Paper I

doi:10.1016/j.jmb.2004.10.025 J. Mol. Biol. (2005) 345, 559–577

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Isocitrate Dehydrogenase from the Hyperthermophile *Aeropyrum pernix*: X-ray Structure Analysis of a Ternary Enzyme–Substrate Complex and Thermal Stability

Mikael Karlström[1]*, Runar Stokke[2], Ida Helene Steen[2] Nils-Kåre Birkeland[2] and Rudolf Ladenstein[1]

[1]*Center for Structural Biochemistry, Department of Biosciences at Novum Karolinska Institute, S-141 57 Huddinge, Sweden*

[2]*Department of Biology University of Bergen, PO Box 7800, Jahnebakken 5, N-5020 Bergen, Norway*

**Corresponding author*

Isocitrate dehydrogenase from *Aeropyrum pernix* (*Ap*IDH) is a homodimeric enzyme that belongs to the β-decarboxylating dehydrogenase family and is the most thermostable IDH identified. It catalyzes the $NADP^+$ and metal-dependent oxidative decarboxylation of isocitrate to α-ketoglutarate. We have solved the crystal structures of a native *Ap*IDH at 2.2 Å, a pseudo-native *Ap*IDH at 2.1 Å, and of *Ap*IDH in complex with $NADP^+$, Ca^{2+} and *d*-isocitrate at 2.3 Å. The pseudo-native *Ap*IDH is in complex with etheno-$NADP^+$ which was located at the surface instead of in the active site revealing a novel adenine-nucleotide binding site in *Ap*IDH. The native and the pseudo-native *Ap*IDHs were found in an open conformation, whereas one of the subunits of the ternary complex was closed upon substrate binding. The closed subunit showed a domain rotation of 19° compared to the open subunit. The binding of isocitrate in the closed subunit was identical with that of the binary complex of porcine mitochondrial IDH, whereas the binding of $NADP^+$ was similar to that of the ternary complex of IDH from *Escherichia coli*. The reaction mechanism is likely to be conserved in the different IDHs. A proton relay chain involving at least five solvent molecules, the 5′-phosphate group of the nicotinamide–ribose and a coupled lysine–tyrosine pair in the active site, is postulated as essential in both the initial and the final steps of the catalytic reaction of IDH. *Ap*IDH was found to be highly homologous to the mesophilic IDHs and was subjected to a comparative analysis in order to find differences that could explain the large difference in thermostability. Mutational studies revealed that a disulfide bond at the N terminus and a seven-membered inter-domain ionic network at the surface are major determinants for the higher thermostability of *Ap*IDH compared to *Ec*IDH. Furthermore, the total number of ion pairs was dramatically higher in *Ap*IDH compared to the mesophilic IDHs if a cutoff of 4.2 Å was used. A calculated net charge of only +1 compared to −19 and −25 in *Ec*IDH and *Bs*IDH, respectively, suggested a high degree of electrostatic optimization, which is known to be an important determinant for increased thermostability.

Keywords: isocitrate dehydrogenase; thermal stability; disulfide; ionic networks; domain movements

Introduction

Enzymes from hyperthermophiles are often highly homologous to their mesophilic counterparts and their catalytic mechanisms are usually identical. They must be stable enough to withstand denaturation at temperatures above 80 °C and simultaneously maintain the flexibility required

Abbreviations used: IDH, isocitrate dehydrogenase; IPMDH, isopropylmalate dehydrogenase; *Ap*IDH, *Aeropyrum pernix* IDH; *Ec*IDH, *Escherichia coli* IDH; *Bs*IDH, *Bacillus subtilis* IDH; *Tm*IDH, *Thermotoga maritima* IDH; *Af*IDH, *Archaeoglobus fulgidus* IDH; *Pf*IDH, *Pyrococcus furiosus* IDH.

E-mail address of the corresponding author: mikael.karlstrom@csb.ki.se

0022-2836/$ - see front matter

for enzymatic activity. There appears to be no single mechanism or structural feature that is responsible for the high thermotolerance of hyperthermostable proteins. The major reason for this is the relatively small free energy difference between the folded and the unfolded state of a protein and the complex way in which the small number of weak forces, determining protein stability, interplay with each other. Upon comparison with the mesophilic homologs, the most common determinants for hyperthermostability are in the first line a statistical prevalence of ionic interactions at the protein surface, increased formation of large ionic networks, electrostatic optimisation and the reduction of repulsive charge–charge interactions.[1–3] Furthermore, a reduction of the hydrophobic accessible surface area, increased hydrogen bonding and structural compactness has been observed.[3,4] The elucidation of the mechanisms and structural determinants responsible for the extreme thermostability of proteins from hyperthermophiles is of importance for understanding the question of why hyperthermophilic proteins show very high thermotolerance in spite of a relatively small gain of free stabilization energy compared to the mesophilic homologs. It is expected that hyperthermostable enzymes will have a great potential in biotechnological and industrial applications in processes at elevated temperatures.

Isocitrate dehydrogenase (IDH) is a metal-dependent (Mg^{2+} or Mn^{2+}) enzyme that catalyses the subsequent dehydrogenation and decarboxylation of isocitrate to α-ketoglutarate using NAD^+ or $NADP^+$ as cofactor.[5] α-Ketoglutarate is an intermediate in the citric acid cycle as well as the skeletal carbon source in the biosynthesis of certain amino acid residues. These pathways are among the first to have evolved in the history of life.[6,7] Consequently, IDH is distributed broadly throughout the three domains of life, Archaea, Bacteria and Eukarya, with diverse primary structures and different oligomeric states.[8]

The IDHs have been distinguished into three subfamilies based on sequence comparisons.[8,9] All of the archaeal and most of the bacterial IDHs are grouped together into subfamily I, eukaryotic homodimeric IDHs and some bacterial IDHs constitute subfamily II, whereas eukaryotic heterooligomeric IDHs form a third subfamily. The members of subfamily II show very little sequence identity with those of subfamilies I and III. There are also monomeric IDHs that show no significant sequence similarity to the other subfamilies.

The crystal structures of IDH from *E. coli*[10] (*Ec*IDH, PDB code 3ICD) and *Bacillus subtilis*[11] (*Bs*IDH, PDB code 1HQS) in subfamily I and of porcine heart mitochondrial IDH[12] (denoted below as porcine IDH, PDB code 1LWD) in subfamily II are all $NADP^+$-dependent homodimers and have revealed a common fold, shared also by the crystal structures of the NAD^+-dependent isopropylmalate dehydrogenases (IPMDH), which belong to the same family of β-decarboxylating dehydrogenases using a substrate that is structurally related to isocitrate.[13–15] The crystal structure of monomeric IDH from *Azotobacter vinelandii* (PDB code 1ITW)[16] is also topologically related to homodimeric IDHs. This appears to have been accomplished by a partial gene duplication, which resulted in a pseudo 2-fold symmetry in one domain corresponding to the other subunit of the dimeric IDHs. IDH and IPMDH share a unique cofactor binding site that is different from the well-known Rossmann fold found in many other dehydrogenases.[10,13,17] Only a few amino acid residues appear to be responsible for the discrimination between NAD^+ and $NADP^+$.[18–21]

The reaction mechanism of IDH has been extensively studied in *Ec*IDH.[22,23] In the proposed mechanism, a proton is removed from the α-hydroxyl group of isocitrate. Subsequently, a hydride ion is transferred in a stereospecific way from the α-carbon atom of the substrate to C-4 of the nicotinamide ring of $NADP^+$, oxidizing isocitrate to oxalosuccinate.[24,25] In a second step, the β-carboxylate group of oxalosuccinate is lost as CO_2,[26] and is replaced by a proton in a stereospecific way to form α-ketoglutarate.[27] During both transition states the negative charge on the hydroxyl oxygen atom of isocitrate is stabilised by a magnesium ion. However, the initial proton abstraction mechanism as well as the final proton donation are still matters of debate (see section The active site).

The regulation of the activity of the different IDHs is diverse. The NAD^+-dependent eukaryotic heterooligomeric IDHs are allosterically regulated by the activators AMP[28] or ADP,[29] whereas *Ec*IDH is regulated by the IDH kinase/phosphatase-mediated phosphorylation of Ser113, which inactivates the enzyme by sterically hindering the binding of isocitrate and by electrostatic repulsion.[30–33] Recently, a self-regulating mechanism of activity was postulated in human cytosolic $NADP^+$-dependent IDH.[34]

The structural homology across the different species in this enzyme family makes them very suitable for comparative studies. We have chosen to study IDH from the strictly aerobic hyperthermophilic archaeon *Aeropyrum pernix* (*Ap*IDH). Previously, we have described *Ap*IDH as the most thermostable IDH characterized, with an apparent melting temperature of 110 °C.[8] It is a Mg^{2+}- and $NADP^+$-dependent dimer with identical subunits of 47.9 kDa and belongs to subfamily I. A short structure notice of *Ap*IDH was published recently[35] where a disulfide-bond at the N terminus and a seven-membered ion pair network at the surface were related to increased thermotolerance, but no stability data were given. The mechanism of regulation of *Ap*IDH has remained unknown. To our knowledge, no gene encoding IDH kinase/phosphatase has been found in the genome of *A. pernix*.[36]

In order to gain more information about the catalytic mechanism and the large domain movements presumably involved in catalysis, we have solved the structure of a ternary complex of *Ap*IDH

with Ca^{2+}–isocitrate–$NADP^+$. The structure was compared with the structure of the ternary complex of *Ec*IDH and with the binary complex of porcine IDH from subfamily II, which has a conserved isocitrate binding site.[12] It is likely that the reaction mechanisms are identical in *Ap*IDH, *Ec*IDH and porcine IDH, since the isocitrate binding sites are conserved both structurally and by sequence.

In order to perform a more thorough analysis of the factors responsible for the high thermotolerance of this enzyme, we have started a mutational approach based on our refined crystal structure of wild-type *Ap*IDH. The structure is reported together with apparent melting temperatures of the mutants and is primarily compared to the structure of *Ec*IDH but to some extent also to *Bs*IDH.

Results and Discussion

Quality and description of the models

In this work, three different *Ap*IDH structures are reported: a native *Ap*IDH solved at 2.2 Å (PDB code 1XGV), a pseudo-native *Ap*IDH solved at 2.1 Å (1TYO) and a ternary complex solved at 2.3 Å (1XKD). The ternary complex was formed in a soaking solution where citrate and Mg^{2+} were exchanged with *d*-isocitrate, Ca^{2+} and $NADP^+$. The pseudo-native *Ap*IDH was soaked with etheno-$NADP^+$ but the etheno-$NADP^+$ molecule did not bind to the active site. However, the etheno-adenine moiety of etheno-$NADP^+$ was located at the surface of *Ap*IDH, making contact also with a symmetry-related molecule, which resulted in a unit cell with the *c* axis elongated by 8.7 Å. This complex is designated as "pseudo-native" because of the unusual binding of etheno-$NADP^+$ outside of the active site and because it is basically identical with native *Ap*IDH and has the same overall conformation. The $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ difference density maps of native *Ap*IDH and pseudo-native *Ap*IDH contained uninterpretable density in the isocitrate binding site. Most likely, the density originated from citrate, since 0.1 M citrate was present in the crystallization buffer.

*Ap*IDH was crystallized as a dimer in the asymmetric unit with a solvent content of 52% corresponding to a Matthews coefficient of 2.6 $Å^3\ Da^{-1}$. The two subunits are referred to as subunit A and subunit B. The space group was $P4_32_12$. Table 1 summarizes the quality of the different models. The quality of the B subunits is somewhat reduced in all models, since the densities of the large domain suffered from disorder, which was probably caused by the few crystal contacts this domain is involved in. Accordingly, less solvent molecules were built into subunit B as compared to the number in subunit A. Arg103 in subunit B of native *Ap*IDH is found in a disallowed conformation, which can be explained by its location in a sharp turn between helix c and strand C. In the ternary complex, Ser120 in subunit A is also found in a disallowed conformation; this is explained by its location at the end of a loop, which is disordered.

Overall fold

In contrast to many other hyperthermophilic enzymes, *Ap*IDH is larger than its mesophilic homologs. Each subunit contains 435 residues, whereas *Ec*IDH and *Bs*IDH have only 416 and 423 residues, respectively. In *Ap*IDH, residues 1–131 and 322–435 belong to the large domain, 132–163 and 206–321 form the small domain and the remaining residues 164–205 form the clasp domain through the subunit interface between two anti-parallel α-helices beneath a four-stranded inter-subunit anti-parallel β-sheet. The large domain is connected to the small domain by a flexible hinge region. Figure 1 shows the final model of the ternary complex of *Ap*IDH.

An alignment based on secondary structural assignment (Figure 2) revealed that most elements of secondary structure were conserved within subfamily I. The superimposed C^α traces showed that the overall topology of *Ap*IDH was almost identical with that of *Ec*IDH and *Bs*IDH (Figure 3). The main differences are the extensions at both termini of *Ap*IDH and the replacement of strand *K* and the preceding loop in the small domain of *Ec*IDH by helix *g2* in *Ap*IDH. In this region, *Bs*IDH has two helices, *g2* and *g3*, which are located in the vicinity of the active site of the neighbouring subunit in the dimer, restricting the access to the phosphorylation site.[37] Strand *L* in this region seems to be conserved in *Ap*IDH, *Ec*IDH and *Bs*IDH, but goes in the opposite direction in *Ap*IDH and *Bs*IDH compared to *Ec*IDH. *Bs*IDH has an insertion of 13 residues in this region with respect to *Ec*IDH, whereas *Ap*IDH has only one more residue. The RMS difference between the large domain of native *Ap*IDH *versus* that of *Ec*IDH and *Bs*IDH was 1.49 Å (using 211 C^α atoms) and 1.74 Å (211 C^α atoms), respectively. For the small domain together with the clasp domain, the RMS difference between *Ap*IDH *versus* *Ec*IDH and *Bs*IDH was 3.28 Å (188 C^α atoms) and 1.43 Å (192 C^α atoms), respectively. The large difference between *Ap*IDH and *Ec*IDH is mainly due to the unique structural feature in the region of helix *g2* in *Ap*IDH. When residues 258–276 in this region were removed from the comparison, the RMS difference between the small and clasp domains of *Ap*IDH *versus* *Ec*IDH and *Bs*IDH was only 0.82 Å (171 C^α atoms) and 0.80 Å (173 C^α atoms), respectively.

In total there are 15 α-helices in *Ap*IDH (39.3% of the residues), of which 13 are conserved, and 15 β-strands (16.3%), of which all are conserved. There are also seven 3_{10} helices (5.1%). The total secondary structure content of 60.7% is only slightly higher than that of *Ec*IDH (57.2%) and reflects the formation of helices *g2*, *n* and three additional 3_{10} helices as well as a few loop deletions (described in the thermostability section) in *Ap*IDH.

Table 1. Data collection and refinement statistics

	Native ApIDH	Pseudo-native ApIDH	Ternary ApIDH
PDB code	1XGV	1TYO	1XKD
Wavelength (Å)	0.996	0.9089	1.089
Resolution limits (Å)	2.2–39.8 (2.20–2.24)	2.1–39.9 (2.15–2.19)	2.3–39.7 (2.30–2.34)
Mosaicity	0.5	0.5	0.6
Unit cell parameters	a=b=107.57, c=171.14 $\alpha=\beta=\gamma=90°$	a=b=107.01, c=179.82, $\alpha=\beta=\gamma=90°$	a=b=107.57, c=171.04, $\alpha=\beta=\gamma=90°$
No. observed reflections	242,940 (10,181)	351,456 (13,556)	375,551 (16,016)
No. unique reflections	51,805 (2522)	57,596 (2836)	45,289 (2259)
Redundancy	4.7 (4.0)	6.1 (4.8)	8.3 (7.1)
Completeness (%)	99.9 (100)	100 (100)	99.9 (100)
I/σ (I)	17.97 (2.91)	17.62 (3.44)	25.42 (5.53)
R_{merge} (%)	5.5 (49.7)	7.6 (52.6)	5.1 (36.7)
Wilson B-factor	40.1	30.3	45.7
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$
Refinement			
Non-hydrogen atoms	6689	6788	6789
Non-hydrogen substrate atoms	0	32 (etheno-NADP)	122 (isocitrate, NADP)
Non-hydrogen ion atoms	0	0	2 (Ca^{2+})
Missing residues	A1-5, B1-6, B114-9, B433-5	A1-5, A433-5, B1-6, B431-5	A1-4, A115-8, B1-5, B66-7 B427-35
Solvent molecules	161	200	195
Resolution range (Å)	2.20–39.8 (2.20–2.26)	2.15–39.9 (2.15–2.21)	2.30–35.0 (2.30–2.36)
R_{cryst} overall (%)	22.5 (25.2)	22.3 (23.0)	22.6 (23.9)
R_{free} (%)	25.1 (28.3)	24.9 (24.4)	24.8 (29.9)
Ramachandran plot (excl Gly and Pro) in subunit A/B			
Most favourable region (%)	89.3/88.8	92.0 /91.1	90.0/86.7
Allowed regions (%)	10.7/10.9	8.0 /8.9	9.7/13.3
Disallowed regions (%)	0/0.3	0/0	0.3/0
RMS deviation from ideal values			
Bond lengths (Å)	0.011	0.012	0.013
Bond angles (deg.)	1.261	1.336	1.928
Average B-factors (Å^2)			
Total	19.5	20.8	27.3
Isocitrate (A/B)			29.6/25.0
$NADP^+$ (A/B)			31.7/31.0
Etheno-$NADP^+$		29.8	

Values in parentheses refer to data in the highest-resolution shell.

Domain movements and differences in orientation between the subunits

Subunit B of the ternary complex had undergone a dramatic closure of the large domain, enclosing the substrate and the cofactor in the active site, which is formed by the domain interface. Subunit A was found still in an open conformation. The relative difference in the orientation of the large domain between subunit A and subunit B was 19°. The domain rotation axis in *Ap*IDH was located in parallel with strand *E* and strand *F* (Figure 1a and b). It coincides with the domain definitions and approximately with the pivot points of the domain movements defined for *Ec*IDH and *Bs*IDH.[11] The hinge loops in *Ec*IDH defined by Doyle *et al*. were not adopted as hinge region.[38]

The RMS difference between the small domains together with the clasp domains of the two subunits of the ternary complex was 0.57 Å for 190 C^α atoms, whereas the RMS difference between the large domains was 1.1 Å using 215 C^α atoms, indicating differences other than just a rigid body movement of the large domain. Three major differences were identified when the domains of both subunits were superimposed; the C^α traces of two NADP-binding loops (residues 322–327 and 343–357) were translated 1.6 Å and 3.2 Å, respectively, and helix *l* (residues 396–400), which is also involved in NADP binding, was translated 2.3 Å.

Both subunits of native and pseudo-native *Ap*IDH were found to be in an open conformation. The relative difference in the rotation of the large domain between the subunits was 2–3° in each dimer. In comparison, the difference between the closed subunits of *Bs*IDH is 2.9°. In *Ec*IDH, the subunits are related by a crystallographic 2-fold axis, thus no differences between the subunits can be observed. However, two different crystal forms of *Ec*IDH have revealed an open and a closed conformation with a relative difference in the rotation of the large domain of 16°.[39] The conformation of native and pseudo-native *Ap*IDH was similar to the open form of *Ec*IDH (Figure 3).

An investigation of the crystal contacts revealed why the conformations of the subunits were so different in the ternary complex. Whereas the large domain of subunit B almost had no crystal contacts, the large domain of subunit A was involved in several contacts that made the closure of the large domain impossible upon substrate binding. It has

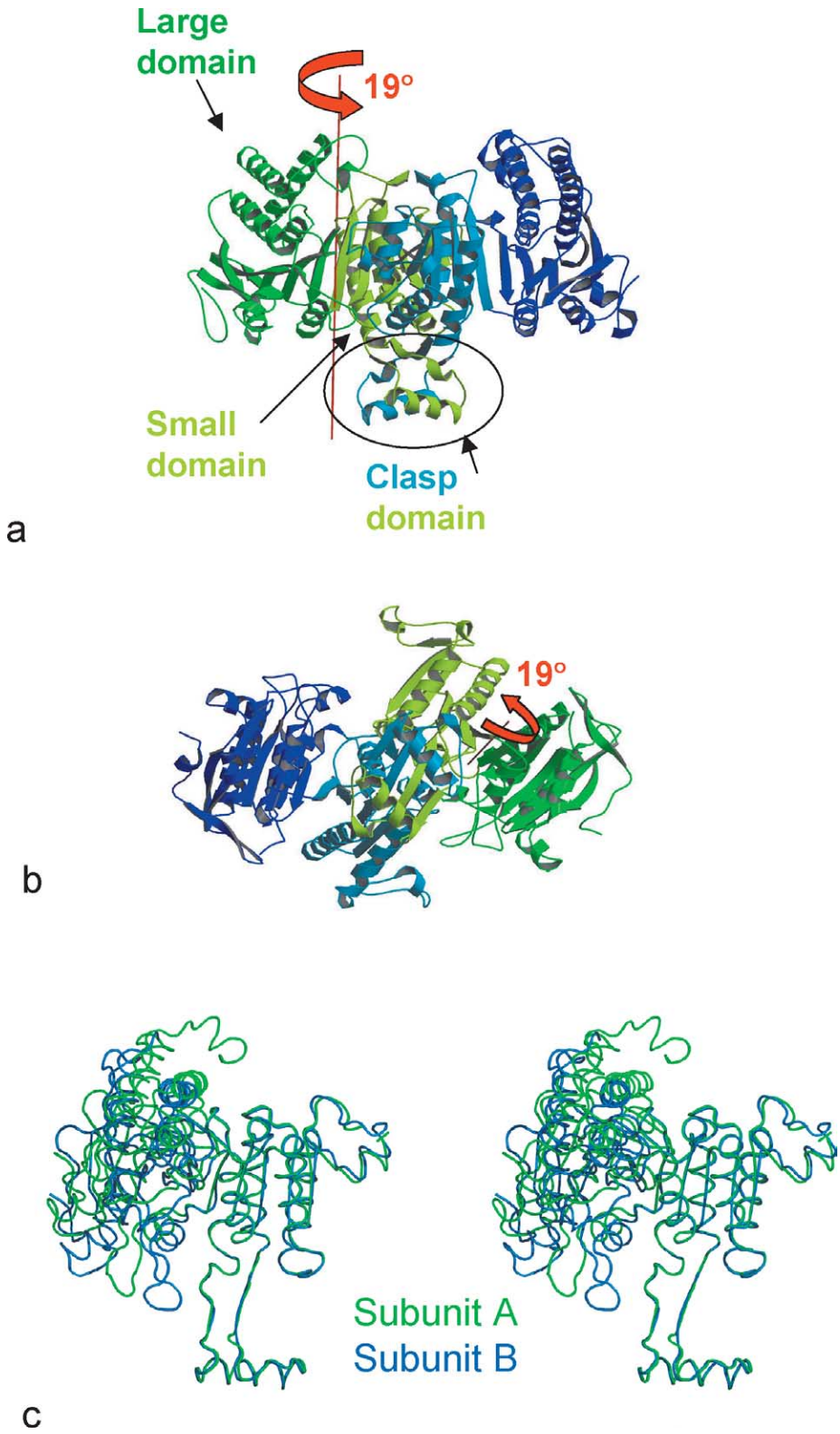

Figure 1. Ribbon representation of the ternary complex of *Ap*IDH showing colour-coded domain definitions and the position of the domain rotation axis (a and b) and a stereoview of an alignment of the open (green) and the closed (blue) subunits of the ternary complex of *Ap*IDH (c). Subunit A was found in an open conformation, whereas subunit B was closed due to a rotation of the large domain towards the small domain upon substrate binding. The difference in rotation of the large domain between the subunits was 19°.

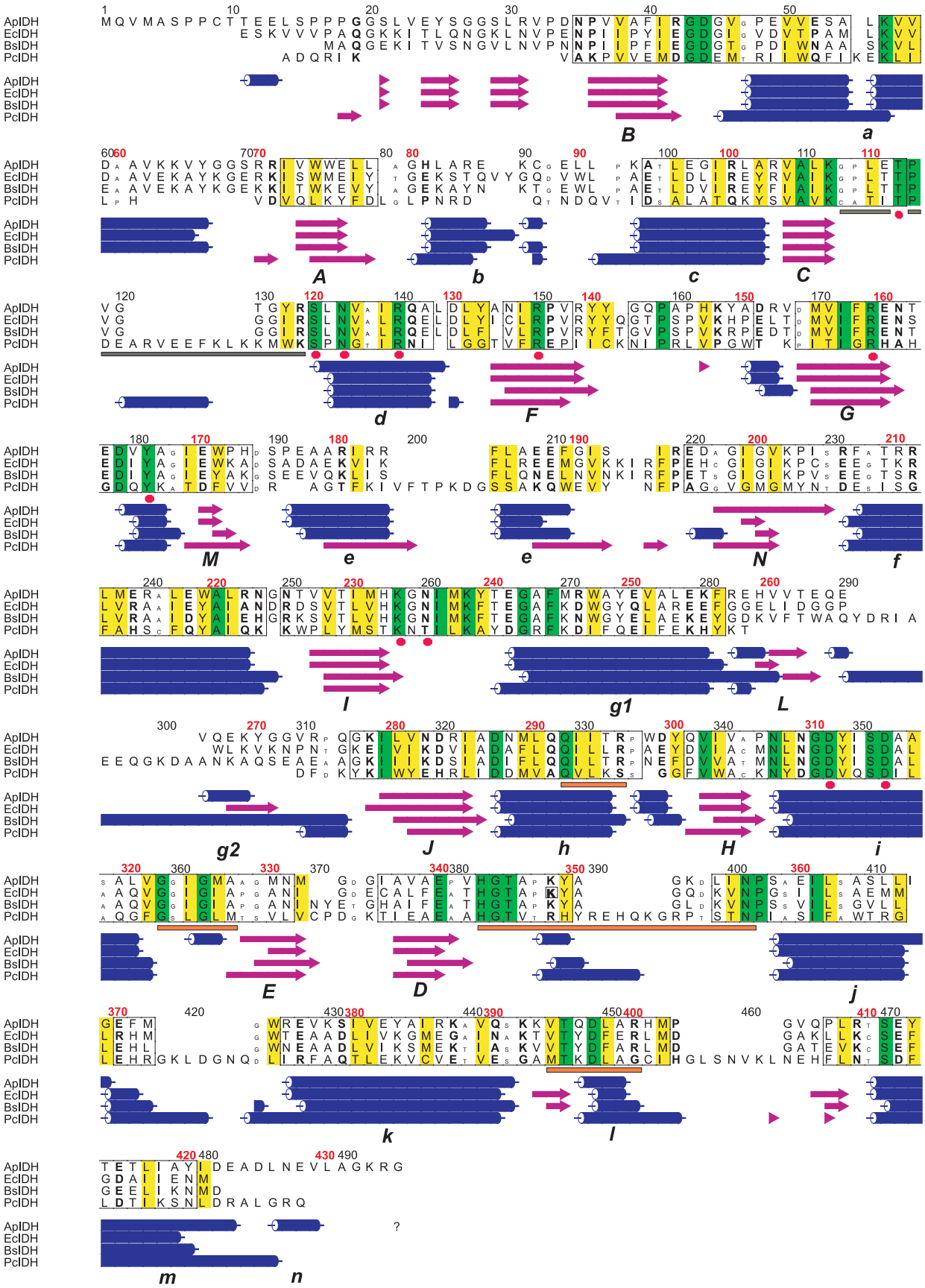

Figure 2. Structure-based sequence alignment of native *Ap*IDH with *E. coli* IDH (*Ec*IDH, PDB code 1CW7), *B. subtilis* IDH (*Bs*IDH, PDB code 1HQS) and porcine IDH (PcIDH, PDB code 1LWD). The residues occurring within structurally equivalent regions are boxed. Helices and strands appear as cylinders and arrows. Conserved residues are coloured

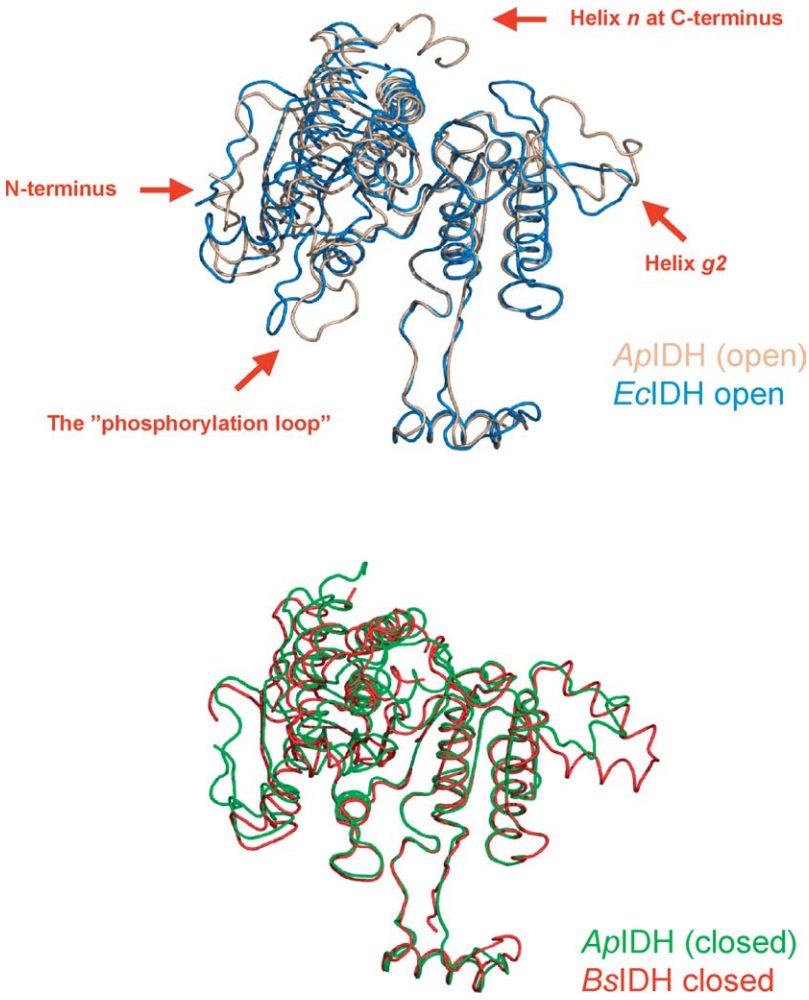

Figure 3. Overlay of native *Ap*IDH (grey) and the open form of *E. coli* IDH (blue, PDB code 1SJS) and of the closed subunit of the ternary complex of *Ap*IDH (green) and *B. subtilis* IDH (red, PDB code 1HQS).

been postulated that *Ec*IDH does not undergo a conformational change upon substrate or cofactor binding but possibly upon product formation.[40] In the case of *Ap*IDH, the crystals easily cracked upon soaking with isocitrate, cofactor and Ca^{2+}, which suggested a conformational change upon substrate binding. It appears that this conformational change was possible only in subunit B, where the lattice packing was weaker, allowing the large domain to move more freely. However, the possibility that the conformational change related to the formation of an ordered ternary complex in one of the subunits effects the same conformational change in the other, cannot be excluded. This raises the question about positive or negative cooperativity. However, whether one of these mechanisms exists, will be investigated in a forthcoming study.

The crystal contacts were also responsible for other differences between the subunits. Especially, the C^{α} positions of helix *g2* (residues 264–269) and the following loop, which forms the unique structural feature in *Ap*IDH mentioned above, were shifted up to 3 Å when the subunits were compared in each dimer. Small differences up to 1.0 Å that might be caused by crystal contacts were found in the loop preceding helix *a* (42–46), in one of the NADP-binding loops (343–357) and in the C terminus end (408–430).

Local differences between the subunits not caused by crystal contacts were found in the so-called "phosphorylation loop" (108–119) preceding the conserved Ser120, equivalent to Ser113 in *Ec*IDH, which is phosphorylated upon inactivation. *Ap*IDH is probably not regulated by phosphorylation, since an IDH kinase/phosphatase has not been found in *A. pernix*. However, the phosphorylation loop is important for the binding of isocitrate in the closed subunit of the ternary complex of *Ap*IDH (see below). In the open subunit, a part of the loop was disordered (see Table 1 for missing

green, positions showing conservation of polar or charged character are in bold, those showing conservation of hydrophobic character are in yellow and residues showing a conservation of small size have smaller font. Sequence numbering according to *Ap*IDH is red. Loops and helices involved in binding of $NADP^+$ are underlined by orange bars. The phosphorylation loop is marked by the grey bar, whereas residues involved in binding of isocitrate are indicated by red dots.

residues). In pseudo-native *Ap*IDH, the phosphorylation loop exhibited different conformations when the subunits were compared and was shifted up to 5 Å, whereas in native *Ap*IDH, the phosphorylation loop was disordered in subunit B.

Dimer association

The interface in *Ap*IDH was found to be similar to that of *Ec*IDH and *Bs*IDH. Except for the formation of the clasp domain, the dimer associates through helices *h* and *i* in both subunits to form a stable four-helix bundle.

The active site

The conformational differences between the subunits of the ternary complex of *Ap*IDH provided a unique possibility to study two different states of isocitrate and cofactor binding in one structure. The active site is located in the cleft between the large and the small domains, and is formed by residues from both of these domains and from both subunits. Thus, the dimer is required for a catalytically active enzyme.

Isocitrate binding

In subunit B of the ternary complex, the α-carboxyl group of isocitrate was bound to three positively charged arginine residues, Arg126, Arg136 and Arg159. The β-carboxylate group was bound by Arg136, Tyr166 and Lys233′ (the prime indicates the neighbouring subunit of the dimer), while Thr112, Ser120 and Asn122 made interactions with the γ-carboxyl group of isocitrate (Figure 4). All of these residues are conserved in other IDHs and the interactions are completely conserved in porcine IDH and mostly in *Ec*IDH. The interaction between isocitrate and the equivalent of Thr112 is not observed in *Ec*IDH.

In subunit A (the open conformation), however, the residues from the large domain, i.e. Thr112, Ser120, Asn122 and Arg126, were all between 6.7 and 7.3 Å away from the substrate, compared to 2.4–3.4 Å in subunit B (Figure 4). Instead, the γ-carboxylate group of isocitrate made interactions with Lys233′ and Asn235′. The remaining interactions with isocitrate were the same as in subunit B. Thr112 and Ser120 belong to the so-called phosphorylation loop and are responsible for substrate specificity, since they interact with the γ-carboxyl group of isocitrate that is replaced by an isopropyl group in isopropylmalate (the substrate for IPMDH). Most likely, the closed conformation is favoured by binding of the correct substrate. The closure was, however, not possible in subunit A due to crystal contacts.

Another difference between the subunits concerning isocitrate binding is that Tyr166 was 4.3 Å away from the substrate in subunit A in contrast to 2.7 Å in subunit B. This residue might be involved in the initial dehydrogenation step as well as in the final protonation of the β-carbanion of oxalosuccinate produced in the decarboxylation step (see below).

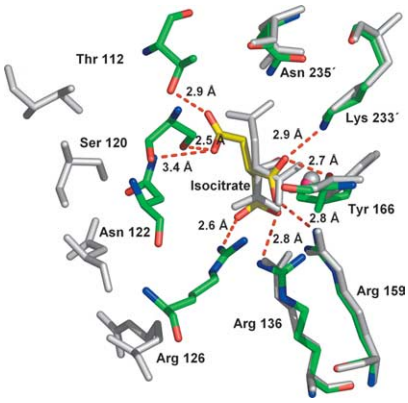

Figure 4. Superposition of bound isocitrate in the open (grey residues) and the closed (green residues) subunits of the ternary complex of *Ap*IDH. The interactions made by Thr112, Ser120 and Asn122 (of the large domain) with the γ-carboxylate group of isocitrate, are responsible for substrate specificity.

Ca^{2+}, which had the strongest density in the map (10σ in subunit A, 6σ in subunit B in a $|F_o|-|F_c|$ density map), was coordinated by eight oxygen ligands in subunit A and six oxygen ligands in subunit B (Figure 5(a) and (b), respectively). The coordination in subunit B is in an octahedral bipyramidal arrangement identical with how Mn^{2+} is bound in the binary isocitrate–porcine IDH complex. In *Ap*IDH, Asp287′, a water molecule (w2 or y4), the α-carboxylate group and the α-hydroxyl group of isocitrate, are the equatorial ligands and Asp311 and another water molecule (w1 or y2) are the axial ligands. However, in subunit A of *Ap*IDH, Asp315 was only 2.4 Å away from the Ca^{2+} (compared to 3.7 Å in subunit B) and is part of the Ca^{2+}-coordination in a way similar to that in the ternary complex of *Ec*IDH where the equivalent aspartate residue is described by the authors as one of the axial ligands together with the α-hydroxyl group of isocitrate.[22] Apparently, the axial plane in the coordination of Mn^{2+} in porcine IDH is interpreted as the equatorial plane in the coordination of Ca^{2+} in *Ec*IDH, which seems possible, since the equatorial water molecule in porcine IDH is absent in *Ec*IDH and an additional water molecule in the axial plane of porcine IDH is present in *Ec*IDH. The coordination of Ca^{2+} in *Ec*IDH is described as octahedral with six ligands but has seven possible ligand candidates if the α-carboxylate group of isocitrate is included.

Cofactor binding

In subunit A, $NADP^+$ was bound exclusively to the large domain, without any interaction with isocitrate, which was bound to the small domain (Figure 6a). In subunit B, however, $NADP^+$ also

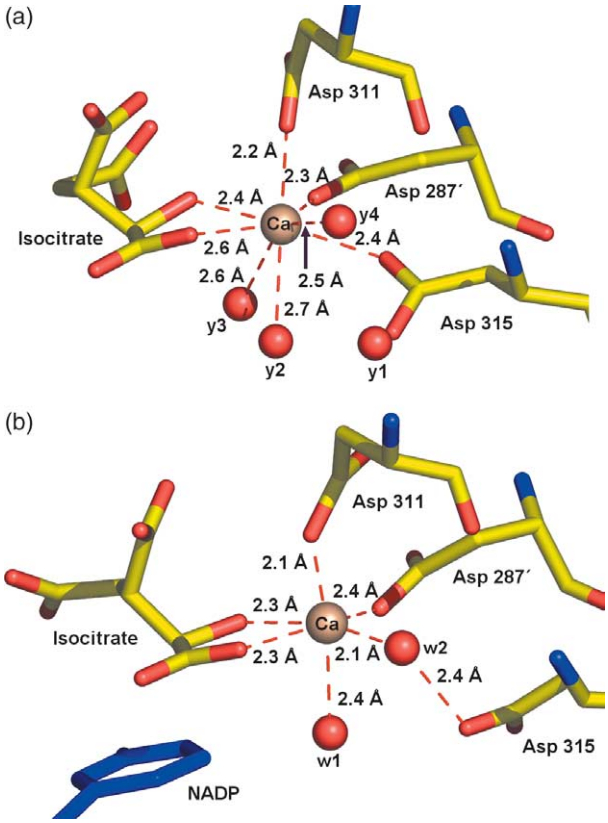

Figure 5. Coordination of Ca^{2+} in subunit A (a) and B (b) of the ternary complex of *Ap*IDH. In subunit A, Ca^{2+} has eight ligands. The water y3 and Asp315 are 2.6 Å and 2.4 Å away, respectively. In subunit B, Ca^{2+} has six ligands, the water corresponding to y3 is absent and Asp315 has moved by 1.3 Å to 3.7 Å from Ca^{2+}.

made interactions with the small domain of the other subunit and with isocitrate, and was closer to the small domain than $NADP^+$ is in the ternary complex of *Ec*IDH (Figures 6b and 7). The residues that interact with the 2′-phosphate group of $NADP^+$ are responsible for the discrimination between $NAD^+/NADP^+$. All of the residues involved in binding of the 2′-phosphate group in *Ec*IDH were conserved in *Ap*IDH except a tyrosine residue (Tyr391 in *Ec*IDH). This tyrosine residue is replaced by Gln396 in *Ap*IDH but did not make any interaction with the 2′-phosphate group. In subunit A of *Ap*IDH, only Tyr349 and Arg400 made conserved 2′-phosphate interactions. In subunit B, the density was poorly defined for Arg400 and the density for Tyr349 was only partially observable. The only clear interaction with the 2′-phosphate group in subunit B was with Gln292′ and Arg296′ from the small domain, 3.0 Å and 3.4 Å away, respectively. In subunit A, these residues were 8–10 Å away because of the open conformation. Lys348 (equivalent to the 2′-phosphate binding Lys344 in *Ec*IDH) was located close to the 2′-phosphate group of $NADP^+$ in subunit A but did not make any interaction, whereas in subunit B, the density for the side-chain was unclear.

Otherwise, the binding of the adenine–ribose moiety of $NADP^+$ in subunit B of *Ap*IDH was very similar to *Ec*IDH, with the adenine ring in the anti conformation with respect to ribose; the ring was lying between the side-chains of His343, Ala346, Ile355, Asp397 and the main chain of Gly325 and Asn356, whereas the adenine ribose moiety was close to Ile324 and Ala346. The 5′-phosphate group of the adenine ribose made interactions with the main chain of residues 344–346 as in *Ec*IDH. The nicotinamide–ribose interacted with the γ-carboxylate group of isocitrate and with the carbonyl group of Thr112 of the phosphorylation loop. In *Ec*IDH, this ribose interacts with the side-chain of the corresponding threonine residue (Thr104). The nicotinamide ring itself was, however, poorly defined. Only an approximate positioning of the nicotinamide was possible. In subunit A, the density of the nicotinamide was also ambiguous, but it is presumably located between Asn122, Glu340 and the carbonyl group of His343. The nicotinamide–ribose interacted with Thr112 as in subunit B.

The catalytic mechanism

Studies on mutants of porcine IDH have revealed that the positively charged Arg110, Arg133 and Lys212′ (equivalent to Arg136, Arg159 and Lys233′ in *Ap*IDH) lower the pK_a of the metal-bound hydroxyl group of isocitrate, promoting the initial proton removal.[41,42] Several base candidates (i.e. aspartate residues) in the enzyme responsible for the initial proton abstraction have been proposed in studies on *Ec*IDH and porcine IDH.[5,43] However, it has also been suggested that two nearby conserved active-site water molecules (w6 and w8) accept this proton as part of a proton relay to solvent where w8 is described as solvent accessible.[12] This was supported by Huang *et al.*,[44] who concluded that none of the aspartate residues functions as a base, since the pH-dependence of V_{max} of porcine IDH cannot be attributed to the ionization of an enzymatic carboxyl group.

The two conserved water molecules are conserved also in subunit B of the ternary complex of *Ap*IDH (w3 and w4, Figure 7). However, w4 in *Ap*IDH, which is equivalent to w8 of the binary porcine IDH complex, is not solvent accessible, since it is covered by the $NADP^+$ molecule. It interacts with the 5′-phosphate group of the nicotinamide–ribose and Asp287′ which is coordinating Ca^{2+}. Moreover, the other conserved water molecule (w3), which is located close to the hydroxyl group of isocitrate, interacts with the strictly conserved Lys233′, which is close to Tyr166 (also strictly conserved).

The mutational studies on porcine IDH have shown that Lys212′ (equivalent to Lys233′ in *Ap*IDH) is essential for decarboxylation of the intermediate oxalosuccinate and that Tyr140

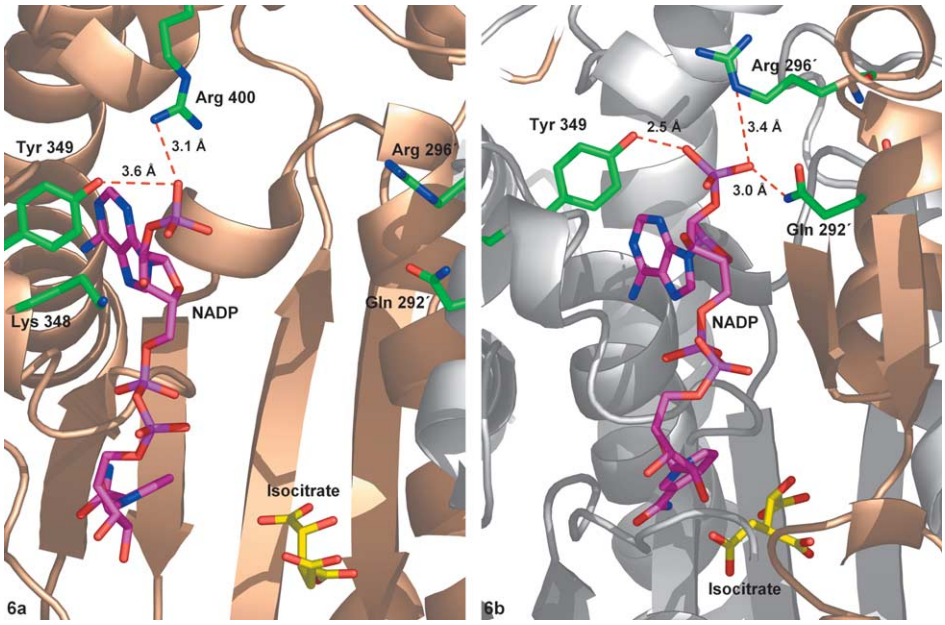

Figure 6. Binding of the 2′-phosphate group of $NADP^+$ in the open form (6a) and in the closed form (6b) of the ternary complex of *Ap*IDH. In the open form, the $NADP^+$ molecule makes interactions with the large domain of subunit A only, whereas in the closed form, the $NADP^+$ molecule makes interactions with the large domain of subunit B and the small domain of subunit A. The distance between the 2′-phosphate group and Arg296′ and Gln292′ of the small domain is about 10 Å in the open subunit (A).

(equivalent to Tyr166 in *Ap*IDH) most probably is the acid that protonates the substrate after decarboxylation.[42] Studies on *Ec*IDH where the equivalent lysine in *Ec*IDH was mutated have also shown that this lysine residue is vital for decarboxylation, whereas a mutant of the equivalent tyrosine in *Ec*IDH was found to have a strong effect on the initial dehydrogenation step.[45]

The results of the studies on *Ec*IDH and porcine IDH may appear as contradictory, but can be explained with the help of the ternary complex of *Ap*IDH by assuming the two conserved water molecules as part of a longer proton relay chain, involving at least five well-defined water molecules, the Lys233′/Tyr166 couple and the $NADP^+$ molecule. Presumably, this chain transfers the proton out to the bulk solvent from the hydroxyl group of isocitrate *via* the water molecules w3 and w4, the 5′-phosphate group of the nicotinamide–ribose, w5, w11 and either of w7 and w8, as well as back to the carbanion of the intermediate oxalosuccinate the same way including Lys233′ and Tyr166 (Figure 7). There are also more water molecules (w1, w6 and w10) that might be part of the chain. In *Drosophila* alcohol dehydrogenase, a coupled tyrosine and lysine pair is part of a proton relay that mediates the dehydrogenation of the alcohol.[46] However, in *Ap*IDH, the lysine/tyrosine pair is presumably part of the proton relay chain in the final proton donation step but is most likely also involved in the initial dehydrogenation step by affecting the pK_a of the hydroxyl group of isocitrate together with the metal ion (Mg^{2+}or Mn^{2+}) and the arginine residues mentioned above (Arg126, Arg136 and Arg159). The ligands of the first two water molecules (w3 and w4) in the chain are completely conserved, the third water molecule (w5) interacts with the 5′-phosphate group, the main chain of Ile285′, which is also conserved, and Asn288′. The end of the chain is very close to the region of helix *g2* which is different in the IDH structures. However, a similar chain can be found at the same location in the ternary complex of *Ec*IDH (PDB code 1AI2) formed by the water molecules w769, w780, w603, w787 and w793, but since the conformation of the ternary complex of *Ec*IDH is slightly more open and the $NADP^+$ molecule is further away from the small domain compared to *Ap*IDH, the environment is slightly different.

Conclusions regarding the active site

Basically, the conserved residues and water molecules involved in binding of the substrates made the same interactions in *Ap*IDH as seen previously in *Ec*IDH and porcine IDH. Thus, the reaction mechanism is very likely to be conserved in these three IDHs. The separate binding of $NADP^+$ to the large domain and of isocitrate to the small domain in the open conformation of *Ap*IDH suggests a random binding order. In *Ec*IDH, the binding order has indeed been shown to be random and the binding of each substrate, isocitrate or $NADP^+$, does not influence the binding of the other.[22,47] The formation of a productive Michaelis–Menten complex, where $NADP^+$ is in position to accept the hydride from isocitrate, appears to be related to the domain rotation observed. The closure of the large domain seems to take place only if the correct substrates have bound, since only

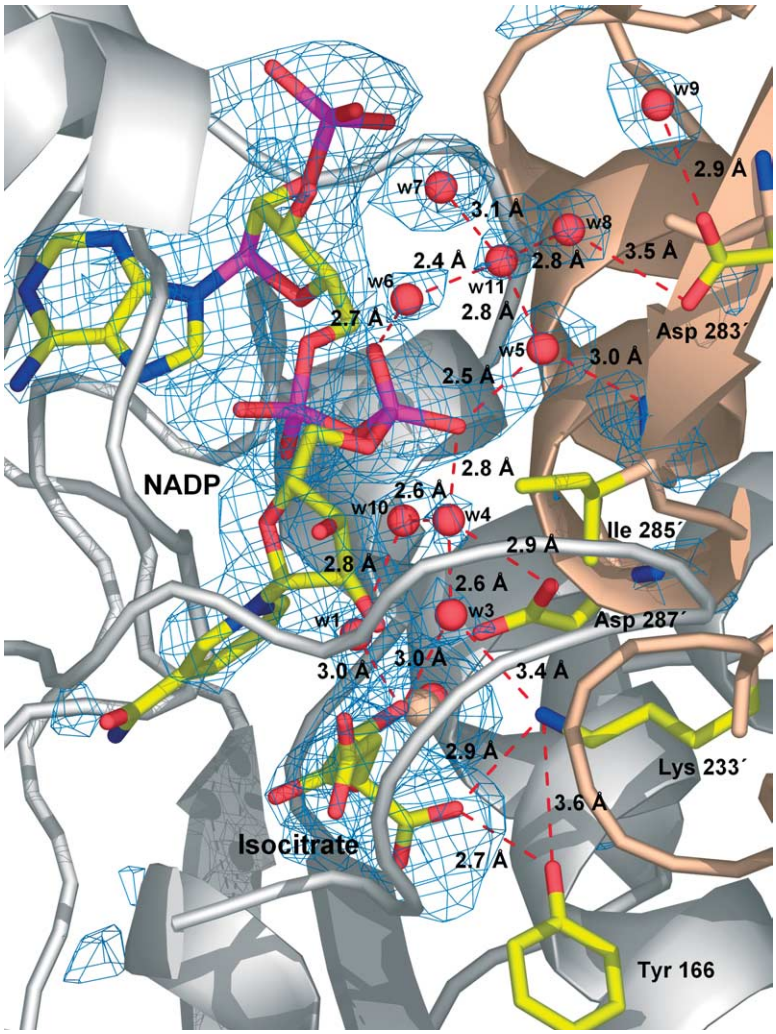

Figure 7. View of the active site of subunit B showing Ca^{2+}, isocitrate, $NADP^+$, bound water molecules and conserved residues. The water molecules and the $NADP^+$ molecule presumably constitute a proton relay chain that might be important for proton translocation. The initial proton abstraction is postulated to occur through a proton transfer from the hydroxyl group of isocitrate *via* the conserved water molecules w3 and w4 to the 5′-phosphate group of the nicotinamide nucleotide, water molecules w5 and w11, which are found close to the bulk solvent. The final proton donation step is assumed to occur *via* the same chain and the coupled lysine/tyrosine pair (Tyr166 and Lys233′). Asp283′, Ile285′ and 287′ are conserved and might be important for the conservation of the water structure in the region. The $F_o - F_c$ omit map ($NADP^+$, Ca^{2+}, isocitrate and water molecules omitted) is contoured at 2σ.

the unique part of the isocitrate molecule interacts with the large domain. It appears also that all of the residues involved in cofactor specificity, making interactions with the 2′-phosphate group, not necessarily have to interact with the 2′-phosphate group of $NADP^+$ simultaneously in the closed form. The initial dehydrogenation and final proton donation are postulated to occur *via* a proton relay to the solvent involving several water molecules, the 5′-phosphate group of the nicotinamide ribose and the coupled residues Lys233′ and Tyr160.

Binding of etheno-$NADP^+$ in pseudo-native ApIDH

The binding of the etheno-adenine moiety of etheno-$NADP^+$ at the surface of the large domain of pseudo-native *Ap*IDH involved two loops that might be part of a novel adenine–nucleotide binding site. The first loop is located between helix *a* and strand *A* and the second loop between helices *j* and *k*. The 2′-phosphate group of etheno-$NADP^+$ was bound to Tyr66, Arg70 and the main chain of Gly373, Arg375 and Glu376 in subunit A. The etheno-adenine ring made an aromatic stacking interaction involving Tyr66, Arg70 and Trp374 of subunit A, and Trp298 and Tyr149 of subunit B of a symmetry-related molecule. The 5′-phosphate group of the nicotinamide–ribosyl interacted with Thr226 and Arg274 of subunit B of the symmetry-related molecule. The nicotinamide–ribosyl moiety itself was not included in the model. All of the residues building up this binding site in subunit A, except Ser69 and Arg375, are conserved or very similar in many IDHs within subfamily I, suggesting a biological function of this binding site. However,

the etheno-$NADP^+$ molecule could not be found at the equivalent site on subunit B.

Thermostability

By sequence comparison, *Ec*IDH and *Bs*IDH are 48% and 45% identical with *Ap*IDH, respectively. Nevertheless, these enzymes are extremely homologous structurally and are therefore good candidates for a comparison with respect to thermostability. The apparent melting temperature (T_m) of *Ec*IDH was determined as 52.6 °C, i.e. 57.3 deg.C lower than the apparent T_m of *Ap*IDH. There is no obvious relation between an apparent T_m and the thermodynamic stability of a protein. However, we intend to relate the large T_m difference of 57.3 deg.C to structural determinants that presumably cause this highly increased thermotolerance of *Ap*IDH. The comparison was made primarily between *Ap*IDH and *Ec*IDH, although data for *Bs*IDH are included in Table 2.

Accessible surface area

There is a clear statistical trend showing that proteins from hyperthermophiles have an increased polar or charged accessible surface area, but a decreased hydrophobic accessible surface area.[1–3,48,49] *Ap*IDH was found to have a significant increase of accessible surface area contributed by polar residues and a slight decrease of surface area contributed by hydrophobic residues compared to *Ec*IDH. The percentage of surface area contributed by charged surface residues was, however, slightly lower in *Ap*IDH. The distribution of charged, polar and hydrophobic surface area of *Ap*IDH, *Ec*IDH and *Bs*IDH is shown in Table 2. The dimer interface of *Ap*IDH buried 2481 Å^2 of the 19,913 Å^2 accessible

Table 2. Characteristics of *Ap*IDH, *Ec*IDH and *Bs*IDH

	*Ap*IDH	*Ec*IDH	*Bs*IDH
PDB code	1TYO	3ICD	1HQS
Apparent melting temperature (°C)	109.9	52.6	Not determined
No. amino acid residues per subunit	435	416	423
Hydrophobic residues[a] (%)	46.2	44.4	43.2
Polar residues[b] (%)	27.6	28.4	30.3
Charged residues[c] (%)	26.2	27.2	26.5
Resolution (Å)	2.1	2.5	1.55
RMSD of C^α *versus* ApIDH (large/small and clasp domain) (Å)	–	1.49/0.82	1.74/0.80
Secondary structure contents (%)	60.7	57.2	64.7
α-Helix content (%)	39.3	36.3	44.9
β-Strand content (%)	14.7	18.0	17.0
3_{10}-Helices content (%)	5.1	2.9	2.8
No. hydrogen bonds	733	666	844
No. hydrogen bonds (SS)[d] per residue	0.085	0.054	0.089
No. hydrogen bonds (SM)e per residue	0.16	0.16	0.21
No. hydrogen bonds (MM)[f] per residue	0.61	0.59	0.68
No. inter-subunit hydrogen bonds	27	24	34
No. ion pairs (6)[g]	73	58	66
No. ion pairs per residue (6)[g]	0.084	0.070	0.078
% of charged residues forming ion pairs (6)[g]	50	42	47
% of ion pairs formed by Arg/Lys/His (6)[g]	63/26/11	41/45/14	47/45/8
% of ion pairs formed by Asp/Glu (6)[g]	45/55	48/52	42/58
% of all Arg forming ion pairs (6)[g]	55	71	89
% of all Lys forming ion pairs (6)[g]	32	32	37
% of all His forming ion pairs (6)[g]	43	60	40
% of all Asp forming ion pairs (6)[g]	79	56	57
% of all Glu forming ion pairs (6)[g]	50	40	46
No. residues forming two ion pairs (6)[g]	23	20	22
No. residues forming three ion pairs (6)[g]	4	0	0
No. 2/3/4 member networks (6)[g]	29/7/1	20/16/2	28/13/1
No. 5/6/7/15 member networks (6)[g]	1/1/3	0/0/0	1/1/0/0
No. inter-subunit ion pairs (6)[g]	7	9	10
Net charge	+1	−19	−25
Accessible surface area of dimer (Å^2)	34812	32465	32974
Buried inter-subunit surface (% of monomer)	12.5	15.4	16.9
Distribution of hydrophobic/polar/charged residues at accessible surface (%)	23.3/28.8/47.9	26.7/23.1/50.2	22.0/27.5/50.5
Distribution of hydro-phobic/polar/charged residues at interface (%)	48.8/26.5/24.7	48.6/19.5/31.9	44.8/25.8/29.4

[a] Hydrophobic residues: A, V, L, I, W, F, P, M.
[b] Polar residues: G, S, T, Y, N, Q, C.
[c] Charged residues: R, K, H, D, E.
[d] SS, side-chain–side-chain hydrogen bonds.
[e] SM, side-chain–main-chain hydrogen bonds.
[f] MM main-chain–main-chain hydrogen bonds.
[g] 4.2 Å cutoff.

surface of the monomer, giving a solvent-accessible surface area of 34,864 Å^2 for the dimer. In comparison, *Ec*IDH dimer buries 2954 Å^2 of the 19,237 Å^2 of the monomer with a total solvent-accessible surface of 32,566 Å^2. In other words, 12.5% of the monomer of *Ap*IDH was buried at the interface compared to 15.4% of the monomer of *Ec*IDH. In *Bs*IDH, 16.9% of the monomer is buried at the interface.

Charged residues and ionic interactions

The total number of charged residues per dimer was basically the same in *Ap*IDH compared to *Ec*IDH: 228 (26.3%) and 226 (27.2%), respectively. Data from genome sequencing shows that the *A. pernix* protein pool actually contains fewer charged residues than do the mesophiles.[50] The charged residues in *Ap*IDH formed about the same total number of ion pairs when a cutoff distance of 4 Å was used. This was surprising, since hyperthermophilic proteins often contain more ion pairs than their mesophilic counterparts do.[1] However, using a cut off of 4.2 Å, resulted in a dramatic increase of ion pairs: 73 ion pairs in *Ap*IDH compared to 58 in *Ec*IDH (see Table 2).

In order to compare the ion pairs qualitatively, the ion pairs were evaluated with respect to involved secondary structure elements. Ion pairs formed by residues that are located on the same secondary structure element (intra-helical, intra-strand or intra-loop ion pairs) probably have a smaller effect on thermostabilization, whereas ion pairs that connect different elements, especially loops, are expected to contribute more to enhanced thermostability.[51] It is, however, also important for the stability of proteins that charged residues have a partner with the opposite charge, and that repulsive or unfavourable contacts are minimized.[3]

In *Ap*IDH, 34 of 115 residues involved in ion pairing in the dimer were located on loops. Eight of them made intra-loop interactions, whereas the remaining residues made interactions with other loops or secondary structure elements. In *Ec*IDH, 34 of 96 residues were located on loops of which none made intra-loop interactions. The number of ion pair residues located on helices or strands was 81 in *Ap*IDH of which 28 were intra-helical or intra-strand. In *Ec*IDH, 63 residues were located on helices or strands of which six were found on the same helix. Thus, there was no significant qualitative difference between the ionic interactions in *Ap*IDH and *Ec*IDH with respect to involved secondary structural elements, that could explain the difference in thermostability.

However, the stabilising contribution of ion pairs is increasing at higher temperatures, since the desolvation penalty is decreased with higher temperature, partially as a result of the decrease of the dielectric constant of water at higher temperatures.[1,50,52]

Moreover, when ionic networks were considered, another significant difference between *Ap*IDH and *Ec*IDH emerged. Ionic networks are assumed to be energetically more favourable than the equivalent number of isolated ion pairs.[48] Due to increased multiplicity the entropic penalty in an ionic network is smaller than for pairwise interactions. Whereas *Ec*IDH contained 16 three-membered and two four-membered networks per dimer, *Ap*IDH contained seven three-membered, one four-membered, one five-membered, one six-membered and three seven-membered networks (Table 2). One ionic network in *Bs*IDH contained six members. Two of the seven-membered networks in *Ap*IDH consisted of Asp130, Arg211, Glu214, Arg215, Glu218, Lys255 and Asp334 in each dimer. The third seven-membered network contained Arg139, Asp154, Arg386, Glu370, Arg375, Lys378 and Glu382 in subunit A. All three networks were located between the large and the small domains at the opposite side of the active site. Using a cut off of 6 Å, resulted in a dramatic extension of one of the seven-membered networks into a network of 23 members in subunit B and the equivalent network in subunit A into 15 members, whereas in *Ec*IDH, a 6 Å cut off resulted in two seven-membered networks only. This result suggests that the cooperation of weaker ionic interactions in networks might also be important for thermostabilization.

In order to make a rough estimation of the degree of electrostatic optimization of the charged residues, the net charge of the dimer was calculated without respect to any cutoff distance. Assuming 50% of the histidine residues as charged at neutral pH resulted in a net charge of +1 in *Ap*IDH, suggesting that excess charges were minimized. In *Ec*IDH, *Bs*IDH and porcine IDH, the respective net charges were determined to −19, −25 and +20, indicating a large fraction of excess charges in the mesophilic IDHs. Most likely, a high content of excess charges is a destabilising factor at neutral pH and normal salt concentration. To confirm this result, the net charges of other hyperthermostable IDHs, with known apparent T_m values, were calculated with the following results: *Pyrococcus furiosus* IDH, +1 (T_m = 103.7 °C);[8] *Thermotoga maritima* IDH, +6 (T_m = 98.3 °C);[8] and *Archaeoglobus fulgidus* IDH, −9 (T_m = 98.5 °C).[8] These results show a trend towards a zero net charge in the IDHs with the highest apparent T_m values (i.e. *Ap*IDH and *Pf*IDH) and a decrease of excess charges in *Tm*IDH and *Af*IDH compared to *Ec*IDH, *Bs*IDH and porcine IDH.

Single-site mutations in the ionic network and at the subunit interface

The mutations D130N, D344N, R211Q and R211M were introduced with the aim of disrupting the seven-membered ionic network. Kinetic characterization of the mutants indicated that they were catalytically active (Table 3). The cofactor affinities were similar to the wild-type recombinant *Ap*IDH, suggesting that the mutations did not cause any large overall structural changes. All of the four

Table 3. Kinetic parameters

	kDa	T_m (°C)	ΔT_m (deg. C)	K_m (μM NADP)	V_{max} (μmol/min/mg)	K_{cat} (s^{-1})	K_{cat}/K_m ($\times 10^7$ M^{-1} s^{-1})
*Ap*IDH$_{NATIVE}$	47.9	109.9	–	10.0	226	181.25	1.81
E188/Q		108.5	−1.4	18.3	262	210.1	1.15
E188/A		107.6	−2.3	16.9	263	210.6	1.25
D130/N		106.4	−3.5	32.9	324	259.6	0.79
D334/N		104.4	−5.5	12.9	200	160.1	1.24
R211/Q		101.2	−8.7	34.2	307	246.2	0.72
R211/M		98.6	−11.3	19.0	277	222.1	1.17
C87/S		100.3	−9.6	25.9	241	193.1	0.75

mutants had lower apparent melting temperatures compared to the wild-type enzyme (Table 3). R211M and R211Q had the largest effects on the stability of the enzyme with apparent melting temperatures of 98.6 °C and 101.2 °C, i.e. 11.3 deg. and 8.7 deg. C less then the wild-type enzyme. R211 is located on helix *f* on the small domain and seems to be a key residue in the seven-membered network, since it is responsible for an inter-domain interaction where both Asp130 and Asp334 of the large domain are involved (Figure 8(a)). The aspartate residues are located on two loops that are stabilized through the ionic interactions. The slightly higher T_m value of the R211Q mutant can be explained by the ability of the glutamine residue to form a hydrogen bond with one of the aspartate residues. An inter-subunit ion pair between R180 and E188 was detected in the clasp domain where two mutations, E188Q and E188A, were introduced with the aim of disrupting this connection. These mutations resulted in apparent melting temperatures close to that of the wild-type enzyme; 108.5 °C and 107.6 °C, respectively, and the conclusion that the inter-subunit ion pair has a minor contribution to the thermostabilization of *Ap*IDH. Taken together, these data prove that the seven-membered network is involved in thermostabilization of *Ap*IDH.

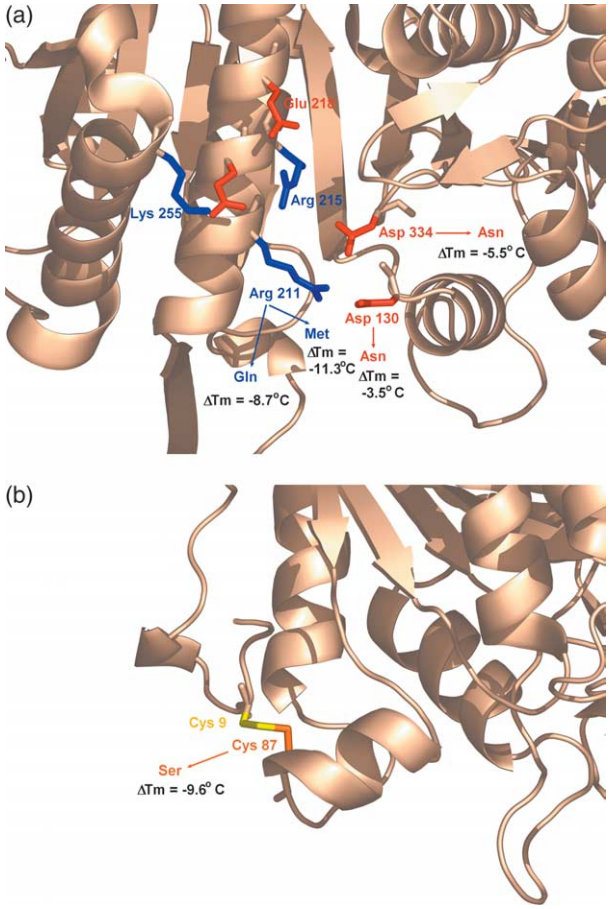

Figure 8. (a) Mutants of the seven-membered ionic network and (b) the disulfide bond at the N terminus end revealed that these features are major determinants of increased thermostability in *Ap*IDH. The mutants R211M and R211Q decreased the apparent melting temperature (T_m) of *Ap*IDH by 11.3 deg. C and 8.7 deg. C, respectively, whereas D334N and D130N decreased the apparent T_m by 5.5 deg. C and 3.5 deg. C, respectively. A mutant of the disulfide, C87S, decreased the T_m by 9.6 deg. C.

Disulfide bond

*Ap*IDH has a disulfide bridge between Cys9 and Cys87 which is likely to anchor the N terminus and prevent it from thermal unfolding. This disulfide is conserved in IDH from *Caldococcus noboribetus*.[53] A mutation Cys87Ser was introduced in order to disrupt the disulfide bond and was found to decrease the apparent melting temperature by 9.6 deg. C (Table 3 and Figure 8(b)), which shows that the disulfide bridge is a major contributing factor to the thermostability of *Ap*IDH.

Disulfides as stabilising and catalytic structure elements seem to be more abundant in proteins from hyperthermophiles than generally accepted. Among the more recent examples found in Archaea are the two catalytic disulfides in *P. furiosus* protein disulfide oxidoreductase[54] and the stabilising disulfide in *Sulfolobus acidocaldarius* Rieske iron–sulfur protein.[55] There are also other examples of disulfides found in crystal structures of intracellular proteins.[56–58] According to a recent study, the abundance of intracellular disulfide bonds is unusually high in *A. pernix* and might be related to the oxygen-tolerant aerobic nature of this organism.[59]

Hydrogen bonds

A small increase in the fraction of hydrogen bonds per residue was observed in *Ap*IDH

compared to *Ec*IDH; a 2.0% increase in main-chain–main-chain hydrogen bond interactions and a 3.1% increase of side-chain–side-chain hydrogen bond interactions (Table 2). However, the determination of hydrogen bonds is highly resolution-dependent, which is reflected by the high content of hydrogen bonds in *Bs*IDH, determined at a resolution of 1.55 Å. Therefore, no conclusions could be made about increased hydrogen bonding as a thermostabilising factor in *Ap*IDH.

Dimer association

A strengthened dimer association is often found in hyperthermophilic proteins. In the clasp domain of *Ap*IDH, the conserved aromatic interaction between Phe184 and Phe184′ on helix *e* was extended to an aromatic cluster by Phe189 and Phe189′, which also involved Trp171 and Trp171′ on strand *M*. This cluster may provide additional stabilization of the interface. Compared to *Ec*IDH, the intersubunit surface was slightly more polar in *Ap*IDH but less charged (see Table 2). In *Bs*IDH, the interface surface was more charged but less hydrophobic compared to *Ap*IDH.

Loop deletions

Shortening of loop regions is commonly seen in hyperthermostable proteins[51] and is sometimes reflected by a higher content of secondary structure. Three small loop deletions were observed in *Ap*IDH as compared to *Ec*IDH: Tyr78-Gly79, Asp135, Lys187–Arg189 in *Ec*IDH (*E. coli* numbering). In addition, helix *g2* replaces the loop between strand *L* and strand *K* in *Ec*IDH.

Conclusions regarding thermostability

Two major determinants conferring the increased thermostability in *Ap*IDH were confirmed by mutational studies; a disulfide bridge at the N terminus and a seven-membered inter-domain ionic network with many neighbouring charged residues extending the network to 23 or 15 members if a cutoff of 6 Å instead of 4.2 Å was used. As a general conclusion, the total number of ion pairs and the size of the ionic networks increased dramatically in *Ap*IDH compared to *Ec*IDH. A further important difference was found in the reduced net charge of only +1 in *Ap*IDH, a factor that presumably contributes to improved electrostatic optimization.

Materials and Methods

Mutations and protein purification

Mutations in *Ap*IDH were introduced by PCR-based mutagenesis using the QuickChange Site-Directed Mutagenesis Kit (Stratagene). The primers used are listed in Table 4 and the mutation sites are underlined. Incorporation of the mutation at the desired position was confirmed by DNA sequence analysis of the entire *idh*-coding region using an ABI 3700 DNA sequencer. Mutated enzyme was produced in *E. coli* strain BL21-Codon-Plus (DE3)-RIL (Stratagene) as described for the wild-type enzyme and purified using heat treatment and Red-Sepharose chromatography.[8]

Differential scanning microcalorimetric measurements

Differential scanning calorimetry (DSC) was carried out with a MicroCal MCS calorimeter controlled by the MCS OBSERVER program (MicroCal). The samples were dialysed against the reference buffer used in the experiment (50 mM potassium phosphate buffer (pH 7.5), 0.1 M NaCl) and degassed for 20 minutes prior to the calorimetric analysis. A protein concentration of 1.3 mg/ml was used. The calorimetric scans were carried out between 20 °C and 120 °C using a scan rate of 1 K/minute. A constant pressure of 2 bar was applied in order to avoid boiling at high temperatures. Each sample was scanned a second time after the actual calorimetric scan to estimate the reversibility of the unfolding transition.

Crystallization

Crystals of native *Ap*IDH were grown as reported.[60] Crystals were soaked in two different solutions: (1) 11% (w/v) PEG 6000, 100 mM *d*-isocitrate-buffer pH 5.6 (D_s-(+)-threo-isocitric acid monopotassium salt, FLUKA 58790), 50 mM $CaCl_2$ and 5 mM $NADP^+$. (2) The original mother solution with 10 mM etheno-$NADP^+$ (nicotinamide 1,N^6-ethenoadenine dinucleotide phosphate, SIGMA N-3630) added. Soaking proceeded for about 48 hours prior to the data collection.

Data collection, processing and characterization of crystals

All datasets were collected at the synchrotron beamline I711 at MAXLAB, Lund, Sweden, except a dataset of a native *Ap*IDH crystal that was collected as described[60] and used in the beginning of the refinement. All other crystals were flash-frozen with boiling nitrogen at 100 K in their respective mother or soaking solution containing 20% (v/v) ethylene glycol as cryoprotectant. An MarCCD

Table 4. Primer sequences

Primer sequence	*Ap*IDH mutant
5′-TTTGCAACTAGGATGCTTATGGAGAGG-3′ 5′-CCTCTCCATAAGCATCCTAGTTGCAAA-3′	R211/M
5′-TTTGCAACTAGGCAGCTTATGGAGAGG-3′ 5′-CCTCTCCATAAGCTGCCTAGTTGCAAA-3′	R211/Q
5′-CCAGGCGCTAAACCTCTACGCTAAC-3′ 5′-GTTAGCGTAGAGGTTTAGCGCCTGG-3′	D130/N
5′-TGAACATGGGCAACGGCATAGC-3′ 5′-GCTATGCCGTTGCCCATGTTCA-3′	D334/N
5′-CCTCGCAGAGCAGTTCGGGATAT-3′ 5′-ATATCCCGAACTGCTCTGCGAGG-3′	E188/Q
5′-CCTCGCAGAGGCTTTCGGGATAT-3′ 5′-ATATCCCGAAAGCCTCTGCGAGG-3′	E188/A
5′-CTCGCCCGGGAGAAGAGCGGCGAGCT GCTTCCC-3′ 5′-GGGAAGCAGCTCGCCGCTCTTCTCCC GGGCGAG-3′	C87/S

165 detector and MarCCD v0.5.33 software were used for data collection. The data were indexed and scaled with the HKL package.[61] The datasets were completed by including all possible *hkl* and R_{free} columns using UNIQUE.[62] Care was taken to keep the same R_{free} flags in all datasets. Structure factors were calculated with the program TRUNCATE.[62] Data collection parameters and processing statistics are given in Table 1. The Matthews coefficient (V_M) was 2.6 Å^3/Da, suggesting a dimer in the asymmetric unit.

Molecular replacement

Molecular replacement was done with the 2.6 Å dataset of the native *Ap*DIH collected.[60] A trimmed model of the structure of *Ec*IDH (PDB code: 3ICD) was used as search model. The sequence identity between IDH from *A. pernix* and *E. coli* is 48%. The orientation of the search model was determined by a fast rotation function calculated in CNS[63] using data between 15 Å and 4 Å. In order to allow for different domain orientations in the model, Patterson correlation refinement in CNS[63] was invoked both before and after the translation search with five groups selected as rigid bodies. The best solution from the cross-rotation search scored 2.29σ above the mean. This peak gave a clear solution in the translation search in space group $P4_32_12$. The monomer was fixed while searching for the second subunit. The highest translation score for the second subunit was found by the 11th cross-rotation solution, which scored only 0.98σ above the mean. This suggested already that the two subunits may show different domain orientations. Later on, a refined dimer of the native *Ap*IDH was used as model in molecular replacement searches for native *Ap*IDH (where the rotation function scored 3.79σ above the mean), pseudo-native *Ap*IDH(4.48σ) and the ternary complex (3.37σ), The same parameters as in the initial search were used.

Crystallographic refinement, density modification and model building

The initial *R*-factor of the native *Ap*IDH was 50.6% (R_{free} 51.0%). The orientations of the two subunits and the domains were refined with rigid body refinement but the *R*-factor was not improved much, since the domains were refined during the Patterson correlation refinement. Simulated annealing in CNS[63] was applied using torsion angles. The starting temperature was 3000 K and was decreased in steps of 25 K. The best results were obtained using tight NCS restraints (300 kcal/(mol/Å^2)) between the small domains and between the clasp domains and loose restraints (50 kcal/(mol/Å^2)) between the large domains. After grouped *B*-factor refinement, using one *B*-factor for each main chain and one for each side-chain, the *R*-factor dropped to 41.6% (R_{free} 45.0%).

Density modification and averaging according to the 2-fold non-crystallographic symmetry was initially done within CNS[63] and revealed continuous density for most of the dimer, showing several of the deleted parts of the native *Ap*IDH model. Additional density showed up gradually during model building and refinement aided by 2-fold averaging in RAVE.[64] NCS operators were improved with the program IMP[64] the mask was generated by MAMA[64] and edited manually in the regions surrounding the missing parts of the model.

Model building was performed with the program O.[65] Each model building session was followed by molecular dynamics refinement in CNS[63] at constant temperature. When subunit A was built, it was rotated over subunit B and adjusted according to the averaged map. However, the refined *B*-factors were alarmingly high, many above 100 Å^2. The *R*-factor at this stage was 26.3% (R_{free} 31.4%). The average *B*-factor for the large domain of subunit B was 42.8 Å^2, compared with 22.2 Å^2 in subunit A. At this stage, a new dataset of native *Ap*IDH with resolution to 2.20 Å as well as data for the pseudo-native *Ap*IDH (2.15 Å) and the ternary complex (2.3 Å) were collected and used in molecular replacement as described.

The initial *R*-factors of the native *Ap*IDH, pseudo-native *Ap*IDH and the ternary *Ap*IDH complex were 37.6% (R_{free} 38.2%), 38.2% (R_{free} 38.2%) and 45.4% (R_{free} 46.4%), respectively. The models were refined with rigid body refinement, simulated annealing and grouped *B*-factor refinement using the same protocol as described above, before adjustments of the models were made in O.[65] The NCS restraints were optimized, taking into account differences due to crystal contacts and averaged maps. The final refinement was done in REFMAC[62] with individual isotropic *B*-factor refinement and parameters for anisotropic thermal motion of rigid groups (TLS refinement), which were defined according to the domains of *Ap*IDH

In the case of the ternary complex, the binding of the cofactor and substrate caused a dramatic closure of the large domain in subunit B. This domain was adjusted or rebuilt residue by residue from 135 to 70 and from 321 to 350. Since the density of the remaining part of this domain was poorly defined, residues 6–69 and 351–426 were added from the large domain of subunit A after a rotation of the whole domain over the newly built residues of subunit B. NCS restraints for the large domains were not applied for the ternary complex. Density modification and averaging of the ternary complex was performed with DMMULTI[62] using one rotation matrix for each domain.

Water molecules were added using the program ARP_WARP[62] and kept only if density was present in both $2|F_o|-|F_c|$ (if above 1 σ) and $|F_o|-|F_c|$ maps (if above 3.0 σ), if a suitable hydrogen bond donor/acceptor was present and if the temperature factor was below 60 Å^2. Some water molecules were added manually. The active site water molecules were confirmed by excluding them one by one followed by investigation of $|F_o|-|F_c|$ difference density maps. In the native and the pseudo-native *Ap*IDH, the water molecules were divided into two separate TLS groups according to the subunits. In the ternary complex, the water molecules on subunit A formed one TLS group, whereas water molecules belonging to subunit B were divided into two TLS groups according to the domains. In addition, the active site water molecules on each subunit were grouped together with the respective isocitrate molecules and Ca^{2+} ions.

Stereochemical quality of the models was checked with the program PROCHECK[66] as implemented in CCP4.[62] Superposition and determination of RMS differences between subunits, domains and homologous structures were done in O.[65]

Domains were considered as rigid bodies. Domain rotations were defined as the rotation angle required to superimpose the large domain in subunit A on the same domain in subunit B, with the small domains of subunit A and B remaining in a fixed superimposed position. Axis direction and rotation angle were calculated from the O rotation matrix by CONVROT (Winfried Meining, unpublished program). No significant translation component was detected.

Sequence alignment was performed with the program STAMP[67] and was based on secondary structural assignments made using the program DSSP.[68]

Surface calculations and analysis of non-covalent interactions

Ion pairs and ionic networks were analysed using the programs CONTACT[62] and IONSTAT (Winfried Meining, unpublished program) with a maximum distance of 4.2 Å[69] or 6 Å. Accessible surface areas were calculated using CNS with Ala, Ile, Leu, Met, Phe, Pro, Trp and Val defined as hydrophobic residues, Asn, Gln, Ser, Thr, Tyr, Cys and Gly treated as polar residues and Asp, Glu, Arg, Lys and His as charged residues. The water probe radius was 1.4 Å and the accuracy of the numerical integration was set to 0.12. Water was excluded from the model.

Net charges were calculated by summing the total number of positive (Arg, Lys and His) and negative (Asp, Glu) charges, assuming 50% of the histidine residues as charged at neutral pH due to the neutral pK_a value of this residue.

Hydrogen bonds were calculated using HBPLUS[70] v3.15 and the following default parameters: maximum distances for D–A, 3.9 Å and for H–A, 2.5 Å; minimum angles for D–H–A, D–A–AA and H–A–AA was 90°.[71] Ion pairs that were counted as hydrogen bonds by the program were excluded.

Figure preparations

Figure 1a and b were generated using the programs MOLSCRIPT[72] and RASTER3D v2.7b.[73] Figures 1c, 3–8 were made using PYMOL.[74] Figure 2 was prepared using the program ALSCRIPT.[75]

Acknowledgements

This work was supported by a PhD grant from Södertörn University College to M.K. We are grateful to Dr Aurora Martinez, Department of Biomedicine, University of Bergen for access to her laboratory facilities and expertise for determining protein melting temperatures of the ion pair mutants, and to Marit Steine Madsen for expression and purification of *Ap*IDH. M.K. thanks Dr Winfried Meining for much-appreciated assistance in how to handle various computer programs.

References

1. Karshikoff, A. & Ladenstein, R. (2001). Ion pairs and the thermotolerance of proteins from hyperthermophiles: a "traffic rule" for hot roads. *Trends Biochem. Sci.* **26**, 550–556.
2. Yip, K. S., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E. *et al*. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure*, **3**, 1147–1158.
3. Spassov, V. Z., Karshikoff, A. D. & Ladenstein, R. (1995). The optimization of protein–solvent interactions: thermostability and the role of hydrophobic and electrostatic interactions. *Protein Sci.* **4**, 1516–1527.
4. Vogt, G. & Argos, P. (1997). Protein thermal stability: hydrogen bonds or internal packing? *Fold. Des.* **2**, S40–S46.
5. Hurley, J. H., Dean, A. M., Koshland, D. E. & Stroud, R. M. (1991). Catalytic mechanism of NADP(+)-dependent isocitrate dehydrogenase: implications from the structures of magnesium-isocitrate and NADP+ complexes. *Biochemistry*, **30**, 8671–8678.
6. Amend, J. P. & Shock, E. L. (1998). Energetics of amino acid synthesis in hydrothermal ecosystems. *Science*, **281**, 1659–1662.
7. Melendez-Hevia, E., Waddell, T. G. & Cascante, M. (1996). The puzzle of the Krebs citric acid cycle: assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *J. Mol. Evol.* **43**, 293–303.
8. Steen, I. H., Madern, D., Karlstrom, M., Lien, T., Ladenstein, R. & Birkeland, N. K. (2001). Comparison of isocitrate dehydrogenase from three hyperthermophiles reveals differences in thermostability, cofactor specificity, oligomeric state, and phylogenetic affiliation. *J. Biol. Chem.* **276**, 43924–43931.
9. Steen, I. H., Lien, T. & Birkeland, N. K. (1997). Biochemical and phylogenetic characterization of isocitrate dehydrogenase from a hyperthermophilic archaeon, *Archaeoglobus fulgidus*. *Arch. Microbiol.* **168**, 412–420.
10. Hurley, J. H., Thorsness, P. E., Ramalingam, V., Helmers, N. H., Koshland, D. E. & Stroud, R. M. (1989). Structure of a bacterial enzyme regulated by phosphorylation, isocitrate dehydrogenase. *Proc. Natl Acad. Sci. USA*, **86**, 8635–8639.
11. Singh, S. K., Matsuno, K., LaPorte, D. C. & Banaszak, L. J. (2001). Crystal structure of *Bacillus subtilis* isocitrate dehydrogenase at 1.55 Å. Insights into the nature of substrate specificity exhibited by *Escherichia coli* isocitrate dehydrogenase kinase/phosphatase. *J. Biol. Chem.* **276**, 26154–26163.
12. Ceccarelli, C., Grodsky, N. B., Ariyaratne, N., Colman, R. F. & Bahnson, B. J. (2002). Crystal structure of porcine mitochondrial $NADP^+$-dependent isocitrate dehydrogenase complexed with Mn^{2+} and isocitrate. Insights into the enzyme mechanism. *J. Biol. Chem.* **277**, 43454–43462.
13. 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.
14. Hurley, J. H. & Dean, A. M. (1994). Structure of 3-isopropylmalate dehydrogenase in complex with NAD+: ligand-induced loop closing and mechanism for cofactor specificity. *Structure*, **2**, 1007–1016.
15. Wallon, G., Kryger, G., Lovett, S. T., Oshima, T., Ringe, D. & Petsko, G. A. (1997). Crystal structures of *Escherichia coli* and *Salmonella typhimurium* 3-isopropylmalate dehydrogenase and comparison with their thermophilic counterpart from *Thermus thermophilus*. *J. Mol. Biol.* **266**, 1016–1031.
16. Yasutake, Y., Watanabe, S., Yao, M., Takada, Y., Fukunaga, N. & Tanaka, I. (2002). Structure of the monomeric isocitrate dehydrogenase: evidence of a protein monomerization by a domain duplication. *Structure (Camb)*, **10**, 1637–1648.
17. Rossmann, M. G., Moras, D. & Olsen, K. W. (1974). Chemical and biological evolution of nucleotide-binding protein. *Nature*, **250**, 194–199.
18. Chen, R., Greer, A. F. & Dean, A. M. (1997). Structural constraints in protein engineering–the coenzyme specificity of *Escherichia coli* isocitrate dehydrogenase. *Eur. J. Biochem.* **250**, 578–582.
19. Hurley, J. H., Chen, R. & Dean, A. M. (1996).

Determinants of cofactor specificity in isocitrate dehydrogenase: structure of an engineered $NADP^+$ NAD^+specificity-reversal mutant. *Biochemistry*, **35**, 5670–5678.
20. Miyazaki, K. & Oshima, T. (1994). Co-enzyme specificity of 3-isopropylmalate dehydrogenase from *Thermus thermophilus* HB8. *Protein Eng.* **7**, 401–403.
21. Steen, I. H., Lien, T., Madsen, M. S. & Birkeland, N. K. (2002). Identification of cofactor discrimination sites in NAD-isocitrate dehydrogenase from *Pyrococcus furiosus*. *Arch. Microbiol.* **178**, 297–300.
22. Stoddard, B. L., Dean, A. & Koshland, D. E. (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.
23. Dean, A. M. & Koshland, D. E., Jr (1993). Kinetic mechanism of *Escherichia coli* isocitrate dehydrogenase. *Biochemistry*, **32**, 9302–9309.
24. Siebert, G., Carsiotis, M. & Plaut, G. W. (1957). The enzymatic properties of isocitric dehydrogenase. *J. Biol. Chem.* **226**, 977–991.
25. Bolduc, J. M., Dyer, D. H., Scott, W. G., Singer, P., Sweet, R. M., Koshland, D. E. & Stoddard, B. L. (1995). Mutagenesis and Laue structures of enzyme intermediates: isocitrate dehydrogenase. [Erratum appears in *Science* 1995 Oct 20; **270**, 365]. *Science*, **268**, 1312–1318.
26. Dalziel, K. & Londesborough, J. C. (1968). The mechanisms of reductive carboxylation reactions. Carbon dioxide or bicarbonate as substrate of nicotinamide-adenine dinucleotide phosphate-linked isocitrate dehydrogenase and malic enzyme. *Biochem. J.* **110**, 223–230.
27. Lienhard, G. E. & Rose, I. A. (1964). The stereochemistry of decarboxylation of isocitrate by isocitric acid dehydrogenase. *Biochemistry*, **170**, 185–190.
28. Hathaway, J. A. & Atkinson, D. E. (1963). The effect of adenylic acid on yeast nicotinamide adenine dinucleotide isocitrate dehydrogenase, a possible metabolic control mechanism. *J. Biol. Chem.* **238**, 2875–2881.
29. Cohen, P. F. & Colman, R. F. (1972). Diphosphopyridine nucleotide dependent isocitrate dehydrogenase from pig heart. Characterization of the active substrate and modes of regulation. *Biochemistry*, **11**, 1501–1508.
30. Garnak, M. & Reeves, H. C. (1979). Phosphorylation of isocitrate dehydrogenase of *Escherichia coli*. *Science*, **203**, 1111–1112.
31. Borthwick, A. C., Holms, W. H. & Nimmo, H. G. (1984). The phosphorylation of *Escherichia coli* isocitrate dehydrogenase in intact cells. *Biochem. J.* **222**, 797–804.
32. LaPorte, D. C. & Koshland, D. E., Jr (1982). A protein with kinase and phosphatase activities involved in regulation of tricarboxylic acid cycle. *Nature*, **300**, 458–460.
33. LaPorte, D. C. (1993). The isocitrate dehydrogenase phosphorylation cycle: regulation and enzymology. *J. Cell. Biochem.* **51**, 14–18.
34. Xu, X., Zhao, J., Xu, Z., Peng, B., Huang, Q., Arnold, E. & Ding, J. (2004). Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *J. Biol. Chem.* **279**, 33946–33957.
35. Jeong, J. J., Sonoda, T., Fushinobu, S., Shoun, H. & Wakagi, T. (2004). Crystal structure of isocitrate dehydrogenase from *Aeropyrum pernix*. *Proteins: Struct. Funct. Genet.* **55**, 1087–1089.
36. Kawarabayasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Haikawa, Y., Jin-no, K. *et al*. (1999). Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1. *DNA Res.* **6**, 83–101 pp. 145–152.
37. Singh, S. K., Miller, S. P., Dean, A., Banaszak, L. J. & LaPorte, D. C. (2002). *Bacillus subtilis* isocitrate dehydrogenase. A substrate analogue for *Escherichia coli* isocitrate dehydrogenase kinase/phosphatase. *J. Biol. Chem.* **277**, 7567–7573.
38. Doyle, S. A., Beernink, P. T. & Koshland, D. E., Jr (2001). Structural basis for a change in substrate specificity: crystal structure of S113E isocitrate dehydrogenase in a complex with isopropylmalate, Mg^{2+}, and NADP. *Biochemistry*, **40**, 4234–4241.
39. Finer-Moore, J., Tsutakawa, S. E., Cherbavaz, D. R., LaPorte, D. C., Koshland, D. E. & Stroud, R. M. (1997). Access to phosphorylation in isocitrate dehydrogenase may occur by domain shifting. *Biochemistry*, **36**, 13890–13896.
40. Stoddard, B. L. & Koshland, D. E. (1993). Structure of isocitrate dehydrogenase with alpha-ketoglutarate at 2.7-Å resolution: conformational changes induced by decarboxylation of isocitrate. *Biochemistry*, **32**, 9317–9322.
41. Soundar, S., Danek, B. L. & Colman, R. F. (2000). Identification by mutagenesis of arginines in the substrate binding site of the porcine NADP-dependent isocitrate dehydrogenase. *J. Biol. Chem.* **275**, 5606–5612.
42. Kim, T. K., Lee, P. & Colman, R. F. (2003). Critical role of Lys212 and Tyr140 in porcine NADP-dependent isocitrate dehydrogenase. *J. Biol. Chem.* **278**, 49323–49331.
43. Grodsky, N. B., Soundar, S. & Colman, R. F. (2000). Evaluation by site-directed mutagenesis of aspartic acid residues in the metal site of pig heart NADP-dependent isocitrate dehydrogenase. *Biochemistry*, **39**, 2193–2200.
44. Huang, Y. C., Grodsky, N. B., Kim, T. K. & Colman, R. F. (2004). Ligands of the Mn^{2+} bound to porcine mitochondrial NADP-dependent isocitrate dehydrogenase, as assessed by mutagenesis. *Biochemistry*, **43**, 2821–2828.
45. Lee, M. E., Dyer, D. H., Klein, O. D., Bolduc, J. M., Stoddard, B. L. & Koshland, D. E., Jr (1995). Mutational analysis of the catalytic residues lysine 230 and tyrosine 160 in the NADP(+)-dependent isocitrate dehydrogenase from *Escherichia coli*. *Biochemistry*, **34**, 378–384.
46. Koumanov, A., Benach, J., Atrian, S., Gonzalez-Duarte, R., Karshikoff, A. & Ladenstein, R. (2003). The catalytic mechanism of Drosophila alcohol dehydrogenase: evidence for a proton relay modulated by the coupled ionization of the active site lysine/tyrosine pair and a NAD+ribose OH switch. *Proteins*, **51**, 289–298.
47. Dean, A. M. & Koshland, D. E., Jr (1990). Electrostatic and steric contributions to regulation at the active site of isocitrate dehydrogenase. *Science*, **249**, 1044–1046.
48. Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T. *et al*. (1998). Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* **255**, 336–346.
49. Auerbach, G., Ostendorp, R., Prade, L., Korndorfer, I., Dams, T., Huber, R. & Jaenicke, R. (1998). Lactate dehydrogenase from the hyperthermophilic

bacterium *Thermotoga maritima*: the crystal structure at 2.1 Å resolution reveals strategies for intrinsic protein stabilization. *Structure*, **6**, 769–781.
50. Vieille, C. & Zeikus, G. J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev. (Washington, DC)*, **65**, 1–43.
51. Russell, R. J., Ferguson, J. M., Hough, D. W., Danson, M. J. & Taylor, G. L. (1997). The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry*, **36**, 9983–9994.
52. Elcock, A. H. (1998). The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* **284**, 489–502.
53. Aoshima, M., Yamagishi, A. & Oshima, T. (1996). Eubacteria-type isocitrate dehydrogenase from an archaeon: cloning, sequencing, and expression of a gene encoding isocitrate dehydrogenase from a hyperthermophilic archaebacterium, *Caldococcus noboribetus*. *Arch. Biochem. Biophys.* **336**, 77–85.
54. Ren, B., Tibbelin, G., de Pascale, D., Rossi, M., Bartolucci, S. & Ladenstein, R. (1998). A protein disulfide oxidoreductase from the archaeon *Pyrococcus furiosus* contains two thioredoxin fold units. *Nature Struct. Biol.* **5**, 602–611.
55. Bonisch, H., Schmidt, C. L., Schafer, G. & Ladenstein, R. (2002). The structure of the soluble domain of an archaeal Rieske iron–sulfur protein at 1.1 Å resolution. *J. Mol. Biol.* **319**, 791–805.
56. Guy, J. E., Isupov, M. N. & Littlechild, J. A. (2003). The structure of an alcohol dehydrogenase from the hyperthermophilic archaeon *Aeropyrum pernix*. *J. Mol. Biol.* **331**, 1041–1051.
57. Isupov, M. N., Fleming, T. M., Dalby, A. R., Crowhurst, G. S., Bourne, P. C. & Littlechild, J. A. (1999). Crystal structure of the glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J. Mol. Biol.* **291**, 651–660.
58. Singleton, M. R., Hakansson, K., Timson, D. J. & Wigley, D. B. (1999). Structure of the adenylation domain of an NAD^+-dependent DNA ligase. *Struct. Fold. Des.* **7**, 35–42.
59. Mallick, P., Boutz, D. R., Eisenberg, D. & Yeates, T. O. (2002). Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds. *Proc. Natl Acad. Sci. USA*, **99**, 9679–9684.
60. Karlstrom, M., Steen, I. H., Tibbelin, G., Lien, T., Birkeland, N. K. & Ladenstein, R. (2002). Crystallization and preliminary X-ray structure analysis of isocitrate dehydrogenase from two hyperthermophiles, *Aeropyrum pernix* and *Thermotoga maritima*. *Acta Crystallog. sect. D*, **58**, 2162–2164.
61. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
62. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
63. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al*. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
64. Kleywegt, G. J. & Jones, T. A. (1994). Halloween…* masks and bones. In *From First Map to Final Model* (Bailey, S., ed.), pp. 59–66, SERC, Daresbury Laboratory, Warrington.
65. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
66. Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**, 1049–1067.
67. Russell, R. B. & Barton, G. J. (1992). Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins: Struct. Funct. Genet.* **14**, 309–323.
68. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
69. Barlow, D. J. & Thornton, J. M. (1983). Ion-pairs in proteins. *J. Mol. Biol.* **168**, 867–885.
70. McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
71. Baker, E. N. & Hubbard, R. E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **44**, 97–179.
72. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
73. Bacon, M. (1997). Raster3D: photorealistic molecular graphics. *Methods. Enzymol.* **271**, 505–524.
74. DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA.
75. Barton, G. J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40.

Edited by R. Huber

(Received 9 July 2004; received in revised form 6 October 2004; accepted 11 October 2004)