Ribitol dehydrogenase of *Klebsiella aerogenes*

Sequence and properties of wild-type and mutant strains

John M. DOTHIE, Jose R. GIGLIO,* Christopher B. MOORE,† Susan S. TAYLOR‡ and
Brian S. HARTLEY§

Department of Biochemistry, Imperial College of Science and Technology, London SW7 2AZ, U.K.

(Received 11 December 1984/15 March 1985; accepted 30 May 1985)

Evidence is presented for the sequence of 249 amino acids in ribitol dehydrogenase-A from *Klebsiella aerogenes*. Continuous culture on xylitol yields strains that superproduce 'wild-type' enzyme but mutations appear to have arisen in this process. Other strains selected by such continuous culture produce enzymes with increased specific activity for xylitol but without loss of ribitol activity. One such enzyme, ribitol dehydrogenase-D, has Pro-196 for Gly-196. Another, ribitol dehydrogenase-B, has a different mutation.

These sequence studies of ribitol dehydrogenase (EC 1.1.1.56) from wild type (RDH-A) and mutant strains (RDH-B, RDH-D) of *K. aerogenes* were carried out in 1970–1976 as part of research on directed evolution of this enzyme towards a novel specificity by continuous culture on a poor substrate, xylitol. Strain A is constitutive for RDH-A synthesis (the wild type is inducible by ribitol) and grows on xylitol by virtue of a weak side specificity of ribitol dehydrogenase which converts this substrate to xylulose (metabolized further by D-xylulokinase). Mutants with increased 'xylitol dehydrogenase' activity grow faster and therefore take over the culture: Fig. 1 shows their genealogy. There are two classes: strains which successively produce larger amounts of the wild-type enzyme (A1, A11 etc.) or strains that produce altered ribitol dehydrogenases with increased activity for xylitol (B, D, etc.).

The early sequence work on RDH-A (Taylor *et al.*, 1974) was with a preparation from strain A, but subsequent work was with enzyme from superproducing strains (Burleigh *et al.*, 1974). The sequence of RDH-D was determined with enzyme from a superproducing strain D1 and partial sequences of RDH-B from strain B (Burleigh *et al.*, 1974). The

substantive conclusions have been summarized previously (Morris *et al.*, 1974; Hartley *et al.*, 1976) but the detailed evidence becomes of interest in view of differences between the protein sequence and that derived from the DNA sequence of the gene (Loviny *et al.*, 1985).

Materials and methods

Microbial strains

K. aerogenes strain A is a spontaneous prototrophic revertant derived by Rigby *et al.* (1974) from strain X1 (*arg*, *gua*, constitutive for synthesis of RDH-A) of Wu *et al.* (1968), which was itself derived from the wild-type *K. aerogenes* 1033 of Neidhart & Magasanik (1956).

Strains A11, A211 and A611 were selected by continuous culture on xylitol and superproduce enzyme which is indistinguishable from wild type (Rigby *et al.*, 1974).

Strain B is a spontaneous prototrophic revertant of strain X2 (*arg*, *gua*, constitutive for synthesis of RDH-B) derived by Wu *et al.* (1968) by nitrosoguanidine mutagenesis of their strain X1. Strain D arose in continuous culture on xylitol following nitrosoguanidine mutagenesis of strain A. A spontaneous mutation (D1) which superproduces the altered enzyme was selected by further continuous culture on xylitol (Hartley *et al.*, 1976).

Reagents

Trypsin (Worthington Biochemical Corp.) was treated with diphenylcarbamoyl chloride (Smyth, 1967) before use or purified by affinity chromatography on a column of soybean trypsin inhibitor

* Present address: Faculdade de Medicina, Universidade de Sao Paulo, Ribeirao Preto, Brazil.

† Present address: Department of Biochemistry, Massey University, Palmerston North, New Zealand.

‡ Present address: Department of Chemistry, University of California San Diego, La Jolla, CA 92037, U.S.A.

§ To whom reprint requests should be addressed.

Abbreviation used: RDH, ribitol dehydrogenase.

Fig. 1. Genealogy of mutants of *K. aerogenes* with increased xylitol dehydrogenase activity

Each event represents a takeover during continuous chemostat culture on xylitol. Strain A is constitutive for synthesis of wild-type ribitol dehydrogenase (RDH-A). Strains A1, A11, A111 etc. represent consecutive events in the same chemostat. Strains A11, A12 and A13 represent separate chemostat cultures of strain A. Strains A1-A111, A2-A2111, etc., produce successively increasing amounts of unchanged RDH-A. Strains D, E, etc. produce altered enzymes with increased specific activity for xylitol. Strains D1, H1, etc., superproduce RDH-D, RDH-H, etc. Mutagenesis by u.v. irradiation *in situ* (u.v.) or by *N*-methyl-*N*-nitro-*N*-nitrosoguanidine (N) was used where indicated. Strain B* was obtained by selection on xylitol plates following chemical mutagenesis (N) by Wu *et al.* (1968).

bound to agarose (Porath & Sundberg, 1971). Chymotrypsin (Worthington), pepsin (Sigma), thermolysin (Calbiochem) and elastase (Whatman Biochemicals) were used without further purification. Iodo[1-¹⁴C]acetic acid and randomly ¹⁴C-labelled amino acid mixture were obtained from The Radiochemical Centre, Amersham, Bucks., U.K.

Purification of ribitol dehydrogenase

Ribitol dehydrogenase-A was purified from strains A, A11, A211 or A611 according to Taylor *et al.* (1974) and ribitol dehydrogenase-B was the sample described by Burleigh *et al.* (1974).

Ribitol dehydrogenase-D was prepared from eight 10-litre cultures of strain D1 according to Taylor *et al.* (1974). All buffers contained 0.1 mM-phenylmethanesulphonyl fluoride to minimize proteolytic cleavage. The fraction from the DE-52 gradient column had a specific activity of 89.4 units/mg, gave a single band on polyacrylamide-gel electrophoresis and was obtained in 21% yield (2.5g from 1 kg of wet cells), indicating that this enzyme represents 23% of the soluble protein in the extract.

Radioactively labelled ribitol dehydrogenase-A was prepared for mutant screening purposes from a 400 ml culture of strain A7 grown to stationary phase on 2% xylitol/M9 medium containing 2.75 mCi of ¹⁴C-labelled amino acids (the effluent ¹⁴CO₂ was adsorbed in three Dreschel bottles containing 2M-NaOH). The cells were spun down

at 2000g for 10 min, washed, resuspended in 40 ml of 0.1M-potassium phosphate, pH7, containing 0.1 mM-phenylmethanesulphonyl fluoride and disrupted in a French pressure cell at 500 kg/cm². Purification according to Taylor *et al.* (1974) yielded 13.5 mg of [¹⁴C]RDH-A containing 8 × 10⁹ c.p.m. (4% of the radioactivity incorporated into the cells) with specific enzyme activity of 81 units/mg and giving a single band on polyacrylamide-gel electrophoresis.

Protein chemistry techniques

These are according to Shotton & Hartley (1973) except as specified. Details of peptide properties, composition and sequence evidence are deposited with the British Library Lending Division, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K. from whom copies can be obtained as described in *Biochem. J.* (1985) **225**, 5, by quoting reference SUP 50128 (60 pages). Mobilities on paper electrophoresis at pH 6.3 (*m*) or pH 2 (*n*) are used to calculate net charge according to Offord (1966).

Trypsin digest of [¹⁴C]carboxymethyl-enzyme. Salt-free [¹⁴C]carboxymethylated enzyme (10 μmol in 60 ml) was digested with trypsin (2 mg) for 6 h at pH 8.1 in a pH-stat and then adjusted to pH 4.5 with 2M-acetic acid. The precipitate was centrifuged and the supernatant freeze-dried.

Trypsin digest of maleylated carboxymethylated enzyme. Carboxymethylated enzyme (10 μmol) from strain A11 was maleylated according to Butler & Hartley (1972), desalted on Sephadex G-

25 in 50mM-NH₄HCO₃ and digested overnight with trypsin (0.1 μmol) in 20ml total volume.

Isolation of N-terminal peptide (Bruton & Hartley, 1970). Maleylated carboxymethylated enzyme (0.5 μmol) was digested for 90 min with thermolysin (0.5 mg) in 10 ml of 0.1 M-NH₄HCO₃, freeze-dried, and dissolved in 2 ml of cold 0.1 M-acetic acid. The digest was separated on a column of Dowex 50 (20 cm × 1 cm diam.) in 0.1 M-acetic acid at 4°C. The breakthrough fraction (50 ml) was neutralized with pyridine, concentrated by rotary evaporation and separated by electrophoresis at pH 6.5 as a 25 cm band on Whatman 3 MM paper. The maleylated N-terminal peptide was located as a ninhydrin-negative, Cl₂/toluidine-positive band, demaleylated on paper in the vapour of pyridine acetate, pH 3.5, at 60°C for 12 h, stitched to Whatman 3 MM paper and isolated by electrophoresis at pH 6.5.

Carboxypeptidase A digestion. Carboxymethylated enzyme (560 nmol) + norleucine (25 nmol) as internal standard was digested with carboxypeptidase A (100 μg) in 5.5 ml of 0.2 M-N-ethylmorpholineacetic acid, pH 8.5. Samples (0.5 ml) were removed at various times, adjusted to pH 2.2 and freeze-dried prior to amino acid analysis.

CNBr cleavage. Carboxymethylated enzyme from strain A211 (10 μmol) and CNBr (2 mmol) were reacted in 15 ml of 70% (v/v) formic acid for 12 h at room temperature. After dilution with water and rotary evaporation (repeated four times) the digest (6 ml) was made 8 M in urea and separated on a Sephadex G-50 (superfine) column (85 cm × 2.1 cm diam.) by upward flow with 8 M-urea/0.1 M-acetic acid. The breakthrough was re-run on Sephadex G-100 (100 cm × 2.5 cm diam.) in the same solvent.

Experimental and results

Sequence of ribitol dehydrogenase-A

N-Terminal and C-terminal sequences. The N-terminal amino acid was found to be methionine by direct dansylation. The N-terminal sequence was established by isolating the peptide with a maleylated N-terminal residue from a thermolysin digest of maleylated carboxymethylated enzyme by the method of Bruton & Hartley (1970). The peptide, when unblocked, was basic with a mobility (*m*) of 0.57 at pH 6.5 and gave the sequence Met-Lys-His-Ser by the dansyl/Edman technique.

The C-terminal sequence was established by measuring the rate of release of amino acids from carboxymethylated enzyme by carboxypeptidase A. Leucine was released rapidly followed by aspartic acid and valine at equivalent rates: serine was released more slowly. The known specificity of

carboxypeptidase A therefore suggests that the enzyme has the C-terminal sequence: -Ser-Val-Asp-Leu.

The presence of peptide T2' (residues 5–15) in variable yield in tryptic digests of RDH-A (see below) correlates with the appearance of 'triplet' bands of the native protein on gel electrophoresis (Burleigh *et al.*, 1974). These suggest proteolysis during purification of the protein since (a) the 'missing' peptide (residues 1–4) was never observed in a tryptic digest, (b) mild subtilisin treatment of samples of native enzyme giving a single band yielded a similar triplet. But since the enzyme is a tetramer, one might expect to generate five electrophoretic bands by such proteolytic nicking. This suggested that only two N-termini might be exposed in the tetramer, so we tested the accessibility of the C-terminal leucine in the native enzyme to carboxypeptidase A. In a timecourse, exactly and only 2 mol of leucine were released/tetramer. Moreover only 2 thiol groups/tetramer are accessible to disulphide exchange (Burleigh *et al.*, 1974). This suggests that the quaternary structure is a 'dimer of dimers' with two N-termini, two C-termini and two thiol groups exposed.

Soluble tryptic peptides from carboxymethylated RDH-A (strain A). A digest of carboxymethylated enzyme (strain A) (160 mg) with diphenylcarbamoylated trypsin (0.6 mg) for 4 h at 37°C, pH 8.3, was separated on Sephadex G-50. The breakthrough fraction was redigested with trypsin and the combined large peptides (Fraction I) or small peptides (Fraction II) were separated by paper electrophoresis and paper chromatography.

Fraction I contained five essentially homogeneous peptides corresponding to AT3, AT5-6-7 (residues 41–53), AT9, AT17 and AT22 in Fig. 2.

Fraction II yielded 18 smaller peptides: AT1, AT2, AT4, AT4' (residues 35–40), AT5, AT5-6 (residues 41–50), AT6, AT6-7 (residues 47–53), AT7, AT11, AT12, AT14, AT15 (Arg), AT16, AT18, AT20 and AT21.

The composition and sequence of all these peptides was firmly established by the dansyl/Edman method, except for the uncertainties indicated in Fig. 2 due to problem amino acids or blocked N-terminus (peptide AT16). Subdigestion by thermolysis of peptide AT3 (giving residues 26–32) and of AT17 (giving residues 185–197 and 197–203) was used to confirm the sequence shown in Fig. 2. Amide groups were assigned unambiguously from the electrophoretic properties except for peptides AT14 (which remained at the origin on paper electrophoresis at pH 6.5) and AT22 (where two Asp and one Asn were indicated).

Soluble chymotryptic peptides from carboxymethylated RDH-A (Strain A). Protein A (160 mg) in 100 ml of 10mM-NH₄HCO₃ was digested for 2 h at

Fig. 2. Evidence for the sequence of ribitol dehydrogenase-A

Peptides AT1, AC1, etc. were from tryptic digests of enzyme from strain A, and peptides T1, etc. and TC26c from strain A11. CNBr fragments (CNBr1, etc.) were from strain A211, and elastase peptides E1, etc. from strain A611. Full lines show complete sequence by dansyl/Edman analysis of the peptide and/or subfragments thereof, and dots show residues inferred from compositions only.

37°C with 0.8 mg of chymotrypsin, freeze-dried and dissolved in 16 ml of 10 mM-acetic acid/8 M-urea. After centrifuging to remove traces of insoluble material, the sample was applied to a column of Sephadex G-50 (78 cm × 2.5 cm diam.) equilibrated and eluted with 10 mM-acetic acid. Paper electrophoresis of test samples guided pooling into four fractions which were freeze-dried and fractionated by high voltage paper electrophoresis and chromatography.

Fig. 2 shows the sequence of those peptides that were purified to a reasonable degree of homogeneity. Peptide AC2 was a mixture of species with and without the *N*-terminal leucine. The *C*-terminal leucine in peptide AC3 was revealed by dansylation of a tryptic digest. The electrophoretic

mobilities confirmed the amide assignments except within peptide AC5, where one amide group must be present. It is noteworthy that there was extensive cleavage at histidine residues in this digest as well as at some asparagine and threonine residues.

Tryptic peptides from carboxymethylated RDH-A (strain A211). The salt-free protein (10 μmol) in 60 ml was digested for 6 h with 2 mg of trypsin in a pH-stat at pH 8.1, 22°C and then adjusted with HCl to pH 3. The acid-insoluble precipitate was centrifuged and the supernatant freeze-dried.

The soluble tryptic peptides were separated on a DE-52 column (400 ml bed volume) eluted with a linear gradient of NH₄HCO₃, pH 8.5 (20–500 mM, 1.5 litres). Peptides were located by digesting small

samples from alternate fractions with chymotrypsin, 'fingerprinting' these by paper electrophoresis and chromatography, and pooling appropriately.

Small basic peptides emerged first: T15 (Arg), T18 and T20, followed by a mixture of larger basic peptides: T1, T7, T12, T4, T21, T16, T14 and T2. Also present in this mixture were significant yields of peptide 134–138 arising from tryptic cleavage at His-133 and peptide T2' (residues 5–15). The latter probably arises from removal of the basic *N*-terminal tetrapeptide by indigenous proteinases during enzyme purification, giving rise to the 'triplet' pattern of enzyme bands on native gel electrophoresis as described by Burleigh *et al.* (1974). The neutral peptides T5 and T11 emerged next together with peptide 129–133, arising from the His-133 cleavage discussed above. All of these peptides were purified by paper electrophoresis and chromatography and sequenced by dansyl/Edman. Subdigestion of peptide T2 with thermolysin (giving peptides 3–4, 5–11 and 12–15) and of peptide T14 (giving peptide 166–169) was used to confirm the sequences shown in Fig. 2.

The remaining fractions containing mainly larger acidic peptides were purified on Sephadex G-25 superfine (1125ml bed volume) in 50mM-NH₄HCO₃, yielding, in order: peptide T3, peptides T17 and T22, peptide T6+Ser-Val-Asp-Leu (derived from T22 by cleavage at Asn-246), peptide T9a (cleavage at Tyr-99 in T9) plus an acidic peptide (*m* = 0.48) which was lost, and finally peptide T9b (also from Tyr-99 cleavage). Fig. 2 summarizes the sequence evidence for each of these peptides. Sub-digestions of T3 with thermolysin (yielding peptides 26–32 and 28–32), of T9a and T9b with chymotrypsin (cleavage at Phe-92 and Trp-112) and of T17 with thermolysin (yielding peptides 188–193, 188–196, 197–203 and 198–203) were used to establish the sequences shown.

The precipitate of acid-insoluble tryptic peptides was dissolved in 15ml of 8M-urea/0.05M-NH₄OH and separated by upward flow on Sephadex G-50 superfine (406ml bed volume) in 0.2M-NH₄HCO₃. Three fractions were pooled, corresponding to *M_r* values 4000–3250 (G1), 3250–2500 (G2), 2500–1500 (G3) and below 1500 (G4), the column having been previously calibrated with peptides of known *M_r*.

The tubes corresponding to fraction G1 were recycled on the same column to yield peptides T8 and T19. Similar recycling for G2 gave peptides T19 plus T3. Fraction G3 contained mainly peptides T9 and T13 plus traces of T5, T3 and T17 and these were purified on a DEAE-cellulose column (62ml) with a linear gradient of 5–600mM-NaCl/25mM-Tris/HCl/8M-urea, pH 8.0 (200ml). The main peptide in fraction G4 was T22, which

was purified by paper electrophoresis at pH 6.5. The sequence of these peptides was determined by the dansyl/Edman method with subdigestion as indicated below.

A chymotryptic digest of peptide T8 gave fragments 54–63, 64–70, 71–83 and 84–88 plus 80–85, 80–83, 84–85, and 86–88. A thermolysin digest gave cleavages *N*-terminal to Leu-58, Phe-63, Leu-65, Leu-69, Val-75, Asn-77, Leu-79, Ile-82 and Leu-85 and the resulting peptides established the sequence except for the amide assignments. The mobilities of thermolysin fragments confirm Glu-57, Glu-60, Asn-61, Gln-80 and Gln-84 and that one of residues 66 or 68 and one of 71, 73 or 74 must be acidic. Mass-spectrometry of various peptides confirms the amide assignments shown (Morris *et al.*, 1974) except that a minor thermolytic fragment indicated Gln-60, a possibility we now reject.

Peptide T9 gave residues 89–102 by dansyl/Edman sequencing, and chymotryptic fragments 89–92, 93–99 and 100–114 established the sequence. This was confirmed by peptic fragments 89–92, 89–93, 93–94, 93–95, 95–99, 96–99, 99–106, 99–114, 100–106, 100–114 and 107–114.

Peptide T13 was difficult to purify, being insoluble at neutral pH. Its sequence was established by dansyl/Edman sequencing (ambiguous residues 153–156 and 163) and from chymotryptic fragments 139–144, 145–160 and 161–164. Subdigestion of these or of the original peptide with thermolysin gave fragments 139–141, 142–145, 148–150, 154–158, 155–158, 159–160, 159–161, 159–164 and 162–164. An error in the sequence quoted by Morris *et al.* (1974) was corrected by sequencing Ile-Trp-Glu-Pro-Val-Tyr-Thr (155–161) and Val-Ile-Trp-Glu-Pro (154–158) using mercaptoethanesulphonic acid hydrolysis of dansyl derivatives to avoid destruction of dansyl-Trp (Giglio, 1977).

Peptide T19 had the *N*-terminal sequence Met-Asp-Glu-Ala- and gave fragments 206–215, 206–227, 216–227 and 228–231 on chymotrypsin digestion. Thermolysin fragments were 206–209, 210–214, 215–218, 219–220, 219–220, 219–224, 221–224, 227–228 and 229–231. A CNBr digest gave fragments 207–216, 217–228 and 229–231. Dansyl/Edman sequencing and electrophoretic mobilities of these established the sequence shown in Fig. 2.

The *C*-terminal peptide T22 was sequenced directly and a thermolysin fragment 240–246 allowed amide group allocation. It is noteworthy that the Asn-246- to- Ser-247 bond was extensively cleaved in this tryptic digest, since fragments 239–246 and 247–249 were found in good yield in the soluble fraction.

Overlaps for tryptic peptides. A tryptic digest of maleylated carboxymethylated RDH-A from

strain A11 (see under 'Materials and methods') was separated by upward flow on a calibrated column of Sephadex G-50 superfine (83 cm $\times$ 2.5 cm diam.) in 250 mM-NH₄HCO₃. All expected fragments T1-T2-T3, T4-T5, T6-T7-T8, T9, T11, T12-T13-T14, T15, T16, T17-T18-T19, T20-T21 and T22 were found with the theoretical partition coefficients (except that T9 was slightly retarded, perhaps because of its high content of aromatic amino acids). Also present was T4-T5-T6-T7-T8 due to incomplete cleavage at Arg-46-Glu-47, but peptide T10 was again missing. After demaleylation (Butler & Hartley, 1972), dansyl/Edman analysis and further tryptic or chymotryptic digestion, the order of the tryptic fragments was established.

A CNBr digest of carboxymethylated RDH-A (10 μ mol) was separated on Sephadex G-50 (see under 'Materials and methods'), giving Fractions I-IV. The breakthrough (I) was re-run on Sephadex G-100 in 0.1 M-acetic acid/8 M-urea to yield essentially homogeneous CNBr3 plus traces of larger peptides. Recycling later tubes (II) gave pure CNBr2. Fraction III was desalted on Sephadex G10 in 0.1 M-NH₄HCO₃, freeze-dried and re-run on Sephadex G-25 in 0.2 M-NH₄HCO₃ giving CNBr and CNBr6 as major components. Fraction IV was similarly desalted, freeze-dried and separated by paper electrophoresis at pH 6.5, giving CNBr1 and CNBr4.

Peptides CNBr1, CNBr4 and CNBr5 were small enough to sequence directly. A tryptic digest of peptide CNBr2 was mapped by paper electrophoresis and chromatography, showing that T3, T4, T5, T6 and T7 were clearly present, together with a fragment of T8 and other spots that were probably due to incomplete cleavage between T5-T6 and T6-T7. Chymotryptic fragments of CNBr2 were isolated and sequenced giving residues 13-28, 29-34, 35-43 and 44-54, confirming that CNBr2 is residues 9-70. Peptide CNBr6 was analysed similarly, and thermolysin fragments 229-234 and 237-239 confirmed that it represents residues 229-249.

Sequence studies with the large CNBr3 fragment were less conclusive. The *N*-terminus was blocked (Gln?) and tryptic fingerprints identified T11, T12, T15, T16 and T18, but the other expected peptides are relatively insoluble. Nevertheless, CNBr3 is probably residues 71-206.

Many attempts to isolate the missing tryptic peptide T10 failed, but the final overlap was establishing by digesting [¹⁴C]carboxymethylated RDH-A from strain A611 (2 μ mol) with elastase (20 nmol) in 20 ml of 0.1 M-NH₄HCO₃ for 16 h at room temperature, and separating by paper electrophoresis and chromatography. One peptide (E1) was Leu-His-Leu-Asn-Ile (residues 116-120); another (E2) had the composition and mobility of

peptide 106-115 and a third (E3) of residues 123-127. A final relevant overlap is peptide TC-26c of Morris *et al.* (1974) which is Val-Leu-His-Leu.

Properties of ribitol dehydrogenase-D

Crude extracts of strain D1 showed a single band of ribitol dehydrogenase activity on polyacrylamide-gel electrophoresis migrating coincident with that of ribitol dehydrogenase-A and the purified protein migrated coincident with the subunits of the wild type enzyme on sodium dodecyl sulphate/polyacrylamide gels. The half-life in crude extracts at 55°C was 30 min compared with 20 min for RDH-A and 6 min for RDH-B (Hartley *et al.*, 1976) so the mutation(s) do not appear to affect the tertiary or quaternary structure in any significant way.

A segregation analysis of strain D1 was performed by the procedure of Rigby *et al.* (1974) to test whether the strain was indeed gene duplicated and whether it also contained copies of the gene for ribitol dehydrogenase-A. Strain D1 was grown to *A*₆₅₀ of 1.0 on casamino acids, and plated on to 0.2% xylitol/M9 at 37°C. Small colonies arose at high frequency (276/18000) and the specific activity of ribitol dehydrogenase was measured in disrupted cell extracts of 20 of these. Four contained only 5-6 units/mg (strain D1 has 30 units/mg), indicating single gene copy segregants at 0.23%. All of these segregants had activity ratios towards 0.5 M-xylitol/50 mM-ribitol identical with that of strain D1 (1.15) and different from that of strain A (0.035). Hence the original mutation must have been in a single copy of the ribitol dehydrogenase gene, and the mutant gene then duplicated in a subsequent event.

The kinetic parameters of pure RDH-D were determined according to Burleigh *et al.* (1974); Table 1 compares these with those previously found for RDH-A and RDH-B. The final column indicates the effect on intracellular pentitol dehydrogenase activity at various presumed intracellular substrate concentrations: 5 mM which is in the range of extracellular xylitol observed in continuous cultures and 500 mM which assumes an extremely efficient 'xylitol pump'. There is a 3-fold increase in activity towards xylitol and a slight increase in activity towards ribitol. The new enzyme activity has been acquired at no expense to the original activity.

Amino acid analysis and peptide mapping. It was expected that RDH-D might have only one or two differences from RDH-A and that these would be in neutral residues, since the enzymes have similar electrophoretic mobility (Hartley *et al.*, 1976). The only notable differences ($\pm 5\%$) in amino acid analyses were Ser (14), Glx (22), Pro (12), Ile (13) in RDH-D and in the RDH-A sequence. However, in

Table 1. Kinetic constants for strains A, B and D with various pentitols

$$\frac{k(e)}{v} = 1 + \frac{K_{P,NAD}}{[P][NAD]} + \frac{K_P}{[P]} + \frac{K_{NAD}}{[NAD]}$$

Substrate	Strain	$K_{P,NAD}$ (mM ²)	K_P (mM)	K_{NAD} (mM)	k (s ⁻¹)	v/v_A	
Ribitol	A	10	8	0.7	299	—	—
	B	6	7	0.4	229	1.0	0.9
	D	10	7	0.2	306	1.2	1.4
Xylitol	A	612	579	0.8	36	—	—
	B	202	273	0.2	65	4.5	3.4
	D	511	394	0.3	86	3.1	3.2
L-Arabitol	A	166	147	0.3	28	—	—
	B	140	170	0.2	41	1.5	1.5

(1) Assuming [NAD] = 1 mM, [pentitol] = 5 mM.

(2) Assuming [NAD] = 1 mM, [pentitol] = 500 mM.

view of the difficulty in determining these residues accurately the differences were not considered significant.

Peptide maps of chymotryptic or tryptic digests of the two proteins appeared identical. Differences were found between peptide maps of elastase digests but in all cases these were found to be due to small differences in the cleavage pattern, and to be consistent with the known sequence of RDH-A. Several of the elastase peptides were purified and their electrophoretic mobilities and amino acid analyses were consistent with the sequences shown in Fig. 3.

Peptide maps of RDH-D containing [¹⁴C]RDH-A. It was clear that a new method was needed to screen for the mutation(s). Hence we studied growth conditions for strains that superproduce RDH-A which would yield enzyme of high specific radioactivity to act as an internal marker in digests of the mutant enzyme. This led to the protocol for synthesis and purification of [¹⁴C]RDH-A described under 'Materials and methods'.

Ribitol dehydrogenase-D (56 mg) plus [¹⁴C]RDH-A (5 mg) was oxidized with performic acid, dissolved in 4 ml of 20 mM-Tris/acetate/8 M-urea, pH 8.0, chromatographed on a DE-52 column (30 cm × 2 cm) equilibrated with the same buffer and eluted with a linear gradient of 500 ml of starting buffer to 500 ml of 500 mM-Tris/acetate/8 M-urea, pH 8.0. The protein was eluted as a single peak containing all the radioactivity, which was dialysed against 50 mM-NH₄HCO₃, pH 8.5.

Trypsin (0.5 mg purified by affinity chromatography on a soybean trypsin inhibitor column) was added and the mixture was incubated at 37°C for 4 h. The digest was applied to a DE-52 column (70 cm × 2 cm) equilibrated with 50 mM-

NH₄HCO₃, pH 8.5, washed with 50 ml of starting buffer and eluted with a linear gradient of 0.05–1.0 M-NH₄HCO₃, pH 8.5 (1000 ml).

Samples from the column fractions were digested with chymotrypsin and subjected to paper electrophoresis at pH 6.5. The neutral bands were cut out and subjected to electrophoresis at pH 2. There was coincidence between the radioactive spots and cadmium/ninhydrin-stained spots except in two regions which correspond to chymotryptic fragments of peptide T17. The electrophoretic mobilities of these peptides suggested that there may be differences in the pattern of chymotryptic cleavage of T17 from RDH-D and RDH-A. Fractions from this column corresponding to eight peaks were pooled and freeze-dried. The small basic peptides in the breakthrough peak were separated by paper electrophoresis at pH 6.5 followed by paper chromatography. In all cases radioactive spots were contiguous with cadmium/ninhydrin-stained spots. Amino acid analyses confirmed their identity with T1, T2, T4, T7, T12, T14, T15, T16, T18, T20 and T21 (Fig. 3).

The acidic peaks were freeze-dried and fractionated on Sephadex G-25 or G-50 columns (75 cm × 1.5 cm) in 50 mM-NH₄HCO₃ yielding pure peptides T3, T5, T6, T8, T9, T11, T17 and T22 plus a fragment (residues 239–245) arising from peptide T22. In all case except that of DT17 the amino acid compositions were consistent with the sequence of RDH-A and where necessary chymotryptic subdigestion confirmed this. Peptide DT17 had a composition Asx (1.8), Thr (1.1), Pro (4.0), Gly (2.0), Ala (1.4), Val (3.1, uncorrected for incomplete hydrolysis), Leu (2.8), Lys (1.1) compared with the expected Asx₂, Thr, Pro₃, Gly₂, Ala₂, Val₄, Leu₃, Lys. Dansyl/Edman analysis gave the sequence: Val-Gly-Ala-Val-Leu-Pro-Gly-Pro-Val-Val-Xaa-Pro-Leu-Leu-Asx-Asx-Xaa-Pro-Xaa. In the same analysis, the [¹⁴C]dansyl spots were analysed to reveal the sequence of the underlying peptide from RDH-A, giving:

Val-Pro-Ala-Val-Leu-Pro-Gly-Pro

To confirm the apparent replacement of Ala-196 by Pro-196 in RDH-D, a thermolysin digest of peptide DT17 was purified by paper electrophoresis at pH 3.5 followed by paper chromatography. One peptide stained with cadmium/ninhydrin but was not radioactive. This was eluted, analysed and sequenced by dansyl/Edman, giving Val-Leu-Pro-Gly-Pro-Val-Val-Thr-Pro. The [¹⁴C]dansyl spots in this analysis gave the sequence Val-Leu-Pro-Gly-Pro-Val-Thr-Ala. Hence one mutation in RDH-D is Pro-196 for Ala-196. Note, however, that the radioactive 'wild-type' enzyme from strain A7 has Pro-186 in place of Gly-186 in RDH-A.

Peptides T10, T13 and T19 are relatively

Fig. 3. Amino acid sequences of RDH-A, RDH-B and RDH-D

The sequence shown is that of RDH-A. Peptides from tryptic digests of RDH-B or RDH-D are labelled BT1, DT1, etc. Peptides labelled DE were purified from an elastase digest of RDH-D. Further digestion of some of these by chymotrypsin or thermolysin is indicated DT3-C3 etc. or DT17-L2 etc. Amino acid compositions consistent with the RDH-A sequence are shown as full lines. Dotted lines indicate identities established by peptide mapping. Dansyl/Edman sequencing of peptide DT17 is shown by arrows.

insoluble and were lost in the above experiment. A repeat yielded DT13 and DT19 but DT10 was again absent. The amino acid composition and

chymotryptic fragments of peptide DT19 were identical with those of the same peptide from RDH-A, but peptide DT13 had a composition

(corrected for hydrolysis errors) of Asx (1.6), Thr (2.0), Ser (3.8), Glx (1.9), Pro (2.1), Gly (2.6), Ala (2.4), Val (4.0), Ile (3.0), Leu (1.0), Tyr (0.9), Phe (0.8), Lys (1.1) compared with (Asp, Thr₂, Ser₂, Glu, Pro₂, Gly₂, Ala₃, Val₅, Ile₄, Tyr, Phe, Trp, Lys) from the RDH-A sequence, which suggested that the peptide was impure. The chymotryptic fragments of peptide DT13 had the expected mobilities, but we cannot exclude the possibility of sequence differences, particularly in the 145–156 region. Therefore a second mutation may lie within peptides DT10 and DT13 but otherwise Pro-196 for Ala-196 is the only mutation in RDG-D.

Partial sequence of ribitol dehydrogenase-B

Attempts to locate the mutation in this enzyme were unsuccessful. A sample (5 μ mol, 135 mg) of the enzyme purified by Burleigh *et al.* (1974) were oxidized with performic acid and chromatographed on a DE-52 column as described for RDH-D. The elution profile was similar and the amino acid analysis of peak tubes was consistent with the wild-type sequence as shown in Fig. 3. The smaller tryptic peptides were identified only by peptide mapping, but the pattern was indistinguishable from that of wild type.

It is therefore probable, though not certain, that the mutation in RDH-B lies between residues 115–125, 139–164 or 206–231. It does not, however, contain the Ala-196→Pro-196 replacement found in RDH-D.

Discussion

Sequence of ribitol dehydrogenase-A

During the sequence studies on ribitol dehydrogenase-A we utilized enzyme from several strains of this organism resulting from many generations of continuous culture on xylitol. Most of the sequence was established with enzyme from strain A11, including the complete sequences of all tryptic peptides (except T10) and overlaps from maleylated enzyme and from mass-spectrometry (Morris *et al.*, 1974). Enzyme from strain A211 was used for CNBr cleavage and in repeating isolation of tryptic peptides, from which the sequence is confirmed in all but peptides T10 (residues 115–125) and T13 (residues 139–164). Enzyme from strain A611 was used for isolation of elastase and chymotryptic peptides overlapping T10. The soluble tryptic and chymotryptic peptides from the wild-type enzyme (strain A) cover the same sequence except for residues 64–85, 139–164, 198–204 and 206–227. Hence with these exceptions we can say that the sequences are identical in all these strains.

The interest arises because the gene for ribitol dehydrogenase from strain A encodes -Ser-Ser- instead of -Ala-Val- at 146–147 and -Asn-Gly- instead of -Asp-Gly- at 212–213 (Loviny *et al.*, 1985; the following paper). The latter is almost certainly a sequencing error due to deamidation during peptide purification, but the former must be a change selected during continuous culture on xylitol leading to strain A11.

The sequence evidence for strain D also speaks to this point, since although the change of Ala-196 to Pro-196 appears responsible for the altered specificity, its sequence appears identical except in the 'missing' peptide DT10 and DT13 which was not fully sequenced. The analytical results for the latter are inconclusive due to apparent impurity but are not inconsistent with -Ser-Ser- for -Ala-Val- at residues 146–147.

The history of the strains used for DNA sequencing will be discussed by Loviny *et al.* (1985), but the most plausible resolution of all the evidence is as follows. The wild-type enzyme appears to be as shown in Fig. 2 except that 146–147 is -Ser-Ser- and 212 is Asn. A mutation to -Ala-Val- at 146–147 occurred during the continuous culture on xylitol leading to strain A11, presumably before the gene-duplication was selected. Similarly a mutation to Pro-186 occurred during the evolution of the superproducing strain A7. Neither of these mutations has a significant effect on the kinetic properties or extracellular stability of the enzyme but may be adaptations to the huge load of intracellular synthesis in these superproducing strains.

In strain D, a mutation of Ala-196 to Pro-196 does cause a significant increase in activity towards xylitol. This preceded a gene duplication giving strain D1 since strains D, D1 and single-copy segregants of the latter had the same xylitol/ribitol activity ratio (results not shown). In strain B a quite different mutation gives an equally improved enzyme, though in this case the thermostability is lower. The other strains E–K must again have different enzymes, indicating a variety of changes that can improve xylitol specificity. The evidence therefore confirms the powerful selection exercised in continuous culture of even minor variants of a gene that limits the growth rate.

References

- Bruton, C. J. & Hartley, B. S. (1970) *J. Mol. Biol.* **52**, 165–178
- Burleigh, B. D., Rigby, P. W. J. & Hartley, B. S. (1974) *Biochem. J.* **143**, 341–352
- Butler, P. J. G. & Hartley, B. S. (1972) *Methods Enzymol.* **25**, 191–199
- Giglio, J. R. (1977) *Anal. Biochem.* **82**, 262–264

- Hartley, B. S., Altosaar, I., Dothie, J. W. & Neuberger, M. S. (1976) in *Structure-Function Relationship of Proteins* (Markham, R. & Horne, R. W., eds.), pp. 191-200, Elsevier/North Holland
- Loviny, T., Norton, P. M. & Hartley, B. S. (1985) *Biochem. J.* **230**, 579-585
- Morris, H. R., Williams, D. H., Midwinter, G. G. & Hartley, B. S. (1974) *Biochem. J.* **141**, 701-713
- Neidhart, F. C. & Magasanik, B. (1956) *Biochim. Biophys. Acta* **21**, 324-334
- Offord, R. E. (1966) *Nature (London)* **211**, 591-593
- Porath, J. & Sundberg, L. (1971) *Protides Biol. Fluids Proc. Colloq.* **18**, 401-410
- Rigby, P. W. J., Burleigh, B. D. & Hartley, B. S. (1974) *Nature (London)* **251**, 200-204
- Shotton, D. M. & Hartley, B. S. (1973) *Biochem. J.* **131**, 643-675
- Smyth, D. G. (1967) *Methods Enzymol.* **11**, 219
- Taylor, S. S., Rigby, P. W. J. & Hartley, B. S. (1974) *Biochem. J.* **141**, 693-700
- Wu, T. T., Lin, E. C. C. & Tanaka, S. (1968) *J. Bacteriol.* **96**, 447-456