Primary Structure of Cholera Toxin β-Chain: A Glycoprotein Hormone Analog?

Abstract. *The completed sequence of the β-chain of cholera toxin (103 amino acid residues) was compared to the β-chains of chorionic gonadotropin, thyrotropin, luteinizing, and follicle stimulating hormones. The overall chemical similarity of the toxin β-chain to the hormones was not statistically different from random; however, a comparison of the first 40 residues of the toxin β-chain to the glycoprotein hormones revealed a segment of the hormones which was significantly chemically similar. The probability was < .003 that the similarity was due to chance.*

Cholera toxin (enterotoxin) is a protein of molecular weight 84,000 which is secreted by *Vibrio cholerae* and is responsible for the clinical manifestations of cholera (*1*). The toxin molecule is composed of three unique polypeptide chains whose molecular weights are α = 24,000, γ = 9,700, and β = 11,590 (*2*). The α- and γ-chains are covalently linked by a disulfide bridge and associate noncovalently with an aggregate of β-chains to give a molecular formula of $\alpha\gamma\beta_n$ (where n = 4, 5, or 6) (*3, 4*). The α- and γ-chains are designated subunit A, and the aggregate of β-chains is designated subunit B (*5*). The B subunit contains the binding site, or sites, for attachment to cell membranes, but does not itself cause any fluid outpouring comparable to intact cholera toxin when applied to intestinal mucosa. During purification of cholera toxin, a component (choleragenoid) of lower molecular weight is also isolated which cross-reacts immunologically with the toxin but is devoid of toxicity (*6*). Chemical and physicochemical studies revealed that cholera toxin and choleragenoid, although immunologically related, differed with regard to molecular weight, amino acid composition, and isoelectric point (*7*). Acrylamide gel electrophoretic studies and amino acid sequence analysis of the native proteins revealed that choleragenoid is comprised solely of an aggregate of β-chains (*2*) and may in fact be dissociated subunit B. This view is supported by the fact that cholera toxin is readily dissociated by mechanical agitation (*6*).

Interest in cholera toxin was aroused when it was shown that the toxin stimulated adenylate cyclase (E.C. 4.6.1.1) and increased the concentration of adenosine 3′,5′-monophosphate (cyclic AMP) in intestinal mucosa as well as in a wide variety of tissues (*8*). Choleragenoid also bound to cell membranes and inhibited binding of cholera toxin but did not stimulate adenylate cyclase. Considerable evidence has accumulated to show that cholera toxin and choleragenoid bind to the G_{M1} ganglioside of the cell membrane (*9*). Experiments with erythrocyte cell lysates and membrane preparations strongly indicate that the adenylate cyclase activity of cholera toxin occurs on the α-chain of the A subunit (*10*).

The effect of cholera toxin on adenylate cyclase activity led many investigators to study the interaction of the toxin with hormone receptors (*11*). That cholera toxin may represent an ancestral hormone was suggested by Hirschhorn and Greenough several years ago (*12*). Recent findings (*13*) revealed that, like the toxin receptor, ganglioside or a ganglioside-like structure is a component of the thyrotropin membrane receptor. Comparison of 42 residues of the NH_2-terminal region of the β-chain of cholera toxin to β-chain portions of human chorionic gonadotropin (CG), bovine thyrotropin (TSH), sheep luteinizing hormone (LH), and human follicle-stimulating hormone (FSH) revealed some preliminary sequence similarities and raised the question of possible structural relatedness (*13, 14*). The molecular weight of the β-chain of cholera toxin is similar to the β-chains of these hormones and, moreover, like the hormones the action of the toxin on whole cells is dependent on the interaction of two protein subunits. In the case of the glycoprotein hormones the β-chain confers the tissue specificity.

The complete sequence of the 103 residues of the β-chain of cholera toxin is given in Fig. 1. This sequence gives a compositional molecular weight of 11,590 which is slightly larger but within reasonable limits of previously reported acrylamide gel electrophoretic and ultracentrifugation estimates (*2, 15*). The β-chain was isolated from purified cholera toxin by gel filtration as previously described (*2*). The amino acid sequence of the polypeptide chain is relatively unremarkable with only a single disulfide bridge formed by half-cystines at opposite ends of the molecule that results in a large loop of 75 residues. By means of both manual and automated methods of sequencing, the sequence was established primarily from tryptic peptides, cyanogen bromide peptides, and from a fragment generated by cleavage of the single tryptophan with 2-(2-nitrophenylsulfenyl)-3-methyl-3-bromoindole (BNPS-skatole). Our results are in complete agreement with the partial sequence data reported by Mendez *et al.* (*16*). Preliminary partial sequences of the β-chain were also presented by others (*4, 17*).

Comparison of the complete primary structure of cholera toxin β-chain to the β-chain amino acid sequences of TSH, LH, FSH, and CG, according to the procedures of either Fitch (*18*) or Needleman and Wunsch (*19*), revealed no statistically significant similarities. However, a comparison of the first 40 amino acids

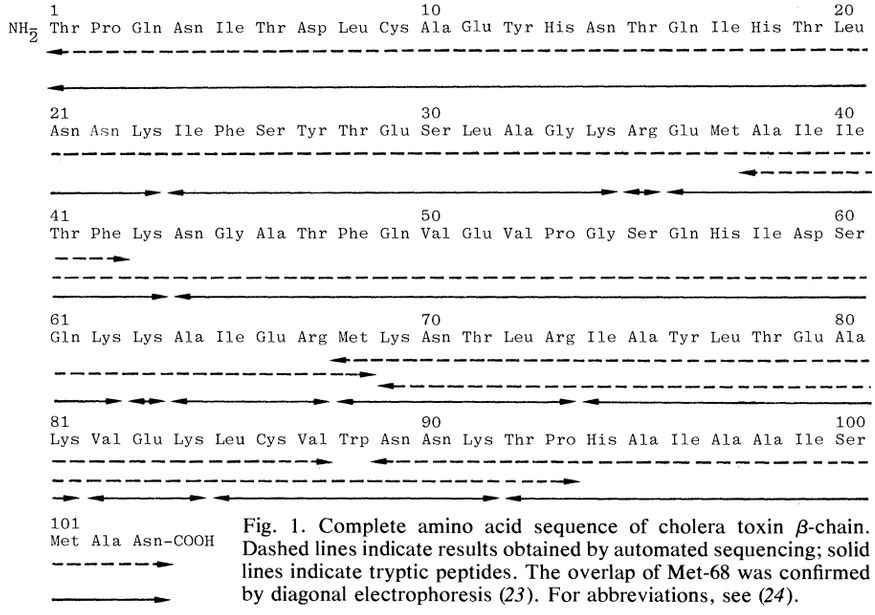

Fig. 1. Complete amino acid sequence of cholera toxin β-chain. Dashed lines indicate results obtained by automated sequencing; solid lines indicate tryptic peptides. The overlap of Met-68 was confirmed by diagonal electrophoresis (*23*). For abbreviations, see (*24*).

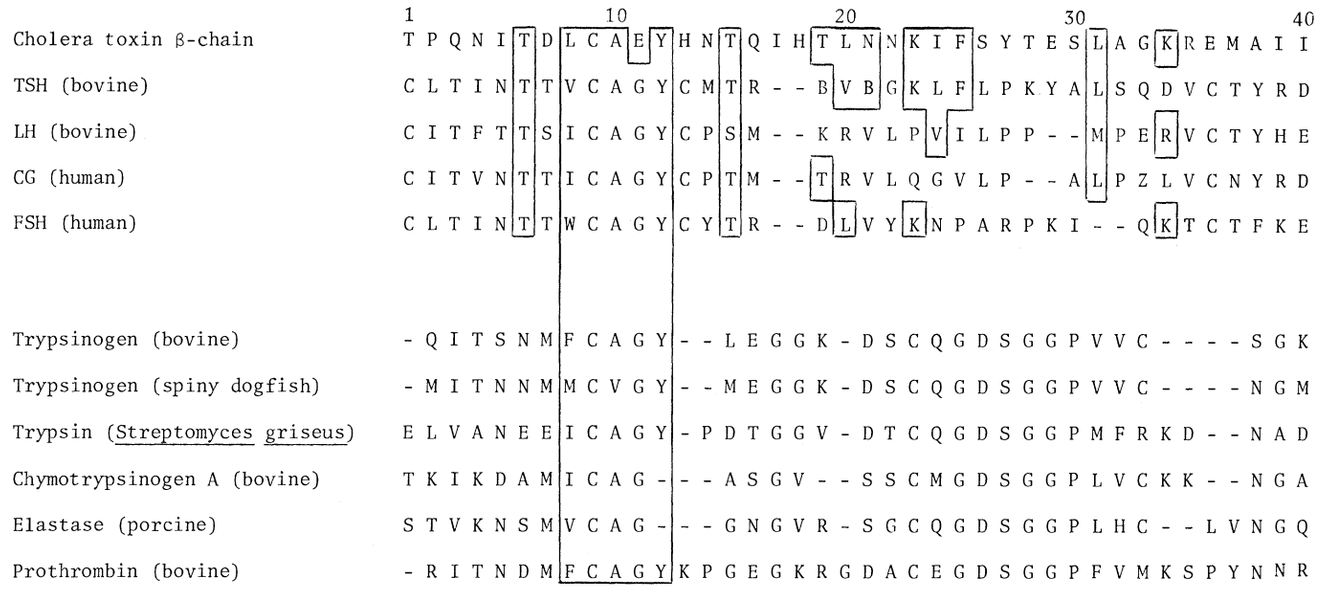

Fig. 2. Comparison of cholera toxin β-chain with glycoprotein hormones and serine proteases. Residue numberings (*24*) are: bovine TSH (19 to 56); bovine LH (26 to 61); human CG (26 to 61); human FSH (20 to 55); and serine proteases (174 to 206, chymotrypsinogen A numbering). Residues 41 to 103 of the toxin do not compare significantly with the hormones.

of the toxin β-chain to the hormone β-chains collectively (Fig. 2) by the Fitch method showed significant similarity to an internal 40-residue segment ($P < .003$ that the similarity was due to chance). The similarity is particularly striking in the region of Cys-9 of the toxin β-chain (hCAGY in the hormones, where h is a hydrophobic residue). The disulfide arrangement of the hCAGY sequence in the hormones is at present unresolved (*20*). The similarity in the Cys-9 region has prompted the hypothesis that hCAGY as well as residues in the immediate spatial proximity may function in the binding of these molecules to receptors on cell surfaces (*14*). On the other hand, since the various β-chains have different binding specificities one might more readily expect differences at the binding site similar to sequence differences in the binding site of other protein families, for example, serine proteases. Moreover, chemical modification of Cys-9 in the toxin β-chain by reduction and alkylation does not result in any loss of binding ability (*21*). Therefore, the Cys-9 region may be only a part of the binding site, critically juxtapositioned with other residues that participate more directly with membrane components (gangliosides) and provide the observed specificities. Additional chemical modification studies will be helpful in further elucidating the nature of the site of interaction of these proteins with membranes. The availability of a complete amino acid sequence will facilitate such studies.

It is of interest that the hCAGY sequence also occurs in the serine protease family of enzymes (Cys-182, chymotrypsinogen numbering) (Fig. 2). However, comparison of the first 40 residues of cholera toxin with the primary structure of serine proteases failed to show any statistically significant similarity. In the serine proteases CAG is in the internal portion of the molecule and is part of the methionyl loop (Cys-168 to Cys-182). This loop is common to all the serine proteolytic enzymes.

An explanation for the common hCAGY region in hormones, cholera toxin, and the serine proteases cannot now be provided with any certainty. In the case of cholera toxin and the hormones there is some functional biological similarity; however, the proteolytic enzymes have no apparent membrane function analogous to the hormones except perhaps the platelet degranulation reaction by thrombin. In the case of cholera toxin and the hormones the similarity may be the result of either limited convergence from unlike proteins (analogy) or extensive divergence from a common ancestor (homology). It is not a logical necessity that the absence of significant similarity over the complete sequence requires that these proteins be analogous rather than homologous in spite of their phylogenetic disparity. Therefore the conclusion by Ledley *et al.* (*14*) that the β-chain of cholera toxin is structurally analogous to the β-chains of certain glycoprotein hormones lacks sufficient evidence. Comparison of the three-dimensional structures, when available, should reveal more about the ancestral relatedness of these proteins.

Finally, since other bacterial enterotoxins also involve adenylate cyclase activity in their mode of action, it is likely that families of homologous toxins exist. The fact that the enterotoxins isolated from *Escherichia coli* and *Salmonella typhimurium* are inactivated by cholera antitoxin supports this hypothesis (*22*).

ALEXANDER KUROSKY
Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch, Galveston 77550

DAIEL E. MARKEL
JOHNNY W. PETERSON
Department of Microbiology, University of Texas Medical Branch

WALTER M. FITCH
Department of Physiological Chemistry, University of Wisconsin, Madison

References and Notes

1. R. A. Finkelstein, in *Critical Reviews In Microbiology* (Chemical Rubber Company, Cleveland, Ohio, 1973), vol. 2, p. 553.
2. A. Kurosky, D. E. Markel, B. Touchstone, J. W. Peterson, *J. Infect. Dis.* **133**, S14 (1976).
3. The number of β-chains comprising subunit B is disputed. Sequence analysis of intact cholera toxin [J. W. Jacobs *et al.* (*4*); Kurosky *et al.* (*2*)] indicated four β-chains, whereas other reports proposed five β-chains [D. M. Gill, *Biochemistry* **15**, 1242 (1976)] and six β-chains [C. Y. Lai, E. Mendez, D. Chang, *J. Infect. Dis.* **133**, S23 (1976)]. The variation in the reported value for the aggregation of the β-chain reflects the possibility that cholera toxin is another example of a dissociating-associating protein complex which could be influenced by such environmental factors as *p*H, ionic strength, or temperature.
4. J. W. Jacobs, H. D. Niall, G. W. G. Sharp, *Biochem. Biophys. Res. Commun.* **61**, 341 (1974).
5. Other nomenclatures are also used to denote the polypeptide chains and subunits of cholera toxin. These are H and L, where H is subunit A and L is the β-chain [I. Lönnroth and J. Holmgren, *J. Gen. Microbiol.* **76**, 417 (1973)]; A_1, A_2, and B subunit, where A_1 is α, A_2 is γ, and the B subunit is the β-chain [R. A. Finkelstein, M. Boesman,

S. H. Neoh, M. K. LaRue, J. J. LoSpalluto, R. Delaney, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **33**, 263 (1974)]. The chain and subunit nomenclature represented herein and elsewhere (*2*) avoids the ambiguity concerning the nature of the β-chain aggregation and is in keeping with the recommendations of the IUPAC-IUB (International Union of Pure and Applied Chemistry–International Union of Biochemistry) Commission on Biochemical Nomenclature [*Biochem. J.* **126**, 769 (1972)]. The use of A_1 and A_2 implies a dimeric form of a polypeptide whereas in fact the α- and γ-chains are unique sequences.
6. R. A. Finkelstein, J. W. Peterson, J. J. LoSpallu to, *J. Immunol.* **106**, 868 (1971); R. A. Finkelstein, M. Boesman, S. H. Neoh, M. K. LaRue, R. Delaney *ibid.* **113**, 145 (1974).
7. J. J. LoSpalluto and R. A. Finkelstein, *Biochim. Biophys. Acta* **257**, 158 (1972); R. Delaney, *Ann. Okla. Acad. Sci.* **4**, 68 (1974).
8. D. E. Schafer, W. D. Lust, B. Sircar, N. D. Goldberg, *Proc. Natl. Acad. Sci. U.S.A.* **67**, 851 (1970); Q. Al-Awquati, J. L. Cameron, M. Field, W. B. Greenough III, *J. Clin. Invest.* **49**, 2a (1970); M. Vaughan, N. F. Pierce, W. B. Greenough III, *Nature (London)* **226**, 658 (1970); W. B. Greenough III, N. F. Pierce, M. Vaughan, *J. Infect Dis.* **121**, S111 (1970); D. V. Kimberg, M. Field, J. Johnson, A. Henderson, E. Gershon, *J. Clin. Invest.* **50**, 1218 (1971); G. W. G. Sharp and S. Hynie, *Nature (London)* **229**, 266 (1971); L. C. Chen, J. E. Rohde, G. W. G. Sharp, *Lancet* **1971-I**, 939 (1971); J. M. Boyle and J. D. Gardner, *J. Clin. Invest.* **53**, 1149 (1974).
9. J. Holmgren, I. Lönnroth, J.-E. Mansson, L. Svennerholm, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2520 (1975).
10. D. M. Gill and C. A. King, *J. Biol. Chem.* **250**, 6424 (1975); M. W. Bitensky, M. A. Wheeler, H. Mehta, N. Miki, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2572 (1975).
11. V. Bennett, L. Mong, P. Cuatrecasas, *J. Membr. Biol.* **24**, 107 (1975).
12. N. Hirschhorn and W. B. Greenough III, *Sci. Am.* **225** (No. 2), 15 (1971).
13. B. R. Mullin, P. H. Fishman, G. Lee, S. M. Aloj, F. D. Ledley, R. H. Winand, L. D. Kohn, R. O. Brady, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 842 (1976).
14. F. D. Ledley, B. R. Mullin, G. Lee, S. M. Aloj, P. H. Fishman, Lois T. Hunt, M. O. Dayhoff, L. D. Kohn, *Biochem. Biophys. Res. Commun.* **69**, 852 (1976). The cholera toxin sequence reported by Mullin *et al.* (*13*) was taken from the abstract of Kurosky *et al.* [A. Kurosky, D. E. Markel, B. Touchstone, J. W. Peterson, *Proc. Am. Soc. Microbiol.* **1975**, 23 (1975)] and included an error in position 40 which appeared in the original abstract and which was subsequently corrected (*2*).
15. C. Y. Lai, E. Mendez, D. Chang, *J. Infect. Dis.* **133**, S23 (1976).
16. E. Mendez, C. Y. Lai, A. Wodnar-Filipowicz, *Biochem. Biophys. Res. Commun.* **67**, 1435 (1975).
17. Y. Nakashima, P. Napiorkowski, D. Schafer, W. Konigsberg, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **35**, 1356 (1976).
18. W. M. Fitch, *J. Mol. Biol.* **16**, 9 (1966); *ibid.* **49**, 1 (1970).
19. S. Needleman and C. D. Wunsch, *ibid.* **48**, 443 (1970).
20. D. C. Chung, M. R. Sairam, C. H. Li, *Int. J. Peptide Protein Res.* **7**, 487 (1975); D. N. Ward, L. E. Reichert, W.-K. Liu, H. S. Nahm, J. Hsia, W. M. Lamkin, N. S. Jones, *Rec. Prog. Horm. Res.* **29**, 533 (1973).
21. I. Lönnroth and J. Holmgren, *J. Gen. Microbiol.* **91**, 263 (1975).
22. P. D. Sandefur and J. W. Peterson, *Infect. Immun.* **14**, 671 (1976). F. Dorner, H. Jaksche, W. Stöckl, *J. Infect. Dis.* **133**, S142 (1976).
23. J. Tang and B. S. Hartley, *Biochem. J.* **102**, 593 (1967).
24. Amino acid sequences were taken from the following reports: bovine TSH, T.-H. Liao and J. G. Pierce, *J. Biol. Chem.* **246**, 850 (1971); bovine LH, G. Maghuin-Rogister and G. Hennen, *Eur. J. Biochem.* **39**, 235 (1973); human CG, R. B. Carlsen, O. P. Bahl, N. Swaminathan, *J. Biol. Chem.* **248**, 6810 (1973); human FSH, B. B. Saxena and P. Rathnam, *ibid.* **251**, 993 (1976); bovine trypsinogen, O. Mikes. V. Holeysovsky, V. Tomasek, F. Sorm, *Biochem. Biophys. Res. Commun.* **24**, 346 (1966); spiny dogfish trypsinogen, C. de Haën, H. Neurath, D. C. Teller, *J. Mol. Biol.* **92**, 225 (1975); *Streptomyces griseus* trypsin, R. W. Olafson, L. Jurášek, M. R. Carpenter, L. B. Smillie, *Biochemistry* **14**, 1168 (1975); bovine chymotrypsinogen A, D. M. Blow, J. J. Birktoft, B. S. Hartley, *Nature (London)* **221**, 337 (1969); porcine elastase, D. M. Shotton, B. S. Hartley, *ibid.* **225**, 802 (1970); bovine prothrombin, S. Magnusson, T. E. Petersen, L. Sottrup-Jensen, H. Claeys, *Cold Spring Harbor Symp. Quant. Biol.*, in press. The amino acids are abbreviated as follows: alanine, Ala, A; arginine, Arg, R; asparagine, Asn, N; aspartic acid, Asp, D; cysteine, Cys, C; glutamine, Gln, Q; glutamic acid, Glu, E; glycine, Gly, G; histidine, His, H; isoleucine, Ile, I; leucine, Leu, L; lysine, Lys, K; methionine, Met, M; phenylalanine, Phe, F; Proline, Pro, P; serine, Ser, S; threonine, Thr, T; tryptophan, Trp, W; tyrosine, Tyr, Y; and valine, Val, V.

25. We thank B. H. Bowman for encouragement and support and B. Touchstone, R. Roberts, and H. D. Kelso for technical assistance. Supported by grants HD 03321 from NICHD, AI 08518 from the National Institute of Allergy and Infectious Diseases, BMS 75-20109 from the National Science Foundation, by the J. W. McLaughlin Fellowship and Faculty funds, and the Robert A. Welch Foundation (H-378).

28 June 1976; revised 8 September 1976.

Leucine 2,3-Aminomutase: A Cobalamin-Dependent Enzyme Present in Bean Seedlings

Abstract. *Leucine 2,3-aminomutase has been demonstrated in extracts of bean seedlings. The activity of this enzyme is stimulated by coenzyme B_{12} and is inhibited by intrinsic factor. The inhibition is removed by the addition of coenzyme B_{12}. This evidence is consistent with the presence of a cobalamin-dependent enzyme in higher plants.*

The catabolism of leucine has long been recognized as proceeding by way of transamination with the subsequent decarboxylation of the resulting α-ketoisocaproic acid. A second catabolic pathway has recently been shown to function in microorganisms (*1*). In this pathway leucine is converted to β-leucine (3-amino-4-methylpentanoic acid) by leucine 2,3-aminomutase (E.C. 5.4.3.–), an enzyme that requires activation by coenzyme B_{12} [adenosylcob(III)alamin]; the β-leucine is deaminated, and the resulting β-ketoisocaproic acid is cleaved to yield acetic acid and isobutyric acid. Leucine 2,3-aminomutase is found in animal livers as well as in human leukocytes. The existence of leucine 2,3-aminomutase activity has now been established in bean seedlings and this may be the first report of a cobalamin-catalyzed activity in higher plants.

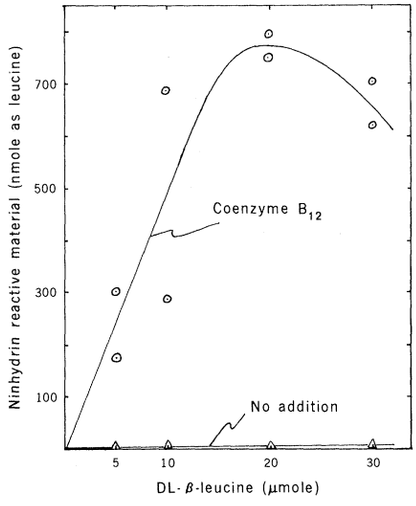

Fig. 1. Production of ninhydrin-reactive material from DL-β-leucine. Reaction conditions were those described in Table 1. After 60 minutes the reaction was stopped by the addition of 0.10 ml of 20 percent (weight to volume) of $HClO_4$. Portions (0.05 ml) of the clear deproteinized solution were taken for ninhydrin assay.

Navy beans (*Phaseolus vulgaris*) were sprouted on a bed of expanded vermiculite kept wet with distilled water and maintained at room temperature in the dark. After 5 days, the etiolated seedlings were harvested. The roots were washed free of vermiculite with distilled water and the plants were drained and weighed. The seedlings, together weighing 354 g, were homogenized in 250 ml of 0.1*M* potassium phosphate buffer, *p*H 7.0, by means of a Waring blender operating at top speed for 1.5 minutes. The homogenate was centrifuged at 16,000*g* for 40 minutes in a refrigerated centrifuge; 500 ml of cell-free extract was reserved for assay. The protein concentration of the extract was 20.3 mg per milliliter as estimated by the method of Lowry *et al.* (*2*).

Leucine 2,3-aminomutase activity was assayed by means of the reverse reaction. The α-leucine produced from the β-leucine substrate was measured by the increase in ninhydrin-reactive material or by gas chromatography after derivatization. Ninhydrin-reactive material was measured by the method of Rosen (*3*), and trimethylsilyl derivatives were made in an oil bath at 150°C, bis(trimethylsilyl)trifluoroacetamide (BSTFA) in acetonitrile (1 : 1 by volume) being used according to the method of Gehrke and Leimer (*4*). The derivatives were chromatographed on a column of 10 percent (by weight) OV-11 on 100/200 mesh Suplecoport (1.8 m by 2 mm) at a constant temperature of 110°C for 5 minutes followed by a temperature increase of 5°C per minute to a maximum temperature of 300°C.

Figure 1 shows the significant increase in the production of ninhydrin-reactive