiae* (mitochond.), yscisodh; *E. coli*, ecoicd.

241 and 245) and ICDH (D283′, 307 and 311) also align with each other. The active site residues Y160 and K230′, which primarily interact with the C6 carboxyl oxygens of isocitrate in ICDH, superimpose with the corresponding Y139 and K185′ in IPMDH.

Structural alignment of the two proteins also reveals their distinctive features. First, the interdomain pocket of IPMDH that contains the active site is significantly larger than that of ICDH (Fig. 3A,B). No large binding-induced conformational changes were observed in the co-crystal structure of ICDH with isocitrate and NADP except for some minor conformational changes in the active site (Stoddard et al., 1993). If IPMDH has the same substrate and coenzyme binding modes as ICDH, more than half of the binding pocket would be occupied by solvent molecules, whereas the ICDH has tight packing of isocitrate and coenzyme with few solvent molecules. Therefore, binding-induced conformational change such as an interdomain hinge motion is likely. In the liver alcohol dehydrogenase (ADH), for example, the binding of coenzyme NAD induces a conformational change of the enzyme in which the protein appears to gain favorable contacts with the coenzyme and reduces the solvent-accessible sur-

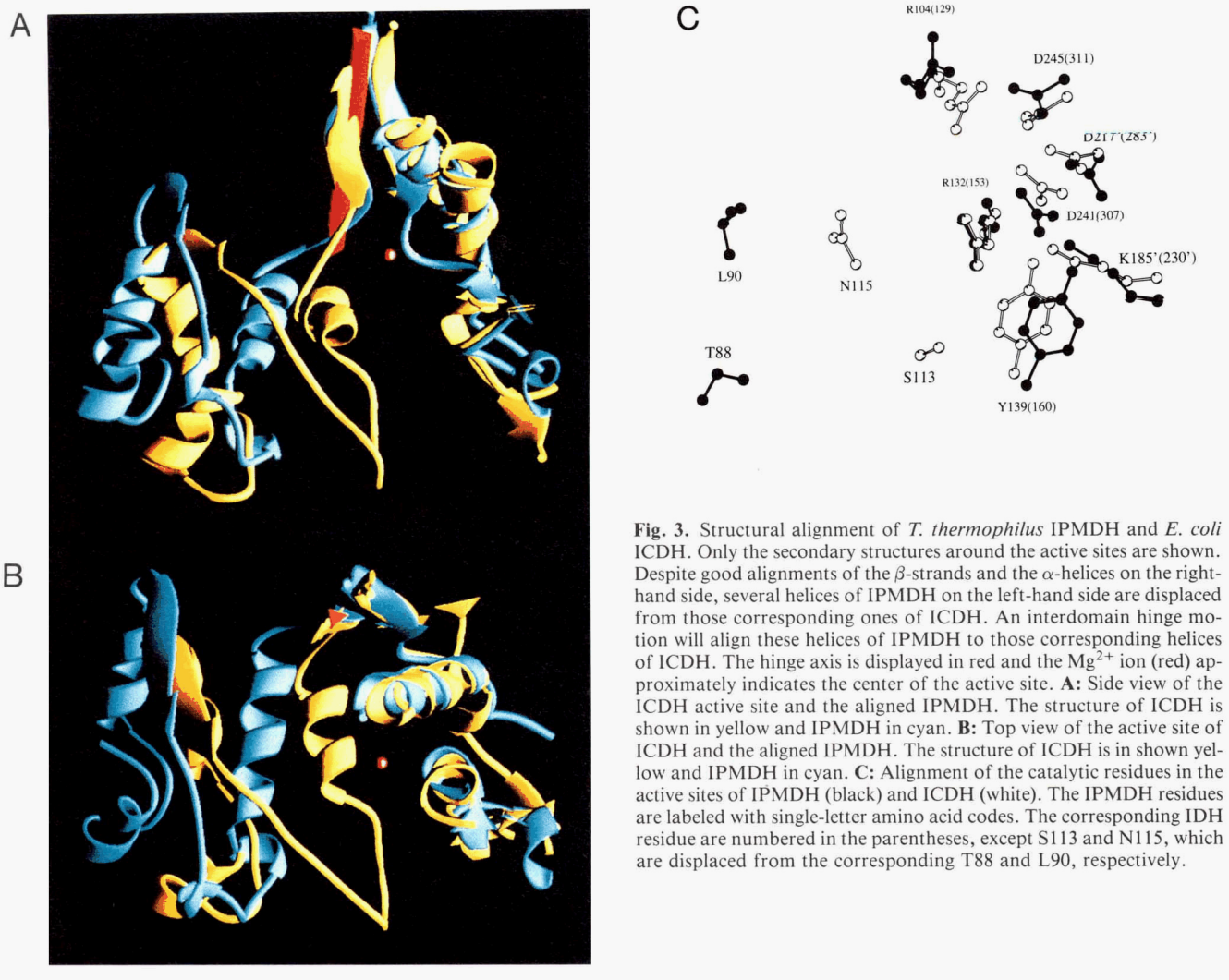


Fig. 3. Structural alignment of *T. thermophilus* IPMDH and *E. coli* ICDH. Only the secondary structures around the active sites are shown. Despite good alignments of the β-strands and the α-helices on the right-hand side, several helices of IPMDH on the left-hand side are displaced from those corresponding ones of ICDH. An interdomain hinge motion will align these helices of IPMDH to those corresponding helices of ICDH. The hinge axis is displayed in red and the Mg^{2+} ion (red) approximately indicates the center of the active site. **A:** Side view of the ICDH active site and the aligned IPMDH. The structure of ICDH is shown in yellow and IPMDH in cyan. **B:** Top view of the active site of ICDH and the aligned IPMDH. The structure of ICDH is in shown yellow and IPMDH in cyan. **C:** Alignment of the catalytic residues in the active sites of IPMDH (black) and ICDH (white). The IPMDH residues are labeled with single-letter amino acid codes. The corresponding IDH residue are numbered in the parentheses, except S113 and N115, which are displaced from the corresponding T88 and L90, respectively.

face area (Eklund et al., 1984). In the case of IPMDH, a similar hinge motion would also close down the interdomain binding pocket (Fig. 3A) and create favorable contact with the substrate and the coenzyme. This hinge motion also would make better alignment of the helices of the outside domains of the two proteins (Fig. 3A,B).

Docking isopropylmalate to IPMDH

To investigate the possible substrate binding modes of IPMDH, isopropylmalate and NAD were docked separately to IPMDH by the automated Monte Carlo docking method of Goodsell and Olson (1990). The "short" and "long" docking schemes produced no significant difference in the docking positions (see Methods). The best docking solutions of isopropylmalate and NAD by the "long" scheme are shown in Figures 4A and 5A, and the calculated energies of the top-ranked docking solutions are listed in the figure legends. The calculated docking energy is, of course, not the absolute binding energy per se, but a relative indication of how favorably a substrate interacts with a protein. The conformations of the bound isocitrate and NADP in ICDH are also shown for comparing the substrate binding modes of the two proteins (Figs. 4B, 5B).

Docking of isopropylmalate with the "long" scheme resulted in 10 different binding positions with calculated energies ranging from −30.6 to −26.8 kcal/mol. The three binding positions with the best calculated energies (−30.6 to −29.5 kcal/mol) are very similar and therefore considered to be one binding mode (Fig. 4A). Because the predominant forces in the docking forcefield were hydrogen bonding and charge–charge interactions, the automated docking resulted in a binding mode that placed the carboxyl and hydroxyl groups of isopropylmalate close to the charged amino acids, i.e., glutamates and arginines. For example, the α-carboxyl group of isopropylmalate makes close contacts with the side chain atoms of R94, N102, R132, L134, and S261. The β-carboxyl group interacts with residues R104, S259, V272, and the nicotinamide ring of the docked coenzyme NAD. The hydroxyl group forms charge–charge interactions with the R132 and a possible bound Mg^{2+} (see below).

The coenzyme NAD was first docked to IPMDH in the absence of bound isopropylmalate molecule. The top docking solutions of both "short" and "long" schemes converged to a single binding mode as shown in Figure 5A. The docked NAD took

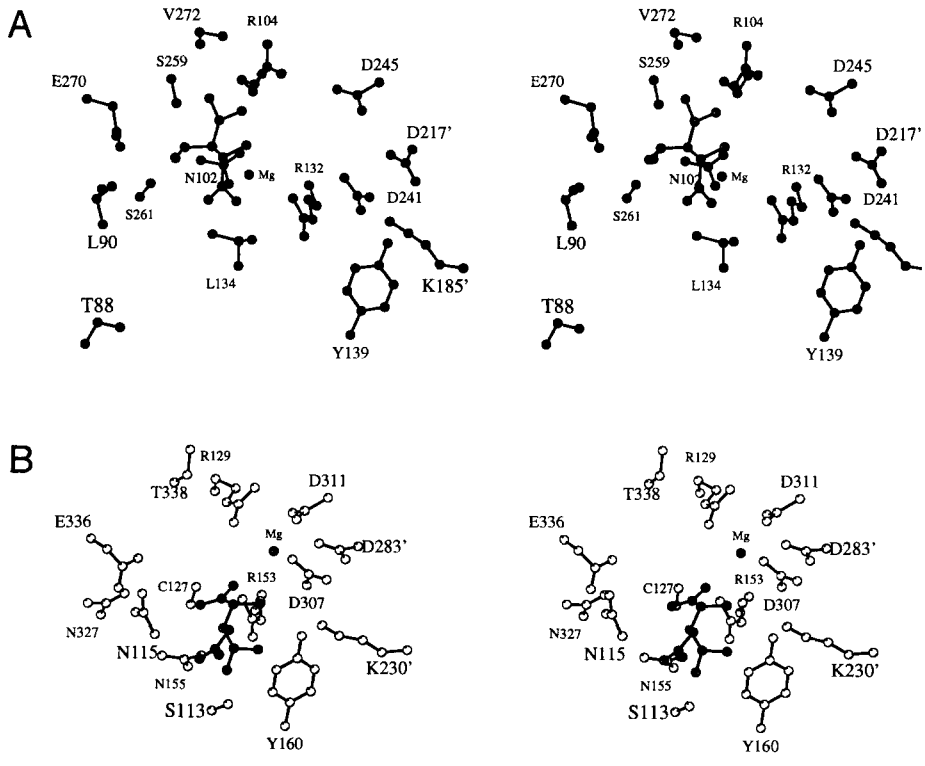


Fig. 4. A: Docked conformation of isopropylmalate (gray) and the surrounding residues of the *T. thermophilus* 3-isopropylmalate dehydrogenase (black). Active site residues are labeled with single-letter amino acid codes. The "long" docking scheme resulted in solutions with calculated energies ranging from −30.6 to −26.8 kcal/mol. The binding mode shown was derived from the top three docking solutions with calculated energies ranging from −30.6 to −29.5 kcal/mol. **B:** Binding conformation of isocitrate (gray) in the active site of *E. coli* isocitrate dehydrogenase (white).

an extended conformation with the nicotinamide ring slightly folding back toward the diphosphoester group. In this binding mode, the nicotinamide ring is close to several conserved residues in IPMDH, including R104, D245, L254, V272, and the β-carboxyl group of the docked isopropylmalate molecule. The hydroxyl groups of the nicotinamide side sugar ring make contact with the residues L90, R94, and E87. The adenine group of NAD makes contact with L91, R94, K95, R132, L134, Y139, and P140. Most of these residues are conserved in IPMDHs from different species (Fig. 2).

The most important feature of this NAD binding mode is that the nicotinamide ring makes close contact with the C_2 atom of the docked isopropylmalate. This close contact facilitates the hydride transfer reaction between the C_2 atom of isopropylmalate and the C_4 atom of the nicotinamide ring of NAD. Such hydride transfer reactions were proposed as essential steps in the catalytic mechanisms of both IPMDH and ICDH. The rest of the docked NAD also makes favorable contact with isopropylmalate but without any overlapping, even though the two molecules were docked independently. Reproducing these essential interactions between substrate and coenzyme by the independent docking of each molecule supports the docking-derived binding modes of substrate and coenzyme in IPMDH.

In addition to the coenzyme NAD, the catalysis of isopropylmalate by IPMDH needs a cation, either a Mg^{2+} or a Mn^{2+} (Imada et al., 1991). The bound Mg^{2+} stabilizes the charged intermediate during the dehydrogenation reaction. In ICDH, the Mg^{2+} coordinates with several aspartate residues and the carboxyl and hydroxyl group of the substrate (isocitrate). Considering the sequence and structural homology between IPMDH and ICDH, we assumed that the cation required in the function of IPMDH played similar role as the Mg^{2+} ion in the catalysis of ICDH. The binding position of the Mg^{2+} was determined by docking to IPMDH–isopropylmalate complex. All docked positions essentially converged to one binding site, which is shown with the docked isopropylmalate and NAD in Figures 4A and 5A. In this binding position, the Mg^{2+} can form coordinating interactions with the conserved catalytic residues D307, D311, and D283′ of IPMDH and the α-,β-carboxyl groups of the docked isopropylmalate. The negatively charged phosphodiester group of the docked NAD molecule may also interact with the Mg^{2+} ion (Fig. 5A). Similar coordinating interactions between the bound cation and the enzyme/substrate were also observed in the active site of ICDH (Figs. 4B, 5B).

The binding of isopropylmalate, NAD, and Mg^{2+} are consistent with the proposed catalytic mechanism of IPMDH. The proposed reaction starts with the base-catalyzed removal of a proton from the hydroxyl group of the isopropylmalate. The base may be one of the nearby aspartates such as D241 or D245. The deprotonated intermediate is stabilized by the bound cation and the surrounding polar residues, e.g., aspartates, arginines, etc. (Fig. 4A). The dehydrogenation induces the transfer of a hydride from the intermediate to the cofactor NAD^+, and the substrate isopropylmalate is then converted to oxaloisoval-

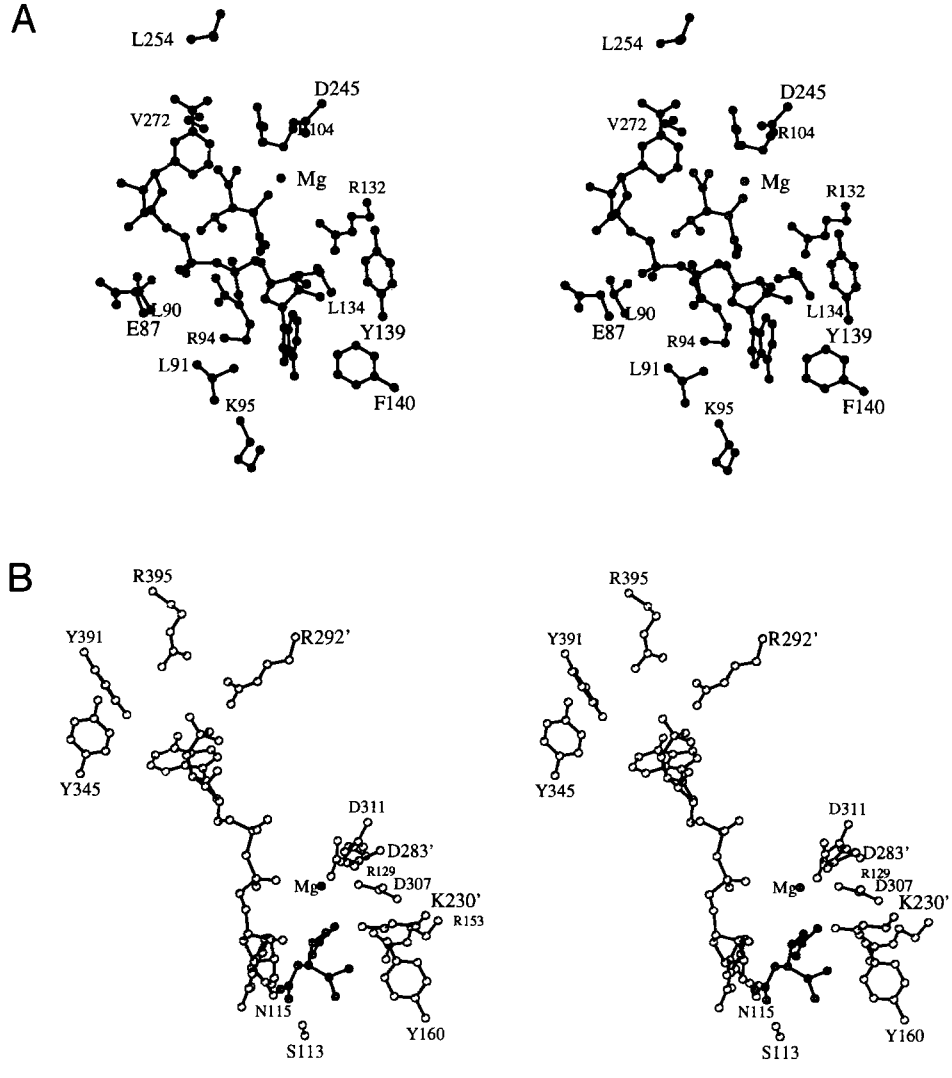


Fig. 5. A: Docked conformation of nicotinamide adenine dinucleotide (NAD, black) and isopropylmalate (gray) in the active site of *T. thermophilus* 3-isopropylmalate dehydrogenase. The "long" docking scheme resulted in solutions with calculated energies ranging from −68.3 to −55.0 kcal/mol. The binding mode shown was derived from the top four docking solutions with calculated energies ranging from −68.3 to −64.1 kcal/mol. **B:** Binding conformation of nicotinamide adenine dinucleotide phosphate (NADP, white) and isocitrate (gray) in the active site of *E. coli* isocitrate dehydrogenase.

erate. The close contact between the nicotinamide ring nitrogen of the NAD and the C_2 atom of isopropylmalate would facilitate this transfer. In the second step, the β-carboxylate of oxaloisovalerate is lost as CO_2, and the concomitant protonation of the β-carbon transforms oxaloisovalerate to α-ketoisocaproate.

Despite these common features, the substrates of IPMDH and ICDH showed key differences in their interactions with the enzymes and coenzymes. Although the nicotinamide rings of both coenzymes make close contact with the C_2 atoms of substrates, in ICDH, the only other contact between NADP and isocitrate involves the nicotinamide ring of NADP and the γ-carboxyl group of isocitrate, and the phosphodiester does not make direct contact with isocitrate. In IPMDH, however, the γ-moiety of the docked isopropylmalate is placed next to the phosphodiester group. The α-carboxyl group faces the adenine ring of NAD. These differences provide an explanation for the different substrate specificities of IPMDH and ICDH. ICDH shows no catalytic activity against isopropylmalate because replacing the γ-carboxyl group with a nonpolar isopropyl group would remove a favorable, specific interaction with ring nitrogen of the nicotinamide of NADP, which orients hydride donor (C_2) on the substrate to the hydride acceptor (C_4) on the coenzyme. The docked isopropylmalate and NAD in IPMDH interacts differently. Because the γ-moiety is placed next to the diphosphoester group of NAD, replacing it with a negatively charged carboxyl group (isopropylmalate → isocitrate) would create repulsive interactions between NAD and isopropylmalate. This repulsive interaction will destabilize the tertiary binding complex and especially alter the orientation of hydride donor on the substrate and acceptor on the substrate. Therefore, the rate of the hydride transfer reaction would be decreased. On the other hand, replacing the γ-moiety with other nonpolar groups seems less destructive to the interaction between isopropylmalate and NAD because the packing interaction around the γ-moiety is not

very specific. Kinetic studies of both ICDH and IPMDH confirmed that the catalysis is less sensitive to the length of alkyl groups at the γ position (Miyazaki et al., 1993). In addition, the β-carboxyl of IPM may interact favorably with the NAD.

IPMDH and ICDH use different coenzymes NAD and NADP, respectively. In ICDH, the extra phospho group of the bound NADP interacts with several arginines and tyrosines residues (R292′, R395, Y345, and Y391). In IPMDH, sequence alignment showed that at least one of the arginines (R395) and the two tyrosines are absent. These lost interactions may explain the distinct binding mode of NAD in IPMDH. On the other hand, the adenine sugar makes close contact with the α-carboxyl group of the bound isopropylmalate and the active site residue R132 (Fig. 4A, 5A) with the average distances of 6 Å and 5 Å, respectively. Adding an extra phospho group to the adenine ribose will create strong interactions with the substrate and R132 and change the relative binding orientations between the substrate and coenzyme.

The binding of isopropylmalate in IPMDH also differs slightly from that of isocitrate in ICDH. In ICDH, the bound isocitrate interacts directly with several aspartates in the active site, such as D283′, D307, and D311. However, the docked isopropylmalate in the active site of IPMDH is displaced from the corresponding aspartates. The distances between the docked isopropylmalate and Y139 and K185′ are also significantly longer than between their counterparts in the active site of ICDH (Y160, K230′, see Fig. 4). This different binding mode of isopropylmalate seems consistent with the recent observation that mutating Y139 to phenylalanine had little effect on binding of isopropylmalate (Miyazaki & Oshima, 1993). The predominant effect of this Y139 → F mutation is on the binding of NAD, which is consistent with the docked position of the coenzyme, where the Y139 is in close contact with the adenine ribose ring (3.3 Å).

Conclusions

A docking algorithm and sequence homology between the *T. thermophilus* IPMDH and *E. coli* ICDH have allowed us to identify the binding site of isopropylmalate. The substrate and coenzyme (NAD) binding modes of IPMDH and ICDH have striking similarities and some important differences. The catalytic residues and metal ion bindings are very similar, but the cofactor bindings are different. The difference in the cofactor binding mode, together with the different side chains of the substrates, lead to the differences identified with substrate specificities. In addition, the contrast between the tightly packed ICDH site and the loosely packed IPMDH site suggests a ligand-induced hinge motion in IPMDH during the enzyme action.

Methods

Sequence and structure alignments

The amino acid sequences of ICDH and IPMDH from different species were translated from the DNA or mRNA sequences reported in GenBank. The homologous alignment of these sequences was done with the Needleman–Wunsch algorithm (Needleman & Wunsch, 1970).

The structural alignment of ICDH and IPMDH was done by using a transformation matrix derived from the nonlinear least-square fitting of selected residues from both proteins. First, the atoms of these selected residues were fitted to each other to derive the transformation matrix. The transformation matrix was then used to superimpose the two protein structures. The residues selected in ICDH are: R129, R153, Y160, K230′, D283′, D307, and D311. These residues were selected because they were shown to play important roles in the enzyme function (Dean & Koshland, 1990, 1993). Based on sequence alignment of ICDH and IPMDH, the corresponding residues in IPMDH are: R104, R132, Y139, K185′, D217′, D241, and D245 (Fig. 2). Coordinates of IPMDH and ICDH were taken from the known crystal structures of the two proteins deposited in the Brookhaven Protein Data Bank (1ipd and 5icd).

Monte Carlo docking

To determine the bound conformation of isopropylmalate on IPMDH, we used the automated docking method of Goodsell and Olson (1990). This method has been described previously and only a brief description is given here (Goodsell & Olson, 1990). In the automated docking method, a flexible substrate performs a random walk around the surface of a static protein to search an optimal binding site. At each search step, the substrate was translated and rotated about its center and each of its torsional angles while its interaction energy with the protein was evaluated. The acceptance or rejection of a configuration was based on the interaction energy and a temperature factor. This large configurational space was sampled by the simulated annealing method. The search starts at high temperature, so the interaction energy has little impact on the acceptance or rejection of a configuration. This allows a relatively unrestricted search of the configurational space. As the simulation proceeds, the temperature is lowered and the impact of interaction energy on the acceptance or rejection is increased. The substrate was therefore restricted to only energetically favorable configurations. Multiple runs are carried out to increase the sampling of this configurational space.

The enzyme–substrate interaction energy was evaluated in terms of a molecular mechanic force field. The force field includes electrostatic, hydrogen bonding, and dispersion/repulsion terms and was able to reproduce crystallographic binding modes of several polar substrates (Goodsell & Olson, 1990). We also tested the force field by docking isocitrate to the *E. coli* ICDH. The best docking solution of isocitrate agrees with known crystallographic structure of ICDH–isocitrate complex. The docking calculation was speeded up by use of a precalculated potential map of a cubic search region. The structure of IPMDH was taken from the 2.2-Å crystal structure of the apo protein (Imada et al. 1991; PDB code: 1ipd). The missing polar hydrogens in the structure were modeled by the H-build routine of Charmm22 (Brooks et al., 1983). The atomic partial charges of the protein assigned according to the Charmm22 extended atom force field. The atomic partial charges of isopropylmalate were determined by fitting the electrostatic potential derived from the semi-empirical quantum mechanical molecular orbital model (PM3).

The search for the binding position of isopropylmalate was confined to the potential active site of IPMDH as identified by structural and sequence homology with ICDH. A "crude" search was first conducted with a 50-Å cubic box of a 0.5-Å grid. A "fine" search was then followed with a 25-Å cubic box of 0.25-Å grid. The torsional angles of each flexible bonds have been kept

flexible in the docking process. Several distance constraints were used to prevent overlapping of atoms that are separated by more than three bonds. The substrates were docked by two different docking schemes. In the "short" docking scheme, substrates were docked with 100 independent docking runs, each with a different random seed number. Each run contains 50 cycles of 3,000 steps accepted or rejected, starting at a high temperature (k_BT = 100 kcal/mol) and decreasing by a factor of 0.9 each cycle. In the "long" docking scheme, the substrates were docked with 10 individual docking runs; each run has 50 cycles of 30,000 steps accepted or rejected while the temperature parameters are kept the same. The top-scored configurations given by both docking schemes were very similar. The results reported in this study were those from the "long" docking scheme (30,000 steps accepted or rejected).

Acknowledgments

This work was supported by the National Science Foundation and the Material Science Division, US Department of Energy, under contract DE-AC-76SF00098. In addition, T.Z. was supported by an NIH postdoctoral fellowship.

References

Brooks B, Bruccoleri R, Olafson B, States D, Swaminathan S, Karplus M. 1983. Charmm: A program for macromolecular energy, minimization and molecular dynamics calculations. *J Comput Chem 4*:187–217.

Colman R. 1973. A glutamyl residue in the active site of triphosphopyridine nucleotide-dependent isocitrate dehydrogenase of pig heart. *J Biol Chem 248*:8137–8143.

Dean A, Koshland DE Jr. 1990. Electrostatic and steric contributions to regulation at the active site of isocitrate dehydrogenase. *Science 249*:1044–1046.

Dean A, Koshland DE Jr. 1993. Kinetic mechanism of *Escherichia coli* isocitrate dehydrogenase. *Biochemistry 32*:9302–9309.

Ehrlich R, Colman R. 1978. Histidine in the nucleotide-binding site of NADP-linked isocitrate dehydrogenase from pig heart. *Eur J Biochem 89*:575–587.

Eklund H, Samama J, Jones T. 1984. Crystallographic investigations of nicotinamide adenine dinucleotide binding to horse liver alcohol dehydrogenase. *Biochemistry 23*:5982–5996.

Goodsell D, Olson A. 1990. Automated docking of substrates to proteins by simulated annealing. *Proteins Struct Funct Genet 8*:195–202.

Grissom C, Cleland WW. 1985. Use of intermediate partitioning to calculate intrinsic isotope effects for the reaction catalyzed by malic enzyme. *Biochemistry 12*:944–948.

Hurley J, Dean A, Koshland DE Jr, Stroud R. 1991. Catalytic mechanism of $NADP^+$-dependent isocitrate dehydrogenase: Implications from the structures of magnesium-isocitrate and $NADP^+$ complexes. *Biochemistry 30*:8671–8678.

Imada K, Sato M, Tanaka N, Katsube Y, Matsuura Y, Oshima T. 1991. Three-dimensional structure of a highly thermostable enzyme, 3-isopropylmalate dehydrogenase of *Thermus thermophilus* at 2.2 Å resolution. *J Mol Biol 222*:725–738.

Kornberg H. 1966. The role and control of the glyoxylate cycle in *Escherichia coli*. *Biochem J 99*:1–11.

Miyazaki K, Eguchi H, Yamagishi A, Wakagi T, Oshima T. 1992. Molecular cloning of the isocitrate dehydrogenase gene of an extreme thermophile, *Thermus thermophilus* hb8. *Appl Environ Microbiol 58*:93–98.

Miyazaki K, Kakinuma K, Terasawa H, Oshima T. 1993. Kinetic analysis on the substrate specificity of 3-isopropylmalate dehydrogenase. *FEBS Lett 11*:35–36.

Miyazaki K, Oshima T. 1993. Tyr 139 in *Thermus thermophilus* 3-isopropylmalate dehydrogenase is involved in catalytic function. *FEBS Lett 11*:37–38.

Needleman S, Wunsch C. 1970. A general method applicable to the search for similarities in the amino acid sequences of two proteins. *J Mol Biol 48*:443–453.

Pirrung M, Han H, Nunn D. 1994. Kinetic mechanism and reaction pathway of *Thermus thermophilus* isopropylmalate dehydrogenase. *J Org Chem 59*:2423–2429.

Stoddard B, Dean A, Koshland DE Jr. 1993. Structure of isocitrate dehydrogenase with isocitrate, nicotinamide adenine dinucleotide phosphate, and calcium at 2.5-Å resolution: A pseudo-Michaelis ternary complex. *Biochemistry 32*:9310–9316.