

UDP-glucose dehydrogenase from bovine liver: Primary structure and relationship to other dehydrogenases

JOHN HEMPEL, JOHN PEROZICH, HANA ROMOVACEK, AMY HINICH,
INGRID KUO, AND DAVID SIDNEY FEINGOLD

Department of Molecular Genetics and Biochemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15261

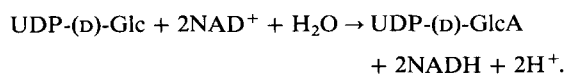
(RECEIVED February 10, 1994; ACCEPTED April 13, 1994)

Abstract

The primary structure of bovine liver UDP-glucose dehydrogenase (UDPGDH), a hexameric, NAD⁺-linked enzyme, has been determined at the protein level. The 52-kDa subunits are composed of 468 amino acid residues, with a free N-terminus and a Ser/Asn microheterogeneity at one position. The sequence shares 29.6% positional identity with GDP-mannose dehydrogenase from *Pseudomonas*, confirming a similarity earlier noted between active site peptides. This degree of similarity is comparable to the 31.1% identity vs. the UDPGDH from type A *Streptococcus*. Database searching also revealed similarities to a hypothetical sequence from *Salmonella typhimurium* and to "UDP-N-acetyl-mannosaminuronic acid dehydrogenase" from *Escherichia coli*. Pairwise identities between bovine UDPGDH and each of these sequences were all in the range of ~26–34%. Multiple alignment of all 5 sequences indicates common ancestry for these 4-electron-transferring enzymes. There are 27 strictly conserved residues, including a cysteine residue at position 275, earlier identified by chemical modification as the expected catalytic residue of the second half-reaction (conversion of UDP-aldehydoglucose to UDP-glucuronic acid), and 2 lysine residues, at positions 219 and 338, one of which may be the expected catalytic residue for the first half-reaction (conversion of UDP-glucose to UDP-aldehydoglucose). A GXGXXG pattern characteristic of the coenzyme-binding fold is found at positions 11–16, close to the N-terminus as with "short-chain" alcohol dehydrogenases. Because the enzyme combines functionalities of alcohol and aldehyde dehydrogenases, it was also of interest to search specifically for other sequence similarities to either of these 2 enzymes, as well as to histidinol dehydrogenase, another 4-electron-transferring dehydrogenase, but none were found.

Keywords: alcohol dehydrogenase; aldehyde dehydrogenase; amino acid sequence; GDP-mannose dehydrogenase; histidinol dehydrogenase; protein evolution; protein family; UDP-glucose dehydrogenase; UDP-N-acetylmannosamine dehydrogenase

UDP-glucose dehydrogenase (UDPGDH, EC 1.1.1.22) is a member of a small group of NAD⁺-linked, 4-electron-transferring oxidoreductases (reviewed by Feingold & Franzen, 1981) and converts UDP-D-glucose to UDP-D-glucuronic acid (UDP-GlcA), performing 2 separate but linked oxidations on the substrate molecule. As such, the product of the first half-reaction, UDP-6-aldehydo-D-glucose, is the substrate for the second. The overall reaction is



Except in plants, UDPGDH provides the only pathway for formation of UDP-GlcA. For proteoglycans, UDP-GlcA provides the D-glucuronosyl and eventual D-iduronosyl moieties of heparan sulfate, dermatan sulfate, and heparin. UDP-GlcA and UDP-N-acetyl glucosamine moieties are alternately incorporated into hyaluronic acid, as found in synovial fluid. Recently, depressed levels of UDPGDH have been observed in cells lining the synovium of rheumatoid arthritic joints (Pitsillides et al., 1993). UDP-GlcA is also the substrate for glucuronosylation of xenobiotics via UDP-glucuronosyl transferases.

The enzyme was first detected in bovine liver in 1954 (Strominger et al., 1954) and was purified to homogeneity 15 years later (Zalitis & Feingold, 1969). It is an apparent homohexamer with 52-kDa subunits (Gainey et al., 1972) and functions as a trimer of dimers (Franzen et al., 1980; Jaenicke et al., 1986). However, despite such data, and despite the availability

Reprint requests to: John Hempel, Department of Molecular Genetics and Biochemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15261; e-mail: hempel@psc.edu.

since 1981 of the sequence of a tryptic peptide containing an active-site cysteine residue (Franzen et al., 1981), there had been no reports concerning the amino acid sequence of this interesting enzyme. Primary structures for 2 other 4-electron-transferring dehydrogenases, GDP-mannose dehydrogenase from *Pseudomonas* (Roychoudhury et al., 1989) and histidinol dehydrogenase from a variety of sources (reviewed by Hinshelwood & Stoker, 1992), have become available within the past several years. For these reasons we undertook determination of the amino acid sequence of bovine liver UDPGDH. The sequence of UDPGDH from *Streptococcus* became available just as we completed our structure (Dougherty & van de Rijn, 1993).

Results

The amino acid sequence of UDPGDH as derived from overlapping peptides is given in Figure 1 with sequences of all supporting peptides. From the free N-terminus, information as far as cycle 43 was obtained directly from the intact protein. At the C-terminus, a peptide from a V8 digest, with no Glx in its amino acid composition, ends with Lys. As well, a negative result was obtained on sequencing material from an endoLys-C digest eluted from a CT-isolator membrane after coupling and anhydrous trifluoroacetic acid (TFA) cleavage, a result compatible with a C-terminal lysine residue.

Peptides are referred to in the text according to the digest code as defined in the legend to Figure 1, followed by the position number of the N-terminal amino acid of the peptide in the complete sequence.

A Ser/Asn microheterogeneity is indicated at position 74 by the sequences of various peptides encompassing this position (cf. Fig. 1). In the case of V8-protease peptides, 2 adjacent HPLC peaks were recovered in approximately equal yield. The sequence of one supported Ser and the other Asn, as indicated in the figure.

In the C-terminal region we found it expedient to take advantage of 2 (acid labile) Asp-Pro linkages, at positions 367/368 and 399/400 in the final structure, to secure one of the final overlaps. Separate redigestions of the smaller of these fragments with V8 protease and endoAsp-N secured confident assignment of Ala-409 and Arg-441 + 442, as indicated in Figure 1. The Ala assignment was particularly problematic due to trailing from Ala-407 combined with severe apparent washout. For confirmation of this residue we coupled ~20 pmol of peptide H⁺V8-402 to an arylamine-derivatized polyvinylidene difluoride (PVDF; Sequelon AA[®], Millipore) membrane through the carboxyls of Asp-405 and Glu-415. Three other peptides were similarly coupled: Ct-71, to support the IKEA sequence (positions 79–82); Ct-323, to support the DKKI sequence at positions 328–329; and Tr-230, to support identifications from the C-terminal half of V8-217 and the CEAT overlap at positions 240–243. This latter peptide proved particularly difficult to isolate from C-18 HPLC; presumably due to its extremely acidic and hydrophobic character, it eluted after all other peptides and in poor relative yield.

A search of the major protein databases, PIR and Swiss-Prot, using the FASTA program (Pearson & Lipman, 1988) revealed 5 sequences with high similarity scores. These included *Pseudomonas* GDP-mannose dehydrogenase (Roychoudhury et al., 1989), *Escherichia coli* UDP-N-acetyl-mannosaminuronic acid dehydrogenase (Daniels et al., 1992; this designation is probably

a misnomer, referring instead to UDP-N-acetyl-mannosamine dehydrogenase [cf. Kawamura et al., 1979]), and a hypothetical sequence from *Salmonella typhimurium* (Waxin H, Kolyva S, Ecobichon C, Popoff MY, 1992, EMBL Data Library, accession S28489). Two other hypothetical sequences, from *Salmonella choleraesuis* and *E. coli* (Bastin DA, Stevenson G, Brown PK, Haase AM, Reeves PR, 1992, EMBL Data Library, accessions S31606 and S31732) were highly similar to the hasB (putative UDPGDH) sequence from group A *Streptococcus*, accessed from the original citation (Dougherty & van de Rijn, 1993).

A multiple alignment of the above sequences, excluding the hypothetical *S. choleraesuis* and *E. coli* sequences to maximize overall diversity, was generated via the GCG Pileup program (Fig. 2). Minor adjustments were introduced manually. Twenty-seven strictly conserved residues are observed, while pairwise identities ranged from 26 to 34% (Table 1). Conserved residues include 3 glycine residues in a GXGXXG pattern at positions 11–16 in the present sequence, 3 other consensus glycine residues just upstream from the single conserved cysteine residue (position 275), and 2 lysine residues (positions 219 and 338).

We used the alignment to generate a substitution matrix (profile) with which to search the databases for other potential relationships (Gribskov et al., 1990), without finding any indication of similarity to the 3 most likely candidates, namely alcohol dehydrogenases, aldehyde dehydrogenases, or histidinol dehydrogenases.

Discussion

Initially this effort was directed toward obtaining peptides of low codon degeneracy in order to design oligonucleotide probes to clone the cDNA for UDPGDH. However, we soon obtained about half of the structure and then undertook a concerted effort to determine the entire structure at the protein level. The sequence represents the first reported primary structure of any eukaryotic UDPGDH that we are aware of.

Sequence

The molecular weight calculated from the monomer structure as determined (52,147) agrees with the earlier estimate of 52,000 (Gainey et al., 1972). Additionally, 12 thiol groups per monomer, as titrated by Gainey et al. (1972), were found in the sequence and amino acid values obtained for the intact enzyme are in good agreement with the integral values expected from the sequence (Table 2). The N-terminal sequence is also compatible with patterns expected for non-N-acetylated proteins with Met-1 residues (Flinta et al., 1986). All peptide alignments are based on overlaps of 2 (cf. positions 95–96 and 338–339) or more residues in addition to sequence compatibility with enzymatic cleavage specificity. The sequence at the C-terminus is supported by the C-terminal Lys of a V8 peptide and is compatible with the negative result from a C-terminal peptide isolator kit.

Tr-267, with the sequence ASVGFGGSCFQK, is the same peptide earlier identified with the catalytic cysteine residue, yet the present peptide sequence differs at its C-terminus from that reported earlier (ASVGFGGSCFZZGK; Franzen et al., 1981). However, allowing for contamination by glycine as frequently encountered with material purified by elution from paper electrophoresis, and overzealous interpretation of the earlier amino

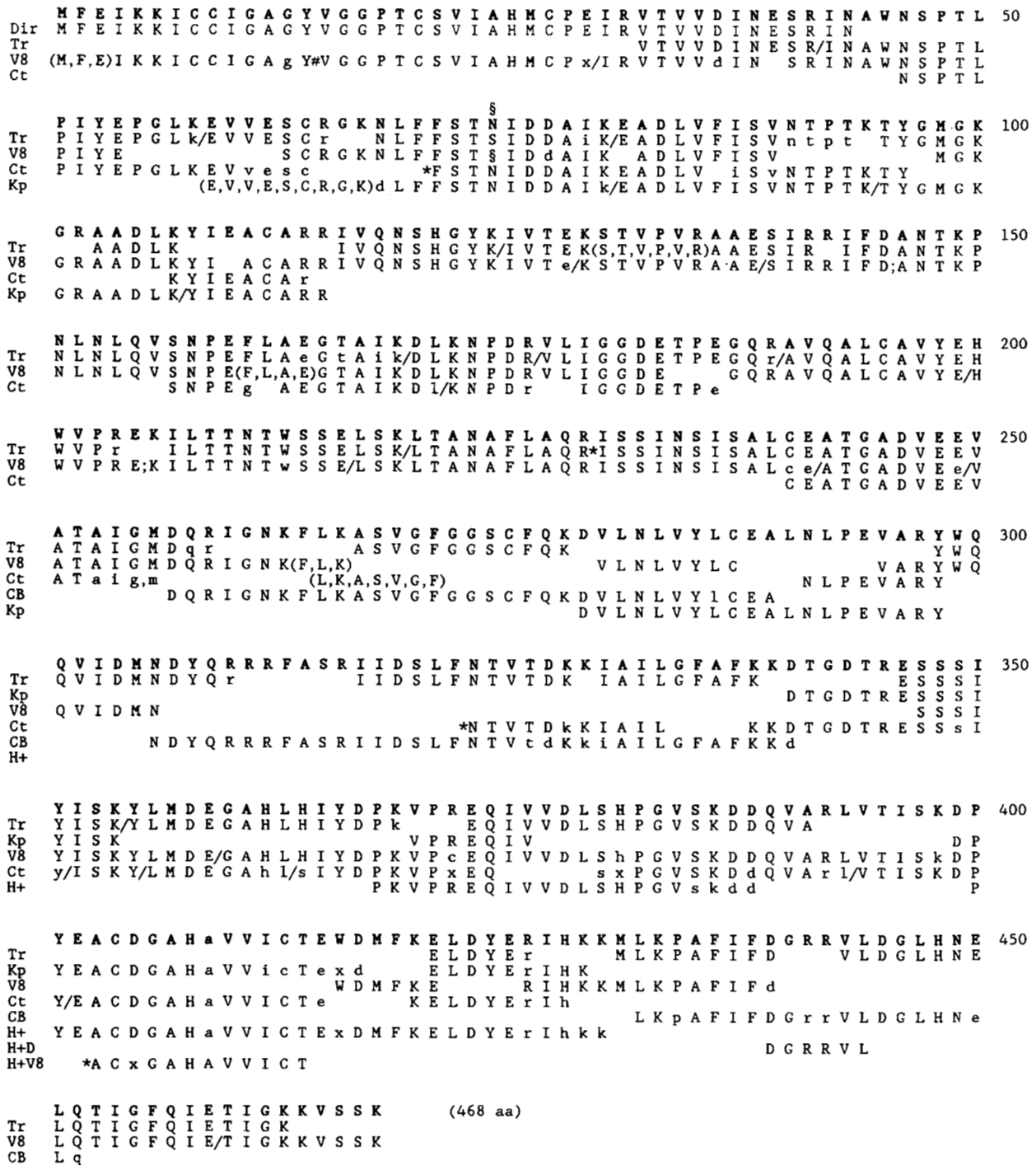


Fig. 1. Diagram of peptides used in assembly of the complete amino acid sequence of bovine liver UDP-glucose dehydrogenase. Peptides are placed on different rows according to the digests they originated from: Tr, trypsin; V8, staphylococcal V8 (Glu-C) protease; Ct, chymotrypsin; Kp, lysine-specific (endoLys-C) protease; CB, cyanogen bromide; H+, acid cleavage in 70% formic acid (preferring Asp-Pro bonds). The completed sequence is shown in bold; periods above every 10th residue (top) and numbers to the right of the sequence enable determination of position numbers. Residues confidently identified by Edman degradation are indicated by capital letters, and tentative identifications by lowercase letters. Boundaries between peptides where their sequences are contiguous in the figure are indicated by /. A few peptides were obtained that contained canonical sites of cleavage that had not been cleaved, as well as a few that represented cleavage at Asp by the V8 protease. Where peptides from both cleaved and noncleaved sites were obtained from the same digest, residue assignments are condensed on the same line and the site is indicated by a semicolon. Commas between residues indicate amino acid compositions of peptides that fit the interpretation but which were not sequenced. Peptides generated by subdigestion with a second protease are indicated by #. The Ser/Asn microheterogeneity at position 74 is noted by §. Sequences obtained after covalent coupling to Millipore arylamine membranes are denoted by * preceding the N-terminal residue.

UDPNAMDH	MSFATISVIGLGLYIGLPTRAFASRQKQ--VIGYDINQHAVDTINRGEIHZVEPDLASVVKTAVEGGFLRASTTP--	72
St-Hypo	MFGIDEVKAIAIGLGLYGLPLAVEFGKSRQ--VVGFQVNNKRILELKNK---VDVNEETTEEELREARYLKFTEIEK	73
SA-UDPGDH	MKIAVAGSGLYGLSL-GVLLSLQNE--VTLYDILPSKVDKINNGLSPTQDEYIEYLLKSKQLS--IKALDLSKA	69
GDPMDH	MRTSIFGLGTYGAVCAGCLSGARCHE--VIGYDVSSTKIDLINQKSPYVEPGLAALQQGRQTRGLSCPTDFKK	72
Bov-UDPGDH	MFEIKKLCCTGAGYVGGTCSVIAHMCPEIRVTVV--VINESRINAWNSPTLTPYEPGLKKEVVEESCRGNLFFST--NIDD	77
UDPNAMDH	--VEADRWLIAVPTPFKGD-----HEPDMTYVESAARSAPV--LKKGALYLESTSPVGGSTEKMAEWLAEMRPDLTFP	142
St-Hypo	I-KECNFYIITVPTPINTY----KQDILTPLIKASETVGTV--LNRGDIYVYESTVYPCCTEEECVPIIARMSGMTFN	144
SA-UDPGDH	AYKEAELVIA--TPTNYNS--RINYRDTQHVETVIKEVLSV---NSHATLIKSTIPIGFIEMRQKFTD-----	134
GDPMDH	AVLSDSVSFTCYCTPSKKN-----DLDLGYIETVCREIGFAIREKSERHTVYVRSIVLPQTVNNVVIPLIEDCSGKKAG	147
Bov-UDPGDH	AIKFADLVFISYNTPTKTYGMKGRAADLKYTEACARRIVQ---NSHGKIKYTERSTVYVRAAESIRRFIDANTKPNLNL	154
UDPNAMDH	QQVGEQADVNIAYCPERVLPGQVMVELIKNDRVIGGMT-----PVCSARASELY-KIFLEGEC--VVTNSRTEEMCKL	212
St-Hypo	QDF-----YVGYSPERINPGDKKHLNLIKITSGST-----AQIAELIDEVYQOIIISAGTY--KAESIKVDEAAKV	209
SA-UDPGDH	-----RIIFSPFELRESKALYDNLPSRIIVSCEENDSPKVKADAERFALLLSAAKKNVPLVIMGASEDAVKL	205
GDPMDH	VDF-----GVGTNPFELRESTAIKDYDFPMTVIC-----ELDKQTGDLLEBIYRELDAPIRKKTVEVEMIKY	211
Bov-UDPGDH	Q-----VSNPFLAEGTAIKDLKNPDRVLIIGDETPEGQAVQA--LCAVYEHVWPREK--ILTTNTWSSLSKRL	220
UDPNAMDH	TENSFRDVIAFANELSLICADQGINVWELIRLANRHPRVNI--LQPGGV--GGHCIAVDLPWFIV--AQNPOQA---RL	283
St-Hypo	IRNTORDLNIALVHELAIIFNRLNIDTEAVLRAAGSKWNFLP---FRPGL-VGGHCIGVDPYYIT--HKSOGIGYYPEI	282
SA-UDPGDH	FANTYLALRVAYFNELDTYAESRKLNSHMI IQGISYDDRIGMHYNNPSPGY--GGYCLPKITKQLL--ANYNNIPQ--TL	279
GDPMDH	TONVWHAAKVTFANETIGNIAKAVGVDGREVMDVICQDHKLNLSRYYMRPGFAFCGSCGLPRDVRATYRASQLDVEH--PM	289
Bov-UDPGDH	TANAFLAQRISINSISALCEATGADVEEVATAIGMDQRIGNKFLKASVGF--GGSCFQKQVLDLNVYLCEALNLPVARY	298
UDPNAMDH	IRTAREVNDH-KPFWVIDQVKAADVCLAATDKRASELKIACFGLAEPNIDDIRESFAMEIAELIAQWHSGETLVVEEN	362
St-Hypo	ILAGRRINDN-MGNYVSEQLIKAMIKKGINVEGS---SVLILGFTKENCPIINTRIIDVVKKEGKY-SCKVDYDPEN	356
SA-UDPGDH	IEAIVSSNNV-RKSYIAKQIINVLKEQESPVK-----VVGVYRLIMKNSDNFRESAIKDVIDLKS-KDKIKIYYEM	351
GDPMDH	LGSIMRSNSNQ---VQK-APD-LITSHDTRK-----VGLLGLSEKAGTDDIRESPLVELAEMIG-KGYEPRDFDRN	355
Bov-UDPGDH	WQQVIDMNDYQRRRFASR-IIDSL-FNTVTDK-----KIAILGFAEKDGTGTRSSSIYISKYEMD-EGALHLYDPEK	369
UDPNAMDH	IH-QLPKKLTGLWYSGAA	379
St-Hypo	VDAEEVRRREYGIIPVSEVKSSHIDA-IIVAVGHQQFKQMGSEDIRGFGKDKHVLYDLKYVLPAEQSDVRL	425
SA-UDPGDH	LNKLESEDQ--SVLVNDLENFKKQANIIVTNRYDNELOQVKNKQVSRDIFGRD	402
GDPMDH	VEYARVHGANKKEYIESKIPHVSS--LLVSDLEVVVA--SSDVLVLGNGDELFDVLDVKNKTPSGKKLVDLVGFMPHHTTAQA	431
Bov-UDPGDH	VPREQIVVDLSHPGVSKDDQVAR---LVTISKDPEYACDGAHAVVICTEWDMFKELDYERIHKKMLKPAFIFDGRRLDGC	446
UDPNAMDH		379
St-Hypo		425
SA-UDPGDH		402
GDPMDH	EGICW	436
Bov-UDPGDH	LHNELQTIGFQIETIGKKVSSK	468

Fig. 2. Multiple alignment of bovine liver UDP-glucose dehydrogenase (Bov-UDPGDH) with GDP-mannose dehydrogenase from *Pseudomonas* (GDPMDH), UDP-glucose dehydrogenase from group A streptococci (SA-UDPGDH), UDP-N-acetylmannosaminuronic acid dehydrogenase from *E. coli* (UDPNAMDH), and a hypothetical sequence from *Salmonella* (St-Hypo). Strictly conserved residues are given in reverse-background and identities in 4 of the 5 sequences are hashed. Pairwise identities range between 26 and 34% (Table 1).

acid analysis, these differences are readily reconciled. This segment was also confirmed in CB-257.

Microheterogeneity

At position 74, an Asn/Ser microheterogeneity was evident. Because the 2 enzyme preparations used in this study each came from a single liver, and HPLC-purified peptides from each preparation variously provided unambiguous evidence for each alternative (Fig. 1), 2 interpretations are possible. Either each liver came from a heterozygote expressing both "Asn-74" and "Ser-74" alleles, or a very recent gene duplication has occurred, leading to expression at 2 loci, with the described microheterogeneity reflecting the only mutational event revealed by the protein sequence. Prior data have always suggested that the enzyme is homohexameric (Gainey et al., 1972; Franzen et al., 1980;

Jaenicke et al., 1986). However, in the case of the second scenario, it cannot be excluded that the functional enzyme requires both Asn-74 and Ser-74 subunits. Although such a situation might serve as the basis for the half-of-the-sites reactivity displayed by bovine UDPGDH, Ser and Asn are of generally similar secondary structural potential (preferring reverse turns; Chou & Fasman, 1978), minimally requiring a single nucleotide base change, and substitution of these 2 residues for one another is frequently observed (Johnson & Overington, 1993).

Relationship to other 4-electron-transferring dehydrogenases

The 5 sequences compared in the multiple alignment (Fig. 2), all 4-electron-transferring nucleotide sugar dehydrogenases, are quite evenly diverged from one another as seen from the ~30% iden-

Table 1. Percent pairwise identities between bovine liver UDP-glucose dehydrogenase and related sequences^a

	UDPNAMDH	St-Hypo	SA-UDPGDH	GDPMDH	Bov-UDPGDH
UDPNAMDH	100	34.3	26.9	30.9	29.0
St-Hypo		100	26.4	26.8	25.9
SA-UDPGDH			100	31.6	31.1
GDPMDH				100	29.6
Bov-UDPGDH					100

^a For abbreviations, see Figure 2.

tity between any pair (Table 1). This suggests that all known 4-electron-transferring nucleotide-sugar dehydrogenases, UDPGDH, GDP-mannose dehydrogenase, UDP-*N*-acetyl-glucosamine dehydrogenase, and UDP-*N*-acetyl-mannosamine dehydrogenase, can be expected to share a common ancestor. In contrast, neither of the other 2 NAD⁺-linked, 4-electron-transferring dehydrogenases, histidinol dehydrogenase and β -hydroxy- β -methylglutaryl-CoA reductase, were identified by either pairwise database-searching programs such as FASTA (Pearson & Lipman, 1988) or Profile Analysis (Gribskov et al., 1990), a strategy based on the multiple alignment. It is thus concluded that these enzymes are of independent evolutionary origin. This appears consistent with the recent finding that neither of the 2 conserved cysteine residues among 4 histidinol dehydrogenases are catalytically important (Teng et al., 1993), in contrast to the strictly conserved Cys-275 of UDPGDH, earlier identified as the catalytic cysteine residue of the second half-reaction (Franzen et al., 1981).

Table 2. Amino acid composition of *S*-carboxymethyl bovine liver UDP-glucose dehydrogenase as determined by amino acid hydrolysis and compiled from the determined sequence

	Hydrolytic values	Derived sequence
Cys(Cm)	10.6	12
Asx	48.2	50 ^a
Thr	25.3	26
Ser	28.8	30
Glx	46.5	45 ^b
Pro	18.9	18
Gly	29.7	29
Ala	36.9	37
Val	31.1	36
Met	7.5	8
Ile	35.3	40
Leu	34.6	33
Tyr	14.5	14
Phe	17.6	18
His	10.4	9
Lys	32.0	33
Trp	N.D. ^c	5
Arg	23.2	25
Total	—	468

^a Asn = 21, Asp = 29.

^b Gln = 14, Glu = 31.

^c N.D. = Not determined.

For a number of the 27 strictly conserved residues that emerged from the multiple alignment, important steric and functional roles are suggested. The 2 conserved proline residues at positions 92 and 159 suggest the locations of fixed main-chain bends. Similarly, the 6 consensus glycine residues suggest the location of special main-chain bends. The GIGYvG pattern (positions 11–16 in the present sequence) fits the GXGXXG pattern frequently seen at the coenzyme binding site of pyridine-nucleotide-dependent dehydrogenases. The other 3 consensus glycine residues just precede the strictly conserved Cys-275. Lys-219 and 338 are the only consensus lysine residues. These are (probably coincidentally) positioned nearly equidistantly in sequence from Cys-275, and one is likely the putative catalytic lysine of the first half-reaction (oxidation of UDP-glucose to UDP-aldehydoglucose).

Relationship to alcohol and aldehyde dehydrogenases

Functionally, UDPGDH combines activities similar to alcohol and aldehyde dehydrogenases, and the potential for a structural relationship to one (or both) of these enzymes has long been entertained. Mechanistically, the role of the catalytic cysteine indicated by chemical modification to be involved in oxidation of UDP-aldehydoglucose to UDP-glucuronic acid (Franzen et al., 1981) exactly parallels that of the catalytic cysteine of aldehyde dehydrogenase (reviewed by Pietruszko et al., 1993).

Other (possibly superficial) similarities are also noted. Aldehyde dehydrogenases are notably sensitive to the dithio compounds disulfiram (diethyl dithiocarbamate, Antabuse[®]) and dithiodipyridine (Kitson, 1984), while bovine UDPGDH is inactivated by the homolog of dithiodipyridine, dithiodinicotinic acid (Franzen et al., 1978). These sensitivities suggest the presence of conformationally vicinal cysteine residues because, with both enzymes, reaction with the dithio compound is followed by slower formation of an intramolecular disulfide, evident from biphasic liberation of 2 equivalents of (5-carboxy) thiopyridone after incubation with less than stoichiometric amounts of dithio compound per subunit (Franzen et al., 1978; Kitson, 1984). Disulfiram also induces an intramolecular disulfide (Vallari & Pietruszko, 1982). Further, aldehyde dehydrogenases display half-of-the-sites reactivity with iodoacetamide (Pietruszko et al., 1993), while UDPGDH displays half-of-the-sites reactivity with iodoacetamide and iodoacetic acid (Franzen et al., 1976). In mammalian liver aldehyde dehydrogenases, Cys-302 reacts with iodoacetamide (Hempel et al., 1985) and disulfiram (Kitson et al., 1991) and is the only consensus residue among a large number of distantly related aldehyde dehydrogenase sequences (Hempel et al., 1993).

The first half-reaction catalyzed by UDPGDH, oxidation of UDP-glucose to UDP-aldehydoglucose (which remains enzyme-bound as a Schiff base and subsequently a thioester intermediate in the mechanism proposed by Ordman & Kirkwood [1977]), is reminiscent of alcohol dehydrogenase insofar as the nature of the substrate and product functionalities. UDPGDH is not known to require a metal cofactor, whereas "classical," or "long-chain" alcohol dehydrogenases require a catalytic Zn atom (reviewed by Jörnvall et al., 1991). However, the independently evolved "short-chain" alcohol dehydrogenases (reviewed by Persson et al., 1991) are not metalloenzymes. The nature of their catalytic mechanism is yet unresolved, but it is likely that strictly conserved Lys and Tyr residues are involved (Tyr-152 and, especially, Lys-156 of $3\alpha,20\beta$ -hydroxysteroid dehydrogenase [cf. Ghosh et al., 1991]). Further, based on conserved glycine residues, the coenzyme-binding fold is also near the N-terminus in members of this family (Persson et al., 1991).

Thus, based on the above superficial similarities of coenzyme-binding fold location within the chain and the suggestion of catalytic lysine residues, speculation regarding a relationship between UDPGDH and "short-chain" alcohol dehydrogenases might be made, just as the thiol reagent reactivities of UDPGDH and aldehyde dehydrogenase tempt anticipation of a structural relationship. However, as noted in the Results, neither relationship could be supported either from direct alignments or by Profile Analysis (Gribskov et al., 1990). The only observation we do note in this regard is that a strictly conserved Gly is found at the third position N-terminal to the catalytic Cys in both the present group of sequences (GGxC, Fig. 2) and the aldehyde dehydrogenase family (GqxC; Hempel et al., 1993).

Materials and methods

UDPGDH was prepared according to Franzen et al. (1976), involving heat treatment, ion exchange, and molecular sieve chromatography. The enzyme was reduced with dithiothreitol and S-alkylated using iodoacetic acid, or (in later work) Krutzsch's reagent (*N*-diisopropyl-iodoacetamide) (Krutzsch & Inman, 1993; this reagent was synthesized as described but is now available from Molecular Probes, Eugene, Oregon).

The alkylated protein, ~900- μ g aliquots, was enzymatically digested with either TPCK-treated trypsin, TLCK-treated chymotrypsin (Worthington), or the Glu-specific protease from staphylococcal strain V8 (Miles), at 2% enzyme-to-substrate ratios in 0.1 M ammonium bicarbonate at 37 °C, 24 h, or with endoLys-C (Boehringer) in 25 mM Tris-HCl/1 mM EDTA, pH 8.5, at 37 °C, 18 h. Chemical cleavage with cyanogen bromide, at ~100-fold excess over protein (w/w) was in 70% formic acid at room temperature, 24 h. Selective hydrolysis directed at Asp-Pro linkages (reviewed by Landon, 1977) took place in 70% formic acid, 37 °C, 72–144 h. Some material resulting from this acid cleavage was recleaved either with endoAsp-N (Boehringer) in 50 mM sodium phosphate, pH 8.0, 37 °C, 18 h, or with the V8 protease (described above).

The entire 468-residue sequence was deduced from overlapping peptides. Digests were prefractionated by ion-exchange HPLC on sulfoethyl aspartamide columns (polyLC, the Nest Group) as described (Alpert & Andrews, 1988). Material from pooled fractions was then separated over C-18 reversed-phase columns (Vydac, the Nest Group). After hydrolysis (6 N HCl,

0.5% phenol, 110 °C, 20 h) and amino acid analysis using a Beckman 6300 amino acid analyzer, selected peaks were submitted to sequence analysis using a Porton 2090E gas-liquid phase sequencer with identification of phenylthiohydantoin amino acid derivatives by HPLC. In a few cases, peptides that were needed for exacting confirmation of certain residues were coupled to arylamine-derivatized PVDF membranes (Sequelon-AA[®], Millipore) using 5 μ L 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide, 10 mg/mL in 100 mM *N*-ethylmorpholine, pH 8.0, 20 min, room temperature.

C-terminal analysis was attempted using a "CT-isolator" (Quality Controlled Biochemicals, Hopkinton, Massachusetts). The methodology of this kit takes advantage of free α -amino groups on all peptides after digestion with endoLys-C and of C-terminal-amino groups on all except the carboxy-terminal peptide. Thus, after coupling an endoLys-C digest to a proprietary (presumably diisothiocyanate glass) membrane and treatment with anhydrous TFA, only the C-terminal peptide is released, and all other peptides remain coupled through ϵ -phenylthiocarbonyl linkages.

Similarity searches were conducted using the FASTA program, and a heuristic multiple alignment was generated using the Pileup program, both in the GCG (Genetics Computer Group, Madison, Wisconsin) package, accessed through the VAX front ends at the Pittsburgh Supercomputer Center. A multiple sequence alignment was also generated using the MSA program, which ran on the Cray C90. Profile analyses (Gribskov et al., 1990) based on these alignments were also run on the Cray.

Acknowledgments

This research was partially funded by the American Cancer Society (IN58Y). The sequencer was acquired through BRSG funding (1 S10 RR05644). Access to the databases and alignment programs via the Pittsburgh Supercomputing Center was supported by the NIH Division of Research Resources, cooperative agreement RR06009 to the PSC. J.P. was supported by a summer internship from the Department of Molecular Genetics and Biochemistry. We thank Dr. Hugh Nicholas for his helpful comments.

References

- Alpert AJ, Andrews PC. 1988. Cation-exchange chromatography of peptides on poly(2-sulfoethyl aspartamide)-silica. *J Chromatogr* 443:85–96.
- Chou PY, Fasman GD. 1978. Empirical predictions of protein conformations. *Annu Rev Biochem* 47:251–276.
- Daniels DL, Plunkett G III, Burland VD, Blattner FR. 1992. Analysis of the *E. coli* genome: DNA sequence from 84.5 to 86.5 minutes. *Science* 257:771–778.
- Dougherty BA, van de Rijn I. 1993. Molecular characterization of *hasB* from an operon required for hyaluronic acid synthesis in group A streptococci: Demonstration of UDP-glucose dehydrogenase activity. *J Biol Chem* 268:7118–7124.
- Feingold DS, Franzen JS. 1981. Pyridine nucleotide-linked four-electron transfer dehydrogenases. *Trends Biochem Sci* 6:103–106.
- Flinta C, Persson B, Jörnvall H, von Heijne G. 1986. Sequence determinants of cytosolic N-terminal protein processing. *Eur J Biochem* 154:193–196.
- Franzen B, Carruba C, Feingold DS, Ashcom J, Franzen JS. 1981. Amino acid sequence of the tryptic peptide containing the catalytic-site thiol of bovine UDP-glucose dehydrogenase. *Biochem J* 199:599–602.
- Franzen JS, Ashcom J, Marchetti P, Cardamone JJ, Feingold DS. 1980. Induced vs. preexisting asymmetry models for half-of-the-sites reactivity in bovine liver uridine diphosphoglucose dehydrogenase. *Biochim Biophys Acta* 614:242–255.

- Franzen JS, Ishman R, Feingold DS. 1976. Half-of-the-sites reactivity of bovine liver UDP glucose dehydrogenase toward iodoacetate and iodoacetamide. *Biochemistry* 15:5665-5671.
- Franzen JS, Marchetti P, Ishman R, Ashcom J, Feingold DS. 1978. Half-sites oxidation of bovine liver UDP glucose dehydrogenase. *Biochem J* 173:701-704.
- Gainey PA, Pestell TC, Phelps CF. 1972. A study of the subunit structure and the thiol reactivity of bovine liver UDP-glucose dehydrogenase. *Biochem J* 129:821-830.
- Ghosh D, Weeks CM, Grochulski P, Duax WL, Erman M, Rimsay RL, Orr JC. 1991. Three-dimensional structure of holo $3\alpha,20\beta$ -hydroxysteroid dehydrogenase. *Proc Natl Acad Sci USA* 88:10064-10068.
- Gribskov M, Lüthy R, Eisenberg D. 1990. Profile analysis. *Methods Enzymol* 183:146-159.
- Hempel J, Kaiser R, Jörnval H. 1985. Mitochondrial aldehyde dehydrogenase from human liver. *Eur J Biochem* 153:13-28.
- Hempel J, Nicholas H, Lindahl R. 1993. Aldehyde dehydrogenases: Widespread structural and functional diversity with a shared framework. *Protein Sci* 2:1890-1900.
- Hinshelwood S, Stoker NG. 1992. Cloning of mycobacterial histidine synthesis genes by complementation of a *M. smegmatis* auxotroph. *Mol Microbiol* 6:2887-2895.
- Jaenicke R, Rudolph R, Feingold DS. 1986. Dissociation and reconstitution of bovine liver UDPGDH. *Biochemistry* 25:7283-7287.
- Johnson MS, Overington JP. 1993. A structural basis for sequence comparisons. *J Mol Biol* 233:716-738.
- Jörnval H, Persson B, Krook M, Hempel J. 1991. Alcohol and aldehyde dehydrogenases. In: Palmer TN, ed. *Molecular pathology of alcoholism*. Oxford, UK: Oxford University Press. pp 130-156.
- Kawamura T, Ishimoto N, Ito E. 1979. Enzymatic synthesis of uridine diphosphate-N-acetyl-D-mannosaminuronic acid. *J Biol Chem* 254:8457-8465.
- Kitson TM. 1984. The time course of the interaction of sheep liver cytoplasmic aldehyde dehydrogenase with 2,2'- and 4,4' dithiodipyridine: A comparison with the action of disulfiram. *Arch Biochem Biophys* 234:487-496.
- Kitson TM, Hill JP, Midwinter GG. 1991. Identification of a catalytically essential nucleophile residue in sheep liver cytoplasmic aldehyde dehydrogenase. *Biochem J* 275:207-210.
- Krutzsch HC, Inman JK. 1993. N-isopropyl-iodoacetamide in the reduction and alkylation of proteins: Use in microsequence analysis. *Anal Biochem* 209:109-116.
- Landon M. 1977. Cleavage at aspartyl-prolyl bonds. *Methods Enzymol* 47:145-149.
- Ordman AB, Kirkwood S. 1977. Mechanism of action of UDP-glucose dehydrogenase. *J Biol Chem* 252:1320-1326.
- Pearson WR, Lipman DJ. 1988. Improved tools for biological sequence comparison. *Proc Natl Acad Sci USA* 85:2444-2448.
- Persson B, Krook M, Jörnval H. 1991. Characteristics of short-chain alcohol dehydrogenases and related enzymes. *Eur J Biochem* 200:537-543.
- Pietruszko R, Abriola DP, Blatter EE, Mukerjee N. 1993. Aldehyde dehydrogenase: Aldehyde dehydrogenation and ester hydrolysis. In: Weiner H, Crabb DW, Flynn TG, eds. *Enzymology and molecular biology of carbonyl metabolism 4*. *Adv Exp Med Biol*, vol 328. New York: Plenum Press. pp 221-231.
- Pitsillides AA, Wilkinson LS, Mehdizadeh S, Bayliss MT. 1993. Uridine diphosphoglucose dehydrogenase activity in normal and rheumatoid synovium. *Int J Exp Pathol* 74:27-34.
- Roychoudhury S, May TB, Gill JF, Singh SK, Feingold DS, Chakrabarty AM. 1989. Purification and characterization of GDP-D-mannose dehydrogenase. *J Biol Chem* 264:9380-9385.
- Strominger JL, Kalckar HM, Axelrod J, Maxwell ES. 1954. Enzymatic oxidation of uridine diphosphate glucose to uridine diphosphate glucuronic acid. *J Am Chem Soc* 76:6411-6412.
- Teng H, Segura E, Grubmeyer C. 1993. Conserved cysteine residues of histidinol dehydrogenase are not involved in catalysis. *J Biol Chem* 268:14182-14188.
- Vallari RC, Pietruszko R. 1982. Human aldehyde dehydrogenase: Mechanism of inhibition by disulfiram. *Science* 216:637-639.
- Zalitis J, Feingold DS. 1969. Purification and properties of UDPGDH from beef liver. *Arch Biochem Biophys* 132:457-465.