The 3,6-dideoxyhexoses, usually confined to the cell wall lipopolysaccharide of gram-negative bacteria, are essential to serological specificity and are formed via a complex biosynthetic pathway beginning with CDP-hexose. In particular, the biosynthesis of CDP-ascarylose, one of the naturally occurring 3,6-dideoxyhexoses, consists of five enzymatic steps, with CDP-6-deoxy-Δ3,4-glucose reductase (E₃) participating as the key enzyme in this catalysis. This enzyme has been previously purified from Yersinia pseudotuberculosis by an unusual procedure (protocol I) including a trypsin digestion step (O. Han, V. P. Miller, and H.-W. Liu, J. Biol. Chem. 265:8033-8041, 1990). However, the cloned gene showed disparity with the expected gene characteristics, and upon expression, the resulting gene product exhibited no E₃ activity. These findings strongly suggested that the protein isolated by protocol I may have been misidentified as E₃. A reinvestigation of the purification protocol produced a new and improved procedure (protocol II) consisting of DEAE-Sephacel, phenyl-Sepharose, Cibacron blue A, and Sephadex G-100 chromatography, which efficiently yielded a new homogeneous enzyme composed of a single polypeptide with a molecular weight of 39,000. This highly purified protein had a specific activity nearly 8,000-fold higher than that of cell lysates, and more importantly, the corresponding gene (ascD) was found to be part of the ascarylose biosynthetic cluster. Presented are the identification and confirmation of the E₃ gene through cloning and overexpression and the culminating purification and unambiguous assignment of homogeneous E₃. The nucleotide and translated amino acid sequences of the genuine E₃ are also presented.

The deoxy sugars, long known as an important class of carbohydrate, are found ubiquitously in nature (7, 16, 47). They are formally derived from common sugars by the displacement of one or more hydroxyl groups with hydrogens. Such a substitution generally induces a dramatic alteration of the biological role of the resulting sugar and is responsible for a fundamental change in the metabolism of the product. Particularly notable are the 3,6-dideoxyhexoses found in the lipopolysaccharides (LPS) of gram-negative bacteria (2). Since LPS is the major surface antigen of the gram-negative cell envelope, this class of dideoxyhexose as the nonreducing end group of LPS has been identified as the key antigenic determinant. In addition, they have also been found to contribute to the serological specificity of many immunologically active polysaccharides (5, 24, 33, 46). Inspired by their specific association with LPS and the intriguing nature of their immunological effects, substantial effort has been devoted to exploring their biosynthetic formation (10, 13). However, although the nature of the precursors of these dideoxy sugars has been well defined and possible routes for their formation have been postulated (17, 26), the only pathway that has been studied at the enzymatic level is the biosynthesis of CDP-ascarylose.

As depicted in Fig. 1, the proposed ascarylose biosynthetic sequence starts with the coupling of glucose-1-phosphate (compound 2) and CTP by d-glucose-1-phosphate cytidylyltransferase (E₂) to give CDP-d-glucose (compound 3). This is followed by an intramolecular oxidation-reduction catalyzed by NAD⁺-dependent CDP-d-glucose 4,6-dehydratase (E₄). The resulting product, CDP-6-deoxy-d-glycero-1-threo-4-hexulose (compound 4), is then converted to 3,6-dideoxy-d-glycero-d-gluco-4-hexulose (compound 7) in two consecutive steps mediated by CDP-6-deoxy-d-glycero-1-threo-4-hexulose-3-dehydrase (E₅) and CDP-6-deoxy-Δ3,4-glucose reductase (E₃), both of which have been isolated from Pasturella pseudotuberculosis (11). The final steps, of which little is known, may be catalyzed by an epimerase and a reductase, a situation similar to that found in the biosynthesis of many 6-deoxyhexoses (40). The culminating step of this biosynthetic sequence is the C-3 dehydration catalyzed by E₆, a pyridoxamine 5'-phosphate (PMP)-linked catalyst, and the NAD(P)H-dependent E₇.

While recent studies of the catalytic role of E₇ have revealed that this dehydrase, despite its having evolved a unique role for the PMP cofactor, retains all the essential elements of catalysis common to other vitamin B₆ phosphate-dependent enzymes (34, 38, 44, 45), the detailed mechanism of E₇ has remained largely unresolved. In light of E₇'s dependence on NAD(P)H, this enzyme is expected to deliver a hydrate from its cofactor to reduce the glucoseon intermediate 5. However, incubation with [4-3H]NADPH resulted in no tritium incorporation in product 7 or the regenerated PMP coenzyme, suggesting an indirect hydride transfer or a stereospecific washout. More-
over, both the 4R and 4S hydrogens of NADPH were found to be labile in this reduction step, making E₃ the only enzyme that lacks prochiral recognition of the two diastereotopic methylene hydrogens on the coenzyme (35). Since the net distance for an effective hydride transfer is approximately 0.55 Å (0.055 nm) (14, 48), for E₃ to directly reduce the E₁-bound intermediate 5, the active sites of these two enzymes must thus be brought into close proximity. This scenario, however, invokes a potentially substantial and unfavorable steric interference. It is also important to note that E₃ is capable of oxidizing NAD(P)H in the presence of oxygen (12, 35). However, unlike other known NAD(P)H oxidases, which are flavin dependent, the reported E₃ does not contain any chromophoric group with an absorption above 300 nm. All of these unusual properties of E₃ are not readily explicable and thus challenge our understanding of this important biological reduction.

In an effort to elucidate the mechanism by which the sugar deoxygenation and/or the coenzyme NADH oxidation are affected, we have purified an E₃ equivalent (14, 15), by monitoring its NADH:dichlororphenolindophenol (DCPIP) oxidoreductase activity, from Yersinia pseudotuberculosis. This enzyme showed no obvious absorption above 300 nm and contained no metal constituents, signifying again the absence of any common cofactors as reported earlier by Gonzalez-Porque and Strominger (12). Since the oxygen metabolite was determined to be H₂O₂, the formation of which is stoichiometrically proportional to the amount of NADH consumed, this enzyme-catalyzed NADH oxidation is clearly an overall two-electron redox process. Our studies also revealed O₂⁻ as the hypothetical reducing intermediate, suggesting that H₂O₂ formation is not a direct two-electron reduction of molecular oxygen but is instead a one-electron reduction followed by dismutation of the nascent superoxide. This finding of E₃'s two-electron/one-electron switching capability provides compelling evidence that it operates through a radical mechanism and also suggests the participation of an enzyme-bound organic cofactor which mediates the obligatory two-electron/one-electron conversion. The mechanistic revision of the reduction step from a hydride transfer to an electron relay process not only explains the lack of direct hydride transfer from NAD(P)H in the reduction of the glucose-end-PM complex 5 but also alleviates the spatial and steric constraints imposed on the delivery of reducing equivalents from E₁ to E₃.

There is little doubt that further mechanistic studies of E₃ require, foremost, the resolution of the structure and function of the putative organic coenzyme. Unfortunately, past attempts to isolate this cofactor were impeded by the limited availability of E₃. In our recent effort to develop a large-scale purification scheme for E₃, we were surprised to discover that the previously isolated E₃ was a mixture of two proteins, with the desired enzyme being a minor component (27). In this paper, we report molecular evidence supporting this unexpected finding, the isolation of the genuine E₃ from Y. pseudotuberculosis, and the cloning, sequencing, and expression of the E₃ gene.

MATERIALS AND METHODS

General. Y. pseudotuberculosis was kindly provided by Otto Lüderitz, Max Planck Institute for Immunobiology, Freiburg, Germany. Escherichia coli DH5α and HB101 were obtained from Bethesda Research Laboratories (Gaithersburg, Md.). DEAE-Sephadex, phenyl-Sepharose, and Sephadex G-100 were purchased from Pharmacia (Piscataway, N.J.). Zeta-Probe blotting membranes were products of Bio-Rad (Richmond, Calif.), and Duralose membranes, λ EMBL3 (and accompanying E. coli host strains), and GigaPack II packaging extracts were from Stratagene (La Jolla, Calif.). The Sequenase version 2.0 DNA sequencing kit, M13 sequencing primers, 7-deaza-dGTP sequencing mixtures, all restriction and DNA-modifying enzymes, and pUC plasmids were purchased from United States Biochemical Corp. (Cleveland, Ohio). Nitrocellulose and Elutip-D minicolumns were products of Schleicher & Schuell (Keene, N.H.), electrophoretic reagents were from Beckman Instruments (Fullerton, Calif.), and [γ-³²P]dATP (6,000 Ci/mmol) as well as [α-³²P]dATP (>1,000 Ci/mmol) were purchased from Amersham (Arlington Heights, Ill.). The E₁d and E₁e used in assay procedures were isolated from the same Y. pseudotuberculosis strain (44, 49) and were kindly supplied by Yuan Yu and Theresa Weigel of this research group. The protease inhibitors used in enzyme purification were prepared as previously described (15). All protease inhibitors, molecular weight standards, Cibacron blue 3GA-agarose (blue A), and most biochemicals were purchased from Sigma (St. Louis, Mo.). All other chemicals were of analytical reagent grade or the highest quality commercially available. Methods and protocols for recombinant DNA manipulations are generally referenced by Ausubel et al. (3) and Maniatis et al. (25).

Enzyme E₃ assays. Four different methods were developed to determine the activity of enzyme E₃. Since the substrate of E₃ is not readily available, it has to be prepared in situ from CDP-d-glucose prior to each assay. The preparation involved the incubation of purified CDP-d-glucose 4,6-dehydratase (7 μg) with CDP-d-glucose (0.25 μmol) and NAD⁺ (0.25 μmol) in 110 μl of 10 mM potassium phosphate buffer (pH 7.5) at 37°C for 30 min. The dehydratase product formation was determined by measuring its characteristic absorption at 320 nm (ε, 6,500 M⁻¹ cm⁻¹) under alkaline conditions (31). An aliquot of this solution was then added to a mixture of E₁ and E₃ with the necessary cofactors to determine the sugar reductase activity of E₃.

(i) Method 1: TBA assay. The sugar reductase activity of E₃
was estimated by a procedure analogous to an E₁ assay developed earlier by us (44). A routine assay involved the incubation of appropriate amounts of E₁ and E₂ with an aliquot of the aforementioned dehydrogenase reaction solution (50 μl), PMP (10 mmol), and NADH (100 nmol) in a total volume of 200 μl of 50 mM potassium phosphate buffer (pH 7.5) at 27°C for 1 h. The resulting product, after reduction (NaBH₄, 0.1 mmol, 30 min), hydrolysis (2 N H₃PO₄, 100°C, 5 min), and degradation (0.025 N periodic acid, 55°C, 20 min), was treated with thiobarbituric acid (TBA) reagent (6% in water, pH 2, 100°C, 15 min), resulting in the development of a characteristic pink chromophore (8, 43) with maximum absorption at 532 nm (ε, 159,200 M⁻¹ cm⁻¹) (29).

(ii) Method 2. GC-MS assay. The E₂ product was generated by using an in vitro system similar to that described above for the TBA assay. After being quenched by NaBH₄ (10 μmol, 30 min), the reduced sugar products were boiled for 10 min in the presence of HCl (pH 2.0), neutralized to pH 7.0 with NaOH, and lyophilized to dryness. The solid residue was redissolved in NH₄OH (0.5 M, 600 μl) and treated with a solution of NaBH₄ in dimethyl sulfoxide (20 mg in 1 ml) at 40°C for 1.5 h, and the reaction was quenched with glacial acetic acid. Acetic anhydride (6 ml) and 1-methylimidazole (2.5 μmol) were added, and the resulting mixture was stirred for 1 h at room temperature. Routine workup followed by evaporation in vacuo gave a residue which was directly subjected to gas chromatography-mass spectrometry (GC-MS) analysis (44).

(iii) Method 3. NADΗ:DCPIP oxidoreductase activity. The NADΗ:DCPIP oxidoreductase activity of E₂ was determined, as previously described (15), by measuring the rate of electron transfer from NADΗ to DCPIP (600 nm; ε, 22,000 M⁻¹ cm⁻¹).

(iv) Method 4. NADΗ oxidase activity. The NADΗ oxidase activity of E₂ was measured by monitoring the rate of the reduction of NADΗ absorption at 340 nm (ε, 6,220 M⁻¹ cm⁻¹).

Protein determination. Protein concentration was determined by the method of Lowry et al. (23), using bovine serum albumin as the standard. A2₈₀ₒ was routinely used to monitor column fractions. This method yielded values ca. 10% higher than those obtained by the Lowry assay on identical samples.

Growth of cells. An overnight culture of Y. pseudotuberculosis was grown in tryptic soy broth medium (3%, 9 liters) in an incubator-shaker (Lab-Line) with vigorous agitation (140 rpm) at 28°C. This inoculum culture was then diluted 12-fold into a 110-liter fermentor (Stainless Steel Products) and grown at 28°C, 100 rpm, pH 7.5. The culture was harvested in the early to mid-logarithmic phase by centrifugation. A typical yield was 400 to 500 g (wet weight) of cells per 110 liters of culture. All cell culture and harvesting procedures were performed at the Biological Process Technology Institute, University of Minnesota.

Enzyme purification. All operations were carried out at 4°C. To preclude enzyme inhibition by exogenous trace metals during purification, all buffers contained 1 mM EDTA unless otherwise specified. The results of the purification are summarized in Table 1. It should be noted that enzyme purified by this new protocol (protocol II) is different from the protein isolated by the old procedure (protocol I) which included a trypsin digestion step (15).

(i) Step 1. Crude extracts. Cells from 110 liters of culture (520 g [wet weight]) were suspended in four times their volume (2.1 liters) of 50 mM potassium phosphate buffer (pH 7.5), and then the protease inhibitor solution was added. The cells were disrupted by sonication for 2 min in batches of 200 ml, at 45-s intervals, with a Virsonic model 300 sonicator at 70% output. The temperature of the extracts was carefully controlled so as not to exceed 5°C during this process. Cellular debris was removed by centrifugation (4,420 × g, 20 min). The supernatant solution was diluted with the same buffer to 2.9 liters and was designated the crude extract.

(ii) Step 2. Streptomycin sulfate treatment. Streptomycin sulfate (5% aqueous solution, 550 ml) was added drop by drop to the crude extract to a final concentration of 0.8%. After standing for an additional hour with stirring, the precipitate was eliminated by centrifugation at 14,000 × g for 2 h. The supernatant solution (3.3 liters) was diluted with 1 M potassium phosphate buffer (pH 7.5, 330 ml) and carried on to the next step.

(iii) Step 3. Ammonium sulfate precipitation. Solid ammonium sulfate was slowly added to the protein solution from step 2 to give a final concentration of 65% saturation. After addition was complete, the cloudy solution was stirred for another 10 h. The precipitated proteins were collected by centrifugation (4,400 × g, 20 min) and were redissolved in a minimum amount of 50 mM potassium phosphate buffer (pH 7.5). This solution was dialyzed against 40 liters of the same buffer for 24 h with four changes of buffer.

(iv) Step 4. DEAE-Sephacel column chromatography. The solution from step 3 (1 liter) was applied to a column of DEAE-Sephacel (4.5 by 40 cm) equilibrated with 50 mM potassium phosphate buffer (pH 7.5). The column was washed with the same buffer (1 liter) and then eluted with a linear gradient of potassium phosphate (50 to 200 mM, pH 7.5, 5-liter total). Fractions of 16 ml were collected throughout. The contents of fractions 105 to 125 were pooled and concentrated to 32 ml via an Amicon ultrafiltration unit (PM-30 membrane).

(v) Step 5. Phenyl-Sepharose column chromatography. The enzyme solution from step 4 was adjusted to 0.9 M KCl by the addition of 2.5 M KCl in 50 mM potassium phosphate buffer.

**TABLE 1. Summary of enzyme E₂ purification from Y. pseudotuberculosis by protocol II**

<table>
<thead>
<tr>
<th>Purification step</th>
<th>Total protein (mg)</th>
<th>Total activity (μmol of DCPIP consumed/min)</th>
<th>Sp act (U/mg of protein)</th>
<th>Purification (fold)</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude extracts⁶</td>
<td>170,000</td>
<td>505</td>
<td>&lt;0.003</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Streptomycin sulfate⁴</td>
<td>ND⁵</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Ammonium sulfate</td>
<td>11,460</td>
<td>91.5</td>
<td>0.042</td>
<td>14</td>
<td>95.3</td>
</tr>
<tr>
<td>DEAE-Sephacel</td>
<td>1,056</td>
<td>16.9</td>
<td>0.13</td>
<td>43</td>
<td>26.7</td>
</tr>
<tr>
<td>Phenyl-Sepharose</td>
<td>19.3</td>
<td>4.8</td>
<td>2.0</td>
<td>667</td>
<td>7.5</td>
</tr>
<tr>
<td>Blue A</td>
<td>1.5</td>
<td>16.5</td>
<td>11.1</td>
<td>3,700</td>
<td>3.3</td>
</tr>
<tr>
<td>Sephadex G-100</td>
<td>0.34</td>
<td>8.0</td>
<td>23.9</td>
<td>7,960</td>
<td>1.6</td>
</tr>
</tbody>
</table>

⁴ Obtained from 520 g (wet weight) of cells.
⁵ ND, not determined.

---

**Downloaded from http://jb.asm.org on October 21, 2017 by guest**
(18 ml). This solution was then applied to a phenyl-Sepharose column (2.5 by 45 cm) which was pre-equilibrated with KCl (0.9 M) in potassium phosphate buffer (50 mM, pH 7.5). The column was washed with 0.9 M KCl buffer (90 ml) and then eluted with a linear gradient of 0.9 to 0 M KCl (500 ml) in the same phosphate buffer. Fractions of 6 ml were collected. The contents of fractions 72 to 88 were pooled and concentrated (YM-10 membrane).

(vi) Step 6. Blue A column chromatography. The material from step 5 was loaded onto a column of blue A (2 by 10 cm) and incubated for 45 min. Elution was then started with a linear gradient between 0 and 0.7 M KCl in 50 mM potassium phosphate buffer (140 ml each). Fractions of 2 ml were collected throughout. Active fractions (fractions 38 to 44) were combined and concentrated to 1.6 ml (YM-10 membrane).

(vii) Step 7. Sephadex G-100 column chromatography. The enzyme solution from step 6 was chromatographed on a column (1.5 by 170 cm) of Sephadex G-100 equilibrated with 50 mM potassium phosphate buffer (pH 7.5). The column was then washed with the same buffer. Fractions of 3 ml were collected, and the desired protein was found in fractions 38 to 46. The active fractions were combined, concentrated (YM-10 membrane), and stored at −85°C.

SDS-PAGE. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was used to monitor purification of E$_5$ and estimate the relative molecular mass of the protein. Electrophoresis was carried out at room temperature at a constant 120 V (−30 mA) with a 6% stacking gel and 13% resolving gel in the discontinuous buffer system of Laemmli (21). Prior to electrophoresis, the samples were incubated at 100°C for 10 min in a solution containing 0.5% SDS and 5% 2-mercaptoethanol. Gels were stained with the preparation of Vesterberg (41) and destained with acetic acid-ethanol-water (15:30:65 by volume).

Molecular weight determination. The subunit molecular weight was determined by SDS-PAGE as described by Laemmli (21). The molecular weight of the native enzyme was determined by gel filtration performed on a column of Sephadex G-100 (1.5 by 170 cm). The column was calibrated by separate chromatographic runs with protein standards. The molecular weight of the purified protein was estimated by the method of Andrews (1).

Amino-terminal analysis. The N-terminal sequence was determined by an Applied Biosystem 470A protein sequencer with an on-line 120A HPLC system. Analyses were carried out at the Microchemical Facility in the Institute of Human Genetics of the University of Minnesota, and the results were confirmed by Theodore Thonnhauser at the Baker Laboratory of Chemistry, Cornell University.

Isoelectric focusing. The isoelectric pH of E$_5$ was determined by a Pharmacia Phast-System instrument. A preformed PhastGel IEF 3-9 plate was used as specified by the manufacturer. The protein bands were visualized by silver staining (30).

Oligonucleotide probes. The mixed oligonucleotide probes used in the hybridization, SFL2 and SFL4, were designed on the basis of the respective amino acid sequence near the amino terminus of each protein purified by either the old method (protocol I [15]) or the new procedure described herein (protocol II) while considering the necessary degeneracies in the genetic code. The synthetic probes SFL2 (5'-GC[CGT|GT]G[ATGC]AC[AG]TA[TGCT|TG][AG|TA-3’]), SFL4 (5’-AA[CT]GT[ATGC]AA[AG][CT][ATGC]C[AG|TG]-3’), and JST1 (5’-AA[AG][ACGG][TG][AG|TG][CT][AG|TA-3’], derived from the N-terminal amino acid sequence of the purified E$_{od}$) were labeled to a specific activity of 7 × 10$^8$ dpm/µg with [γ-32P]ATP and T4 polynucleotide kinase prior to hybridization.

Genomic DNA isolation and hybridization. The Y. pseudotuberculosis genomic DNA, isolated by a modified procedure of Ausubel et al. (3), was digested with restriction enzymes (HindIII for YPT1 and BamHI for YPT2), electrophoresed through a 1% agarose gel, and transferred to a Zeta-Probe membrane. Membranes containing blotted DNA were incubated with the labeled DNA probe overnight at 42°C in 1 mM EDTA-7% SDS-5× Denhardt’s solution-0.5 M Na$_2$HPO$_4$ (pH 7.2) containing 0.5 mg of denatured salmon sperm DNA per ml. The membranes were then washed twice for 30 min each time in 2× SSC (1× SSC is 0.15 M sodium chloride plus 0.015 M sodium citrate)-1% SDS at 23°C, washed twice for 30 min each time in 0.1× SSC-0.1% SDS at 42°C, and subjected to autoradiography (−8°C).

Subgenomic library construction. The subgenomic library was constructed by completely digesting the genomic DNA of Y. pseudotuberculosis with HindIII and then size selecting the 3- to 4-kb fragments, which were identified by Southern hybridization with the mixed probe SFL4. After purification and treatment with alkaline phosphatase, the inserts (300 ng) were ligated with HindIII-digested λ ZAP II vector (1 µg) overnight at 4°C. The resulting mixture was packaged by using Gigapack II Plus packaging extracts and plated onto bacterial lawns as instructed by the manufacturer. The resulting plaques were then replicated to nitrocellulose membranes and screened by using labeled SFL4 probe. After transblotting, the filter was washed twice at room temperature (30 min each time) in 1% SDS-2× SSC and twice at 45°C (45 min each time) in 1% SDS-0.1× SSC. The isolated recombinant was then replated, rescreened, and found to be identical to a subgenomic library, pYPT1, constructed on the basis of the hybridization with JST1 derived from E$_{od}$ (27, 40).

Analogously, the subgenomic library YPT2 was constructed by size selection of 15- to 23-kb fragments of BamHI-digested genomic DNA of Y. pseudotuberculosis, which was identified by Southern hybridization with the mixed probe SFL2. The subsequent ligatation with commercially digested BamHI λ EMBL3 arms (1 µg) and packaging by using Gigapack II Plus packaging mix were done as instructed by the manufacturer. The resulting plaques were then replicated to Duralose membranes and screened by using the SFL2 probe. The isolated recombinant was then replated, rescreened, and labeled pYPT2.

Restriction analysis and plasmid construction. Restriction mapping of the cloned DNA insert was accomplished through the analysis of DNA size patterns on 0.8% agarose gels subsequent to single, double, or triple digestion with various restriction endonucleases. Comparison of patterns with known reference points within the vector DNA allowed for the construction of a linear map of the restriction sites within the cloned insert. The relative position of the β-gene (lacZ) within the cloned insert was determined by Southern blotting with radiolabeled SFL4 and DNA sequencing. The expression construct pSFL28 was built by directionally subcloning the insert of pYPT1 into the BamHI-SacI sites of pUC19. Standard recombinant DNA techniques were used for all plasmid constructions.

DNA sequencing. Plasmids were isolated from E. coli by the alkali method (4). Nucleotide sequences were determined directly from double-stranded templates by the dideoxynucleotide chain termination method of Sanger et al. (36), using Sequenase version 2.0. The nucleotide analog 7-deaza-dGTP was substituted for dGTP in all sequencing manipulations. Sequencing was completed by using nested deletions (18) of pSFL28 generated by exonuclease III, allowing the use of both
M13 forward and reverse universal sequencing primers as well as designed synthetic primers to complete desired regions. Computer analysis of the resulting sequence information was performed with Intelligenetics software (release 5.4).

**Nucleotide sequence accession number.** The ascD sequence will appear in the EMBL/Genbank/DDJB nucleotide sequence data libraries under accession number L25594.

**RESULTS**

Cloning and overexpression of the gene coding for the protocol I-purified E3. Purification of E3 by a previously reported procedure (protocol I) yielded amounts of protein sufficient only for preliminary mechanistic studies (15). To facilitate further characterization of this enzyme, cloning and overexpression of the particular gene were pursued. Oligonucleotide (SFL2) screening of 4,000 plaques of YPT2 resulted in the isolation of 65 positive clones (1.63%), and analysis of 5 randomly selected positive plaques revealed identical restriction fragment patterns of the inserted DNA. Further Southern blotting of this DNA led to the isolation of a 3.2-kb KpnI fragment containing the gene coding for the putative E3 which was subsequently ligated into pUC18 (pSFL31) for exonuclease III digestion. Ligation of these deleted pSFL31 inserts into pUC19 followed by transformation and SDS-PAGE analysis revealed one particular construct, pYL36, which readily overexpressed the desired gene product. However, the expressed protein showed no E3 activity by either the GC-MS or NADH:DCPIP oxidoreductase assay. The negative results in the assay for E3 activity from the overexpressed E3 gene clone suggested that the protein isolated by protocol I may have been misidentified as E3.

**Enzyme purification.** Although E3 had been previously isolated by an unusual purification sequence (protocol I), the cloning results reported above strongly suggested that the resulting enzyme, which deceivingly appeared homogeneous at first, was still a mixture of two proteins in which E3 was only a minor component. To further purify the desired enzyme, chromatography on several affinity reagents that often bind to nicotinamide-utilizing enzymes, 2',5'-ADP-Sepharose (39) and NAD⁺-agarose (22), was attempted, albeit with no success. Fortunately, it was found that the dye ligand blue A successfully bound and selectively eluted E3. Thus, a blue A column following phenyl-Sepharose was incorporated in the purification sequence. Such a modification rendered the tryptic digestion step unnecessary and, in effect, greatly simplified the purification procedure. As shown in Table 1 and Fig. 2, purification by blue A affinity chromatography followed by gel filtration effected a 7,960-fold enrichment of E3, which is now truly homogeneous. In retrospect, a close examination by SDS-PAGE of the protein purified by the earlier protocol (15) occasionally revealed the presence of two protein bands differing by ~1 to 2 kDa (Fig. 2, lane F). This phenomenon was, however, inconsistent among preparations and was most often barely discernible. Chromatography on blue A apparently removed the major contaminant that was previously misidentified as E3.

**Properties of E3.** (i) Molecular weight. The molecular weight of the newly purified and homogeneous E3 was estimated by gel filtration to be 39,000. SDS-PAGE also showed a single band with a molecular weight of 39,000. Thus, the native enzyme is a monomeric protein consisting of a single polypeptide chain.

(ii) Isoelectric point. Purified E3 was subjected to electrophoresis on an analytical isoelectric focusing gel as described in Materials and Methods. The pI of E3 was determined to be 4.7.

**(iii) UV-visible spectrum.** The electronic spectrum of the newly purified E3 has significant adsorption above 300 nm, suggesting the existence of a flavin and possibly an iron-sulfur center.

(iv) Substrate specificity. The competence of NADPH as an alternate substrate for E3 was tested by replacing NADPH with NADH in the NADH:DCPIP oxidoreductase activity assay. The rate of the latter was less than 5% of that occurring when NADPH was used under identical conditions, indicating that the E3 enzyme preferentially utilizes NADH as a substrate.

(v) Kinetic parameters for the NADH:DCPIP oxidoreductase activity. As an NADH:DCPIP oxidoreductase, E3 displayed normal Michaelis-Menten saturation kinetics. The $K_m$ of 37 μM for NADH and $V_{\text{max}}$ of 47 μmol min⁻¹ mg⁻¹ for this catalysis were determined by plotting the data according to the method of Lineweaver and Burk.

(vi) Alternate electron acceptors. Enzyme E3 can utilize a variety of alternate electron acceptors for the oxidation of NADH. As shown in Table 2, potassium ferricyanide, ferriochrome c, and DCPIP were the most efficient electron acceptors tested. Interestingly, the quinones menadione and ubiquinone-0 are also good electron acceptors. Flavins, riboflavin, flavin mononucleotide, and flavin adenine dinucleotide are poor electron acceptors, suggesting that if E3 contains a flavin cofactor, it is tightly bound and virtually remains in the active site throughout the purification procedure.

**Gene expression and protein purification of recombinant E3.** A previously produced clone pYPT1 was found to contain the genuine E3 gene (ascD), as judged from its consistent hybridization with SFLA. The relative position of the E3 gene within the insert of pYPT1 was located by Southern blotting and DNA sequencing. The initiation codon of the ascD gene is only 270 bp away from the vector sequence of pYPT1, and the insert has an opposite orientation to the lac promoter of the vector. To match the orientation of the ascD gene to the direction of transcription of the lac promoter, the ascD gene was inverted by directionally subcloning the insert of pYPT1 into the BamHI-SacI sites of pUC19 (Fig. 3). The resulting construct (pSFL28), after transformation into E. coli DH5α, readily expressed the desired gene product, as observed by SDS-PAGE. The recombinant E3 expressed in E. coli, under
TABLE 2. Utilization of alternate electron acceptors by homogeneous E3 during NADH oxidation

<table>
<thead>
<tr>
<th>Added electron acceptor</th>
<th>Conc (µM)</th>
<th>Wavelength observed (nm)</th>
<th>Redox potential (E^\prime_{\text{red}}) (mV)</th>
<th>Relative velocity (µmol of 2e(^-)/min/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O(_2)</td>
<td>1,300(^a)</td>
<td>340(^b)</td>
<td>820(^b)</td>
<td>1</td>
</tr>
<tr>
<td>K(_2)Fe(CN)(_6)</td>
<td>50</td>
<td>340</td>
<td>420(^b)</td>
<td>41</td>
</tr>
<tr>
<td>Cytochrome c</td>
<td>50</td>
<td>550</td>
<td>250</td>
<td>84</td>
</tr>
<tr>
<td>DCPIP</td>
<td>50</td>
<td>600</td>
<td>220</td>
<td>84</td>
</tr>
<tr>
<td>Menadione</td>
<td>50</td>
<td>340</td>
<td>0.00(^h)</td>
<td>59</td>
</tr>
<tr>
<td>Ubiquinone-0</td>
<td>50</td>
<td>340</td>
<td>100</td>
<td>52</td>
</tr>
<tr>
<td>Ascorbic acid</td>
<td>530</td>
<td>340</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>Riboflavin</td>
<td>50</td>
<td>340</td>
<td>-200</td>
<td>26</td>
</tr>
<tr>
<td>Flavin adenine dinucleotide</td>
<td>50</td>
<td>340</td>
<td>-180</td>
<td>9</td>
</tr>
<tr>
<td>Flavin mononucleotide</td>
<td>50</td>
<td>340</td>
<td>-220</td>
<td>9</td>
</tr>
<tr>
<td>Glutathione</td>
<td>530</td>
<td>340</td>
<td>-230</td>
<td>0</td>
</tr>
<tr>
<td>Methylviologen</td>
<td>40</td>
<td>601</td>
<td>-550</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^a\) A typical assay contained 44 pmol of the purified enzyme, 100 nmol of NADH, and the indicated amount of added electron acceptor in 1.5 ml of 50 mM potassium phosphate buffer (pH 7.5).
\(^b\) Taken from Walsh (42) unless specified otherwise.
\(^c\) See Materials and Methods for assay conditions.
\(^d\) Buffer saturated with O\(_2\) at 30°C was used. The value was recalculated from Sendroy et al. (37).
\(^e\) Reaction was monitored by the consumption of NADH.
\(^f\) Potential corresponds to the reduction of O\(_2\) to H\(_2\)O\(_2\).
\(^g\) Taken from O'Reilly (32).
\(^h\) Taken from Dawson et al. (9).

control of the lac promoter, is produced at levels nearly 30 times higher than those in Y. pseudotuberculosis, as deduced from amounts of purified proteins. A summary of the five-step purification for the recombinant protein is presented in Table 3. This simplified procedure typically yields >10 mg of purified protein from 6 liters of cell culture of the E. coli DH5\(\alpha\) (pSFL28) system.

Properties of the overexpressed E3. Characterization of the overexpressed E3 from E. coli revealed properties identical to those of the wild-type enzyme purified from Y. pseudotuberculosis.

Nucleotide and amino acid sequences of E3. The nucleotide sequence of the E3 gene (ascD) and the deduced amino acid sequence of the corresponding protein are presented in Fig. 4. The DNA sequence predicts a protein of 328 amino acids with a molecular weight of 36,160. This predicted molecular weight is in agreement with the molecular weight of 39,000 determined by SDS-PAGE and gel filtration. While translation is expected to start from ATG, the mature protein must have been posttranslationally modified, since a serine residue instead of a methionine was found as the N-terminal amino acid of the purified E3. Although expression of the protein is under the control of the lac promoter in E. coli (DH5\(\alpha\)), analysis of the ascD gene revealed a sequence, 5'-TAATAT-3', which resembles the E. coli consensus Pribnow box (5'-TATAAT-3') and is positioned 16 to 21 bp upstream from the translational start codon. A sequence resembling the -35 motif of consensus E. coli promoters (5'-TTGACA-3') could also be identified about 20 bp further upstream (5'-TTGAGT-3'). An inverted repeat sequence, located immediately after the stop codon, could act as a transcriptional termination signal or perhaps an mRNA stabilizer. The latter appears to be more likely because of the presence of another open reading frame (ORF) beginning 10 to 20 bp from the stop codon of the ascD gene.

DISCUSSION

Although E3 has been isolated by Gonzalez-Porque and Strominger from P. pseudotuberculosis (12), its purification has always been complicated by the lack of a facile and sensitive assay to determine its activity. To circumvent this problem, we have developed a few convenient and sensitive methods to assay E3 activity. Among these, the particular method exploiting E3's capability to utilize DCPIP as an electron acceptor for the oxidation of NADH affords the simplest means of assaying this enzyme's activity. Purification of E3, described herein relied upon the convenient NADH:DCPIP oxidoreductase assay. However, the catalytic role of the isolated enzyme was further analyzed by two additional methods which assay the sugar reductase activity by monitoring the product formation directly. While the TBA assay is very sensitive in detecting the E1-E3 product, it will also respond to 2- and 3-deoxyhexoses. Hence, a definitive verification involves conversion of the E1-E3 product (product 7) to the corresponding alditol acetates, which can be separated and identified by the GC-MS assay. Because product formation by incubation with purified E3 and E1 was confirmed by this assay, enzyme E3 purified on the basis of the NADH: DCPIP oxidoreductase activity was unequivocally demonstrated to be the desired sugar reductase.

As previously reported, initial E3 purification efforts in our group began with the procedure of Gonzalez-Porque and Strominger (12), but this was soon modified to a procedure (protocol I) which led to a 3,600-fold overall purification of E3 (15). While the protein purified by protocol I was valuable for preliminary mechanistic studies, characterization of possible cofactor(s) requires substantial amounts of homogeneous protein which was unattainable by the original protocol. Cloning of the gene coding for this protein was achieved by standard methodology. Although the nucleotide-derived gene product size (39,000 Da) closely matched that found for the putative E3.
TABLE 3. Summary of E₃ purification from E. coli DH5α(pSFL28)

<table>
<thead>
<tr>
<th>Purification step</th>
<th>Total protein (mg)</th>
<th>Total activity (µmol of DCPIP consumed/min)</th>
<th>Sp act (U/mg of protein)</th>
<th>Purification (fold)</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonium sulfate*</td>
<td>5,549</td>
<td>1,332</td>
<td>0.24</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>DEAE-Sephacel</td>
<td>295.9</td>
<td>1,391</td>
<td>4.7</td>
<td>19.6</td>
<td>104</td>
</tr>
<tr>
<td>Phenyl-Sepharose</td>
<td>35.1</td>
<td>636</td>
<td>18.1</td>
<td>75.4</td>
<td>47.7</td>
</tr>
<tr>
<td>Matrex Blue A</td>
<td>24.9</td>
<td>443</td>
<td>17.8</td>
<td>74.2</td>
<td>33.3</td>
</tr>
<tr>
<td>Sephadex G-100</td>
<td>13.2</td>
<td>428</td>
<td>32.4</td>
<td>135</td>
<td>32.1</td>
</tr>
</tbody>
</table>

* Obtained from 33 g (wet weight) of cells.

(41,000 Da by SDS-PAGE and gel filtration), inconsistencies began to accumulate upon closer examination. First, with the program FASTDB (6), a comparison of the nucleotide-derived amino acid sequence of the putative E₃ with all sequences available in protein data bases failed to reveal significant homology with any NAD(P)H oxidases or other similar enzymes. Second, a detailed search of the putative E₃ gene failed to give any of the common NAD(P)H binding consensus sequences typically found in NAD(P)H-utilizing enzymes. Third, since the E₃ gene was expected to be within close proximity of the E₃ start codon, positive hybridization with the isolated YPT2 insert was expected with the E₃ probe (JST1);

however, despite many attempts, no hybridization was observed. Finally, definitive proof of misidentification was found by overexpression of the putative E₃ gene product and subsequent assay for activity. No activity was evident by either the GC-MS or NADH:DCPIP oxidoreductase assay. Thus, the protein isolated by protocol I is not E₃, and the desired E₃ may actually be a minor contaminant of the apparent homogeneous protein obtained from protocol I. This information prompted a thorough reinvestigation of the existing E₃ purification protocol.

Modification of the former purification procedure by incorporation of a dye ligand chromatography (protocol II) surpris-
ingly resulted in the isolation of a new protein. This 8,000-fold-purified enzyme is a monomeric protein with a molecular weight of 39,000 and, most importantly, has a higher specific activity. It was suspected, therefore, that this newly purified enzyme was the genuine E₃. To confirm this conclusion, the N-terminal amino acid sequence of the protein purified by protocol II was used to design an oligonucleotide primer (SFL4) which was used to screen and construct a clone that was found to be identical to a previously constructed clone, pYPT1, known to carry the entire genes for Eᵣ and Eₒ. The fact that the gene coding for this newly purified protein is part of the ascarylase gene cluster provides convincing evidence that the enzyme purified by protocol II is the genuine E₃. This conclusion was further supported by the fact that the catalytic properties of the ascD gene product purified from E. coli are identical to those of newly purified wild-type E₃. Interestingly, high sequence homology (51%) was found between ascD and a previously unidentified gene (ORF 7.6) within the abequose biosynthetic cluster (rfb) of Salmonella typhimurium (20). Since compound 5 has been suggested as the common intermediate for the biosynthesis of most 3,6-dideoxyhexoses (except for colitose), identification of the gene function of ascD has therefore also allowed the assignment of ORF 7.6 in the Salmonella O antigen biosynthetic cluster as the corresponding E₃ in the abequose biosynthetic pathway (40).

As an NADH:DCPIP oxidoreductase, this newly purified enzyme exhibits a strong preference for NADH over NADPH, and its calculated Kₘ and Vₘₐₓ are substantially higher than those previously reported. In contrast to the protein purified by protocol I, the genuine E₃ shows strong absorption above 300 nm, which is characteristic of a flavoprotein. Thus, the above-baseline absorption at higher wavelength observed for the previously purified enzyme is not due to contaminants but rather is an intrinsic property of E₃ and suggests the existence of a flavin and/or other cofactors. Apparently, earlier conceived notions on the mystery of E₃ catalysis can now be attributed to the complexity of its purification arising mainly from a contaminant whose molecular weight and chromatographic behaviors are similar to those of the genuine E₃. With the help of modern biochemical and genetic techniques, we have finally obtained the correct homogenous enzyme, the purification of which had proved elusive for nearly two decades. Now that the pure enzyme is at hand, we can commence work on the mechanistic details of E₃ catalysis.

ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health grant GM 35906 (to H.-W.L.). H.-W.L. is the recipient of National Institutes of Health Research Career Development Award GM 00559. We are indebted to Otto Lüderitz of the Max Planck Institute for the gift of Y. pseudotuberculosis. We are also grateful to Yuan Yu for CDP-6-deoxy-β-pyrophosphorylase and Theresa Weigl for Eₒ. A special note of thanks is extended to Peter Donahue and Yang-nan Liu at the Biomedical Research Institute, St. Paul Children’s Hospital, for their instructive assistance in gene cloning and DNA sequencing.

REFERENCES

27. Miller, V. P., and H.-W. Liu. 1992. Biosynthesis of 3,6-dideoxyhex-


