Amino Acid Sequence of Alcohol Dehydrogenase from the Thermophilic Bacterium *Thermoanaerobium brockii*

Moshe Peretz and Yigal Burstein*,‡
Department of Organic Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel
Received February 27, 1989; Revised Manuscript Received April 19, 1989

ABSTRACT: The complete amino acid sequence of alcohol dehydrogenase of *Thermoanaerobium brockii* (TBAD) is presented. The S-carboxymethylated protein was cleaved at methionine residues (with cyanogen bromide) to provide a set of 10 nonoverlapping fragments accounting for 90% of the sequence. These fragments were then overlapped and aligned, and the sequence was completed by using peptides generated by proteolytic cleavage at lysine residues (with *Achromobacter* protease I). The protein subunit contained 352 amino acid residues corresponding to a molecular weight of 37 652. The sequence showed about 35% identity with that of the prokaryotic *Alcaligenes eutrophus* alcohol dehydrogenase and about 25% identity with any one of the eukaryotic alcohol/polyol dehydrogenases known today. Of these, only 18 residues (5%) are strictly conserved: 11 Gly, 2 Asp, and 1 each of Cys, His, Glu, Pro, and Val.

Alcohol dehydrogenases have been isolated from different sources, and the amino acid sequences of several such enzymes are known. Most of the alcohol dehydrogenases depend on zinc for their biological activity. One group of such enzymes includes tetrameric alcohol dehydrogenases from yeast (Jörnvall, 1977; Wills & Jörnovall, 1979; Bennetzen et al., 1982; Russell & Hall, 1983; Russell et al., 1983; McKnight et al., 1985; Young & Pilgrim, 1985) and from the bacterium *Alcaligenes eutrophus* (Jendrossek et al., 1988) and the tetrameric sorbitol dehydrogenase from sheep liver (Jeffrey et al., 1984). Another group of enzymes consists of dimeric alcohol dehydrogenases from the following: higher plants (Brändén et al., 1984; Dennis et al., 1985; Chang & Meyerowitz, 1986; Llewellyn et al., 1987; human liver class I α, β, and γ subunits [reviewed in Jörnvall et al. (1987)], class II π subunit (Höög et al., 1987), and class III ξ subunit (Kaiser et al., 1988); rat liver class I (Crabb & Edenberg, 1986) and class III (Julia et al., 1988); mouse liver (Edenberg et al., 1985) and horse liver class I (Jörnvall, 1970). Insect alcohol dehydrogenases or bacterial ribitol and glucose dehydrogenases lack zinc (Jörnvall et al., 1981, 1984; Jeffery et al., 1981), and other alcohol dehydrogenases use iron for activation (Scopes, 1983; Neale et al., 1986; Williamson & Paquin, 1987). The horse liver enzyme is the only alcohol dehydrogenase for which a detailed three-dimensional structure has been determined by X-ray crystallography (Eklund et al., 1976). An evolutionary relationship between the long-chain zinc-dependent alcohol dehydrogenases has been suggested, in the form of a flow scheme (Jörnvall et al., 1987). This scheme outlined dimeric and tetrameric alcohol dehydrogenases from eukaryotic species (including mammals, higher plants, fungi, and yeasts) since amino acid sequences of known prokaryotic alcohol dehydrogenases were then not yet available.

Thermoanaerobium brockii alcohol dehydrogenase (TBAD)[1] is an NADP-linked enzyme that was first described by Lamed and Zeikus (1980). The enzyme was later purified to "near homogeneity" and partially characterized (Lamed & Zeikus, 1981). TBAD, a tetramer with a molecular weight of 40 000 per subunit, is stable at high temperatures (up to 85 °C), tolerates organic solvents well, and shows broad specificity and high reactivity toward secondary alcohols and low reactivity toward primary alcohols (Lamed & Zeikus, 1981; Lamed et al., 1981; Keinan et al., 1986).

We report here the purification to homogeneity and the amino acid sequence of an alcohol dehydrogenase from the bacterium *T. brockii*. This is the first complete protein sequence of a prokaryotic, of an NADP-linked, and of a thermophilic alcohol dehydrogenase.

Materials and Methods

T. brockii strain HTD4 was a generous gift of Dr. R. Lamed, Tel-Aviv University. The bacterium was cultured according to the method of Zeikus et al. (1979). Cells were harvested in late exponential growth phase and were collected by continuous-flow centrifugation. Cell extracts were prepared from which the enzyme TBAD was purified by a modification of a procedure described by Lamed and Zeikus (1981). TBAD was purified 76-fold in six steps as summarized in Table I. The crude extract (500 mL, 20 mg of protein/mL) in 25 mM Tris-HCl (pH 7.3) was treated with phenylmethanesulfonyl fluoride (10 mL of a 3% solution in ethanol) and with DNase (75 μg/g of protein) for 30 min at room temperature. The solution was then heated at 85 °C for 2.5 min, rapidly cooled in an ice–water bath, and centrifuged. Then the supernatant was passed through a ZetaPrep 15-DEAE disk or cartridge (AMF, Inc.) onto the first Red Sepharose column (Pharmacia, 3 × 30 cm). The protein was eluted from the affinity column by using a solution of $NaClO_4$ (0.2 M) at pH 8 (25 mM Tris-HCl), dialyzed against 25 mM Tris-HCl (pH 8), and then passed through a second Red Sepharose minicolumn (0.5 × 4 cm). The protein fractions eluting from the column were combined, and aliquots containing 20 mg of protein were subjected to gel filtration through a column of Sephacryl S-300 (Pharmacia, 3 × 180 cm). The column was eluted with a solution of 50 mM Tris-HCl/100 mM NaCl (pH 7.3) at a

‡ Maynard Wishner Professor of Bioorganic Chemistry.

[1] Abbreviations: TBAD, *Thermoanaerobium brockii* alcohol dehydrogenase; HLAD, horse liver alcohol dehydrogenase; AEAD, *Alcaligenes eutrophus* alcohol dehydrogenase; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; HPLC, high-performance liquid chromatography; CM, carboxymethyl; RP, reversed phase; Tris, tris(hydroxymethyl)aminomethane; API, *Achromobacter* protease I; DIDS, 4,4′-diisothiocyano-2,2′-stilbenedisulfonate; OPA, *o*-phthalaldehyde; TFA, trifluoroacetic acid.

0006-2960/89/0428-6549$01.50/0

Table I: Purification of TBAD

no.	fraction	volume (mL)	protein[a] (mg)	act.[b] (units)	sp act. (units/mg)	yield (%)	purification (x-fold)
1	crude extract	500	10000	11 200	1.1	100	1
2	heat treatment	520	2630	10 450	4.0	93	4
3	ZetaPrep-DEAE	520	1880	10 340	5.5	92	5
4	Red Sepharose I	50	165	9 460	57.3	85	52
5	Red Sepharose II	54	155	9 410	60.7	84	55
6	Sephacryl S-300	110	98	8 250	84.2	74	76

[a] Protein concentration was determined according to the method of Bradford (1967). [b] One unit of enzyme catalyzes the oxidation of 1 μmol of 2-butanol/min at 40 °C.

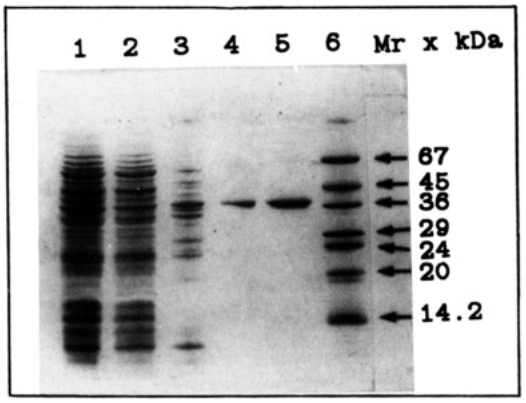

FIGURE 1: SDS–PAGE analysis of TBAD during different purification steps. Electrophoresis was performed according to the method of Laemmli (1970), on 12% gel slabs. Samples were denatured in the presence of 2-mercaptoethanol. (Lane 1) Crude enzyme after heat treatment; (lane 2) fraction from anion-exchange column; (lane 3) fraction from second Red Sepharose column; (lanes 4 and 5) purified protein, 2 and 8 μg, respectively, after Sephacryl S-300 (see text). The gel was stained with Coomassie Brilliant Blue. Positions of molecular weight markers are indicated.

flow rate of 16 mL/h, and fractions containing TBAD were pooled. Under these conditions TBAD displayed an apparent molecular weight of 160 000 when compared to elution volumes of the protein standards catalase (240 000), aldolase (158 000), and bovine albumin 67 000) (data not shown). On sodium dodecyl sulfate–polyacrylamide gel electrophoresis, the purified enzyme migrated as a single band with an apparent molecular weight of 38 000 per subunit (Figure 1).

Staphylococcus aureus V8 protease was a product of Miles Laboratories; clostripain was purchased from Sigma; *Achromobacter* protease I (API), which specifically cleaves C-lysyl peptide bonds, was a generous gift of Dr. Koiti Titani, Fujita-Gakuen Health University, Japan; carboxypeptidase Y was a product of Merck.

Reduction and S-carboxymethylation were performed by a modification of the method of Cresfield et al. (1963) using 6 M Gdn·HCl and dithiothreitol (Cleland, 1964) in placed of urea and 2-mercaptoethanol. For S-carboxymethylation, 2-iodoacetic acid or iodo[2-^{14}C]acetic acid (Amersham, 2 μCi/μmol) was used. The alkylated protein was separated from excess reagents and salts by dialysis against 50 mM ammonium bicarbonate in the dark. Methylation of methionine residues and isolation and characterization of methionine-containing peptides followed the procedure of Sasagawa et al. (1983).

Amino acid sequence of peptides was determined with a Model 470A protein sequencer equipped with an on-line PTH analyzer, Model 120A (Applied Biosystems). Data were collected and processed by using the chromstation-AT program of Spectra-Physics.

The procedure used for blocking primary amines by *o*-phthalaldehyde (OPA) during automatic gas-phase microsequencing was a modification of the method of Brauer et al. (1984) and involved the following steps: (1) stopping the analysis at the end of the cycle preceding a proline residue; (2) disassembling the cartridge and exposing the filter carrying the sample; (3) pipetting 30 μL of a freshly prepared solution of 1% OPA and 1% 2-mercaptoethanol in chlorobutane (S3) onto the sample at the glass fiber filter (the solvent dries immediately); (4) reassembling the cartridge; (5) running steps 1–18 of cycle 03COPA (see below); (6) repeating steps 2–4 of this procedure; (7) running a full cycle 03COPA; (8) resuming normal sequence procedure using cycle 03CPTH.

Cycle 03COPA is a modified 03CPTH cycle from which delivery of R1 was omitted and washes with S2 were extended. The changes were (a) in steps 2, 3, 9, and 12, cartridge function was changed from "deliver R1" to "deliver R2", and (b) in step 17 the time was extended to 480 s.

Complex mixtures of peptides were usually first separated by gel filtration on Sephadex G-50 in 5% formic acid and then purified by RP-HPLC on Aquapore RP-300 (C8) or PLRP-S 300 (polystyrene/divinylbenzene) in a Spectra-Physics modular chromatograph Model SP8800/8780/8500/8773/4290 equipped with a Barspec Model 1937 Chrom-A-Scope fast scan detector. In each case, an acetonitrile gradient was used in dilute aqueous solutions of trifluoroacetic acid (TFA) or sodium perchlorate/phosphoric acid.

For amino acid composition, samples of peptides were analyzed with Dionex D-500 and D-400 amino acid analyzers. Protein concentration was determined according to the method of Bradford (1976). Enzymic activity of TBAD was assayed by following the formation of NADPH from NADP at 334 nm (ϵ_{334} = 6.1 mM^{-1} cm^{-1}) in a reaction mixture containing NADP (0.5 mM), 2-butanol (150 mM), and Tris-HCl (100 mM, pH 7.8) at 40 °C. One unit of TBAD was defined as the amount of enzyme that catalyzes the oxidation of 1 μmol of 2-butanol/min at 40 °C.

Cleavage at methionyl peptide bonds followed the general procedure of Gross (1967). The protein (10 mg/mL) in 70% formic acid was treated with CNBr (10 mg/mL) at room temperature for 16–24 h. The reaction mixture was then diluted with water and lyophilized. Lysyl peptide bonds were selectively cleaved with API in 50 mM Tris-HCl/3 M urea (pH 9) for 5 h at 37 °C, S/E = 300; arginyl peptide bonds were cleaved with clostripain in 50 mM Tris-HCl (pH 8) for 6 h at 37 °C, S/E = 100, after activation of the enzyme in 1 mM $CaCl_2$ 1 mM 2-mercaptoethanol; glutamyl peptide bonds were cleaved with *S. aureus* V8 protease in ammonium bicarbonate pH 7.8 for 24 h at 37 °C, S/E = 500.

Carboxyl-terminal analyses of polypeptides were performed with carboxypeptidase Y according to the method of Klemm (1984) and by hydrazinolysis (Schroeder, 1972).

Aminosilylation of glass fiber filters was performed by dipping TFA-activated glass fiber filters (ABI 400379) in a solution of (3-aminopropyl)triethoxysilane (2% v/v) in acetonitrile/water (97:1) for 2 min at room tempeature followed by extensive washing of the filters with acetonitrile and drying at 100 °C for 1 h.

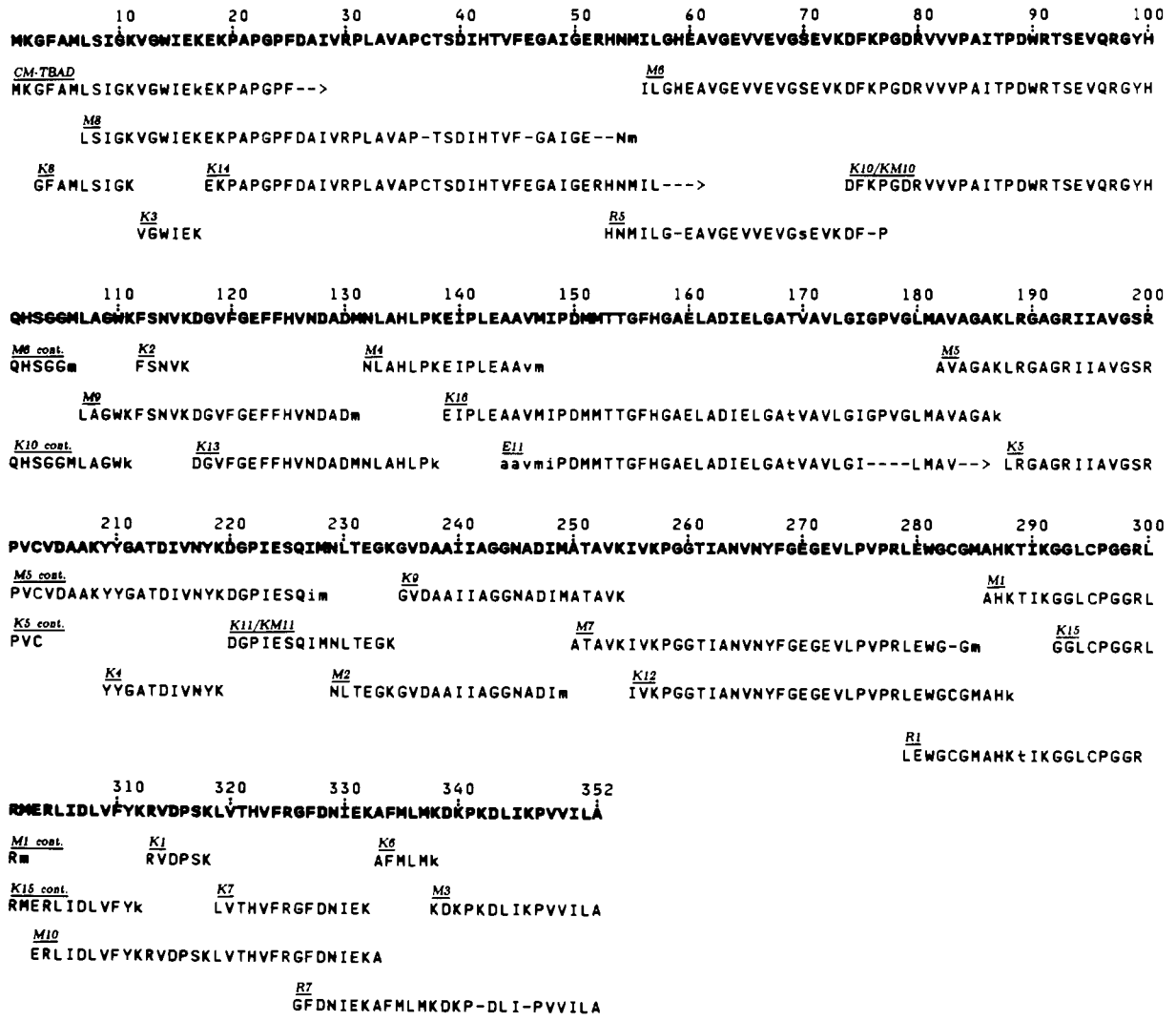

FIGURE 2: Detailed summary of the proof of sequence of TBAD. The proven sequences of specific peptides, with underlined names, are given in one-letter code below the summary sequence (bold type). Prefixes M, K, R, and E denote peptides generated by cleavage of CM-TBAD at methionyl, lysyl, arginyl, and glutamyl peptides bonds, respectively. The prefix KM denotes peptides of methylated CM-TBAD cleaved with API at lysine residues. Peptide sequences written in upper case letters were proven by Edman degradation; those in lower case letters indicate tentative identification or sequences deduced from amino acid compositions. Those not identified are shown by dashes or by long arrows, which indicate a long unidentified sequence.

Coupling of DIDS to peptides in solution was performed in a 1.5-mL capped Eppendorf tube. After the peptide was dissolved in 30 μL of an aqueous solution of 8 M Gdn·HCl and heated to 56 °C under argon, 40 μL of an aqueous solution of 1% trimethylamine and 0.5 mg of solid DIDS were added and allowed to react for 10 min at 56 °C under argon. Ammonia (20 μL of a solution of 25% in water) was added, and the reaction was continued for an additional 5 min and then terminated by the addition of 50 μL of 25% aqueous TFA. This solution of the modified peptide was concentrated to about 30 μL with argon and then applied to the gas-phase sequencer.

Coupling of peptides to a glass fiber filter via DIDS was performed in a similar way. Aminosilylated glass fiber filter (see above) was dipped in 0.5 mL of an aqueous solution of 1% trimethylamine and then reacted with 1 mg of solid DIDS at 56 °C under argon for 15 min. The filter was extensively washed with methanol and, after dipping in an aqueous solution of 1% trimethylamine, transferred to a small Eppendorf tube, dried under argon, and heated to 56 °C. After the peptide was loaded (1–10 nmol in 30 μL of an aqueous solution of 8 M Gdn·HCl) onto the filter, 20 μL of 1% aqueous trimethylamine was added, and then the coupling reaction was performed at 56 °C for 15 min under argon. To modify any remaining free amines, 300 μL of a solution of DIDS in 1% trimethylamine (2 mg/mL) was added, and the reaction was continued for 10 min. Ammonia (20 μL of a solution of 25% in water) was added for an additional 5 min, and the reaction was terminated with a solution of 25% TFA. The filter was extensively washed with methanol and dried.

Results

General Strategy of Sequence Analysis. Most of the sequence of TBAD was obtained by analysis of the amino terminus of the CM protein (residues 1–25, Figure 2) and of 10 of its nonoverlapping cyanogen bromide fragments. The remainder of the sequence and alignment of the fragments were obtained by analysis of peptides isolated from cleavages at lysyl residues by *Achromobacter* protease I, at arginyl residues by clostripain, and at glutamyl residues by *S. aureus* V8 protease. Complete details of the proof of the sequence are summarized in Figure 2. The results provide the complete sequence of *T. brockii* alcohol dehydrogenase.

Cyanogen Bromide Peptides. Reduced and radiolabeled [^{14}C]CM-TBAD (4 mg) was cleaved by cyanogen bromide, and an aliquot (0.5% of the total mixture of peptides) was analyzed by RP-HPLC (Figure 3A). Ten peptides could be identified (M1–10), four of which contained radioactivity of [^{14}C]CM-cysteine (M1, M5, M7, and M8). Figure 3A shows

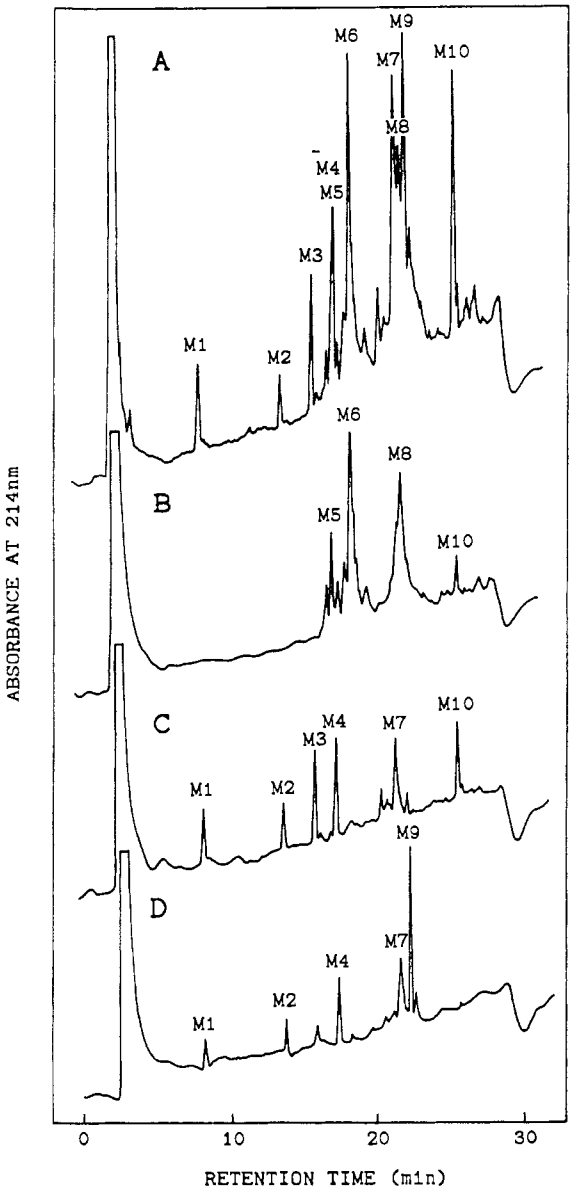

FIGURE 3: RP-HPLC separations of cyanogen bromide digest of CM-TBAD. The digest (4 mg) was first separated on a column of Sephadex G-50 superfine (1.5 × 100 cm) in 5% formic acid at a flow rate of 0.5 mL/min. Eleven 8 mL peptide containing fractions were collected (data not shown). An aliquot from each fraction was analyzed by RP-HPLC using an analytical column (Aquapore RP-300, 4.6 × 250 mm) and a gradient of acetonitrile (15–45% in 0.1% TFA for 30 min) at a flow rate of 1 mL/min. (A) Crude cyanogen bromide digest (20 μg) before gel filtration; (B, C and D) aliquots (0.5%) of fractions 3, 7, and 9 of the Sephadex G-50 column, respectively. Purified peptides are identified by the prefix M, as in Figure 2.

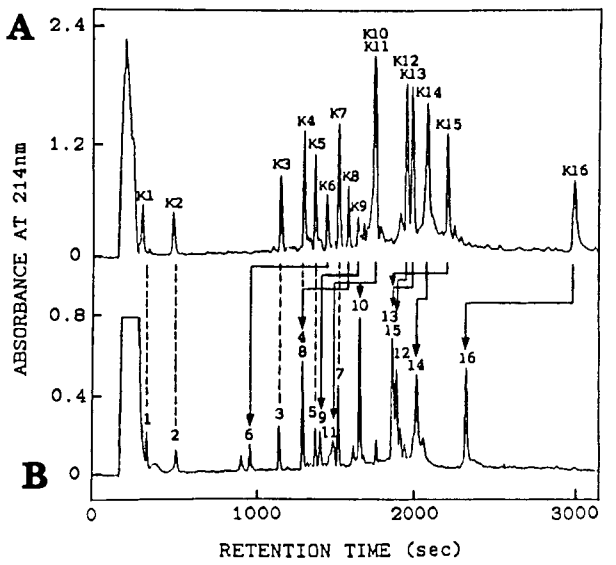

FIGURE 4: RP-HPLC separations of API digests of CM-TBAD and methylated CM-TBAD. The carboxymethylated proteins were digested by API in 3 M urea, and aliquots of the digests (250 μg) were separated on an Aquapore RP-300 column (4.6 × 250 mm) by using identical gradients of acetonitrile (5–55% in 0.1% TFA for 50 min) at a flow rate of 1 mL/min. Peptides were monitored at 214 nm. (A) Elution pattern of API digest of nonmethylated CM-TBAD; purified peptides are identified by the prefix K and a number reflecting their order of elution from the column. (B) Elution pattern of API digest of methylated CM-TBAD; purified peptides are identified by numbers (no prefix), identical with those in (A). Solid arrows point to the new elution position of the methylated peptides; dotted lines mark original and new unchanged positions of nonmethylated peptides.

that RP-HPLC separation of the 10 cyanogen bromide peptides was not complete; therefore, it was necessary to perform a preliminary separation of the peptide mixture using gel filtration on a column of Sephadex G-50 superfine. The 11 fractions thus obtained (data not shown) were further purified by RP-HPLC, an example of which is shown in Figure 3B–D. By this procedure, 10 pure cyanogen bromide fragments were isolated. The amino acid sequences of the peptides are shown in Figure 2, and the labeled peaks in the chromatograms of Figure 3 correspond to the amino acid sequence and composition given in Figure 2.

Lysyl Peptides. Ten milligrams of [^{14}C]CM-TBAD was cleaved at lysyl residues by API at pH 9 in the presence of 3 M urea, and the resulting peptide mixture was separated by RP-HPLC. Most peptides were obtained in pure form after this step. Figure 4A shows that 14 fractions corresponding to pure peptides K1–9 and K12–16 and one fraction containing a mixture of peptides K10 + K11 were isolated in reasonable yields (30–65%). Radioactivity of [^{14}C]CM-cysteine was detected in peptides K5, K12, K14, and K15. The use of a fast-scan detector in the HPLC separation enabled fast identification of Tyr- and Trp-containing peptides according to their absorption at 260–310 nm. Trp absorption was detected in peptides K3, K10, and K12, and Tyr absorption was detected in peptides K4, K10, K12, and K15. The amino acid sequences of the peptides are shown in Figure 2, and the labeled peaks of the chromatograms of Figure 4 correspond to the amino acid sequence and composition given in Figure 2.

Chemical Modification of the Mixture of Peptides K10 + K11 and of Peptide K16. Peptides K10 and K11 were obtained in a mixture that contained two methionine residues. Since the peptides in the mixture could not be resolved by RP-HPLC, an attempt was made to try to separate them following mild chemical modification. This was achieved by introducing into the peptides positively charged sulfonium groups at their methionine residues, thus modifying the "charge density" of the peptides in the mixture. The general procedure of Sasagawa et al. (1983) was used, except that methylation of methionine residues of [^{14}C]CM-TBAD was performed with nonlabeled methyl iodide. Figure 4 shows the separation patterns of the API digests of methyl-Met-[^{14}C]CM-TBAD (B) and of nonmethylated [^{14}C]CM-TBAD (A). Two groups of peptides could be identified, one of which consisted of six peptides (K1–5 and K7) whose retention times were not affected by methylation since they were devoid of methionine residues. The second group consisted of the methionine-containing lysyl peptides K6 and K8–16, in which the retention times were decreased in the methylated form and increased in the nonmethylated form. The change in the mobility of each

S-methylmethionine lysyl peptide (in RP-HPLC) was proportional to the change in its net charge density: the peptides K6, K8, and K16 that gained one charge per 3, 9, and 10 residues, respectively, showed the largest increase in mobility; peptides K11, K9, K15, and K13 that gained one charge per 15, 20, 21, and 22 residues, respectively, showed an intermediate increase in mobility; the smallest increase in mobility was recorded for peptides K12, K10, and K14 that gained one charge per 34, 39, and 56 residues, respectively. The mixture K10 + K11, the only unresolved lysyl peptides in Figure 4A, was well resolved after methylation (Figure 4B), and the two methylated lysyl peptides, KM10 and KM11, were isolated and characterized (Figure 2). In this experiment it was convenient to use nonradioactive methyl iodide and radiolabeled protein [rather than vice versa as in Sasagawa et al. (1984)] since methyl iodide is volatile, and special precautions should be taken when [^{14}C]methyl iodide is used.

K16, a fragment about 50 residues long (according to its amino acid composition), could be sequenced through residue 13 and was then washed out of the sequencer. To retain the peptide in the sequencer during the analysis, K16 was covalently bound to an aminosilylated glass fiber filter disk through the water-soluble bifunctional phenyl isothiocyanate reagent, DIDS (see Materials and Methods). "Solid-phase" sequence analysis of the DIDS-bound peptide (in the gas-phase sequencer) yielded the sequence of its residues 2–48, thus providing the complete sequence of peptide K16.

Arginyl and Glutamyl Peptides. Following clostripain cleavage of [^{14}C]CM-TBAD, 14 arginyl peptides were obtained by RP-HPLC and were identified by their amino acid compositions. Three such peptides, namely, R1, R5, and R7, were isolated and characterized. R5 yielded the partial sequence of residues 53–77, and R1 corresponded to residues 279–299 of TBAD (Figure 2). R7 was composed of 27 amino acids (residues 326–352, Figure 2), of which 20 residues would be extremely hydrophobic during Edman degradations [15 hydrophobic plus 5 (phenylthiocarbamyl)lysines]. Indeed, R7 could be sequenced only through residue 16 and was then washed out of the sequencer in the same manner as K16. Therefore, R7 was coupled with DIDS as described earlier, and the modified peptide was retained on the filter during the analysis and was sequenced through its C-terminal residues. The N-terminal glycine and the five lysines could not be identified in the analysis of the DIDS-modified R7.

Cleavage of [^{14}C]CM-TBAD with *S. aureus* V8 protease yielded 11 peptide fractions (isolated by RP-HPLC). Peptide E11 was formed by partial cleavage at Glu-143 and spanned through glutamic acid residues 160 and 165 and most probably also through 224. Analysis of the first cycles of the Edman degradation of E11 showed that this peptide was isolated in a mixture composed of about 40% E11 and 60% uncleaved or partially cleaved TBAD, and proline of E11 was identified in cycle 6 of the mixture. To block the contaminating sequences at cycle 6, the peptide mixture was reacted with OPA, a compound that reacts with primary amines but not with the imino group of proline (Roth, 1971; Brauer et al., 1984). Peptide mixture E11 (1 nmol) was subjected to automatic gas-phase microsequence analysis; the analysis was interrupted after cycle 5, OPA was pipetted onto the sample (double coupling), and after appropriate washings (see Materials and Methods), the analysis was resumed for a total of 40 cycles. Only one unique sequence, that of peptide E11, was detected and yielded the sequence of residues 6–32 and 37–40 of this peptide, corresponding to residues 148–175 and 180–183 of TBAD (Figure 2).

C-Terminal Analyses of TBAD. Hydrazinolysis of CM-TBAD yielded 1 mol of alanine/mol of protein. Carboxypeptidase Y digestion of CM-TBAD released 2 mol of valine and 1 mol each of alanine, leucine, and isoleucine at almost identical rates. It was therefore postulated that alanine was the C-terminal residue of TBAD and was preceded by four hydrophobic amino acid residues (Ile, Val, Val, Leu). M3 was the only peptide of the 10 cyanogen bromide fragments isolated and analyzed that lacked homoserine, indicating that it originated from the carboxyl terminus of the protein. These results were confirmed by sequencing of peptide M3.

Alignment of Peptides and the Complete Sequence. As indicated in Figure 2, the alignment of overlapping peptides showed that 95% of the sequence of TBAD was proven unequivocally by data obtained from at least two different cleavages. The remaining 5% of the sequence represented 17 residues scattered throughout the molecule that were not identified in two different sets of peptides but rather in replicate analyses of one set of peptides. These included residues 1, 2, 37, 46, 52, 59, 176–179, 204–208, 342, and 346 (Figure 2). Overall, the data provide a unique alignment and a complete sequence of the 352 amino acid residues of *T. brockii* alcohol dehydrogenase. No evidence was found for any microheterogeneity in the primary structure of TBAD, and it is likely that the tetramer of this enzyme is composed of identical subunits of 352 amino acid residues. The molecular weight of the subunit of TBAD calculated from the amino acid sequence (37 652) agrees with the molecular weight determined by amino acid composition analyses of the protein, as well as by SDS–PAGE analysis performed in this study (38 000, Figure 1) and previously by others (40 000; Lamed & Zeikus, 1981).

Discussion

The alignment of alcohol dehydrogenase from different sources essentially as suggested by Jörnvall et al. (1987) is given in Figure 5. The primary structure of TBAD is aligned with the partial protein sequence of the *Bacillus stearothermophilus* enzyme (residues 1–50; Bridgen et al., 1973; Jeck et al., 1979), with the translated nucleotide sequence of alcohol dehydrogenase of *A. eutrophus* (Jendrossek et al., 1988) and with the horse liver enzyme (Jörnvall, 1970) (the only alcohol dehydrogenase for which a three-dimensional structure is known; Eklund et al., 1976). Gaps were introduced to optimize identity and homology among the sequences. Some common parameters and major differences among TBAD and other alcohol dehydrogenases are discussed. TBAD is an NADP-linked tetrameric alcohol dehdyrogenase. Alignment of the sequence of TBAD with that of the horse liver enzyme shows that the bacterial enzyme is also a long-chain alcohol dehydrogenase but, like the *B. stearothermophilus* enzyme, lacks eight residues at the amino terminus (Figure 5). Comparison with the N-terminus of the *A. eutrophus* enzyme cannot be done at present, since the N-terminal residue of the enzyme is still unidentified at the protein level (Jendrossek et al., 1988). Five gaps exist in the TBAD molecule. A long gap composed of 18 residues is found at position 95 (of HLAD), and four small gaps of 1–2 residues at positions 39, 183, 246, and 283 (numbering of HLAD) are found. Such or similar gaps are also found in other alcohol dehydrogenases, although they may vary in size and may be shifted a few residues (Jörnvall et al., 1987). These gap sites are located at or very close to intron positions in the genes of alcohol dehydrogenases [Brändén et al., 1984; reviewed in Jörnvall et al. (1987)].

Overall, TBAD is very different from the other two enzymes HLAD and AEAD. Only 96 of 357 positions in TBAD are

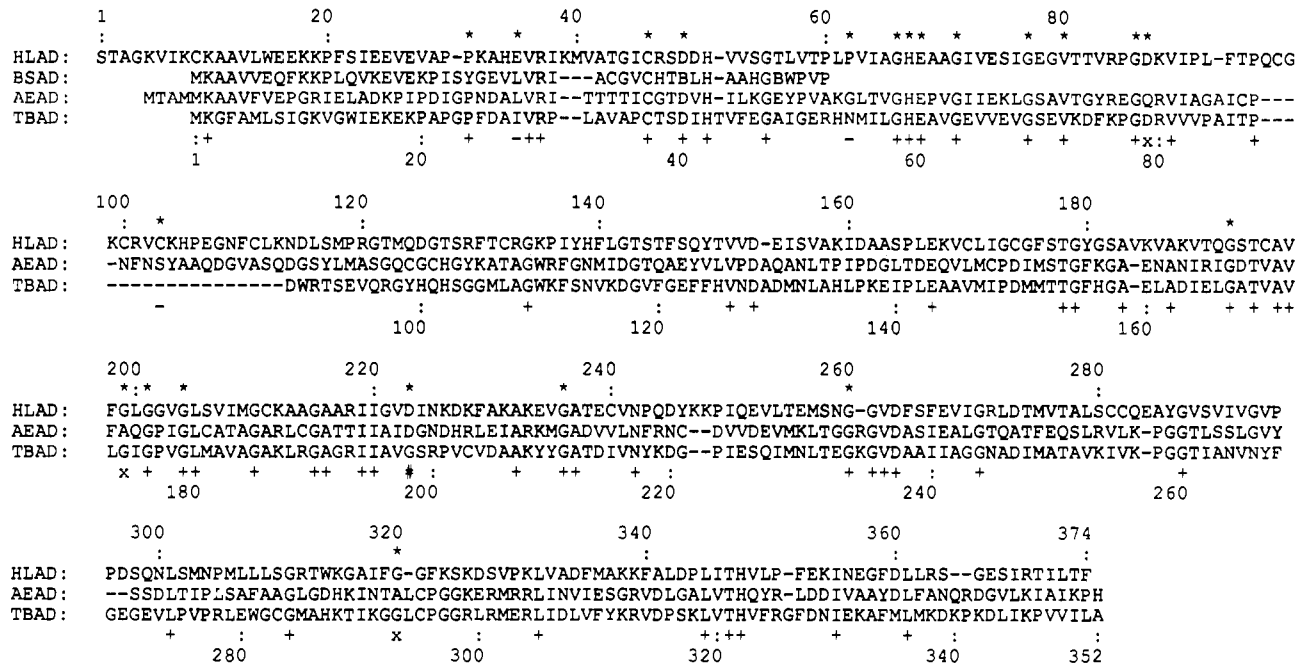

FIGURE 5: Comparison of the primary structures of alcohol dehydrogenases from horse liver (HLAD, EE enzyme) and from the bacteria *B. stearothermophilus* (BSAD, partial sequence), *A. eutrophus* (AEAD, translated from the nucleotide sequence), and *T. brockii* (TBAD). Gaps were introduced to optimize homology between the sequences. Residue numbering of HLAD is above the sequence, and that of TBAD is below the sequence. HLAD sequence is from Brändén et al. (1975); BSAD partial sequence is from Bridgen et al. (1973), corrected by Jeck et al. (1979); AEAD sequence is from Jendrossek et al. (1988); TBAD sequence is from Figure 2. (+) Identical residues in TBAD, AEAD, and HLAD; (*) strictly conserved residues in 16 different eukaryotic alcohol dehydrogenases [data from Jörnvall et al. (1987)]; (x) conserved residues that have been preserved in TBAD and replaced in AEAD; (−) conserved residues that have been replaced in both TBAD and AEAD; (#) conserved residue in all alcohol dehydrogenases with the exception of TBAD.

identical with those of the eukaryotic HLAD (27% identity), while in the other prokaryotic alcohol dehydrogenase, AEAD, 124 of 354 positions compared are identical with those of TBAD (35% identity). However, the percent of identity varies within different regions of the molecules, regions between the gap sites of TBAD. The N-terminal region of TBAD (region 1, residues 1–88 of this enzyme) shows 36% identity with HLAD and 32% identity with AEAD. The second region (residues 89–159) shows 17% identity with HLAD and 39% identity with AEAD. The third region (residues 160–221) shows 43% identity with both HLAD and AEAD, and the C-terminal region (region 4, residues 222–352) shows 23% identity with HLAD and 36% identity with AEAD. The greatest similarity between TBAD and the other two alcohol dehydrogenases is seen in region 3, where the vast majority of the coenzyme binding residues are located.

Most of the coenzyme binding residues of HLAD are conserved or conservatively replaced in TBAD, especially residues that interact with the nicotinamide, the nicotinamide ribose, and the pyrophosphate moieties of the coenzyme [Figure 5 and Eklund et al. (1981)]. Some differences are observed in residues that interact with the adenine moiety, but the major difference is observed in the interactions of the protein with the adenosine ribose moiety of the coenzyme. Asp-223 (of HLAD) is a strictly conserved residue in all NAD-linked alcohol dehydrogenases (Jörnvall et al., 1987). It forms an important hydrogen bond with O-2′ of the adenosine ribose moiety of NAD, which is essential for proper orientation and binding of the coenzyme to the enzyme. NADP has an extra phosphate group bound to O-2′ of the ribose moiety and cannot be properly oriented in HLAD in the same way as NAD, due to interference of the extra phosphate as well as charge repulsion between phosphate and Asp-223 (Brändén et al., 1975; Eklund et al., 1981). Indeed, horse liver alcohol dehydrogenase showed very little activity with NADP (Sund & Theorell, 1963; Fawcett & Kaplan, 1962). TBAD is an NADP-linked enzyme, and to accommodate an extra phosphate on the adenosine ribose, Asp-223 would have to give way to an amino acid with a much smaller side chain, such as Gly-198 of TBAD. Furthermore, a charged or a polar interaction between the phosphate group and the protein would be expected. The side chain of residue 225 of HLAD is oriented in the proper direction, is in close proximity, and would be a likely candidate for such interactions (Eklund et al., 1981). This side chain of Asn-225 is replaced in TBAD by the side chain of Arg-200; consequently, the positive charge of the arginine side chain could form a charge contact with the extra phosphate of NADP.

The horse liver enzyme contains two zinc atoms, one catalytic and one structural (Brändén et al., 1975). Preliminary results with TBAD indicated that the bacterial enzyme contained only one zinc per subunit and was reversibly inactivated by zinc chelating compounds, such as *o*-phenanthroline (Peretz et al., unpublished results). The sequence data revealed a stretch of 18 amino acids (residues 96–113) in HLAD that was missing in TBAD (Figure 5), which contained all 4 cysteine ligands of the structural zinc of the horse liver enzyme (cysteines 97, 100, 103, and 111). As for the ligands of the catalytic zinc atom, only two were preserved in TBAD: Cys-46 and His-67 (Figure 5). A similar situation was observed in sorbitol dehydrogenase, a tetrameric NAD-linked zinc enzyme from sheep liver (Jeffery et al., 1981), which lacks the structural zinc and three of its cysteine ligands (the fourth one, Cys-103, is preserved) as well as one of the catalytic zinc ligands, Cys-174 (Jeffery et al., 1984). Model studies (Eklund et al., 1985) have suggested that a glutamic acid residue replaces Cys-174 in HLAD. In TBAD, Asp-150 is in the homologous position of Cys-174 (of HLAD), which could serve as the third catalytic zinc ligand. The fourth ligand in HLAD is a water molecule or a hydroxyl ion (Brändén et al., 1975).

Most of the strictly conserved residues of other alcohol dehydrogenases have been preserved in TBAD as well. Of the 22 conserved residues of the eukaryotic alcohol dehydrogenases (Jörnvall et al., 1987), 18 were preserved in TBAD, including 11 glycines, 2 aspartic acids, and 1 each of Pro, Cys, His, Glu, and Val. In AEAD 3 of these residues, 2 Gly residues and 1 Asp, were replaced by Ala residues and 1 Gln, respectively, thus reducing the strictly conserved residues of the alcohol dehydrogenases to 15.

ACKNOWLEDGMENTS

We thank Drs. Hans Jörnvall and Michael Levitt for helpful discussions and suggestions, Dr. Rafi Lamed for a generous gift of the bacterium *Thermoanaerobium brockii*, Irit Lehavot for excellent technical assistance with the sequence analyses, and Dr. Virginia Buchner for helpful remarks during manuscript preparation.

SUPPLEMENTARY MATERIAL AVAILABLE

Two figures showing the general strategy for the determination of the sequence of TBAD and the relative orientation of the peptides, as well as the chromatographic separations (by RP-HPLC) of arginyl and glutamyl peptides of CM-TBAD, and two tables containing the amino acid compositions of CM-TBAD and of peptides obtained from its cleavage with cyanogen bromide and *Achromobacter* protease I (4 pages). Ordering information is given on any current masthead page.

REFERENCES

Bennetzen, J. L., & Hall, B. D. (1982) *J. Biol. Chem. 257*, 3018–3025.
Bradford, M. (1976) *Anal. Biochem. 72*, 248–254.
Brändén, C.-I., Jörnvall, H., Eklund, H., & Furugren, B. (1975) *Enzymes (3rd Ed.) 11A*, 103–190.
Brändén, C.-I., Eklund, H., Cambillau, C., & Pryor, A. J. (1984) *EMBO J. 3*, 1307–1310.
Brauer, A. W., Oman, L. C., & Margolies, M. N. (1984) *Anal. Biochem. 137*, 134–142.
Bridgen, J., Kolb, E., & Harris, J. I. (1973) *FEBS Lett. 33*, 1–3.
Chang, C., & Meyerowitz, E. M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1408–1412.
Cleland, W. W. (1964) *Biochemistry 3*, 480–482.
Crab, D. W., & Edenberg, H. J. (1986) *Gene 48*, 287–291.
Cresfield, A. M., Moore, S., & Stein, W. H. (1963) *J. Biol. Chem. 238*, 622–627.
Dennis, E. S., Sachs, M. M., Gerlach, W. L., Finnegan, E. J., & Peacock, W. J. (1985) *Nucleic Acids Res. 13*, 727–743.
Edenberg, H. J., Zhang, K., Fong, K., Bosron, W. F., & Li, T.-K. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2262–2266.
Eklund, H., Nordstrom, B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T., Soderberg, B.-O., Tapia, O., Brändén, C.-I., & Akeson, A. (1976) *J. Mol. Biol. 102*, 27–59.
Eklund, H., Samama, J.-P., Wallen, L., Brändén, C.-I., Akeson, A., & Jones, A. (1981) *J. Mol. Biol. 146*, 561–587.
Eklund, H., Horjales, E., Jörnvall, H., Brändén, C.-I., & Jeffery, J. (1985) *Biochemistry 24*, 8005–8012.
Fawcett, C. P., & Kaplan, N. O. (1962) *J. Biol. Chem. 237*, 1709–1715.
Gross, E. (1967) *Methods Enzymol. 11*, 238–255.
Höög, J.-O., von Bahr-Lindstrom, H., Heden, L.-O., Holmquist, B., Larsson, K., Hempel, J., Vallee, B. L., & Jörnvall, H. (1987) *Biochemistry 26*, 1926–1932.
Jeck, R., Woenckhaus, C., Harris, J. I., & Runswick, M. J. (1979) *Eur. J. Biochem. 93*, 57–64.
Jeffery, J., Cummins, L., Carlquist, M., & Jörnvall, H. (1981) *Eur. J. Biochem. 120*, 229–234.
Jeffery, J., Cederlund, E., & Jörnvall, H. (1984) *Eur. J. Biochem. 140*, 7–16.
Jendrossek, D., Steinbuchel, A., & Schlegel, H. G. (1988) *J. Bacteriol. 170*, 5248–5256.
Jörnvall, H. (1970) *Eur. J. Biochem. 16*, 25–40.
Jörnvall, H. (1977) *Eur. J. Biochem. 72*, 425–442.
Jörnvall, H., Persson, M., & Jeffery, J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4226–4230.
Jörnvall, H., von Bahr-Lindstrom, H., & Jeffery, J. (1984) *Eur. J. Biochem. 140*, 17–23.
Jörnvall, H., Persson, B., & Jeffery, J. (1987) *Eur. J. Biochem. 167*, 195–201.
Julia, P., Pares, X., & Jörnvall, H. (1988) *Eur. J. Biochem. 172*, 73–83.
Kaiser, R., Holmquist, B., Hempel, J., Vallee, B. L., & Jörnvall, H. (1988) *Biochemistry 27*, 1132–1140.
Keinan, E., Hafeli, E. K., Seth, K. K., & Lamed, R. J. (1986) *J. Am. Chem. Soc. 108*, 162–169.
Klemm, P. (1984) in *Methods in Molecular Biology, Proteins* (Walker, J. M., Ed.) Vol. 1, pp 255–259, Humana Press, Clifton, NJ.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
Lamed, R. J., & Zeikus, J. G. (1980) *J. Bacteriol. 141*, 1251–1257.
Lamed, R. J., & Zeikus, J. G. (1981) *Biochem. J. 195*, 183–190.
Lamed, R. J., Keinan, E., & Zeikus, J. G. (1981) *Enzyme Microb. Technol. 3*, 144–148.
Llewellyn, D. J., Finnegan, E. J., Ellis, J. G., Dennis, E. S., & Peacock, W. J. (1987) *J. Mol. Biol. 195*, 115–123.
McKnight, G. L., Kato, H., Upshall, W., Parker, M. D., Saari, G., & O'Hara, P. J. (1985) *EMBO J. 4*, 2093–2099.
Neale, A. D., Scopes, R. K., Kelly, J. M., & Wettenhall, R. E. H. (1986) *Eur. J. Biochem. 154*, 119–124.
Roth, M. (1971) *Anal. Chem. 43*, 880–882.
Russel, D. W., Smith, M., Williamson, V. M., & Young, E., T. (1983) *J. Biol. Chem. 258*, 2674–2682.
Russell, P. R., & Hall, B. D. (1983) *J. Biol. Chem. 258*, 143–149.
Sasagawa, T., Titani, K., & Walsh, K. A. (1983) *Anal. Biochem. 128*, 371–376.
Schroeder, W. (1972) *Methods Enzymol. 25*, 138–143.
Scopes, R. K. (1983) *FEBS Lett. 156*, 303–306.
Sund, H., & Theorell, H. (1963) *Enzymes, 2nd Ed. 7*, 25–83.
Williamson, V. M., & Paquin, C. E. (1987) *Mol. Cell. Biol.* (in press).
Wills, C., & Jörnvall, H. (1979) *Eur. J. Biochem. 99*, 323–331.
Young, E. T., & Pilgrim, D. (1985) *Mol. Cell. Biol. 5*, 3024–3034.
Zeikus, J. G., Hegge, P. W., & Anderson, M. A. (1979) *Arch. Microbiol. 122*, 41–48.