Isocitrate dehydrogenase from bovine heart: primary structure of subunit 3/4

Yan ZENG, Cipora WEISS, Tsai-Tse YAO, Jianmin HUANG, Linda SICONOLFI-BAEZ, Patty HSU and Julie Ivory RUSHBROOK*
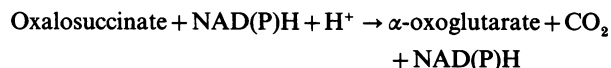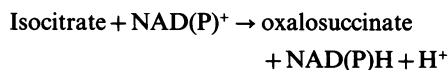Department of Biochemistry, State University of New York Health Science Center at Brooklyn, 450 Clarkson Ave., Brooklyn, NY 11203, U.S.A.

Bovine NAD^+-dependent isocitrate dehydrogenase was shown previously to contain four subunits of approx. 40 kDa (subunits 1–4) possessing different peptide maps and electrophoretic properties [Rushbrook and Harvey (1978) Biochemistry **17**, 5339–5346]. In this study the heterogeneity is confirmed using enzyme purified by updated methods and from single animals, ruling out allelic variability. Subunits 1 and 2 were differentiated from each other and from subunits 3 and 4 by N-terminal amino acid sequencing. Subunits 3 and 4 (subunits 3/4) were identical in sequence over 30 residues. The N-terminal residues of subunits 1 and 2 were homologous but not identical with the β- and γ-subunits respectively of the comparable pig heart enzyme. Subunits 3/4 were identical over 30 residues with the N-terminus of the pig heart α-subunit. Full-length sequence, including that for mitochondrial import, is presented for a protein with the processed N-terminus of subunits 3/4, deduced from cloned cDNA obtained utilizing the N-terminal sequence information.

The derived amino acid sequence for the mature protein contains 339 amino acids and has a molecular mass of 36685 Da. Complete identity with N-terminal and Cys-containing peptides totalling 92 residues from the α-subunit of the pig heart enzyme [Huang and Colman (1990) Biochemistry **29**, 8266–8273] suggests that maintenance of a particular three-dimensional structure in this subunit is crucial to the function of the enzyme. An electrophoretic heterogeneity within the pig heart α-subunit, similar to that shown by bovine subunits 3/4, was demonstrated. One reordering of the Cys-containing peptides of the pig heart α-subunit is indicated. Sequence comparison with the distantly related $NADP^+$-dependent enzyme from Escherichia coli, for which the three-dimensional structure is known [Stoddard, Dean and Koshland (1993) Biochemistry **32**, 9310–9316] shows strong conservation of residues binding isocitrate, Mg^{2+} and the NAD^+ moiety of $NADP^+$, consistent with a catalytic function.

INTRODUCTION

Isocitrate dehydrogenase [IDH; $(2R,3S)$-isocitrate:$NAD(P)^+$ oxidoreductase (decarboxylating); EC 1.1.1.41(42)] catalyses the reaction

$$\text{Isocitrate} + NAD(P)^+ \rightarrow \text{oxalosuccinate}$$
$$+ NAD(P)H + H^+$$
$$\text{Oxalosuccinate} + NAD(P)H + H^+ \rightarrow \alpha\text{-oxoglutarate} + CO_2$$
$$+ NAD(P)H$$

with the participation of a bivalent metal ion, usually Mg^{2+}. Eukaryotes possess both NAD^+- and $NADP^+$-dependent IDHs. NAD^+-IDH, encoded by the nuclear genome, is located exclusively in the mitochondrial matrix and is a key regulatory enzyme in the Krebs cycle. Separate eukaryotic $NADP^+$-IDHs, for which the biological functions are not well defined, are found in the cytosol and mitochondria. Prokaryotes, such as Escherichia coli, have a single IDH with $NADP^+$ as a cofactor.

NAD^+-IDHs have been purified and studied extensively from bovine and pig hearts and from yeast. The enzymes function as tetramers or multiples of tetramers with subunits of about 40 kDa [1–6]. Activity is regulated by multiple allosteric effects. The enzyme from each source exhibits positive co-operativity for isocitrate [6–8]. Negative co-operativity has been reported for NAD^+ in the enzyme from bovine heart [9]. In the mammalian enzymes, ADP acts as a positive heterotropic effector whereas ATP, NADH and NADPH have negative heterotropic effects [4,10]. The enzyme from yeast is simpler, with AMP acting as a positive heterotropic effector [6].

Despite intensive study over many years illuminating many aspects of the enzyme's action, the structural information required to understand the catalytic mechanism and complex allosteric effects has been slow to emerge. Heterogeneity among the subunits is present consonant with the degree of allosterism. The enzyme from bovine heart was reported to contain four non-identical subunits of approx. 40 kDa (subunits 1–4 in order of increasing mobility on SDS/PAGE), apparently present in equimolar amounts, possessing distinctive peptide maps [11]. Concurrently, the enzyme from pig heart was found to possess first two subunits, by an SDS/PAGE system different from that used for the bovine heart [12], and then, by isoelectric focusing, three subunits, α, β and γ, present in the ratio 2:1:1 [13]. Sequencing of the N-terminal region and of Cys-containing peptides of the three subunits of the pig heart enzyme established their uniqueness [14]. The discrepancy in the number of different subunits in the enzymes from bovine and pig hearts remained unresolved.

Isolated α-, β- and γ-subunits from pig heart contained low catalytic activity, with the α-subunit being the most active [15]. Subunits β and γ were inactive when combined, but each conferred enhanced activity when added to α. The highest activity was found only in native enzyme containing all three species. Formation of dimers using chemically modified subunits indicated that the isocitrate-binding sites of the α-subunits but not of β or γ were crucial for activity [16]. These data suggest a catalytic function for the α-subunit. The partial sequence obtained for the pig heart γ-subunit N-terminal and Cys-containing peptides [14] provided information leading recently to the sequencing of full-

length γ-subunit clones from rat epididymis and monkey testis tissue and to additional sequence from a pig heart clone [17].

The simpler enzyme from yeast was shown early on to contain two non-identical subunits, IDH1 and IDH2, present in equimolar ratio [5]. Genomic DNA encoding each of these has recently been cloned and sequenced, facilitated by the absence of introns [18,19]. The homology between the two subunits was 42 %. Site-directed mutagenesis suggested that IDH2 functioned catalytically, whereas IDH1 had a regulatory role [20]. The subunits had 40–45 % homology with the mammalian γ-subunits [17].

Recent structural work on the distantly related $NADP^+$-dependent enzyme from *E. coli* [21–23], a homodimer regulated solely by phosphorylation which prevents isocitrate binding [24], has provided information that serves as a valuable reference point for investigating the catalytic action of the eukaryotic NAD^+-dependent enzymes. The *E. coli* enzyme lacks homology with eukaryotic $NADP^+$-dependent enzymes but has low (32 %) but significant homology with sequenced subunits of eukaryotic NAD^+-dependent enzymes (yeast IDH1, IDH2 and mammalian γ). High-resolution three-dimensional structures obtained by X-ray crystallography revealed that its topology and substrate- and cofactor-binding structures were similar to those of isopropyl-malate dehydrogenase, another decarboxylating dehydrogenase [25], but different from those of other known dehydrogenases, suggesting a distinct evolutionary pathway for decarboxylating dehydrogenases [26]. Key residues binding isocitrate, Mg^{2+} and NAD^+ were identified. Those binding isocitrate are conserved in both of the yeast subunits and the mammalian γ-subunit. Residues involved in binding NAD^+ are additionally conserved in the yeast IDH2 subunit, supporting studies indicating its importance in catalysis [20].

Although studies of the enzyme from lower organisms can provide valuable information on the mechanism of catalysis and, eventually, in the case of yeast, an understanding of co-operativity for isocitrate and the allosteric effect exhibited by AMP, the more complex allosteric interactions involving the multiple effectors regulating the mammalian enzymes can only be elucidated by direct study of these enzymes.

In this report we confirm the presence of the four non-identical subunits (subunits 1–4) previously described for the enzyme from bovine heart [11] and characterize them by N-terminal sequencing. Subunits 1 and 2 are the counterparts of pig heart subunits β and γ respectively. Subunits 3 and 4 are identical over their first 30 residues and homologous with the pig heart α-subunit. The pig heart α-subunit is shown to possess heterogeneity on electrophoresis, yielding two species similar to bovine heart subunits 3 and 4. The full-length sequence, including that for mitochondrial import, is presented for a protein containing the processed N-terminus of bovine heart subunits 3 and 4, deduced from cDNA obtained and cloned by utilizing N-terminal sequence information. Homology comparisons with partial sequence reported for the pig heart α-subunit, the γ-subunits of rat epididymis and monkey testis, the subunits of the NAD^+-dependent enzyme from yeast, and with the $NADP^+$-dependent enzyme from *E. coli*, suggest a catalytic function for this subunit.

MATERIALS AND METHODS

Materials

Bovine hearts, purchased from Max Insel Cohen, Newark, NJ, U.S.A., were transported packed in ice for protein purification and used immediately. For mRNA purification, the ventricular tissue from a single heart was cut into small pieces immediately on removal from the animal. The pieces were quickly frozen in liquid nitrogen for transportation and stored at -70 °C. Chemicals for protein purification and activity determination were reagent grade or better. Reagents other than pH buffers for acrylamide gel analysis of proteins and nucleic acids were obtained from Bio-Rad Laboratories, Hercules, CA, U.S.A. HPLC-grade water was used for HPLC. T4 DNA ligase and restriction enzymes were obtained from Promega (Madison, WI, U.S.A.) except for *Eco*RI which was from New England Biolabs (Beverly, MA, U.S.A.). *Taq* polymerase was from Perkin–Elmer (Norwalk, CT, U.S.A.) or Promega. *E. coli* strain JM109 was from Promega, and DH5α from Clontech (Palo Alto, CA, U.S.A.). Agarose was from FMC (Rockland, ME, U.S.A.). Other chemicals for procedures involving nucleic acids were reagent grade or better. Sources for chromatography columns and kits for nucleic acid procedures are given with the appropriate procedure.

Protein assays

Protein and enzyme activity were determined as described by Plaut [27]. One unit of enzyme activity is defined as the amount of enzyme that catalyses the formation of 1 μmol of NADH/min at 24 °C.

Bovine heart NAD^+-IDH purification

The enzyme was prepared to the stage of $(NH_4)_2SO_4$ fractionation of an extract of lyophilized mitochondria as described previously [11]. Chromatography on Sepharose 6B was replaced by fractionation on a preparative hydrophobic interaction column (5PW; Bio-Rad), equilibrated with buffer A [1.7 M $(NH_4)_2SO_4$, 0.1 M sodium phosphate, pH 7, 0.1 mM ADP, and 0.1 mM dithioerythritol (DTE)]. The enzyme was applied in a solution containing 0.6 M $(NH_4)_2SO_4$, 0.1 M sodium phosphate, pH 7, 0.1 mM ADP, and 0.1 mM DTE. Elution was accomplished with a gradient from 60 to 100 % buffer B in 15 min, where buffer B was identical to buffer A except that it lacked $(NH_4)_2SO_4$. The active fractions were concentrated by $(NH_4)_2SO_4$ precipitation and dialysed against buffer C (50 mM imidazole hydrochloride, pH 7, 0.1 mM DTE and 0.1 mM ADP). The sample was applied to an anion-exchange column (Mono Q; Pharmacia, Piscataway, NJ, U.S.A.) equilibrated with buffer C lacking ADP. Elution was carried out with a gradient from 0 to 35 % buffer D (identical to buffer C with the addition of 1 M KCl) in 20 min. Active fractions were combined and dialysed against buffer A. The final step was rechromatography on an analytical hydrophobic interaction column (5PW) using buffers A and B and a gradient of 10–100 % buffer B in 20 min. Enzyme purified in this way was used for most of the studies described. Despite the high purity of the final product, as indicated by electrophoretic analysis, final specific activities of 16 units/mg rather than the 20–40 obtained by earlier procedures were obtained, because of loss of activity on the otherwise effective Mono Q column. In one preparation, the preparative hydrophobic interaction column was replaced by conventional chromatography on Superose 6 (preparation grade; Pharmacia) equilibrated with 10 mM sodium phosphate, pH 7.2, containing 5 % glycerol, 0.1 mM DTE and 10 %-satd. $(NH_4)_2SO_4$. This was followed by the preparative hydrophobic interaction column and the Mono Q column. Sequencing studies employing electroblotting utilized the active fraction from the final step of either procedure.

Pig heart NAD⁺-IDH purification

Pig hearts were cut into small pieces, ground and homogenized in 0.01 M Tris/citrate buffer, pH 7.2, [28] to which was added 10 % glycerol, 2 mM $MnSO_4$ and 0.1 mM dithiothreitol [13]. The homogenate was centrifuged at 23300 g [29] and the supernatant filtered through cheesecloth. Proteins precipitating between 22 and 34 %-satd. $(NH_4)_2SO_4$ [29] were subjected to gel filtration on Superose 6 equilibrated as described by Shen et al. [2]. Active fractions were applied to the preparative hydrophobic interaction column (5PW), as described for the bovine heart enzyme. This was followed by chromatography on the Mono Q anion-exchange column using the conditions developed for the bovine heart enzyme. The most active fraction had a specific activity of 16 units/mg.

Two-dimensional electrophoretic analysis of purified bovine heart enzyme

Electrophoresis in the first dimension under non-denaturing conditions and in the second in the presence of SDS was carried out as described previously [11].

N-Terminal sequencing of enzyme subunits

The bovine and porcine subunits, separated by SDS/PAGE [11], were electroblotted to poly(vinylidene difluoride) (PVDF) paper (Immobilon from Millipore, Milford, MA, U.S.A. or Problott from Applied Biosystems, Foster City, CA, U.S.A.) according to the procedure of Matsudaira [30]. Modifications employed in studying the pig heart subunits included aging the gel by preparing the separating gel the night before electrophoresis, using the free-radical scavenger mercaptoacetic acid in the running buffer and omitting acetic acid from blot stain and destain solutions. Edman degradations and phenylthiohydantion amino acid analyses were carried out using standard cycles on a model 470A gas-phase protein sequencer connected on-line to a microbore HPLC phenylthiohydantion-amino acid analyser (model 120), both from Applied Biosystems. Sequences were evaluated visually and quantification was carried out with the aid of a model 900 data-processing module from Applied Biosystems.

mRNA isolation

mRNA for procedures leading to 3'-RACE (a system for rapid amplification of cDNA ends) was obtained by the method of Auffray and Rougeon [31], modified for muscle (S. Benoff, personal communication) and followed by elution from an oligo(dT)-cellulose column [32]. Frozen bovine heart tissue bathed in liquid nitrogen was pulverized using a mortar and pestle to yield a powder which was homogenized in a Dounce homogenizer in 2 M LiCl/4 M urea (8 ml/g tissue) at 37 °C. The homogenate was allowed to stand at 4 °C for 48 h or longer. A pellet was collected by centrifugation at 12620 g for 30 min and resuspended in 100 mM Tris/HCl, pH 8, containing 150 mM NaCl, 2 mM EDTA and 0.5 % SDS. The solution was extracted with 1 vol. of phenol/chloroform/3-methylbutan-1-ol (25:24:1, by vol.) three times, and finally with 1 vol. of chloroform/3-methylbutan-1-ol (24:1, v/v). Then 0.1 vol. of 2 M sodium acetate, pH 5.5, and 2 vol. of 100 % ethanol were added and the solution was left overnight at −20 °C or several hours on solid CO_2 to precipitate RNA. The pellet was collected (12062 g for 30 min) and dissolved in 10 mM Tris/HCl, pH 7.5, or water. Polyadenylated [poly(A)⁺] RNA was obtained by chromatography on oligo(dT)-cellulose (Gibco/BRL, Grand Island, NY, U.S.A.).

PCR with degenerate primers and cloning

All PCRs utilized the GeneAmp PCR System 9600 (Perkin–Elmer). Poly(A)⁺ RNA was first reverse-transcribed using Perkin–Elmer's RNA PCR kit with random hexamers as the primer. The concentrations of Mg^{2+} and dNTP were 5 mM and 1 mM respectively. Reaction conditions were 42 °C for 15 min followed by 99 °C for 5 min. The degenerate primers for PCR, 905S and 903A, with BamHI and EcoRI sites respectively added at their 5' ends, were based on the N-terminal amino acid sequence of the bovine heart NAD⁺-IDH 3/4 subunit (see under 'Primers' and 'Sequencing of subunits 3/4 using PCR technology' in the Results section). Mg^{2+} (5 mM final concentration), buffer, water, 2.5 units of Taq polymerase and 1 μg of each primer were added to the reverse-transcription mixture to a final volume of 100 μl. PCR conditions were 95 °C for 1.5 min, followed by 45 cycles of 95 °C for 15 s, 37 °C for 1 min and 63 °C for 2 min. The reaction was evaluated at 30 cycles by running one-tenth of the PCR mixture on a 10 % polyacrylamide gel. The PCR product was extracted, precipitated, and digested with EcoRI and BamHI overnight at 37 °C. The band obtained at about 90 bp on a 10 % polyacrylamide gel was excised and the DNA electroeluted. The DNA was then ligated to EcoRI–BamHI-cut pGEM7Zf(+) (Promega) and transformed into competent E. coli cells. Selected colonies were grown and the plasmid DNA was isolated and cleaved with EcoRI and BamHI. Ten clones with inserts were subjected to sequencing. Despite the fact that several nucleotide positions within the primer regions differed, the internal amplified sequences of all ten tested clones were the same, and the peripheral variability could be attributed to primer degeneracies and a low annealing temperature during PCR.

3'-RACE

First-strand cDNA was synthesized from poly(A)⁺ RNA using the 3'-RACE system for rapid amplification of cDNA ends [33] from Gibco/BRL as described in the Gibco protocol except that 0.5 μl (10 units) of RNase inhibitor from the Perkin–Elmer RNA PCR kit was added before reverse transcription. A small amount (1/20 to 1/600) of the cDNA was amplified by PCR, using the universal amplification primer (UAP; Gibco/BRL 3'-RACE kit) and the 5'-gene-specific primer Z1 (see under 'Primers') and employing the hot start PCR procedure (Perkin–Elmer). Buffer, dNTPs (0.2 mM each), Mg^{2+} (2.5 mM), primers (200 nM each) and wax were heated for 5 min at 80 °C and cooled on ice. Taq polymerase (2.5 units) and cDNA were then added and the tubes heated at 95 °C for 1.5 min. Thirty cycles of 95 °C for 15 s, 55 °C (or 60 °C) for 30 s, and 72 °C for 3.5 min took place, followed by a final extension of 14 min at 72 °C. Products were evaluated by agarose-gel electrophoresis. The PCR product was extracted, precipitated and cut with MluI (UAP has an MluI site) and EcoRI (Z1 has an EcoRI site) in MluI buffer overnight at 37 °C. DNA greater in size than 1400 bp was isolated by electroelution after PAGE, ligated to pGEM7Zf(+) and introduced into E. coli cells. Colonies were tested for the presence of inserts by miniprep plasmid DNA preparations and PAGE analysis of restriction digests. Four independent clones were sequenced, one completely on both strands, the others partially.

5'-RACE

Poly(A)⁺ RNA for 5'-RACE was obtained using the PolyATract System 1000 (Promega) according to the manufacturer's instruc-

tions. Briefly, after homogenization of bovine heart tissue, poly(A)$^+$ RNA was hybridized to a biotinylated oligo(dT) probe and captured by Streptavidin MagneSphere Paramagnetic Particles. The particles were washed with 0.0075 M sodium citrate/0.075 M NaCl (0.5 × SSC). Poly(A)$^+$ RNA was eluted with water, precipitated with ethanol and dissolved in a small amount of water.

cDNA for PCR was prepared using the 5′-AmpliFINDER RACE kit (Clontech) [34]. Poly(A)$^+$ RNA was primed with random hexamers, and cDNA was synthesized. The RNA was then hydrolysed with NaOH. After neutralization, the cDNA was separated from the primers using GENO-BIND (Clontech) and precipitated. Ligation to the AmpliFINDER anchor employed T4 RNA ligase in a 20 h incubation at room temperature. 1% of the cDNA was amplified by the hot start PCR procedure using anchor primer and Z12, an IDH subunit 3/4-specific primer (see under 'Primers'). The conditions were 95 °C for 20 s, 35 cycles of 94 °C for 15 s, 55 °C for 15 s, and 72 °C for 75 s, and finally 7 min at 72 °C. Electrophoresis revealed a major band larger than 600 bp in the PCR mixture. The product was cloned, using EcoRI and BamHI sites incorporated in the primers, into pGEM7Zf(+), as described for 3′-RACE. Six positive clones were analysed. One was sequenced in both directions from the 5′ end of the insert past the end of the leader sequence using the M13 forward primer and CIP3/4SA199 (see under 'Primers'). The rest were sequenced in the forward direction.

Sequencing of cloned products

The Magic and Wizard miniprep kits from Promega were used to generate plasmid DNA for sequencing according to the manufacturer's technical bulletin. DNA was precipitated by ethanol and resuspended in 8–20 μl of water. Sequencing was performed based on the enzymic method [35] using the silver sequencing system from Promega. Each procedure used 4 μl (about 2–4 μg) of DNA and 4.5 pmol of primer. Initial denaturation occurred at 95 °C for 2 min. For M13 forward primer (24-mer) sequencing, the PCR sequencing conditions were 60 cycles of 95 °C for 10 s, and 70 °C for 25 s; for M13 reverse primer (17-mer) and other synthetic primers the conditions were, 95 °C for 10 s, 42 °C for 10 s and 70 °C for 20 s. Sequencing plates were prepared as indicated in the Promega manual. Products were separated on a 6% polyacrylamide/8 M urea gel (National Diagnostics, Atlanta, GA, U.S.A.).

To sequence the internal region of the 1600 bp DNA, three strategies were employed. (a) A 1100 bp XbaI/NsiI-excised fragment of clone 102 which included the 5′-portion of the insert was subcloned into pGEM7Zf(+). (b) The ExoIII method (Erase-a-Base system of Promega) was used. Plasmid DNA was released from overnight bacterial culture using 0.2 M NaOH and 1% SDS, extracted with phenol/chloroform/3-methylbutan-1-ol (25:24:1, by vol.) twice and chloroform/3-methylbutan-1-ol (24:1, v/v) once, and dissolved in water after precipitation. Nicked DNA was removed by acid/phenol extraction [36]. The remaining DNA was cut with BbuI and XbaI and purified by phenol/chloroform extraction. 5′-Deletion was then carried out according to the manufacturer's instructions and the products were religated into pGEM7Zf(+) and cloned. The size of the insert was determined by MluI and ApaI digestion of plasmid DNA isolated from bacterial culture followed by PAGE. (c) Synthetic primers based on newly obtained sequence were ordered from Ransom Hill Bioscience (Ramona, CA, U.S.A.) with trityl-specific purification and used for sequencing.

Sequence analysis

Sequences were analysed using programs in PC/Gene (IntelliGenetics, Mountain View, CA, U.S.A.).

Primers

905S (5′-CA<u>GGATCC</u> GCT GGN GGN GTN AA(A/G) AC-3′) codes for AGGVKT, the N-terminal six amino acids of the 30-amino acid sequence determined for subunits 3/4, with all possible codons for GGVK and the most frequently used codon for alanine (see 'Sequencing of subunits 3/4 using PCR technology' in the Results section). Underlined is the BamHI site. 903A (5′-TC<u>GAATTC</u> GC (A/G)TC (A/G)AA (A/T/G)AT (T/C)TT CAT-3′) codes for MKIFDA, residues 24–29 of the 30-amino acid sequence determined for subunits 3/4 (see Results section), with all possible codons (antisense). Underlined is the EcoRI site. These primers were synthesized by the DNA Synthesis and Purification Laboratory of the Department of Microbiology and Immunology of SUNY Health Science Center at Brooklyn. Adaptor primer [AP, 5′-GGCCACGCGTCGACTAGTAC-(T)$_{17}$-3′] and UAP (5′-CUACUACUACUAGGCCACGCGT-CGACTAGTAC-3′) are included in the Gibco 3′-RACE kit. AP was used as the reverse-transcription primer; UAP was used in the PCR step of 3′-RACE.

Z1 (5′-TC<u>GAATTC</u>CCAGAAATTTCTGCTGCAG-3′) (see Figure 8) was used in 3′-RACE PCR as the 5′-sequence specific primer. Underlined is the EcoRI site. Z2 (5′-CATCAGCC-TGTGTTAGA-3′) (sense, bases 1191–1207; see Figure 10), Z5 (5′-ACGATCACATGCTCGAT-3′) (antisense; bases 509–525; see Figure 10), Z6 (5′-GCCAAATTTGTATGGAC-3′) (sense; bases 796–812; see Figure 10), Z7 (5′-GTTGGCTCCAATG-TTGC-3′) (antisense; bases 870–886; see Figure 10), Z8 (5′-GTAAATATCATGATACAGA-3′) (sense; bases 1473–1491; see Figure 10) and Z9 (5′-CGTAAACCAAGAATATC-3′) (antisense; bases 1533–1549; see Figure 10) were based on newly obtained sequence. Z1–Z9 were synthesized and purified (trityl-specific) by Ransom Hill Bioscience. Z2, Z6 and Z8 were used to sequence the insert of clone 102 in the sense direction; Z5, Z7 and Z9 were used to sequence it in the antisense direction.

For 5′-RACE, Clontech provided the AmpliFINDER anchor, 5′-CAC<u>GAATTC</u>ACTATCGATTCTGGAACCTTCAGAG-G-NH$_2$ (the EcoRI site is underlined), and anchor primer, 5′-CTGGTTCGGCCCACCTCTGAAGGTTCCAGAATCGA-TAG-3′.

Z12 (5′-CA<u>GGATCC</u>GGTGGTTGTTCCGTGC-3′) (antisense; bases 602–617; see Figure 10: underlined is the BamHI site) was ordered from Ransom Hill Bioscience without trityl-specific purification and used as the sequence-specific primer in 5′-RACE.

CIP3/4SA199 (5′-GGCTTCTGGAGGGATCATCCACTT-3′) (antisense; bases 272–295; see Figure 10), a trityl-specific sequence from Ransom Hill Bioscience, was used in reverse sequencing of 5′-RACE products.

RESULTS

Subunit heterogeneity

Four non-identical subunits were previously identified in bovine heart NAD$^+$-IDH [11], a result that has remained controversial because only three species were found in the related enzyme from pig heart [13]. The finding for bovine heart is further documented here with enzyme purified by updated methods from a single animal. As the purity of the preparation increased, the four subunit bands previously identified as components of the active
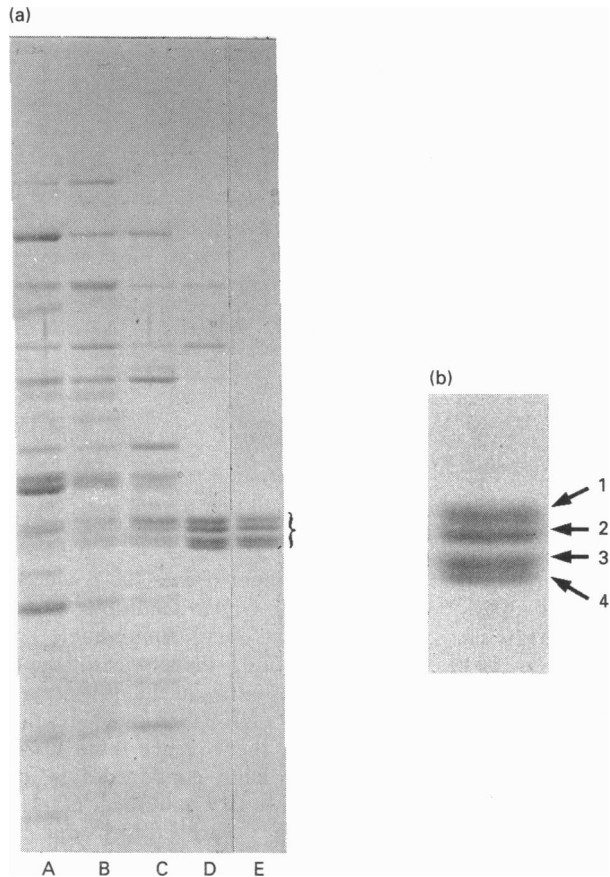
Figure 1   SDS/PAGE of fractions obtained during purification of bovine heart NAD⁺-IDH

(a) Lane A, mitochondrial extract; lane B, 30–60%-satd. (NH$_4$)$_2$SO$_4$ cut; lane C, active fractions from the preparative 5PW hydrophobic interaction column; lane D, active fractions from the Mono Q column; lane E, active fractions from the analytical 5PW column; 4–10 μg of protein was applied. The bracket shows the location of the four subunits of the enzyme. (b) Enlargement of the subunit region of lane E in (a), with the subunits numbered.

Figure 2   One- and two-dimensional electrophoresis of NAD⁺-IDH purified from a single bovine heart by updated procedures

(a) Non-denaturing gels stained for protein (lane A) and activity (lane B) (6 μg of protein). x locates the active band; y and z denote two bands migrating faster than the active band. (b) A gel such as that in (a) containing 11 μg of protein was electrophoresed into an SDS/polyacrylamide gel. The region indicated by (B) contains the two-dimensional pattern obtained. A stained one-dimensional gel is shown at the top for reference, with the active band, x, and bands y and z indicated. The region indicated by (A) contains enzyme applied directly to the SDS/polyacrylamde gel (2 μg of protein).

	1	10	20
Subunit 1:	A S S R T Q G E D V R V E G A F P V T M L P		
Subunit 2:	F S Q Q T I P P S A K Y G A R H T V T M I P		
Subunits 3 and 4:	A G G V K T V T L I P G D G I G P E I S A A		
	23	30	
Subunits 3 and 4:	V M K I F D A A		

Figure 3   N-Terminal sequences of the four subunits of bovine NAD⁺-IDH

enzyme emerged on SDS/PAGE (Figure 1). The subunits are numbered 1–4 in order of increasing electrophoretic mobility.

On electrophoresis under native conditions (not isoelectric focusing), the purified enzyme stained for protein shows a strong major band, band x (Figure 2a, gel A), which is followed immediately by a diffusely staining region and then two distinct bands, y and z. Band x is the only band to exhibit enzyme activity (Figure 2a, gel B). When a similar unstained native gel is subjected to second-dimension electrophoresis in the presence of SDS, the pattern shown in region (B) of Figure 2(b) is obtained. The active band in the first dimension, band x, is shown to contain each of the four species, 1, 2, 3 and 4, which are the major bands when the enzyme is analysed directly by SDS/PAGE (as in Figure 2b, region A). The diffusely staining material migrating just ahead of band x in the first dimension gives rise to a spot co-migrating with subunit 1, and bands y and z in the first dimension contain proteins with the mobilities of subunits 4 and 3 respectively. 'Trails' extend from subunit components 3 and 4 of the active band x to the spots derived from first-dimension bands z and y respectively, and Coomassie Brilliant Blue-staining material similarly trails from active band subunit 2 to a position correlating with the top of the first-dimension gel. Active band

subunit 1 is too close to its partner, derived from the diffusely stained material in the first dimension, to show such a trail.

As proposed previously [11], this pattern is consistent with the enzyme being in a state of partial dissociation at the initiation of first-dimension native PAGE. The dissociation continues slowly during native PAGE. Time-dependent incubation of the active band, cut out after an initial period of electrophoresis, followed by continuing native PAGE, generated the distinctive fast-moving bands y and z, in support of this analysis [11]. Missing from the two-dimensional pattern of Figure 2(b) (region B) is a spot corresponding to dissociated subunit 2 at the commencement of native PAGE. Previous work [11] suggests that, once dissociated with the enzyme, subunit 2 tends to aggregate and not enter the gel.

Whereas the earlier work utilized enzyme prepared from pooled bovine hearts, the above analysis was carried out with enzyme from a single heart. The two other single-heart preparations made also showed the same four-band pattern on SDS/PAGE. The four species present in the active band thus do not originate in allelic differences. This is noted particularly for subunits 3 and 4 (see below). The pattern did not depend on whether the enzyme purification included gel filtration or not. The constancy of the observation, whether or not a cocktail of protease inhibitors

Subunit1	ASSRTQGEDV-----RVEGAFP-VTMLP---------------------	22
Pigβ	AS-RSQAEDV-----RVEGAFPTVTMLPGDGVGPELMAAV----------	34
Subunit2	--FSQQTIPP----SAKYGARHTVTMIP---------------------	22
Pigγ	--FSQQTIPP----SAKYGGILTVTMSPGDGDGPELMLTV----------	34
Subunits3/4	AG----------------GVKTVTLIPGDGIGPEISAAVMKIFDAA---	30
Pigα	AG----------------GVKTVTLIPGDGIGPEISAAVMKIFDAAKAP	33
Yeast2	ATVKQPSIGRYTGKPNPSTGKYTVSFIEGDGIGPEISKSVKKIFSAANVP	50
Yeast1	ATAAQAE----RTLPKKYGGRPTVTLIPGDGVGKEITDSVRTIFEAENIP	46

Figure 4   Comparison of N-terminal sequences from the subunits of NAD⁺-IDH from bovine and pig hearts and from yeast (subunit1, subunit2, subunits3/4 from bovine heart; pigα, pigβ, pigγ from pig heart [14]; yeast1 and yeast2 from yeast [18,19]

The alignment utilized the Clustal program of PC/Gene. An asterisk indicates perfect alignment, and a bold point that a position is well conserved. A consensus sequence is underlined.
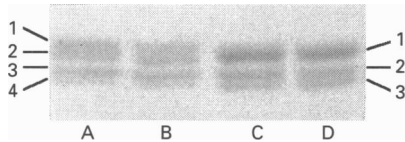
Figure 5   Comparison of NAD⁺-IDH from bovine and pig hearts by electrophoresis in the Laemmli SDS/PAGE system

Lanes A and B, purified enzyme from bovine heart; lanes C and D, purified enzyme from pig heart. Approx. 4 μg/lane was used. The 40 kDa region of the gel is shown enlarged. Both bovine and pig heart bands are numbered in order of increasing mobility.

was present in the mitochondrial extraction solution, strongly suggests that artefactual proteolysis is not involved in the origin of subunits 3 and 4 (see below).

Nature of differences in the four subunits: N-terminal sequencing of subunits of the enzyme

The four subunits were electroblotted from a one-dimensional SDS/polyacrylamide gel on to PVDF paper, the bands excised and subjected to N-terminal sequencing (Figure 3). Extended sequence data were obtained for subunits 1, 3 and 4. Direct sequence data for subunit 2 showed contamination with subunit 1. The sequence shown for subunit 2 was therefore obtained by sequencing subunits 1 and 2 together and subtracting out the previously obtained unambiguous sequence found for subunit 1.

Subunits 1 and 2 clearly differ in sequence from each other and from subunits 3 and 4 which are identical over their first 30 residues. Figure 4 compares these sequences with N-terminal peptide sequences from the three subunits identified for the enzyme from pig heart, α, β and γ [14], and with similar stretches of N-terminal sequence taken from the full sequences of the yeast IDH1 and IDH2 subunits [18,19].

Bovine heart subunit 1 has strong homology with pig heart subunit β, subunit 2 with pig heart subunit γ. The bovine sequences from subunits 3 and 4 are identical over the length obtained with that of pig heart α. It is notable that subunits 3 and 4 are homologous with pig heart subunit α which is present in the intact enzyme in the ratio 2α:1β:1γ. If four non-identical subunits are indeed present in the pig heart enzyme, it is from within α that we would expect the additional species to appear. Although the sequences compared are short, the complete identity, across species lines, of subunits 3/4 and α hints at the particular importance of these subunits to the function of the enzyme. Comparison of the mammalian sequences with those of yeast shows considerable homology. Of note is a consensus

Figure 6   N-terminal sequencing of the three bands shown in Figure 5 for pig heart NAD⁺-IDH

Band 1:	A S R S - A E D V R V E G	(pig heart β)
	F S Q Q T I P P S A K Y G	(pig heart γ)
Band 2:	A G G V K T V T	(pig heart α)
Band 3:	A G G V K T V T L I P G D	(pig heart α)

The top band contained a mixture of two sequences which could be identified by comparison with published work as those of β and γ [14]. Each of the bottom bands had the N-terminal sequence published for the α-subunit. -, No residue identifiable.

sequence, P/E G D G V/I/D G P/K E I/L, common to yeast and mammalian subunits.

Electrophoretic heterogeneity in the pig heart α-subunit

Purified pig heart enzyme was compared with the bovine heart enzyme by SDS/PAGE (Figure 5). Under conditions where the bovine heart enzyme migrated as four distinguishable bands, that of the pig heart gave three bands, the slowest of the three staining about twice as strongly as either of the other two. The pig heart bands were blotted to PVDF paper and N-terminally sequenced (Figure 6). Each cycle of the sequence obtained for the stronger-staining slowest-moving band contained two amino acids, which on comparison with the published N-terminal peptides of the pig heart subunits [14] proved to be due to the presence of subunits β and γ. Each of the two lower bands possessed the N-terminal sequence of the pig heart α-subunit. Thus the heterogeneity represented by bovine heart subunits 3 and 4 (apparent after electrophoresis under native conditions, in the presence of urea alone or SDS, in the Laemmli procedure [11]) is also revealed in the homologous pig heart subunit on examination in the Laemmli SDS/PAGE procedure.

Sequencing of subunits 3/4 using PCR technology

Subunits 3 and 4 were selected to be studied first, as they have the added interest of the origin of their structural differences. A non-redundant 5'-sense primer, to be used subsequently in PCR to yield a full-length cDNA from poly(A)⁺mRNA, was obtained as follows: reverse translation of the protein sequence obtained for subunits 3 and 4 yields an oligonucleotide with the degeneracies:

```
1              10              20              30

A G G V K T V T L I P G D G I G P E I S A A V M K I F D A A

4 4 4 4 2 4 4 4 4 6 3 4 4 2 4 3 4 4 2 3 6 4 4 4 1 2 3 2 2 4 4
```

Degenerate primers were selected for reverse transcription and amplification of a region of mRNA containing the central portion of this sequence:

```
5'-sense primer:                    3'-antisense primer:

        1         5                  24              29
        A  G  G  V  K  T             M   K   I   F   D   A
5'-CAGGATCC - GCT GGN GGN GTN AAA AC-3'   3'-TAC TTT TAA AAA CTA CG - CTTAAGCT-5'
        G                                      C   T   G   G
                                                   G
```

The 5'- and 3'-primers have BamHI and EcoRI restriction enzymes sites respectively at their 5' ends (underlined).

PCR using these primers, after reverse transcription of bovine
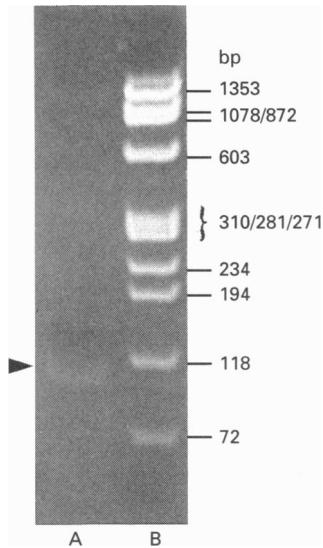
Figure 7   PAGE (10% acrylamide) of products stained with ethidium bromide after reverse transcription-PCR using degenerate primers derived from N-terminal sequencing of subunits 3/4

Lane A, PCR product (one-tenth of reaction mixture); lane B, molecular-mass markers.

```
               A   G   G   V   K   T   V   T   L   I   P   G

5'-GGATCC GCT GGT GGT GTT AAG AC A GTA ACT TTA ATT CCA GGA

3'-CCTAGG CGA CCA CCA CAA TTC TG T CAT TGA AAT TAA GGT CCT


   D   G   I   G   P   E   I   S   A   A   V   M   K   I

GAT GGA ATT GGC CCA GAA ATT TCT GCT GCA GTT ATG AAA ATA/C

CTA CCT CAA CCG GGT CTT GAA AGA CGA CGT CAA TAC TTT TAT/G


   F   D   A

TTC GAT GC GAATTC-3'

AAG CTA CG CTTAAG-5'
```

Figure 8   Sequence of the cloned PCR product encoding the N-terminal region of subunits 3/4

The original primer regions are underlined. The sequence-specific sense primer selected for 3'-RACE is shown in bold. An *Eco*RI site was added to its 5' end.

heart poly(A)$^+$ RNA, yielded a band with the predicted length (86 bp + 16 bp in linkers) when examined by electrophoresis on a 10% polyacrylamide gel (Figure 7). The sequence of the cloned product encoded the N-terminus of subunits 3/4 (Figure 8). A sense primer, Z1, for use as the sequence-specific primer in 3'-RACE, was selected from within the non-degenerate central region (Figure 8, sequence in bold). Bovine heart poly(A)$^+$ RNA was reverse-transcribed using the oligo(dT) region of AP and the resultant cDNA amplified using the UAP of AP as antisense primer and Z1 as sense primer. Electrophoresis of the products revealed several bands, only one of which was greater than 1 kb (Figure 9). Cloning of this band yielded five positive clones. One
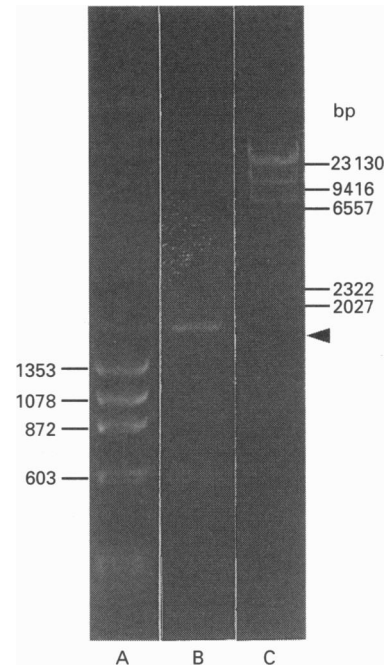
Figure 9   3'-RACE product [lane B (one-fifth of PCR reaction mixture)] run on a 1% agarose gel and stained with ethidium bromide

Lanes A and C, molecular-mass markers.

clone was sequenced completely on both strands, the others partially. The sequence of the 5'-untranslated region and the leader sequence were obtained after cloning of 5'-RACE products as described in the Materials and methods section.

The complete cDNA and deduced amino acid sequences are shown in Figure 10. Of six 5'-RACE products examined, the two longest gave the sequence shown. A typical poly(A)$^+$ signal, AATAAA, is present near the cDNA 3' end. The sequence GTG GAC GCG ATG G, enclosing the initiation codon ATG at bases 44–46, is identical at nine of 13 positions with the translation initiation consensus sequence, GCC GCC G/ACC ATG G [37]. The identity includes the strong preference for G at positions 1, 4, 7 and 13.

The molecular mass of the predicted mature protein (339 amino acids) is 36685 Da. The subunit contains sequences identical with those of the N-terminal and 11 Cys-containing peptides from the pig heart α-subunit totalling 92 residues [14] (Figure 11). These sequences are distributed throughout the length of the bovine heart subunit. The high degree of homology continues that remarked on for the N-terminal sequence. The bovine sequence calls for a rearrangement of the previous ordering of certain of the pig heart Cys-containing peptides. Specifically, pig heart peptides II, III and IV (indicated in Figure 11) immediately precede peptide IX, rather than overlapping by two residues the C-terminus of peptide Va as proposed by Huang and Colman [14]. Peptide Va is located some distance from the C-terminus. This revision increases the number of known Cys residues in the pig heart enzyme from 6 to 7. The bovine heart enzyme contains an additional Cys residue at position 236, which was not detected in the pig heart study. If present, it is some distance (position 37) into the long tryptic fragment that begins with peptide IX.

The bovine sequence is additionally aligned in Figure 11 with the unique subunit of NADP$^+$-IDH from *E. coli* [38], with the

(a)

(b)

```
         10        20        30        40        50        60        70
          |         |         |         |         |         |         |
GACTACTGCCGCCGCTGTCGCCGGGGTGAGAGGA[GTGGACGCGATGG]CAGGGCCCGCGTGGATCTCCAAGGT
                                              M  A  G  P  A  W  I  S  K  V
```

```
         80        90       100       110       120       130       140
          |         |         |         |         |         |         |
CTCTCGGCTGCTGGGGGCATTCCACAACCAAAAACAGGTGACCAGAGGTTTTGCTGGTGGTGTTAAGACAGT
 S  R  L  L  G  A  F  H  N  Q  K  Q  V  T  R  G  F  A  G  G  V  K  T  V
```

```
        150       160       170       180       190       200       210
          |         |         |         |         |         |         |
AACTTTAATTCCAGGAGATGGAATTGGCCCAGAAATTTCTGCTGCAGTTATGAAGATTTTTGATGCTGCCAA
 T  L  I  P  G  D  G  I  G  P  E  I  S  A  A  V  M  K  I  F  D  A  A  K
```

```
        220       230       240       250       260       270       280
          |         |         |         |         |         |         |
AGCACCTATTCAGTGGGAGGAGCGAAACGTCGCCGCCATTCAAGGTCCGGGAGGCAAGTGGATGATCCCTCC
 A  P  I  Q  W  E  E  R  N  V  A  A  I  Q  G  P  G  G  K  W  M  I  P  P
```

```
 290       300       310       320       330       340       350       360
  |         |         |         |         |         |         |         |
AGAAGCCAAAGAGTCCATGGATAAGAACAAGATGGGCTTGAAAGGCCCTTTAAAAACACCAATAGCTGCTGG
 E  A  K  E  S  M  D  K  N  K  M  G  L  K  G  P  L  K  T  P  I  A  A  G
```

```
        370       380       390       400       410       420       430
          |         |         |         |         |         |         |
TCACCCATCTATGAATTTATTGCTGCGTAAAACATTTGACCTTTATGCAAATGTCCGACCATGTGTCTCAAT
 H  P  S  M  N  L  L  L  R  K  T  F  D  L  Y  A  N  V  R  P  C  V  S  I
```

```
        440       450       460       470       480       490       500
          |         |         |         |         |         |         |
CGAAGGCTATAAAACCCCTTACCATGATGTAAATATTGTCACCATTCGAGAGAACACGGAAGGAGAATACAG
 E  G  Y  K  T  P  Y  H  D  V  N  I  V  T  I  R  E  N  T  E  G  E  Y  S
```

```
        510       520       530       540       550       560       570
          |         |         |         |         |         |         |
TGGAATCGAGCATGTGATCGTGGATGGAGTTGTGCAGAGTATCAAGCTCATCACCGAGGCAGCGAGCAAGCG
 G  I  E  H  V  I  V  D  G  V  V  Q  S  I  K  L  I  T  E  A  A  S  K  R
```

```
 580       590       600       610       620       630       640
  |         |         |         |         |         |         |
CATCGCGGAGTTTGCATTTGAATACGCACGGAACAACCACCGGAGCAACGTCACAGCCGTGCACAAAGCCAA
 I  A  E  F  A  F  E  Y  A  R  N  N  H  R  S  N  V  T  A  V  H  K  A  N
```

```
 650       660       670       680       690       700       710       720
  |         |         |         |         |         |         |         |
CATCATGCGAATGTCAGATGGGCTTTTTCTGCAAAAATGCAGGGAAGTTGCAGAAAACTGTAAAGATATTAA
 I  M  R  M  S  D  G  L  F  L  Q  K  C  R  E  V  A  E  N  C  K  D  I  K
```

```
        730       740       750       760       770       780       790
          |         |         |         |         |         |         |
ATTTAATGAGATGTACCTTGATACAGTATGTTTGAAATATGGTCCAAGACCCGTCCCAGTTTGATGTTCTTGT
 F  N  E  M  Y  L  D  T  V  C  L  N  M  V  Q  D  P  S  Q  F  D  V  L  V
```

```
 800       810       820       830       840       850       860
  |         |         |         |         |         |         |
TATGCCAAATTTGTATGGAGACATCCTTAGTGACCTGTGTGCGGGATTGATTGGAGGTCTTGGTGTGACACC
 M  P  N  L  Y  G  D  I  L  S  D  L  C  A  G  L  I  G  G  L  G  V  T  P
```

```
 870       880       890       900       910       920       930
  |         |         |         |         |         |         |
AAGCGGCAACATTGGAGCCAACGGAGTTGCAATCTTCGAGTCGGTTCACGGGACCGCCCCGGACATTGCGGG
 S  G  N  I  G  A  N  G  V  A  I  F  E  S  V  H  G  T  A  P  D  I  A  G
```

```
 940       950       960       970       980       990      1000
  |         |         |         |         |         |         |
CAAGGACATGGCCAATCCCACAGCCCTCCTGCTCAGCGCCGTGATGATGCTGCGCCACATGGGGCTTTTTGA
 K  D  M  A  N  P  T  A  L  L  L  S  A  V  M  M  L  R  H  M  G  L  F  D
```

```
1010      1020      1030      1040      1050      1060      1070      1080
  |         |         |         |         |         |         |         |
CCACGCTGCGAAGATTGAGACGGCCGTGTTTTGCTACAATTAAGGATGGAAAGAGCTTAACCAAAGATTTGGG
 H  A  A  K  I  E  T  A  C  F  A  T  I  K  D  G  K  S  L  T  K  D  L  G
```

```
1090      1100      1110      1120      1130      1140      1150
  |         |         |         |         |         |         |
AGGCAATTCAAAGTGCTCGGACTTCACAGAAGAAATCTGTCGCCGAGTAAAAGATTTAGATTAAGGCTTCTA
 G  N  S  K  C  S  D  F  T  E  E  I  C  R  R  V  K  D  L  D
```

```
1160      1170      1180      1190      1200      1210      1220
  |         |         |         |         |         |         |
CGGATGGTATCTGCATCAGTCACTCCTAGTGGACGCCACATCAGCCTGTGTTAGAATACTTCCCATTATGTA
```

```
1230      1240      1250      1260      1270      1280      1290
  |         |         |         |         |         |         |
TGCATTTATGCATTTGTTGCTTGTTTCTTGACAGAGTACATTTTTAGACCTGGCCTTTTTCTTAAAAATCTG
```

```
1300      1310      1320      1330      1340      1350      1360
  |         |         |         |         |         |         |
TGCAGAGGGTGCAGGTAATGTCCCCAGGCCTGCTTTCAAAGGACTTCTTCCAAGTGCTTGTTTTATTTATTA
```

```
1370      1380      1390      1400      1410      1420      1430      1440
  |         |         |         |         |         |         |         |
AACGTCTATCTGGTATACATTTTTTTGTAAACTCTCAAGTGAACGGTTATCATTTGCTGTTGTTAACCATTT
```

```
1450      1460      1470      1480      1490      1500      1510
  |         |         |         |         |         |         |
TATGCTTCAGTTAAAATAGTTTTTCTTTCACTGTAAATATCATGATACAGAAATACTAAGAGAAAACATCTA
```

```
1520      1530      1540      1550      1560      1570      1580
  |         |         |         |         |         |         |
ACTTTTTAAAGAAAAACACTGATATTCTTGGTTTACGAAAAATATAATGGAGATAAAGTAGAATGTTGGCAT
```

```
1590      1600      1610      1620      1630      1640      1650
  |         |         |         |         |         |         |
AATAGCATGAGAGGATATTCTTATAAAACTAAATGGGCACAAGAGAAATTTCCTGGGAAAGTTTACATCAAA
```

```
1660      1670      1680      1690      1700      1710      1720
  |         |         |         |         |         |         |
AACAGTGAATATGGTATATTTCTTAGTGATATGGAAAGTAAAGACTAATTTGTACTTTACATGAATTACTGA
```

```
1730      1740      1750      1760      1770
  |         |         |         |         |
GTATAAATAAAGACTTTGGATCGGAGAATAAAAAAAAAAAAAAAAA
```

Figure 10   Bovine heart NAD⁺-IDH subunit 3/4 cDNA sequence and deduced amino acid sequence

(a) Sequencing strategy for clone 102, a product of 3′-RACE. Sequences 1 and 6 were obtained from the original clone using M13 forward and reverse primers respectively. The Erase-a-Base method deleted sequence before sequence 2, which was then obtained with the M13 forward primer. Sequence 8 was obtained after NsiI digestion and subcloning (M13 reverse primer). Sequences 3, 4, 5, 7, 9 and 10 were obtained using synthetic primers (Z6, Z2, Z8, Z9, Z7 and Z5 respectively) based on newly obtained sequences. (b) cDNA for subunit 3/4 (5′-untranslated region, regions coding for mitochondrial leader sequence and mature protein, and 3′-untranslated region). The N-terminal amino acid sequence determined independently is underlined and the poly(A)⁺ signal AATAAA is double underlined. The translation initiation site consensus sequence is boxed [37]. The DNA sequence has been deposited in the GenBank Database under the accession number U07980.

two types of subunits present in yeast NAD⁺-IDH [18,19] and with the γ-subunits from monkey testis and rat epididymis [17]. The comparison groups subunit 3/4 with yeast IDH2, ahead of the other sequences, reflecting the higher identity between the mature forms of these proteins (52%) than between subunit 3/4 and yeast IDH1 (44%), rat or monkey γ (42%) and E. coli (30%). It should be pointed out that identity plus well-conserved residues in the comparison range from 79% (subunit 3/4–yeast IDH2) to 62% (subunit 3/4–IDH of E. coli). A P-loop-type binding site ([AG]-X-X-X-X-G-K-[ST]) common to the regions of a number of proteins that bind ATP or GTP [39] is found only in the mammalian subunits, 3/4 and γ (shown shadowed). The IDH and isopropylmalate dehydrogenases signature sequence (N-[LIMVFY]-X-G-D-[LIMVFY]-X-[SG]-[DN]-X-X-[SA]-X₃ₒᵣ₄-G-[SG]-[LIVM]-G-[LIVMF]), a Gly-rich stretch located in the C-terminal regions of the E. coli NADP⁺-IDH and iso-propylmalate dehydrogenase (PC/Gene; Prosite analysis), is recognized in subunit 3/4 and in yeast IDH2, but not yeast IDH1 or in the mammalian γ-subunits (shown in bold).

The three-dimensional structure of the E. coli enzyme has been determined at high resolution, and residues involved in the interactions with isocitrate, Mg²⁺ and NADP⁺ identified [21–23]. All residues interacting with isocitrate and Mg²⁺ are conserved in the bovine heart subunit 3/4 sequence and five of eight residues involved in binding the NAD⁺ portion of NADP⁺ (Figure 11). Thus although the identity of bovine and E. coli subunits is only 30%, regions of particular interest are highly conserved.

DISCUSSION

This study has confirmed previous work identifying four electrophoretically distinct subunits in bovine NAD⁺-IDH. Heterogeneity due to allelic variability or to artefactual proteolysis has been ruled out. N-Terminal sequencing of isolated subunits

N-terminal peptide

```
Bovine IDH3/4  MAGPAWISKV------SRLLGAPHNQKQVTRGFAG----------GVK  32
Yeast IDH2     MLRNTFFRNT------SRRFLATVKQPSIGR-YTGKPNPST-----GKY  37
Rat IDHγ       AVAAGSAAKAIFKPALLCRPWEVLAAHEAPRRSISSQQTIPPSAKYGGRH  50
Monkey IDHγ    --------------------------------ISSQQTIPPSAKYGGRH  17
Yeast IDH1     MLNRTIAKRT-----------LATAAQAERTLPK--------KYGGRF  29
E. coli IDH    MESKV---------------VVPAQGKKITLQNGKLNVPENP------  27


Bovine IDH3/4  TVTLIPGDGIGPEISAAVMKIFDAAKAP-------IQWEERNVAAIQGP-  74
Yeast IDH2     TVSFIEGDGIGPEISKSVKKIFSAANVP-------IEWESCDVSPIFVN-  79
Rat IDHγ       TVTMIPGDGIGPELMLHVKSVFRHACVP-------VDFEEVHVSSNADE-  92
Monkey IDHγ    TVTMIPGDGIGPELMLHVKSVFRHACVP-------VDFEEVHVSSNADE-  59
Yeast IDH1     TVTLIPGDGVGKEITDSVRTIFEAENIP-------IDWETINIKQTDHK-  71
E. coli IDH    IIPYIEGDGIGVDVTPAMLKVVDAAVEKAYKGERKISWMEIYTGEKSTQV  77
               ... * ***.*  ..   ...  .      . . . .
                  ■


Bovine IDH3/4  --GGKWMIPPEAKESMDKNKMGLKGPLKTP--IAAGHPSMNLLLRKTFDL  120
Yeast IDH2     --GLT-TIPDPAVQSITKNLVALKGPLATP--IGKGHRSLNLTLRKTFGL  124
Rat IDHγ       --EDIR----NAIMAIRRNRVALKGNIETNHDLPPSHKSRNNILRTSLDL  136
Monkey IDHγ    --EDIR----NAIMAIRRNRVALKGNIETNHNLPPSHKSRNNILRTSLDL  103
Yeast IDH1     --EGVY----EAVESLKRNKIGLKGLWHTPAD-QTGHGSLNVALRKQLDI  114
E. coli IDH    YGQDVW-LPAETLDLIREYRVAIKGPLTTP--VGGGIRSLNVALRQELDL  124
               .    .  ...**   *    .  * * **   ...
                                          ◆ ◆ ◆


                            Vc
Bovine IDH3/4  YANVRPCVSIEGYKTPYHD---VNIVTIRENTEGEYSGIEHVIVDGVVQ-  166
Yeast IDH2     PANVRPAKSIEGFKTTYEN---VDLVLIRENTEGEYSGIEHIVCPGVVQ-  170
Rat IDHγ       YANVIHCKSLPGVVTRHKD---IDILIVRENTEGEYSSLEHESVAGVVE-  182
Monkey IDHγ    YANVIHCKSLPGVVTRHKD---IDILIVRENTEGEYSSLEHESVAGVVE-  149
Yeast IDH1     YANVALFKSLKGVKTRIPD---IDLIVIRENTEGEFSGLEHESVPGVVE-  160
E. coli IDH    YICLRPVRYYQGTPSPVKHPELTDMVIFRENSEDIYAGIEWKADSADAEK  174
               .  .     *  .  .  .... ***.*. ....*     .  ...
                  ◆          ◆   ◆ ◆


Bovine IDH3/4  --------------------SIKLITEAASKRIAEFAFEYARNNHRSN  194
Yeast IDH2     --------------------SIKLITRDASERVIRYAFEYARAIGRPR  198
Rat IDHγ       --------------------SLKIITKAKSLRIAEYAPKLAQESGRKK  210
Monkey IDHγ    --------------------SLKIITKAKSLRIAEYAPKLAQESGRKK  177
Yeast IDH1     --------------------SLKVMTRPKTERIARFAPDFAKKYNRKS  188
E. coli IDH    VIKFLREEMGVKKIRFPEHCGIGIKPCSEEGTKRLVRAAIEYAIANDRDS  224
                ..*    . .*.  *.. * * .


                          II/III/IV
Bovine IDH3/4  VTAVHKANIMRMSDGLFLQKCREVA-ENC-----------------KD  224
Yeast IDH2     VIVVHKSTIQRLADGLFVNVAKELS-KEY-----------------PD  228
Rat IDHγ       VTAVHKANIMKLGDGLFLQCCREVAAR-Y-----------------PQ  240
Monkey IDHγ    VTAVHKANIMKLGDGLFLQCCREVAAR-Y-----------------PQ  207
Yeast IDH1     VTAVHKANIMKLGDGLFRNIITEIGQKEY-----------------PD  219
E. coli IDH    VTLVHKGNIMKFTEGAFKDWGYQLAREEFGGELIDGGPWLKVKNPNTGKE  274
               *. ***..*  ....* *  . .      .
                  ◆


                    IX
Bovine IDH3/4  IKFNEMYLDTVCLNMVQDPSQFD--VLVMPNLYGDILSDLCAGLIGG-LG  271
Yeast IDH2     LTLETELIDNSVLKVVTNPSAYTDAVSVCPNLYGDILSDLNSGLSAGSLG  278
Rat IDHγ       ITFDSMIVDNTTMQLVSRPQQFD--VMVMPNLYGNIVNNVCAGLVGGP-G  287
Monkey IDHγ    ITFENMIVDNTTMQLVSRPQQFD--VMVMPNLYGNIVNNVCAGLVGGP-G  254
Yeast IDH1     IDVSSIIVDNASMQAVAKPHQFD--VLVTPSMYGTILGNIGAALIGGP-G  266
E. coli IDH    IVIKDVIADAFLQQILLRPAEYD--VIACMNLWGDYISDALAAQVGG-IG  321
               . ...  *.    .  * *  .   . . *.   .  . .     .* *
                   ▲                         ▲ ▲          ■■


Bovine IDH3/4  VTPSGNIGANGVAIFE-SVHGTAPDIAGKDMANPTALLLSAVMMLRHMGL  320
Yeast IDH2     LTPSANIG-HKISIFE-AVHGSAPDIAGQDKANPTALLLSSVMMLNHMGL  326
Rat IDHγ       LVAGANYG-HVYAVFETATRNTGKSIANKNIANPTATLLASCMMLDHLKL  336
Monkey IDHγ    LVAGANYG-HVYAVFETATRNTGKSIANKNIANPTATLLASCMMLDHLKL  303
Yeast IDH1     LVAGANFG-RDYAVFEPGSRHVGLDIKGQNVANPTAMILSSTLMLNHLGL  315
E. coli IDH    IAPGANIG-DECALFE-ATHGTAPKYAGQDKVNPGSIILSAEMMLRHMGW  369
               .....* *   ..**  .  .  .     ...**.. .*.. .** *.
                         ■ ■           ■■


                VI                    Va
Bovine IDH3/4  FDHAAKIETACPATIKDG-KSLTKD--LGGNS--KCSDFTEEICRRVKDL  365
Yeast IDH2     TNHADQIQNAVLSTIASGPENRTGD--LAGTA--TTSSFTEAVIKRL---  369
Rat IDHγ       HSYATSIRKAVLASMDNE-NMHTPD--IGGQG--TTSQAIQDIIRHIRII  381
Monkey IDHγ    HSYATSIRKAVLASMDNE-NMHTPD--IGGQG--TTSEAIQDIIRHIRVI  348
Yeast IDH1     NEYATRISKAVHETIAEG-KHTTRD--IGGSS--STTDFTNEIINKLSTM  360
E. coli IDH    TEAADLIVKGMEGAINAKTVTYDFERLMDGAKLLKCSEFGDAIIENM---  416
               . *. *  ..  ...  .      ...*     .  . ... ..
                            ■


Bovine IDH3/4  D------  366
Yeast IDH2     -------  369
Rat IDHγ       NGRAVEA  388
Monkey IDHγ    NGRAVEA  355
Yeast IDH1     -------  360
E. coli IDH    -------  416
```

revealed that subunits 1, 2 and 3 differ in primary structure. Subunit 4, although differing electrophoretically from the other subunits, was identical in sequence with subunit 3 over the first 30 residues.

Comparison of the above sequences with those of the related enzyme from pig heart [14] revealed that subunits 1 and 2 are homologous, but not identical, with subunits β and γ respectively of the pig heart enzyme. Subunits 3 and 4 are identical, over the first 30 residues, with pig heart subunit α. Since the pig heart subunit α is present in the ratio $2\alpha:1\beta:1\gamma$ in the intact enzyme, the electrophoretic heterogeneity resulting in subunits 3 and 4 of the bovine heart enzyme raised the possibility of a comparable heterogeneity in the pig heart α-subunit. Such a heterogeneity was indeed found after electrophoresis of the pig heart subunits under the conditions that yielded the bovine heart heterogeneity.

Utilization of the N-terminal sequence for subunits 3/4 led to a full-length cDNA through reverse transcription-PCR, cloning and sequencing. The sequence was extended in the 5' direction by anchored PCR, yielding a leader sequence containing the complete mitochondrial import sequence, as indicated by the consensus sequence for the translation initiation start site. Further work will be required to determine whether the 5'-untranslated region is complete.

The amino acid sequence for the mature protein contained 339 amino acids and had a molecular mass of 36685 Da, comparable with the 39000 Da and 39600 Da obtained for subunits 3 and 4 respectively by SDS/PAGE [11]. The degree of identity of subunits 3/4 with derived protein sequences for the γ-subunits from rat epididymal and monkey testis issue was relatively low (42%) although both species possess a P-loop-type ATP/GTP-binding site near the C-terminus and have conserved residues indicated in *E. coli* to be required for isocitrate binding. The IDH/isopropylmalate dehydrogenase signature sequence is present in subunits 3/4 but this region in the γ-subunits is sufficiently altered that it is not recognized by PC/Gene. In contrast, the 92 residues reported for the N-terminal region and Cys-containing peptides of the pig heart α-subunit [14] were completely conserved. This identity across species line suggests that maintenance of a particular sequence and therefore structure for subunits 3/4/α may be crucial to the enzyme's function.

Subunits 3/4 retain all residues that were shown in the distantly related *E. coli* NADP$^+$-dependent enzyme to bind the substrates Mg^{2+} and isocitrate and five of eight residues interacting with NAD$^+$. The γ-subunits from rat and monkey possess six of seven residues binding isocitrate but fewer than half those involved with either Mg^{2+} or NAD$^+$ [17]. These observations are consistent with evidence indicating that the homologous pig heart α-subunits are important catalytically [15,16], whereas the β- and γ-subunits act in a regulatory fashion. The higher homology of subunits 3/4 with the yeast IDH2 subunit (indicated to be

Figure 11 Alignment of amino acid sequences of eukaryotic NAD$^+$-IDH subunits and *E. coli* NADP$^+$-IDH: bovine heart IDH subunits 3/4, rat epididymal and monkey testicular γ-subunits [17], yeast IDH1 and IDH2 [18,19] and *E. coli* IDH [38]

Pig heart N-terminal and Cys-containing peptides of the α-subunit [14], identical with sequences in the bovine subunit 3/4, are underlined and their peptide numbers shown above. Residues that are perfectly conserved are denoted by an asterisk and those well conserved by a bold point. The P-loop-type ATP/GTP-binding site is shown shadowed in the proteins that possess it. The IDH and isopropylmalate dehydrogenase signature sequence is shown in bold. The amino acid residues shown in the *E. coli* enzyme to bind isocitrate are indicated by ◆, those binding Mg^{2+} by ▲ and those binding the adenine and 5'-phosphate of NADP$^+$ by ■ [21–23].

catalytic) than with IDH1 (indicated to be regulatory) [19] fits with these conclusions.

The origin of the heterogeneity in subunits 3 and 4 in the bovine heart enzyme remains to be determined. Possibilities include the existence of two unique gene products, limited alternate splicing of a single primary transcript or a post-translational modification(s). Although only a single full-length coding sequence was obtained after 3'-RACE, this may have been because of low cloning efficiency (only five clones were obtained after several rounds of PCR and transformation). Of possible post-translational modifications, phosphorylation is of interest because of its inhibitory effect on isocitrate binding in the *E. coli* enzyme [40]. Numerous potential phosphorylation sites can be located within subunits 3/4 using prediction programs such as Prosite from PC/Gene. Ser-109, homologous with the phosphorylatable Ser-113 of *E. coli*, does not have the characteristics of such a site. A single potential N-glycosylation is present at residue 167.

The complexity of NAD$^+$-IDH, evident in the mammalian enzymes by multiple homotropic and heterotrophic effects, is apparent even in the *E. coli* enzyme in the involvement of the two subunits of the homodimer in each of the two Mg^{2+}-isocitrate-binding sites. The conservation of key residues for Mg^{2+}-isocitrate binding in the bovine subunits 3/4 indicates that adjacent subunits contribute to the active site in the mammalian enzymes. Since the minimal size for the active enzyme is a tetramer, it is very possible that an active site comprises a contribution from a catalytic subunit, such as subunits 3/4 or α, and a 'regulatory' subunit which may be subunit 1/β or 2/γ. Whether this is so must await determination of the quaternary structure of the mammalian enzymes. Such an arrangement, where 'regulatory' subunits are required for significant activity, would explain the less than clear cut results found when catalytic and regulatory functions for specific subunits were investigated in yeast and porcine enzymes [15,16,20].

Recently, functional yeast NAD$^+$-IDH was shown to bind with high affinity and specificity to the 5'-untranslated regions of all major mitochondrial mRNAs [41–43]. Such a finding is quite unexpected, placing IDH into a growing family of RNA-binding enzymes that include glutamate dehydrogenase and thymidylate synthase [44]. Location of the RNA-binding domain in the enzyme, characterization of the RNA-binding element and determination of the regulation and biological role of the enzyme–RNA interaction are of interest. It is hypothesized that the enzymes that bind both (di)nucleotide and RNA and other proteins that bind only RNA originate from a common ancestral enzyme with (di)nucleotide-binding capacity [44]. For yeast NAD$^+$-IDH, binding to mRNA may affect translation, thereby linking the mitochondrial energy-production function to mitochondrial biogenesis [42]. Whether bovine NAD$^+$-IDH possesses such an activity remains an open question.

REFERENCES

1 Georgio, N. A. J., Yip, A. T., Fleming, J. and Plaut, G. W. E. (1970) J. Biol. Chem. **245**, 5469–5477
2 Shen, W. C., Mauck, L. and Colman, R. F. (1974) J. Biol. Chem. **249**, 7942–7949
3 Fan, C. C., Lin, J. P. F. and Plaut, G. W. E. (1975) J. Biol. Chem. **250**, 2022–2027
4 Chen, R. F. and Plaut, G. W. E. (1963) Biochemistry **2**, 1023–1032
5 Illingworth, J. A. (1972) Biochem. J. **129**, 1119–1124
6 Barnes, L. D., Kuehn, G. D. and Atkinson, D. E. (1971) Biochemistry **10**, 3939–3944
7 Plaut, G. W. E., Schramm, V. L. and Aogaichi, T. (1974) J. Biol. Chem. **249**, 1848–1856
8 Ehrlich, R. S. and Colman, R. F. (1981) J. Biol. Chem. **256**, 1276–1282
9 Rushbrook, J. I. (1978) Ph.D. Dissertation, Rutgers the State University of New Jersey
10 Ehrlich, R. S. and Colman, R. F. (1982) J. Biol. Chem. **257**, 4769–4774
11 Rushbrook, J. I. and Harvey, R. A. (1978) Biochemistry **17**, 5339–5346
12 Ramachandran, N. and Colman, R. (1978) Proc. Natl. Acad. Sci. U.S.A. **75**, 252–255
13 Ramachandran, N. and Colman, R. (1980) J. Biol. Chem. **255**, 8859–8864
14 Huang, Y.-C. and Colman, R. F. (1990) Biochemistry **29**, 8266–8273
15 Ehrlich, R S. and Colman, R. (1983) J. Biol. Chem. **258**, 7079–7086
16 Ehrlich, R. S. and Colman, R. (1984) J. Biol. Chem. **259**, 11936–11942
17 Nichols, B. J., Hall, L., Perry, A. C. F. and Denton, R. M. (1993) Biochem. J. **295**, 347–350
18 Cupp, J. R. and McAlister-Hern, L. (1991) J. Biol. Chem. **266**, 22199–22205
19 Cupp, J. R. and McAlister-Hern, L. (1992) J. Biol. Chem. **267**, 16417–16423
20 Cupp, J. R. and McAlister-Henn, L. (1993) Biochemistry **32**, 9323–9328
21 Stoddard, B. L. and Koshland, D. E., Jr. (1993) Biochemistry **32**, 9317–9322
22 Stoddard, B. L., Dean, A. and Koshland, D. E., Jr. (1993) Biochemistry **32**, 9310–9316
23 Hurley, J. H., Dean, A. M., Koshland, D. E., Jr. and Stroud, R. M. (1991) Biochemistry **30**, 8671–8678
24 Laporte, D. C. and Koshland, D. E., Jr. (1983) Nature (London) **305**, 286–290
25 Imada, K., Sato, M., Tanaka, N., Katsube, Y., Matsuura, Y. and Oshima, T. (1991) J. Mol. Biol. **222**, 725–738
26 Hurley, J. H., Thorsness, P. E., Ramalingam, V., Helmers, N. H., Koshland, D. E., Jr. and Stroud. R. M. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 8635–8639
27 Plaut, G. W. E. (1969) Methods Enzymol. **13**, 34–42
28 Ehrlich, R. S., Hayman, S., Ramachandran, N. and Colman, R. (1981) J. Biol. Chem. **256**, 10560–10564
29 Cohen, P. F. and Colman, R. F. (1971) Biochim. Biophys. Acta **242**, 325–330
30 Matsudaira, P. (1987) J. Biol. Chem. **262**, 10035–10038
31 Auffray, C. and Rougeon, F. (1980) Eur. J. Biochem. **107**, 303–314
32 Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. U.S.A. **69**, 1408–1412
33 Frohman, M. A., Dush, M. K. and Martin, G. R. (1985) Proc. Natl. Acad. Sci. U.S.A. **85**, 8998–9002
34 Edwards, J. B. D. M., Delort, J. and Mallet, J. (1991) Nucleic. Acids Res. **19**, 5227–5232
35 Sanger, F., Nicklen, S. and Coulsen, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
36 Zasloff, M., Ginder, G. D. and Felsenfeld, G. (1978) Nucleic Acids Res. **5**, 1139–1152
37 Kozak, M. (1987) Nucleic Acids Res. **15**, 8125–8132
38 Thorsness, P. E. and Koshland, D. E., Jr. (1987) J. Biol. Chem. **262**, 10422–10425
39 Saraste, M., Sibbald, P. R. and Wittinghofer, A. (1990) Trends Biochem. Sci. **5**, 430–434
40 Dean, A. M., Lee, M. H. and Koshland, D. E., Jr. (1989) J. Biol. Chem. **264**, 20482–20486
41 Papadopoulou, B., Dekker, P., Blom, J. and Grivell, L. A. (1990) EMBO J. **9**, 4135–4143
42 Elzinga, S. D. J., Bednarz, A. L., von Oosterum, K., Dekker, P. J. T. and Grivell, L. A. (1993) Nucleic Acids Res. **21**, 5328–5331
43 Dekker, P. J. T., Stuurman, J., van Oosterum, K. and Grivell, L. A. (1992) Nucleic Acids Res. **20**, 2647–2655
44 Hentze, M. W. (1994) Trends Biochem. Sci. **19**, 101–103