THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 264, No. 15, Issue of May 25, pp. 8509–8519, 1989
Printed in U.S.A.
© 1989 by The American Society for Biochemistry and Molecular Biology, Inc.

The Primary Structure of Rabbit Liver Mitochondrial Serine Hydroxymethyltransferase*

(Received for publication, October 26, 1988)

Filippo Martini‡, Bruno Maras‡, Paola Tanci‡, Sebastiana Angelaccio‡, Stefano Pascarella‡, Donatella Barra‡, Francesco Bossa‡, and Verne Schirch§¶

From the ‡Dipartimento di Scienze Biochimiche and Centro di Biologia Molecolare del Consiglio Nazionale delle Richerche, Università di Roma La Sapienza, Piazzale Aldo Moro 5, 00185 Roma, Italy and the §Department of Biochemistry, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298

The complete amino acid sequence of mitochondrial serine hydroxymethyltransferase from rabbit liver was determined. The sequence was obtained from analysis of peptides isolated from chymotryptic, cyanogen bromide, and limited acid cleavages of the protein. The enzyme consists of four identical subunits, each of 475 residues, *i.e.* 8 residues shorter than the subunit of the corresponding cytosolic isoenzyme. The sequences of the two rabbit proteins are easily aligned, provided a gap of 5 residues near the amino terminus and a gap of 3 resìdues near the carboxyl terminus are included in the mitochondrial sequence. The overall degree of identity between the two isoenzymes is 61.9%, whereas the structural identity of each eukaryotic isoenzyme with the corresponding *Escherichia coli* enzyme is about 40%. The rabbit isoenzymes are about 70 residues longer than the *E. coli* enzyme, with one-half of these residues accounted for by insertions in both isoenzymes near their carboxyl terminus. Predictions of secondary structure and calculations of hydropathy profiles are also presented, suggesting an even more extensive degree of identity in the three-dimensional folding of the three proteins, in accord with the known similarity of their catalytic properties. Evidence was obtained for the existence of additional molecular forms of the mitochondrial protein, differing in the absence of some amino acid residues at the amino terminus of the polypeptide chain.

Serine hydroxymethyltransferase is a pyridoxal phosphate-containing enzyme that catalyzes the conversion of serine and tetrahydrofolate to glycine and 5,10-methylenetetrahydrofolate. It is a key enzyme in the biosynthesis of purines, lipids, hormones, and several other cell components. The enzyme is present in prokaryotic organisms, plants, and eukaryotic cells. Eukaryotic cells have been shown to contain both cytosolic and mitochondrial forms of the enzyme (1).

Both the cytosolic and mitochondrial isoenzymes from rabbit liver have been purified to homogeneity and compared with respect to reaction and substrate specificity (2–8). A study of cysteine-containing peptides from tryptic digests shows clearly that these two isoenzymes have different primary structures (5–7). Both isoenzymes catalyze the cleavage of many different 3-hydroxyamino acids, and several reactions characteristic of other pyridoxal phosphate-containing enzymes, *i.e.* decarboxylation, transamination, and racemization (8–10). Each isoenzyme is a tetramer of identical subunits, and both have isoelectric points near 7.2 (2).

Recently, the complete amino acid sequence of the cytosolic isoenzyme was determined (4). The primary structure of the corresponding *Escherichia coli* enzyme was deduced from the sequence of the *glyA* gene (11). Studies comparing the two eukaryotic isoenzymes and the *E. coli* enzyme suggest that all three enzymes have the same mechanism of action (8, 12).

The purpose of this paper is to present the complete primary structure of the mitochondrial isoenzyme from rabbit liver and to compare this sequence with those of the corresponding cytosolic and bacterial proteins, as a basis for establishing structure-function relationships in this system.

EXPERIMENTAL PROCEDURES AND RESULTS[1,2]

DISCUSSION

The complete amino acid sequence of rabbit liver mitochondrial serine hydroxymethyltransferase is reported in Fig. 1. The enzyme subunit has 475 amino acid residues, yielding a molecular weight of 52,647, which is in good agreement with the previously reported value determined from sodium dodecyl sulfate-polyacrylamide gel electrophoresis (3). Moreover, there is good agreement between the amino acid composition determined from acid hydrolysis of the protein and that deduced from sequence data (Table I).

The sequence was deduced following the isolation and the identification of an almost complete set of chymotryptic peptides which were first tentatively aligned with the sequence of the cytosolic enzyme on the basis of homology considerations and subsequently definitively ordered with the help of

* This work was supported in part by grants from Ministero Pubblica Istruzione and Consiglio Nazionale delle Ricerche (Grant 87.00056.04 (to F. B.)) and by Grant GM 28143 from the National Institutes of Health (to V. S.). Part of this work will be submitted in partial fulfillment of the requirements of the degree of Ph.D. at the Università La Sapienza, Roma (by S. P.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence and reprint requests should be addressed.

[1] Portions of this paper (including "Experimental Procedures," "Results," Tables II–XX, and Figs. 3–5) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

[2] The abbreviations used are: CmCys, *S*-carboxymethylcysteine; Hse, homoserine; Hsl, homoserine lactone; hplc, high performance liquid chromatography.

This is an Open Access article under the CC BY license.

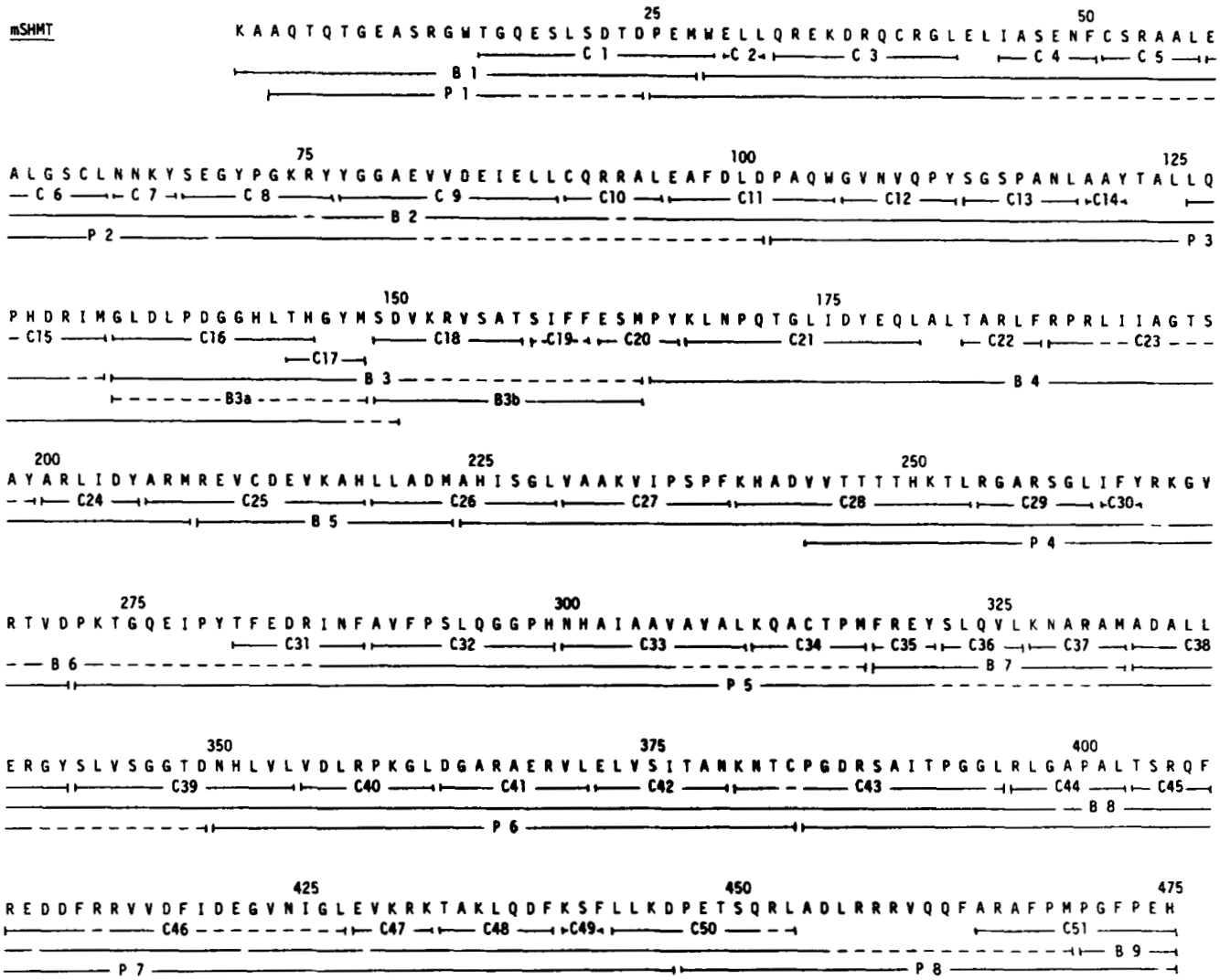


FIG. 1. **Complete amino acid sequence of mitochondrial serine hydroxymethyltransferase from rabbit liver.** ——, extent of the various fragments used to reconstruct the sequence; – – –, sequence inferred from amino acid compositions. *C*, chymotryptic peptides; *B*, cyanogen bromide peptides; *P*, limited acid hydrolysis peptides.

overlapping peptides produced from cyanogen bromide and partial acid cleavages.

Peptide purification was performed by gel filtration followed by high performance liquid chromatography on macroporous reverse phase columns. Most of the small and medium size peptides were analyzed by dansyl-Edman degradation, while for the larger peptides a gas-phase automated Sequencer was used. The latter type of analysis, by allowing quantitative evaluation of the progress of Edman degradation, was also essential for collecting sequence information from fragments, such as those generated after partial acid cleavage, which were not obtained in pure form. For identification of the carboxyl-terminal peptide, a selective procedure was applied (13).

Sequence analysis of both the entire protein and amino-terminal peptides showed that the enzyme preparation is heterogeneous, containing forms with both an amino-terminal Lys and Ala. The sequencing results suggest that the longest enzyme form contains an amino-terminal Lys and accounts for about 60% of the protein. Another 30% of the protein is missing the amino-terminal Lys, and 10% is missing the amino-terminal Lys-Ala (Figs. 1 and 2). Such microheterogeneity, not found in the cytosolic isoenzyme, which has the amino-terminal alanine acetylated, has also been observed in the case of another pyridoxal-P-dependent mitochondrial enzyme, ornithine aminotransferase from rat liver (14). This phenomenon may be an indirect demonstration that mitochondrial serine hydroxymethyltransferase is a protein coded by the nuclear genome that, simultaneously to or after translocation into the mitochondrion, undergoes a post-translational proteolytic process to remove a leader peptide (15). However, it may also be due to proteolytic processing during purification.

The amino acid sequence of rabbit liver mitochondrial serine hydroxymethyltransferase is compared in Fig. 2 with both that of the corresponding cytosolic isoenzyme and that of the protein coded by the *glyA* gene of *E. coli* (4, 11). In order to optimize the alignment, it was necessary to insert a number of gaps into the sequences. These gaps account in part for the larger size of the mammalian proteins (483 and 475 residues for the cytosolic and mitochondrial isoenzymes, respectively) compared to the bacterial protein (417 residues). The extra residues of the cytosolic enzyme with respect to the

TABLE I

Amino acid composition of mitochondrial serine hydroxymethyltransferase from rabbit liver

Acid hydrolyses were performed on unmodified serine hydroxymethyltransferase for 24, 48, and 82 h. The values of threonine and serine were obtained by extrapolation to zero time of hydrolysis. Values of valine and isoleucine were from 82-h hydrolysates.

Amino acid	Amino acid analysis	Sequence
	residues/subunit	
Aspartic acid	42.6	42[a]
Threonine	25.4	26
Serine	25.6	26
Glutamic acid	49.3	48[b]
Proline	25.8	25
Glycine	36.2	36
Alanine	49.3	50
Half-cystine	6.5[c]	7
Valine	28.3	28
Methionine	8.2	9
Isoleucine	17.8	18
Leucine	51.8	52
Tyrosine	13.3	15
Phenylalanine	18.9	19
Histidine	9.9	11
Lysine	21.8	21
Arginine	38.4	39
Tryptophan	ND[d]	3

[a] 29 aspartic acid and 13 asparagine residues.
[b] 28 glutamic acid and 20 glutamine residues.
[c] Determined as cysteic acid after performic acid oxidation.
[d] ND, not determined.

mitochondrial one appear to be mainly clustered at the amino- and carboxyl-terminal extremities of the chain, where the two eukaryotic proteins also show major structural differences with the *E. coli* enzyme. Fig. 2 shows the positions which are occupied by either the same residues in the three proteins (a total of 169 positions, *i.e.* 34.3% identity) or only in the two eukaryotic isoenzymes (299 residues, 61.9% identity). Also, the residues which are found selectively conserved between the bacterial and either the cytosolic or the mitochondrial enzyme are marked. It is interesting to note that more residues are selectively conserved in the bacterial and cytosolic proteins than in the bacterial and mitochondrial pair (37 and 25 residues, respectively).

The distribution of conserved and substituted regions is clearly not uniform throughout the three molecules. For example, the 10-residue long sequence encompassing the lysyl residue which binds pyridoxal-P (Lys-256) represents a more extended tract of structure identical for the three serine hydroxymethyltransferases. On the other hand, the amino- and carboxyl-terminal extremities are the protein regions which exhibit a lower degree of identity. For both the mitochondrial and cytosolic isoenzymes, about 70% of their increased length, with respect to the *E. coli* enzyme, is the result of large insertions at their amino and carboxyl terminus. One insertion begins at residue 20 and the other at residue 412 (Fig. 2). Even though both the rabbit isoenzymes have larger inserts at these locations, there is little homology between these inserts in the two isoenzymes. For both isoenzymes, these two large inserts are hydrophilic, suggesting they are on

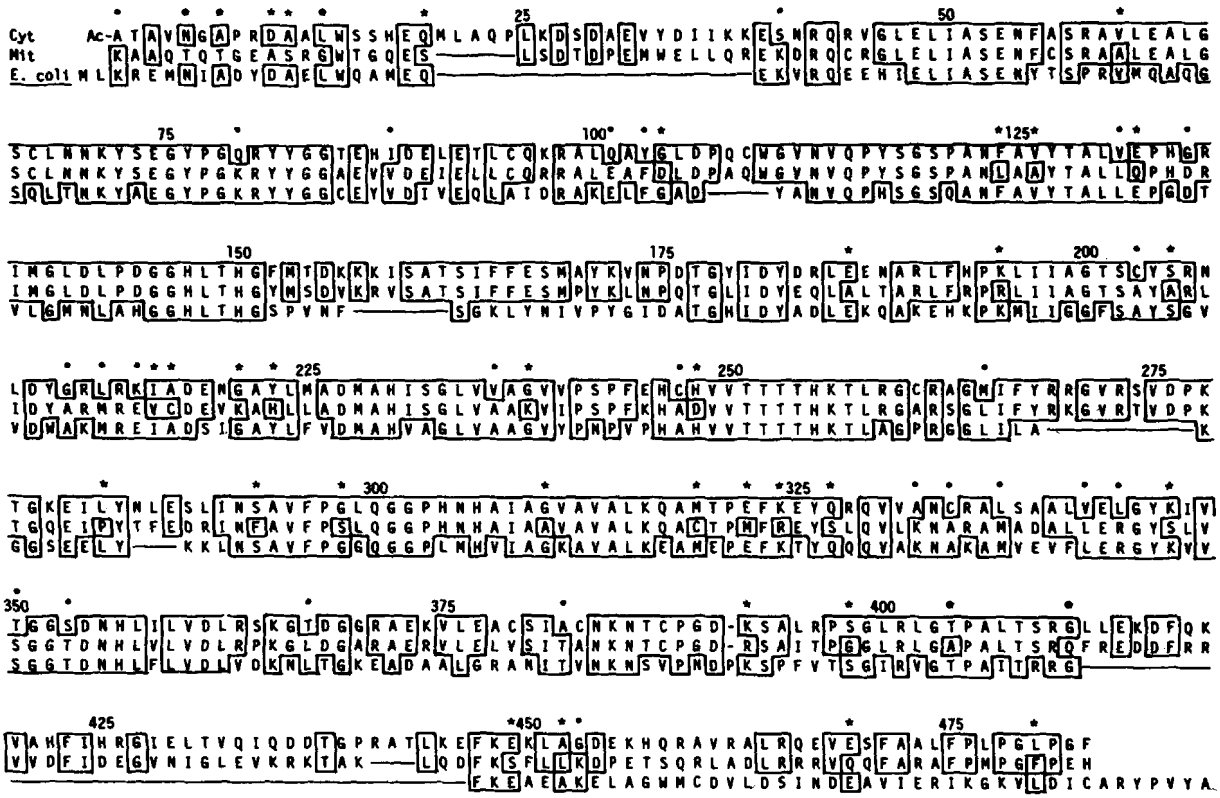


FIG. 2. **Comparison of the amino acid sequence of *E. coli* serine hydroxymethyltransferase with those of the corresponding cytosolic and mitochondrial isoenzymes from rabbit liver.** *Gaps* (——) in the sequences have been introduced to maximize the homology. *Boxes* indicate positions at which residues are identical. Residues selectively conserved between the bacterial and either the cytosolic or the mitochondrial enzyme are marked with * or ·, respectively. *Numbering* refers to the cytosolic isoenzyme.

the surface of the protein. For the cytosolic isoenzyme, we have previously shown that digestion with trypsin results in 35 residues being lost from the amino-terminal portion of the protein with a concomitant 2-fold increase in k_{cat} (16). The peptide being lost includes essentially all of the insert at the amino terminus. Although similar experiments with the mitochondrial and *E. coli* enzymes do not show this loss of amino-terminal residues and increase in activity when incubated with trypsin, the results with the cytosolic enzyme suggest that this portion of the molecule is exposed to solvent and does not play a significant role in the catalytic function of the enzyme.

If the two large inserts in the cytosolic and mitochondrial isoenzymes have a function, it is not for increased catalytic efficiency, since the *E. coli* enzyme has the largest k_{cat}/K_m value for the three enzymes (12). Other possible functions for the two inserted regions could be either to generate an allosteric site or to generate a site which interacts with some other cellular component. However, presently no allosteric effectors have been found for either the cytosolic or mitochondrial isoenzymes. Also, no evidence is currently available that either isoenzyme interacts with other proteins or membranes in the cell. The mitochondrial enzyme is transported from the cytosol to the matrix of the mitochondria, and it is possible that its inserted regions play some role in this process.

The overall degree of identity between the cytosolic and mitochondrial isoenzymes (61.9%) is significantly higher than that found for another pair of pyridoxal-P isoenzymes, *i.e.* the cytosolic and mitochondrial aspartate aminotransferases from pig heart, where the degree of identity is 47.3% (17). Moreover, comparison of the cytosolic and mitochondrial serine hydroxymethyltransferase isoenzymes with the corresponding *E. coli* enzyme shows a degree of identity of 41.8% and 40.1%, respectively. Similar values are found when comparing the two eukaryotic aspartate aminotransferase isoenzymes with the corresponding prokaryotic enzyme (39.1% for the cytosolic enzyme and 40.4% for the mitochondrial enzyme). We have also demonstrated the residues which are conserved in all three forms of both serine hydroxymethyltransferase and aspartate aminotransferase. If one excludes the insertions in the eukaryotic forms of these enzymes, the extent of identity among the three serine hydroxymethyltransferases is 41.5%, while for the three aspartate aminotransferases the identity is only 30.3%. The greater retention of residues in serine hydroxymethyltransferase suggests more extensive requirements for conservation of catalytically and/or conformationally relevant residues.

Similarity of the three-dimensional folding of the three proteins is suggested by the conservation of glycyl and prolyl residues. Out of the 44 glycyl and 23 prolyl residues present in the cytosolic enzyme, 23 and 11, respectively, occupy the same position in the mitochondrial and *E. coli* proteins. Prediction of secondary structure shows (Fig. 4) that the three enzymes are 44.7% identical by this criterion. Furthermore, the calculated hydrophilicity profile of the inserted segments (Fig. 5) is compatible with their location on the surface of the protein molecule, suggesting that the inserted segments should cause no large deviation from a catalytically competent three-dimensional folding. Thus, the common origin of the three enzymes is supported not only by the primary structures, but also by the predicted secondary structures.

General relationships based on primary structures of various eukaryotic isoenzymes and their prokaryotic counterparts have been recently reviewed by McAlister-Henn (19). The case of the serine hydroxymethyltransferase system is similar to that of citrate synthase (20), where the mitochondrial isoenzyme is also much more closely related to the cytoplasmic isoenzyme than to the *E. coli* enzyme.

REFERENCES

1. Schirch, L. (1982) *Adv. Enzymol. Relat. Areas Mol. Biol.* **53,** 83–112
2. Schirch, L. (1975) *J. Biol. Chem.* **250,** 1939–1945
3. Schirch, L., and Peterson, D. (1980) *J. Biol. Chem.* **255,** 7801–7806
4. Martini, F., Angelaccio, S., Pascarella, S., Barra, D., Bossa, F., and Schirch, V. (1987) *J. Biol. Chem.* **262,** 5499–5509
5. Schirch, L., Slagel, S., Barra, D., Martini, F., and Bossa, F. (1980) *J. Biol. Chem.* **255,** 2986–2989
6. Gavilanes, F., Peterson, D., and Schirch, L. (1982) *J. Biol. Chem.* **257,** 11431–11436
7. Gavilanes, F., Peterson, D., Bullis, B., and Schirch, L. (1983) *J. Biol. Chem.* **258,** 13155–13159
8. Shostak, K., and Schirch, V. (1988) *Biochemistry* **27,** 8007–8014
9. Palekar, A. G., Tate, S. S., and Meister, A. (1973) *J. Biol. Chem.* **248,** 1158–1167
10. Schirch, L., and Jenkins, W. T. (1964) *J. Biol. Chem.* **238,** 3801–3807
11. Plamann, M., Stauffer, L., Urbanowski, M., and Stauffer, G. (1983) *Nucleic Acids Res.* **11,** 2065–2075
12. Schirch, V., Hopkins, S., Villar, E., and Angelaccio, S. (1985) *J. Bacteriol.* **163,** 1–7
13. Barra, D., Martini, F., Angelaccio, S., Pascarella, S., Bossa, F., and Schirch, L. (1987) in *Proteins, Structure and Function* (L'Italien, J. J., ed) pp. 765–770, Plenum Publishing Corp., New York
14. Simmaco, M., Edwards, W. D., Barra, D., Bossa, F., and John, R. A. (1988) *J. Protein Chem.* **7,** 287–288
15. Schatz, G. (1987) *Eur. J. Biochem.* **165,** 1–6
16. Schirch, V., Schirch, D., Martini, F., and Bossa, F. (1986) *Eur. J. Biochem.* **161,** 45–49
17. Doonan, S., Martini, F., Angelaccio, S., Pascarella, S., Barra, D., and Bossa, F. (1986) *J. Mol. Evol.* **23,** 328–335
18. Kondo, K., Wakabayashi, S., Yagi, T., and Kagamiyama, H. (1984) *Biochem. Biophys. Res. Commun.* **122,** 62–67
19. McAlister-Henn, L. (1988) *Trends Biochem. Sci.* **13,** 178–181
20. Rosenkrantz, M., Alam, T., Kim, K. S., Clark, B. J., Srere, P. A., and Guarente, L. P. (1986) *Mol. Cell Biol.* **6,** 4509–4515
21. Barra, D., Schininà, M. E., Simmaco, M., Bannister, J. V., Bannister, W. H., Rotilio, G., and Bossa, F. (1984) *J. Biol. Chem.* **259,** 12595–12601
22. Bossa, F., Barra, D., Martini, F., Schirch, L. V., and Fasella, P. (1976) *Eur. J. Biochem.* **70,** 397–401
23. Landon, M. (1977) *Methods Enzymol.* **47,** 145–149
24. Giglio, J. R. (1977) *Anal. Biochem.* **82,** 262–264
25. Simmaco, M., Barra, D., and Bossa, F. (1985) *J. Chromatogr.* **349,** 99–103
26. Hargrave, P. A., and Wold, F. (1973) *Int. J. Peptide Protein Res.* **5,** 85–89
27. Garnier, J., Osguthorpe, D. J., and Robson, B. (1978) *J. Mol. Biol.* **120,** 97–120
28. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132
29. Pascarella, S., and Bossa, F. (1987) *CABIOS* **3,** 325–331
30. Margolies, M. N., Brauer, A. W., Kwong, R. F., and Matsueda, G. R. (1987) in *Proteins, Structure and Function* (L'Italien, J. J., ed) pp. 333–341, Plenum Publishing Corp., New York

SUPPLEMENTAL MATERIAL TO:
THE PRIMARY STRUCTURE OF RABBIT LIVER MITOCHONDRIAL
SERINE HYDROXYMETHYLTRANSFERASE

BY

FILIPPO MARTINI, BRUNO MARAS, PAOLA TANCI, SEBASTIANA ANGELACCIO,
STEFANO PASCARELLA, DONATELLA BARRA, FRANCESCO BOSSA AND VERNE SCHIRCH

EXPERIMENTAL PROCEDURES

Materials . Mitochondrial serine hydroxymethyltransferase was isolated from rabbit liver according to the method of Schirch and Peterson (3). Trypsin (code TRTPCK), chymotrypsin (code CDI), carboxypeptidase B (COBPMS) were from Worthington Biochemical Co.; carboxypeptidase Y from Boehringer GmbH; Staphylococcus aureus V8 protease from Miles. Iodo (2-^{14}C)acetate was from Radiochemical Centre, Amersham, and guanidine-HCl (from Carlo Erba) was recrystallized from methanol. Cyanogen bromide, glycinamide and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide were from Fluka. The liquid chromatography solvents, hplc grade, were from Carlo Erba.

S-carboxymethylation of the protein . Alkylation with iodo (2-^{14}C) acetate, was performed as previously described (21).

Enzymatic and chemical cleavages . A sample (50 mg) of enzyme reduced with sodium borohydride (22) and then S-carboxymethylated was suspended in 5 ml of 0.1 M ammonium bicarbonate and incubated at 37° C for 2.0 h after addition of 1 mg of chymotrypsin. The cleavage reaction was stopped by addition of glacial acetic acid to 10% final concentration. A second sample of 60 mg of S-carboxymethylated protein was dissolved in 2 ml of 70% (v/v) formic acid and incubated with 100 mg of CNBr for 20 h at room temperature in the dark. A last sample (140 mg) of alkylated protein was incubated in 15 ml of 70% formic acid plus 8 M urea for 72 h at 37° C in order to cleave the Asp-Pro bonds (23).

Digestion of the largest peptide fractions was performed with trypsin, chymotrypsin or S. aureus protease at an E/S ratio 1/30 for 3 h at 37°C in 0.1 M ammonium bicarbonate. To better solubilize the substrates to be digested, urea or acetonitrile were added to a final concentration of 1-2 M or 10-20% (v/v), respectively.

Peptide purification . The chymotryptic peptides, dissolved in 10% (v/v) acetic acid, were first fractionated by gel-filtration on a Sephadex G-25 superfine column (2.3 x 120 cm) in 10% acetic acid. The various fractions collected after Sephadex G-25 chromatography were further purified using a Beckman model 332 high performance liquid chromatography system (hplc), on macroporous reverse-phase columns (Aquapore RP-300, 4.6 or 7.0 x 250 mm, 10 um, Brownlee Labs) eluted with gradients from 0% to 70% acetonitrile in 0.2% (v/v) trifluoroacetic acid, at a flow rate of 1.0 or 3.0 ml/min. Elution of the peptides was monitored on a Beckman 165 Spectrophotometer at 220 and 280 or 325 nm. The products of the secondary fragmentation procedures were purified by hplc under similar conditions.

Peptides obtained after CNBr or limited acid hydrolysis were first fractionated on a column (2.5 x 140 cm) of Sephadex G-50 superfine in 10% acetic acid. Some fractions were digested with trypsin, chymotrypsin or with S. aureus protease either directly or after further purification by hplc as described above. The digestion mixtures were generally purified by a two-step procedure, utilizing pre-fractionation by gel-filtration (Sephadex G-25 superfine, 2.3 x 120 cm; 10% acetic acid) and final purification by reverse-phase hplc. In particular, the purification of the largest peptides from the limited acid hydrolysis was performed on an Altex column, (Ultrapore RPSC; 4.6 x 75 mm) at 1.0 ml/min flow rate.

Analytical techniques . Amino acid analyses and sequence determinations by manual dansyl-Edman procedure were performed as described in (15). Dansyl-Trp was identified after 5 h-hydrolysis at 110°C with 4 N methanesulfonic acid containing 0.2% 3-(2-aminoethyl) indole, other conditions being identical to those described in (24). The amidation states of Glu and Asp were assigned by direct identification by hplc of the phenylthiohydantoin (PTH) derivatives released during dansyl-Edman sequence analysis. The same procedure was used for the identification of the PTH-derivative of carboxymethyl-cysteine (25). In some cases, the position of carboxymethylcysteine in the sequence was confirmed by measuring the radioactivity of an aliquot of the ethyl acetate extracts after each cycle of degradation. The largest peptides were sequenced automatically by using an Applied Biosystems model 470A gas-phase sequencer equipped with an Applied Biosystems model 120A PTH analyzer for the on-line detection of PTH-amino acids.

Isolation of the C-terminal peptide . A modified version (13) of the original procedure reported by Hargrave and Wold (26) was followed. An aliquot (25 mg) of carboxymethylated enzyme was dissolved in 2.5 ml of 8 M urea solution containing 1 M glycinamide and 0.1 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide. The reaction mixture was adjusted to and maintained at pH 4.75 for 3 h at room temperature. The protein was then extensively dialyzed against 1 mM acetic acid and lyophilized. The modified protein was suspended in 2.0 ml of 0.1 M ammonium bicarbonate and first digested for 4 h at 37°C with 2.5 mg of trypsin and then with 0.4 mg of carboxypeptidase B overnight. After repeated lyophilization, the trypsin-carboxypeptidase B digest was chromatographed on a column (2 x 25 cm) of QAE-Sephadex A-25 equilibrated and eluted with 1 mM NaOH. The unretained fraction was concentrated by lyophilization and subjected to a final purification step by hplc under conditions described in the Peptide purification section.

Peptide nomenclature . The peptides were numbered retrospectively according to their location in the sequence, starting from the N-terminus. Chymotryptic peptides were designated with a C, CNBr peptides with a B and limited acid cleavage peptides with a P. All fragments obtained from digestions of B or P peptides were designated by a second letter indicating the proteolytic enzyme used (T = trypsin, C = chymotrypsin, S = S. aureus protease) and numbered according to their position in the sequence, starting from the N-terminus.

Structure predictions . Secondary structures of the three proteins were predicted according to the method of Garnier et al. (27). The decision constants were set at -75 for alpha-helix, -87.5 for beta-sheet and the others to zero. The hydropathic profiles were calculated according to Kyte and Doolittle (28). The plots utilize a span setting of 7. In both cases calculations were performed by means of simple programs written in Applesoft Basic on an Apple IIe (29).

RESULTS

N-terminal sequence of the protein . The results obtained after automated Edman degradation of a sample from a typical preparation of the protein are reported in Table II. These data, and those from the analysis of pertinent peptides, are in accord with the presence in the protein preparations of various molecular species with the following N-terminal structures:

Lys-Ala-Ala-Gln-Thr-Gln-Thr-Gly-Glu..... (60%)
Ala-Ala-Gln-Thr-Gln-Thr-Gly-Glu..... (30%)
Ala-Gln-Thr-Gln-Thr-Gly-Glu..... (10%)

Approximate proportions of the various species are indicated in brackets.

C-terminal sequence of the protein . Final hplc purification of the peptide fraction obtained after using the modified procedure of Hargrave and Wold (13) allowed the isolation of two basic peptides (Fig. 3). Analysis of one of these (peak 1) revealed the sequence: Leu-Phe-Arg-Pro-Arg. This is identical to the sequence of the polypeptide chain from position 186 to 190, suggesting that this peptide is not from the C-terminus. Automated Edman degradation of the second basic peptide (peak 2) gave the following structure: Ala-Phe-Pro-Met-Pro-Gly-Phe-Pro-Xxx-His-Gly. PTH-glycine, released at the 11th cycle of Edman degradation, originated from the glycinamide moiety added in the course of the above described procedure, thus explaining the basicity of the isolated peptide and allowing its assignment to the C-terminus of the protein.

Chymotryptic peptides . Analytical data of the almost complete set of the chymotryptic peptides are summarized in Tables III-V.

Cyanogen bromide peptides . Analytical data on CNBr peptides are reported in Tables VI-VIII. Amino acid composition and sequence information on fragments obtained after digestion of some CNBr peptides with various proteolytic enzymes are reported in Tables IX-XII.

Peptide B1 was sequenced by automated Edman degradation up to residue 7 (Table VIII). An aliquot was digested with trypsin and fragments T1 and T2, corresponding to residues 2-12 and 13-27 respectively, were purified and analyzed (Tables IX and X). Sequence information on peptide B1 is summarized in Table VII.

Peptide B2 was sequenced by automated Edman degradation up to residue 58 (Table VIII). An aliquot was digested with trypsin and the fragments were purified and analyzed. For the sake of clarity, only analytical data on fragments T1, T2, and T3, corresponding to residues 54-74, 76-92, 94-131, are reported in Tables IX-XI. Sequence information on peptide B2 is summarized in Table VII.

Peptide B3 was sequenced by automated Edman degradation up to residue 151 (Table VIII). Peptide B3b was completely sequenced (same Table) thus allowing the complete reconstruction of peptide B3.

Peptides B4 and B5 were completely sequenced by automated Edman degradation (Table VII and VIII).

Peptide B6 was sequenced by automated Edman degradation up to residue 263 (Table VIII). An aliquot was digested with trypsin and fragments T1, T2, and T3, corresponding to residues 265-268, 269-273 and 286-311, were isolated and analyzed (Tables IX and X). Sequence information on peptide B6 is summarized in Table VII.

Peptide B7 was sequenced by manual dansyl-Edman degradation up to residue 331.

Peptide B8 was sequenced by automated Edman degradation up to residue 374 (Table VIII). An aliquot was digested with S. aureus protease and fragments S1, S2, S3, S4, S5, S6, and S7, corresponding to residues 339-356, 373-398, 399-409, 410-421, 422-428, 429-448, 449-455, respectively, were isolated and analyzed (Tables IX and XII). Another aliquot of B8 was digested with trypsin. The amino acid composition of peptide T1 (position 441-452), which was sequenced by manual dansyl-Edman degradation, is reported in Table IX. Sequence information on peptide B8 is summarized in Table VII.

Peptide B9 was completely sequenced by manual dansyl-Edman degradation.

Peptides resulting from limited acid hydrolysis . The experimental procedure described above was not so specific as to cleave only the Asp-Pro peptide bonds. In fact, we have observed good cleavage at the CmCys-Pro bond in position 383-384. Moreover, cleavages to a minor extent were observed at Asp-Val (150-151 and 243-244), Asp-Asn (349-350), Asp-Leu (356-357) and Asp-Gly (363-364) bonds. It must be noted that only the peptides required to get overlapping sequence data and/or to integrate the structural information obtained from the protein fragmentation pattern reported above, were analyzed. Due to the complexity of the mixture and to the large size of the peptides, in some cases it was not possible to obtain pure fractions. For this reason in Table XIII only the amino acid compositions of peptides P1, P2, P4, P5, P7 and P8, which were obtained in pure form, are reported. From sequence analysis, peptide P3 appeared to be contaminated by P7 and peptide P6 by P2. However, sequence data were obtained also for peptides P3 and P6, by subtracting at each cycle the PTH-derivative originated from the known sequence of the contaminating peptide. Moreover, in the course of the automated Edman degradation of peptide P3, after the 9th cycle, the filter was treated with o-phthalaldehyde in order to block primary amines, thus interrupting the sequencing of all the peptide molecules except P3, which has proline as N-terminus in that position (30).

Analytical data for limited acid cleavage peptides are reported in Tables XIII-XVIII. Analytical data for fragments obtained after digestion of some P peptides with various proteolytic enzymes are reported in Tables XIX and XX.

Peptide P1. Only a short version of peptide P1 from the N-terminus (3-24) was purified and sequenced up to residue 17 (Table XV).

Peptide P2 was sequenced by automated Edman degradation up to residue 46 (Table XV). An aliquot was digested with S. aureus protease and peptides S1, S2 and S3, corresponding to residues 58-69, 70-81 and 86-96, were isolated and analyzed (Tables XIX and XX). Sequence information on peptide P2 is summarized in Table XIV.

Peptide P3 was sequenced by automated Edman degradation up to residue 147 (Table XVI).

Peptide P4 was sequenced by automated Edman degradation up to residue 267 (Table XV). An aliquot was digested with chymotrypsin and the analytical data of fragment C1 (position 264-271) are reported in Tables XIX and XX. Sequence information on peptide P4 is summarized in Table XIV.

Peptide P5 was sequenced by automated Edman degradation up to residue 321 (Table XVII). An aliquot was digested with trypsin to isolate peptide T1 (position 331-339), which was necessary to provide the overlap between peptides B7 and B8. Analytical data on this peptide are reported in Tables XIX and XX. Sequence information on peptide P5 is summarized in Table XIV.

Peptide P6 was completely sequenced by automated Edman degradation (Table XV).

Peptide P7 was sequenced by automated Edman degradation up to residue 431 (Table XVIII). An aliquot was digested with S. aureus protease and the fragment S1, corresponding to residues 429-446, was isolated and analyzed (Table XIX and XX). Sequence information on peptide P7 is summarized in Table XIV.

Peptide P8 was sequenced by automated Edman degradation up to residue 474 (Table XV).

Structure predictions. Prediction of the secondary structure and the hydropathy profiles for E. coli and rabbit liver serine hydroxymethyltransferases are shown in Figs. 4 and 5, respectively. The three sequences were aligned according to Fig. 2. The predicted secondary structure identity among the three proteins is 36.9%. The structural identity between either the mitochondrial or the cytosolic and the bacterial enzymes is 49.4% or 47.9%, respectively. Between the two eukaryotic proteins the identity is 63.1%. In all cases, the predicted secondary structure identity is higher than the corresponding sequence identity. The predicted content in alpha-helix and extended conformation for the mitochondrial enzyme is 41.9% and 39.2%, respectively.

TABLE II

Automated Edman Degradation of Native Serine Hydroxymethyltransferase (0.5 nmol)

Cycle no.	PTH-aa. Yield (pmol)					
	Lys	Ala	Gln	Thr	Gly	Glu
1	60	30	5	NQ	18	8
2	17	87	13	NQ	14	12
3	9	75	28	6	14	12
4	10	30	46	NQ	17	12
5	9	22	25	13	22	13
6	9	20	36	6	17	16
7	10	18	21	11	21	17
8	10	17	18	7	34	17
9	9	18	10	5	24	29

N-terminal sequence of the protein: Lys-Ala-Ala-Gln-Thr-Gln-Thr-Gly-Glu...

TABLE III

Amino Acid Composition of Chymotryptic Peptides

The composition from sequence analysis of each peptide is indicated by the numbers in parentheses

Peptide	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	C 10	C 11
Residue nos.	15-28	29-31	32-42	45-50	51-56	57-63	64-67	68-76	77-89	90-95	96-105
CmCys			0.9(1)		0.8(1)	1.0(1)				0.8(1)	
Asp	2.0(2)		1.2(1)	1.0(1)			1.9(2)		1.0(1)		1.9(2)
Thr	1.7(2)										
Ser	1.9(2)			1.0(1)	1.2(1)	0.8(1)		0.9(1)			
Glu	3.1(3)	1.0(1)	2.9(3)	1.1(1)		1.2(1)		1.1(1)	3.0(3)	1.1(1)	2.1(2)
Pro	1.1(1)							1.0(1)			1.1(1)
Gly	1.2(1)		1.3(1)			1.0(1)		2.1(2)	2.0(2)		
Ala				1.1(1)	1.9(2)	1.1(1)			1.1(1)	1.1(1)	2.0(2)
Val									1.6(2)		
Met	0.9(1)										
Ile				1.0(1)					0.9(1)		
Leu	1.1(1)	1.9(2)	0.9(1)		1.0(1)	2.1(2)			1.8(2)	1.0(1)	1.0(1)
Tyr							0.9(1)	1.8(2)	0.9(1)		
Phe				0.9(1)							0.9(1)
His											
Lys			1.0(1)				1.0(1)	1.0(1)			
Arg			2.7(3)		1.0(1)			1.2(1)		2.0(2)	
Trp	+ (1)										+ (1)
Yield %	58.3	15.6	3.1	14.2	3.5	18.1	58.9	3.7	10.1	12.0	25.0
N-terminal residue	Thr	Glx	Gln	Ile	CmCys	Glu	Asn	Ser	Tyr	CmCys	Glu

Peptide	C 12	C 13	C 14	C 15	C 16	C 17	C 18	C 19	C 20	C 21	C 22
Residue nos.	106-112	113-119	120-122	126-133	134-145	144-148	149-157	158-161	162-166	167-180	183-187
CmCys											
Asp	1.0(1)	1.1(1)		1.1(1)	2.2(2)		1.1(1)			2.0(2)	
Thr					1.0(1)	0.8(1)	1.0(1)			1.0(1)	0.7(1)
Ser		1.8(2)					1.5(2)	1.0(1)	1.0(1)		
Glu	1.1(1)			1.1(1)					1.0(1)	3.3(3)	
Pro	1.0(1)	1.0(1)		1.1(1)	1.1(1)				0.8(1)	1.1(1)	
Gly	1.0(1)	1.1(1)			3.0(3)	1.1(1)					1.1(1)
Ala		1.1(1)	2.0(2)				1.0(1)				0.9(1)
Val	1.9(2)						1.4(2)				
Met				0.6(1)		0.7(1)			1.0(1)		
Ile				0.9(1)				1.0(1)		0.8(1)	
Leu		1.0(1)		1.2(1)	2.8(3)					3.0(3)	1.0(1)
Tyr	0.6(1)		0.8(1)			1.0(1)			0.7(1)	0.8(1)	
Phe								2.2(2)			1.0(1)
His				0.9(1)	2.0(2)	1.1(1)					
Lys							1.1(1)			1.0(1)	
Arg				0.9(1)			1.0(1)			1.0(1)	
Trp											
Yield %	20.3	20.1	6.4	33.1	53.3	2.9	3.5	1.5	5.0	13.6	8.2
N-terminal residue	Gly	Ser	Ala	Leu	Gly	Thr	Ser	Ser	Glu	Lys	Thr

Peptide	C 23	C 24	C 25	C 26	C 27	C 28	C 28a	C 29	C 30	C 31
Residue nos.	188-199	200-205	206-218	219-229	230-239	240-253	248-253	254-260	261-263	281-288
CmCys			0.8(1)							
Asp		1.1(1)	1.0(1)	1.0(1)		1.0(1)				1.9(2)
Thr	0.8(1)					4.3(5)	2.5(3)			1.1(1)
Ser	0.8(1)			0.8(1)	1.1(1)			0.8(1)		
Glu			2.1(2)							1.1(1)
Pro	0.7(1)				1.9(2)					
Gly	1.1(1)			1.1(1)				2.1(2)		
Ala	2.3(2)	1.0(1)	1.7(2)	1.8(2)	2.1(2)	1.0(1)		1.0(1)		
Val			1.7(2)		1.8(2)	1.3(2)				
Met			0.7(1)	1.0(1)						
Ile	1.4(2)	1.1(1)		0.8(1)	0.9(1)				0.9(1)	1.0(1)
Leu	0.7(1)	0.9(1)		3.1(3)		0.9(1)	1.0(1)	1.0(1)		
Tyr	1.3(1)	0.6(1)							1.1(1)	
Phe					1.1(1)				1.0(1)	2.0(2)
His			0.8(1)	0.8(1)		2.0(2)	1.1(1)			
Lys			0.9(1)		1.1(1)	1.6(2)	0.9(1)			
Arg	2.2(2)	1.1(1)	1.7(2)					2.1(2)		1.0(1)
Trp										
Yield %	32.0	13.3	2.0	10.0	15.1	3.3	3.4	20.5	4.8	60.0
N-terminal residue	Arg	Ala	Ala	Leu	Val	Lys	Thr	Arg	Ile	Thr

TABLE III (continued)

Peptide	C 32	C 33	C 34	C 35	C 36	C 37	C 38	C 39	C 40	C 41
Residue nos.	289-299	300-310	311-317	318-321	322-326	327-332	333-341	342-354	355-362	363-371
CmCys			0.9(1)							
Asp		1.0(1)				1.0(1)	1.0(1)	2.0(2)	1.0(1)	0.9(1)
Thr			1.0(1)					0.9(1)		
Ser	0.9(1)				0.9(1)			1.6(2)		
Glu	0.9(1)		1.0(1)	1.0(1)	1.1(1)		1.1(1)			1.1(1)
Pro	1.6(2)		0.9(1)						0.9(1)	
Gly	1.8(2)						1.0(1)	2.0(2)	1.2(1)	1.1(1)
Ala	1.1(1)	4.7(5)	1.1(1)			2.2(2)	1.7(2)			2.0(2)
Val	0.8(1)	2.0(2)			0.9(1)			2.0(2)	0.9(1)	1.0(1)
Met			0.9(1)			0.8(1)				
Ile		1.0(1)								
Leu	1.0(1)	1.1(1)			2.0(2)		2.0(2)	2.7(3)	1.9(2)	1.0(1)
Tyr				0.9(1)			0.6(1)			
Phe	0.7(1)			0.7(1)						
His	0.7(1)	1.0(1)						0.9(1)		
Lys			0.9(1)			1.0(1)			0.9(1)	
Arg				0.7(1)		1.1(1)	1.1(1)		1.1(1)	2.0(2)
Trp										
Yield %	52.0	4.6	25.6	6.8	20.9	30.3	40.3	18.0	35.0	19.8
N-terminal residue	Ala	Asn	Lys	Phe	Ser	Lys	Ala	Ser	Val	Asp

Peptide	C 42	C 43	C 44	C 45	C 46	C 47	C 48	C 49	C 50	C 51
Residue nos.	372-379	380-395	396-402	403-407	408-427	428-432	433-439	440-442	443-453	464-475
CmCys		1.1(1)								
Asp	1.0(1)	2.2(2)			5.0(5)		1.1(1)		1.0(1)	
Thr	0.9(1)	1.9(2)		1.0(1)			0.8(1)		1.0(1)	
Ser	1.0(1)	1.0(1)		0.8(1)				1.0(1)	0.9(1)	
Glu	1.0(1)			1.0(1)	2.6(2)	1.2(1)	1.0(1)		2.1(2)	1.0(1)
Pro		2.1(2)	1.1(1)						1.0(1)	2.7(3)
Gly		3.2(3)	1.1(1)		2.2(2)					1.1(1)
Ala	1.1(1)	1.0(1)	2.0(2)				1.1(1)			2.0(2)
Val	0.8(1)				2.6(3)	0.9(1)				
Met										0.9(1)
Ile	0.7(1)	0.9(1)			1.4(2)					
Leu	1.1(1)	1.1(1)	1.9(2)		1.0(1)		1.0(1)		2.7(3)	
Tyr										
Phe				0.8(1)	2.0(2)		0.8(1)	1.1(1)		2.0(2)
His										0.9(1)
Lys		1.0(1)				2.1(2)	1.0(1)	1.0(1)	1.1(1)	
Arg		1.0(1)	1.1(1)	1.0(1)	3.0(3)	1.0(1)			1.1(1)	1.0(1)
Trp										
Yield %	8.2	5.8	10.7	28.1	6.2	5.8	16.8	17.1	22.4	13.8
N-terminal residue	Glx	Lys	Arg	Thr	Arg	Glu	Thr	Lys	Leu	Ala

Presence of tryptophan was indicated by absorbance at 280 nm.

TABLE IV

Summary of Sequence Studies on Chymotryptic Peptides

C 1	(15-28)	*GQESLSDTDPEMW
C 2	(29-31)	Z(L,L)
C 3	(32-42)	QREKDRQCRGL
C 4	(45-50)	TASENF
C 5	(51-56)	CSRAAL
C 6	(57-63)	EALGSCL
C 7	(64-67)	NNKY
C 8	(68-76)	SEGYPGKRY
C 9	(77-89)	YGGAEVVDEIELL
C 10	(90-95)	CQRRAL
C 11	(96-105)	EAFDLDPAQW
C 12	(106-112)	GVNVQPY
C 13	(113-119)	SGSPANL
C 14	(120-122)	AAY
C 15	(126-133)	LQPHDRIM
C 16	(134-145)	GLDLPDGGHLTH
C 17	(144-148)	THGYM
C 18	(149-157)	SDVKRVSAT
C 19	(158-161)	S([,F,F]
C 20	(162-166)	ESMPY
C 21	(167-180)	KLNPQTGLIDYEQL
C 22	(183-187)	TARLF
C 23	(188-199)	RPRLIIAGTSAY
C 24	(200-205)	ARL:BY
C 25	(206-218)	ARMREVCDEVKAH
C 26	(219-229)	LLADMAHISGL
C 27	(230-239)	VAAKVIPSPF
C 28	(240-253)	KHADVVTTTTHKTL
C 28a	(248-253)	TTHKTL
C 29	(254-260)	RGARSGL
C 30	(261-263)	I([F,Y)
C 31	(281-288)	TFEDRINF
C 32	(289-299)	AVFPSLQGGPH
C 33	(300-310)	NHAIAAVAVAL
C 34	(311-317)	KQACTPM
C 35	(318-321)	FREY
C 36	(322-326)	SLQVL
C 37	(327-332)	KNARAM
C 38	(333-341)	ADALLERGY
C 39	(342-354)	SLVSGGTBBHLVL
C 40	(355-362)	VDLRPKGL
C 41	(363-371)	DGARAERVL
C 42	(372-379)	ELVSITAN
C 43	(380-395)	KBTCPGDRSAITPGGL
C 44	(396-402)	RLGAPAL
C 45	(403-407)	TSRQF
C 46	(408-427)	RZBBFRRVVBFIBZGVBIGL
C 47	(428-432)	EVKRK
C 48	(433-439)	TAKLQDF
C 49	(440-442)	KSF
C 50	(443-453)	LLKDPETSZRL
C 51	(464-475)	ARAFPMPGFPEH

The residues above the arrow were identified by dansyl-Edman degradation (→) or by automated Edman degradation on a gas-phase sequencer (⇝).

TABLE V

Automated Edman Degradation of Chymotryptic Peptides

Cycle No.	C 1 (1 nmol) PTH-aa	Yield (pmol)	C 3 (0.5 nmol) PTH-aa	Yield (pmol)	C 9 (0.5 nmol) PTH-aa	Yield (pmol)	C 10 (1 nmol) PTH-aa	Yield (pmol)	C 16 (0.5 nmol) PTH-aa	Yield (pmol)
1	Thr	231	Gln	129	Tyr	133	CmCys	399	Gly	41
2	Gly	203	Arg	46	Gly	125	Gln	513	Leu	58
3	Gln	186	Glu	92	Gly	133	Arg	403	Asp	64
4	Glu	193	Lys	26	Ala	170	Arg	435	Leu	46
5	Ser	125	Asp	51	Glu	84	Ala	447	Pro	64
6	Leu	170	Arg	24	Val	117	Leu	215	Asp	43
7	Ser	116	Gln	42	Val	129			Gly	35
8	Asp	106	CmCys	NQ[a]	Asp	63			Gly	44
9	Thr	99	Arg	17	Glu	65			His	19
10	Asp	96	Gly	35	Ile	84			Leu	11
11	Pro	129	Leu	14	Glu	19			Thr	10
12	Glu	98			Leu	43			His	6
13	Met	100			Leu	47				
14	Trp	33								
Repetitive yield %	(Asp 8-Asp 10) 95.2		(Arg 6-Arg 9) 89.1		(Glu 5-Glu 9) 93.8		(Gln 2-Ala 5) 95.5		(Asp 3-Asp 6) 87.6	
Position	15-28		32-42		77-89		90-95		134-145	

Cycle No.	C 18 (0.5 nmol) PTH-aa	Yield (pmol)	C 20 (0.5 nmol) PTH-aa	Yield (pmol)	C 25 (0.5 nmol) PTH-aa	Yield (pmol)	C 28a(0.5 nmol) PTH-aa	Yield (pmol)	C 42 (1 nmol) PTH-aa	Yield (pmol)
1	Ser	54	Glu	76	Ala	28	Thr	189	Glu	211
2	Asp	35	Ser	71	Arg	23	Thr	100	Leu	322
3	Val	56	Met	66	Met	25	His	51	Val	178
4	Lys	39	Pro	47	Arg	25	Lys	39	Ser	283
5	Arg	44	Tyr	40	Glu	25	Thr	59	Ile	203
6	Val	40			Val	20	Leu	41	Thr	26
7	Ser	25			CmCys	18			Ala	194
8	Ala	21			Asp	10			Asn	67
9	Thr	11			Glu	12				
10					Val	12				
11					Lys	4				
12					Ala	7				
13					His	NQ[a]				
Repetitive yield %	(Val 3-Val 6) 89.3		(Glu 1-Met 3) 93.2		(Val 6-Val 10) 88.0		(Thr 2-Thr 5) 83.9		(Leu 2-Ile 5) 85.7	
Position	149-157		162-166		206-218		248-253		372-379	

[a]NQ: observed but not quantitated.

TABLE VI

Amino Acid Composition of CNBr Peptides

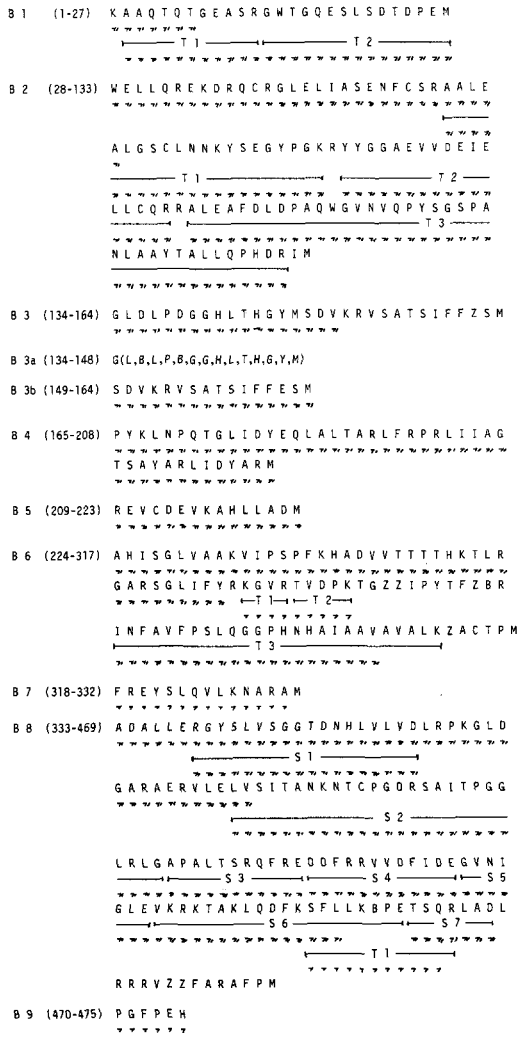
The composition from sequence analysis of each peptide is indicated by the numbers in parentheses

Peptide	B 1	B 2	B 3	B 3a	B 3b	B 4	B 5	B 6	B 7	B 8	B 9
Residue nos.	1-27	28-133	134-164	134-148	149-164	165-208	209-223	224-317	318-332	333-469	470-475
CmCys		2.6(4)					0.5(1)	0.6(1)		0.7(1)	
Asp	2.0(2)	10.0(10)	3.0(3)	2.0(2)	1.0(1)	3.0(3)	1.6(2)	5.4(5)	1.2(1)	14.8(16)	
Thr	3.1(4)	0.7(1)	1.7(2)	0.9(1)	0.7(1)	2.5(3)		6.4(9)		6.8(7)	
Ser	2.6(3)	5.7(6)	3.7(4)		3.2(4)	1.1(1)		3.8(4)	0.8(1)	5.8(7)	
Glu	6.0(6)	17.3(16)	1.6(1)		1.4(1)	3.8(3)	1.9(2)	5.8(5)	2.6(2)	12.4(12)	1.1(1)
Pro	0.8(1)	4.4(5)	1.2(1)	0.9(1)		3.0(3)		7.2(7)		6.5(6)	1.8(2)
Gly	3.1(3)	8.2(8)	3.8(4)	3.6(4)		2.2(2)		7.0(7)		11.1(11)	1.0(1)
Ala	2.5(3)	12.0(12)	1.2(1)		0.8(1)	5.4(6)	2.1(2)	11.4(12)	2.0(2)	11.7(12)	
Val		3.4(4)	2.0(2)		1.7(2)		2.0(2)	7.6(9)	1.0(1)	9.3(10)	
Ile		3.3(3)	1.0(1)		0.8(1)	2.9(4)		4.8(6)		3.6(4)	
Leu	1.0(1)	13.5(14)	3.1(3)	3.1(3)		6.3(7)	2.1(2)	5.4(5)	2.1(2)	16.4(18)	
Tyr		4.8(6)	0.6(1)	0.7(1)		3.1(4)		1.2(2)	0.6(1)	1.0(1)	
Phe		2.1(2)	1.4(2)		1.6(2)	1.1(1)		3.8(5)	0.7(1)	5.8(7)	0.8(1)
His		1.1(1)	1.8(2)	1.7(2)			0.5(1)	5.2(5)		0.6(1)	0.9(1)
Lys	0.7(1)	3.5(3)	1.2(1)		0.8(1)	1.0(1)	0.9(1)	4.8(6)	0.8(1)	6.2(7)	
Arg	0.9(1)	7.9(8)	1.1(1)		0.9(1)	5.2(5)	1.0(1)	5.2(5)	1.8(2)	14.5(16)	
Trp	+ (1)	+ (2)									
Hse/Hsl	+ (1)	+ (1)	+ (2)	+ (1)	+ (1)	+ (1)	+ (1)	+ (1)	+ (1)	+ (1)	
Yield %	12.5	35.8	2.1	10.6	2.9	66.0	48.0	34.0	10.0	45.0	18.7
N-terminal residue	ND	Trp	Gly	Gly	Ser	Pro	Arg	Ala	Phe	Ala	Pro

Presence of tryptophan was indicated by absorbance at 280 nm.

TABLE VII

Summary of Sequence Studies on CNBr Peptides



The residues above the arrow were identified by dansyl-Edman degradation (→) or by automated Edman degradation on a gas-phase sequencer (⇒). Subfragments obtained after digestion with trypsin (T) and S. aureus V 8 protease (S) are indicated by solid lines.

TABLE VIII

Automated Edman Degradation of Cyanogen Bromide Peptides

Cycle No.	B 1 (1 nmol) PTH-aa	B 1 (1 nmol) Yield (pmol)	B 2 (1 nmol) PTH-aa	B 2 (1 nmol) Yield (pmol)	B 3 (1 nmol) PTH-aa	B 3 (1 nmol) Yield (pmol)	B 3b (1.5 nmol) PTH-aa	B 3b (1.5 nmol) Yield (pmol)	B 4 (1 nmol) PTH-aa	B 4 (1 nmol) Yield (pmol)	B 5 (1.5 nmol) PTH-aa	B 5 (1.5 nmol) Yield (pmol)	B 6 (1 nmol) PTH-aa	B 6 (1 nmol) Yield (pmol)	B 8 (1 nmol) PTH-aa	B 8 (1 nmol) Yield (pmol)
1	Lys	143	Trp	163	Gly	102	Ser	698	Pro	216	Arg	408	Ala	266	Ala	125
2	Ala	194	Glu	143	Leu	136	Asp	397	Tyr	287	Glu	638	His	110	Asp	96
3	Ala	181	Leu	195	Asp	72	Val	434	Lys	228	Val	560	Ile	154	Ala	134
4	Gln	147	Leu	192	Leu	91	Lys	398	Leu	223	CmCys	594	Ser	127	Leu	156
5	Thr	31	Gln	115	Pro	98	Arg	363	Asn	116	Asp	435	Gly	118	Leu	142
6	Gln	114	Arg	57	Asp	35	Val	348	Pro	194	Glu	500	Leu	142	Glu	108
7	Thr	22	Glu	74	Gly	59	Ser	193	Gln	157	Val	388	Val	98	Arg	29
8			Lys	118	Gly	64	Ala	402	Thr	141	Lys	335	Ala	108	Gly	91
9			Asp	43	His	13	Thr	216	Gly	138	Ala	477	Ala	132	Tyr	77
10			Arg	59	Leu	32	Ser	372	Leu	160	His	232	Lys	120	Ser	29
11			Gln	107	Thr	41	Ile	138	Ile	100	Leu	274	Val	65	Leu	72
12			CmCys	NQ[a]	His	NQ[a]	Phe	121	Asp	87	Leu	293	Ile	98	Val	53
13			Arg	41	Gly	41	Phe	138	Tyr	148	Ala	351	Pro	103	Ser	26
14			Gly	85	Tyr	11	Glu	85	Glu	69	Asp	140	Ser	96	Gly	41
15			Leu	92	Met	7	Ser	63	Gln	113	Hse	87	Pro	82	Gly	46
16			Glu	31	Ser	9	Hse	37	Leu	104			Phe	50	Thr	22
17			Leu	106	Asp	12			Ala	140			Lys	43	Asp	29
18			Ile	57	Val	4			Leu	98			His	22	Asn	19
19			Ala	63					Thr	69			Ala	43	His	12
20			Ser	32					Ala	105			Asp	26	Leu	19
21			Glu	28					Arg	63			Val	26	Val	12
22			Asn	29					Leu	70			Val	38	Leu	29
23			Phe	24					Phe	55			Thr	22	Val	12
24			CmCys	6					Arg	46			Thr	26	Asp	17
25			Ser	16					Pro	53			Thr	27	Leu	24
26			Arg	9					Arg	34			Thr	26	Arg	NQ[a]
27			Ala	29					Val	51			His	10	Pro	22
28			Ala	35					Ile	64			Lys	12	Lys	14
29			Leu	32					Ile	74			Thr	14	Gly	22
30			Glu	9					Ala	89			Leu	22	Leu	22
31			Ala	15					Gly	48			Arg	10	Asp	14
32									Thr	38			Gly	19	Gly	19
33									Ser	26			Ala	17	Ala	22
34									Ala	56			Arg	9	Arg	10
35									Tyr	46			Ser	7	Ala	22
36									Ala	58			Gly	12	Glu	10
37									Arg	33			Leu	15	Arg	NQ[a]
38									Leu	34			Ile	7	Val	12
39									Ile	28			Phe	10	Leu	12
40									Asp	24			Tyr	5	Glu	7
41									Tyr	33					Leu	19
42									Ala	41					Val	10
43									Arg	23						
44									Hse	8						
Repetitive yield		93.3 (Ala 2-Ala 3)		90.8 (Glu 7-Glu 16)		86.1 (Gly 7-Gly 13)		92.9 (Val 3-Val 6)		94.6 (Leu 4-Leu 10)		94.9 (Glu 2-Glu 6)		92.1 (Ala 8-Ala 19)		89.5 (Leu 4-Leu 11)
Position		1-27		28-133		134-164		149-164		165-208		209-223		224-317		333-469

[a]NQ: observed but not quantitated.

TABLE IX

Amino Acid Composition of Fragments Obtained after Digestion of Cyanogen Bromide Peptides

The composition from sequence analysis of each peptide is indicated by the number in parentheses

	B 1		B 2			B 6		
Peptide	T 1	T 2	T 1	T 2	T 3	T 1	T 2	T 3
residue nos.	2-12	13-27	54-74	76-92	94-131	265-268	269-273	286-311
CmCys			0.5(1)	0.7(1)				
Asp		2.0(2)	1.9(2)	1.0(1)	4.6(5)		1.1(1)	2.0(2)
Thr	1.5(2)	2.0(2)			1.1(1)		0.8(1)	
Ser	1.1(1)	2.0(2)	1.8(2)		1.7(2)			0.8(1)
Glu	3.0(3)	3.7(3)	2.1(2)	3.5(4)	4.2(4)			1.1(1)
Pro		0.8(1)	1.0(1)		3.5(4)		1.2(1)	2.1(2)
Gly	1.1(1)	1.8(2)	3.1(3)	2.0(2)	2.1(2)	1.0(1)		2.2(2)
Ala	2.8(3)		3.0(3)	1.2(1)	6.2(7)			6.2(6)
Val				0.8(1)	2.0(2)	0.8(1)	1.0(1)	2.5(3)
Ile				0.7(1)				1.4(2)
Leu		1.0(1)	2.8(3)	1.7(2)	4.4(5)			2.0(2)
Tyr			1.7(2)	1.4(2)	1.2(2)			
Phe					1.1(1)			1.7(2)
His					0.8(1)			1.6(2)
Lys			1.9(2)			0.9(1)	1.0(1)	0.9(1)
Arg	1.1(1)			0.9(1)	1.0(1)	1.0(1)		
Trp		+ (1)			+ (1)			
Hse/Hsl		+ (1)						
Yield %	19.5	38.0	24.6	10.0	26.0	10.0	11.0	9.5
N-terminal residue	Ala	Gly	Ala	Tyr	Ala	Lys	Thr	Ile

	B 8							
Peptide	T 1	S 1	S 2	S 3	S 4	S 5	S 6	S 7
residue nos.	441-452	339-356	373-398	399-409	410-421	422-428	429-448	449-455
CmCys			0.7(1)					
Asp	1.0(1)	3.0(3)	3.0(3)		3.6(4)	1.1(1)	2.0(2)	1.0(1)
Thr	0.9(1)	0.8(1)	2.7(3)	0.7(1)			0.9(1)	0.8(1)
Ser	1.7(2)	2.0(2)	1.9(2)	0.9(1)			0.8(1)	1.0(1)
Glu	2.1(2)			1.9(2)	1.1(1)	0.9(1)	2.1(2)	1.1(1)
Pro	1.0(1)		2.0(2)	1.0(1)			1.1(1)	
Gly		2.9(3)	3.8(4)			2.2(2)		
Ala			2.0(2)	2.0(2)			1.0(1)	0.7(1)
Val		2.8(3)	0.9(1)		1.5(2)	1.0(1)	0.9(1)	
Ile			1.3(2)		1.0(1)	0.8(1)		
Leu	2.0(2)	3.0(3)	3.0(3)	0.8(1)		0.9(1)	3.0(3)	0.8(1)
Tyr		0.7(1)						
Phe	0.9(1)			0.8(1)	2.0(2)		1.8(2)	
His		0.8(1)						
Lys	0.9(1)		0.7(1)	1.8(2)			4.5(5)	
Arg	1.0(1)	0.8(1)	1.6(2)		2.1(2)		0.9(1)	0.8(1)
Yield %	16.5	11.3	3.7	15.0	10.0	11.2	4.5	7.2
N-terminal residue	Ser	Arg	Leu	Ala	Asx	Gly	Val	Thr

Presence of tryptophan was indicated by absorbance at 280 nm.

TABLE X

Automated Edman Degradation of Fragments Obtained after Digestion of Cyanogen Bromide Peptides (B 1, B 2 and B 6)

Cycle No.	B 1-T 1 (1 nmol) PTH-aa	B 1-T 1 (1 nmol) Yield (pmol)	B 1-T 2 (1 nmol) PTH-aa	B 1-T 2 (1 nmol) Yield (pmol)	B 2-T 1 (0.5 nmol) PTH-aa	B 2-T 1 (0.5 nmol) Yield (pmol)	B 2-T 2 (0.5 nmol) PTH-aa	B 2-T 2 (0.5 nmol) Yield (pmol)	B 6-T 3 (0.5 nmol) PTH-aa	B 6-T 3 (0.5 nmol) Yield (pmol)
1	Ala	262	Gly	135	Ala	43	Tyr	64	Ile	65
2	Ala	187	Trp	57	Ala	35	Tyr	60	Asn	41
3	Gln	178	Thr	91	Leu	31	Gly	56	Phe	41
4	Thr	160	Gly	100	Glu	29	Gly	59	Ala	37
5	Gln	127	Gln	74	Ala	36	Ala	45	Val	36
6	Thr	148	Glu	90	Leu	26	Glu	39	Phe	48
7	Gly	86	Ser	21	Gly	24	Val	42	Pro	41
8	Glu	72	Leu	46	Ser	24[a]	Val	46	Ser	42
9	Ala	72	Ser	28	CmCys	NQ[a]	Asp	40	Leu	41
10	Ser	60	Asp	35	Leu	22	Glu	25	Leu	48
11	Arg	12	Thr	32	Asn	18	Ile	43	Gln	26
12			Asp	42	Asn	20	Glu	24	Gly	38
13			Pro	77	Lys	12	Leu	37	Gly	38
14			Glu	46	Tyr	17	Leu	39[a]	Pro	21[a]
15			Hse	25	Ser	13	CmCys	NQ[a]	His	NQ[a]
16					Glu	11	Gln	21[a]	Asn	22[a]
17					Gly	9	Arg	NQ[a]	His	NQ[a]
18					Tyr	5			Ala	14
19					Pro	4			Ile	7
20					Gly	4			Ala	12
21					Lys	2			Ala	12
22									Val	7
23									Ala	7
Repetitive yield %	87.3 (Ala 2-Ala 9)		92.0 (Glu 6-Glu 14)		94.3 (Leu 3-Leu 6)		89.5 (Glu 6-Glu 10)		93.7 (Ala 4-Ala 18)	
Position	2-12		13-27		54-74		76-92		286-311	

[a]NQ: observed but not quantitated.

TABLE XI

Automated Edman Degradation of Subfragment B 2-T 3 (1 nmol)

Cycle	PTH-aa	Yield (pmol)	Cycle	PTH-aa	Yield (pmol)
1	Ala	167	20	Ser	30
2	Leu	149	21	Gly	26
3	Glu	124	22	Ser	26
4	Ala	143	23	Pro	22
5	Phe	99	24	Ala	22
6	Asp	100	25	Asn	20
7	Leu	99	26	Leu	19
8	Asp	93	27	Ala	22
9	Pro	78	28	Ala	26
10	Ala	93	29	Tyr	13
11	Gln	68	30	Thr	11
12	Trp	21	31	Ala	17
13	Gly	61	32	Leu	12
14	Val	60	33	Leu	15
15	Asn	54	34	Gln	9
16	Val	54	35	Pro	8
17	Gln	47	36	His	5
18	Pro	42	37	Asp	6
19	Tyr	32	38	Arg	3

Repetitive yield: 93.1% (Ala 4-Ala 10)
Position: 94-131

TABLE XII

Automated Edman Degradation of Fragments Obtained after Digestion of Cyanogen Bromide Peptide (B 8)

Cycle No.	B 8-S 1 (1 nmol) PTH-aa	Yield (pmol)	B 8-S 2 (0.5 nmol) PTH-aa	Yield (pmol)	B 8-S 3 (1 nmol) PTH-aa	Yield (pmol)	B 8-S 4 (1 nmol) PTH-aa	Yield (pmol)	B 8-S 5 (0.5 nmol) PTH-aa	Yield (pmol)	B 8-S 6 (0.5 nmol) PTH-aa	Yield (pmol)	B 8-S 7 (0.5 nmol) PTH-aa	Yield (pmol)
1	Arg	116	Leu	111	Ala	434	Asp	208	Gly	60	Val	37	Thr	79
2	Gly	277	Val	81	Pro	250	Asp	236	Val	50	Lys	23	Ser	36
3	Tyr	269	Ser	49	Ala	352	Phe	349	Asn	46	Arg	10	Gln	43
4	Ser	134	Ile	68	Leu	244	Arg	290	Ile	44	Lys	20	Arg	39
5	Leu	264	Thr	36	Thr	180	Arg	324	Gly	48	Thr	23	Leu	48
6	Val	192	Ala	76	Ser	134	Val	312	Leu	50	Ala	26	Ala	28
7	Ser	109	Asn	48	Arg	177	Val	309	Glu	8	Lys	18	Asp	8
8	Gly	146	Lys	44	Gln	131	Asp	142			Leu	24		
9	Gly	144	Asn	42	Phe	118	Phe	279			Gln	22		
10	Thr	95	Thr	32	Arg	130	Ile	258			Asp	18		
11	Asp	57	CmCys	9	Glu	27	Asp	105			Phe	15		
12	Asn	85	Pro	34			Glu	40			Lys	14		
13	His	NQ[a]	Gly	29							Ser	5		
14	Leu	90	Asp	14							Phe	8		
15	Val	63	Arg	7							Leu	7		
16	Leu	89	Ser	11										
17	Val	58	Ala	19										
18	Asp	8	Ile	16										
19			Thr	11										
20			Pro	15										
21			Gly	15										
22			Gly	17										
23			Leu	16										
24			Arg	NQ[a]										
25			Leu	14										
26			Gly	6										
Repetitive yield %	89.1 (Gly 2-Gly 8)		88.2 (Ala 6-Ala 17)		90.1 (Ala 1-Ala 3)		96.3 (Phe 3-Phe 9)		94.6 (Gly 1-Gly 5)		94.1 (Lys 2-Lys 6)		86.7 (Glu 3-Ala 6)	
Position	339-356		373-398		399-409		410-421		422-428		429-448		449-455	

[a] NQ: observed but not quantitated.

TABLE XIII

Amino Acid Composition of Limited Acid Cleavage Peptides

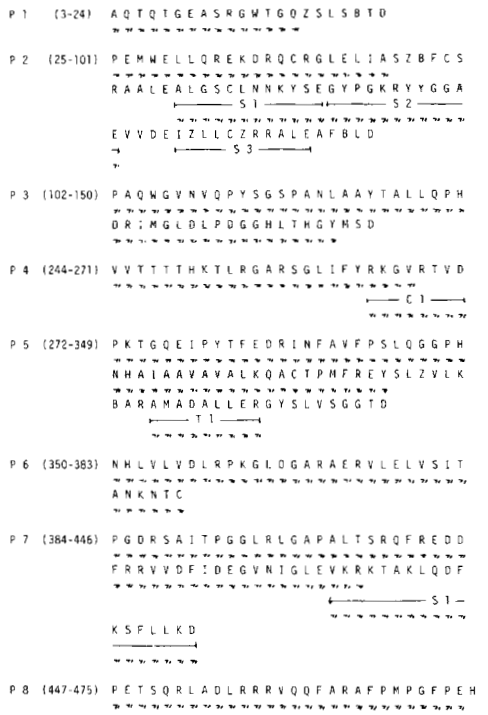
The composition from sequence analysis of each peptide is indicated by the numbers in parentheses

Peptide	P 1	P 2	P 4	P 5	P 7	P 8
Residue nos.	3-24	25-101	244-271	272-349	384-446	447-475
CmCys		2.2(4)		0.6(1)		
Asp	2.0(2)	6.5(7)	1.0(1)	5.8(6)	7.4(8)	1.0(1)
Thr	3.4(4)		4.4(6)	3.6(4)	2.4(3)	0.9(1)
Ser	2.9(3)	4.2(4)	0.9(1)	3.7(4)	2.6(3)	0.6(1)
Glu	5.1(5)	14.2(14)		7.6(8)	4.7(5)	5.1(5)
Pro		2.1(2)		5.1(5)	3.2(3)	2.4(4)
Gly	3.2(3)	6.0(6)	3.1(3)	6.0(6)	5.8(6)	1.3(1)
Ala	2.0(2)	7.1(7)	1.0(1)	10.6(11)	4.0(4)	2.9(3)
Val		1.9(2)	3.5(4)	4.5(5)	3.4(4)	0.9(1)
Met		0.8(1)		1.6(2)		1.0(1)
Ile		1.8(2)	0.8(1)	2.6(3)	2.7(3)	
Leu	1.0(1)	10.6(11)	2.0(2)	7.1(7)	7.2(7)	1.6(2)
Tyr		2.8(4)	0.7(1)	2.8(3)		
Phe		1.7(2)	0.9(1)	3.6(4)	4.6(5)	2.5(3)
His			0.9(1)	1.7(2)		1.0(1)
Lys		3.1(3)	1.6(2)	3.2(3)	4.8(5)	
Arg	0.9(1)	6.4(7)	3.8(4)	3.6(4)	7.6(7)	4.6(5)
Trp	+ (1)	+ (1)				
Yield % N-terminal residue	20.0 Ala	9.3 Pro	12.7 Val	5.4 Pro	29.3 Pro	13.2 Pro

Presence of tryptophan was indicated by absorbance at 280 nm.

TABLE XIV

Summary of Sequence Studies on Asp-Pro Peptides



The residues above the arrow were identified by automated Edman degradation on a gas-phase sequencer (▾). Subfragments obtained after digestion with trypsin (T), chymotrypsin (C) or *S. aureus* V 8 protease (S) are indicated by solid lines.

TABLE XV

Automated Edman Degradation of Limited Acid Cleavage Peptides

Cycle No.	P 1 (0.5 nmol) PTH-aa	Yield (pmol)	P 2 (1.5 nmol) PTH-aa	Yield (pmol)	P 4 (1.5 nmol) PTH-aa	Yield (pmol)	P 6 (0.5 nmol) PTH-aa	Yield (pmol)	P 8 (1 nmol) PTH-aa	Yield (pmol)
1	Ala	77	Pro	397	Val	444	Asn	74	Pro	252
2	Gln	36	Glu	270	Val	435	His	42	Glu	213
3	Thr	23	Met	272	Thr	202	Leu	79	Thr	48
4	Gln	28	Trp	120	Thr	146	Val	64	Ser	92
5	Thr	20	Glu	211	Thr	109	Leu	72	Gln	147
6	Gly	38	Leu	233	Thr	93	Val	76	Arg	56
7	Glu	34	Leu	251	His	33	Asp	74	Leu	147
8	Ala	24	Gln	96	Lys	30	Leu	76	Ala	158
9	Ser	14	Arg	57	Thr	46	Arg	53	Asp	95
10	Arg	9	Glu	99	Leu	38	Pro	48	Leu	132
11	Gly	25	Lys	81	Arg	13	Lys	33	Arg	53
12	Trp	12	Asp	74	Gly	47	Gly	50	Arg	84
13	Thr	NQ[a]	Arg	69	Ala	36	Leu	46	Arg	100
14	Gly	16	Gln	85	Arg	14	Asp	41	Val	72
15	Gln	8	CmCys	40	Ser	16	Gly	43	Gln	86
16			Arg	59	Gly	36	Ala	44	Gln	106
17			Gly	59	Leu	19	Arg	32	Phe	62
18			Leu	95	Ile	17	Ala	38	Ala	71
19			Glu	53	Phe	18	Glu	27	Arg	39
20			Leu	87	Tyr	12	Arg	17	Ala	65
21			Ile	49	Arg	10	Val	34	Phe	25
22			Ala	79	Lys	7	Leu	32	Pro	22
23					Gly	14	Glu	24	Met	15
24					Val	10	Leu	28	Pro	27
25							Val	29	Gly	14
26							Ser	17	Phe	16
27							Ile	23	Pro	19
28							Thr	8	Glu	9
29							Ala	18		
30							Asn	16		
31							Lys	12		
32							Asn	14		
33							Thr	NQ[a]		
34							CmCys	NQ[a]		
Repetitive yield %	91.2 (Gly 6-Gly 11)		92.8 (Leu 6-Leu 18)		93.6 (Gly 12-Gly 16)		94.7 (Leu 3-Leu 13)		96.5 (Leu 7-Leu 10)	
Position	3-24		25-101		244-271		350-383		447-475	

[a] NQ: observed but not quantitated.

TABLE XVI

Automated Edman Degradation of Peptide P 3

(2 nmol)

Cycle	PTH-aa	Yield (pmoles)	Cycle	PTH-aa	Yield (pmoles)
1	Pro	759	24	Leu	57
2	Ala	675	25	Leu	64
3	Gln	488	26	Gln	30
4	Trp	116	27	Pro	36
5	Gly	316	28	His	17
6	Val	248	29	Asp	35
7	Asn	175	30	Arg	39
8	Val	260	31	Ile	25
9	Gln	165	32	Met	13
10	Pro[a]	176	33	Gly	29
11	Tyr	91	34	Leu	48
12	Ser	50	35	Asp	37
13	Gly	78	36	Leu	55
14	Ser	51	37	Pro	27
15	Pro	58	38	Asp	33
16	Ala	102	39	Gly	33
17	Asn	37	40	Gly	35
18	Leu	61	41	His	NQ[b]
19	Ala	124	42	Leu	57
20	Ala	111	43	Thr	10
21	Tyr	36	44	His	NQ
22	Thr	22	45	Gly	31
23	Ala	68	46	Tyr	12

Repetitive yield: 94.4% (Ala 16-Ala 23)
Position: 102-150

[a]After the 9th cycle the filter was treated with o-phthalaldehyde (see text).
[b]NQ: observed but not quantitated

TABLE XVII

Automated Edman Degradation of Peptide P 5

(1 nmol)

Cycle	PTH-aa	Yield (pmoles)	Cycle	PTH-aa	Yield (pmoles)
1	Pro	143	26	Gly	28
2	Lys	142	27	Pro	10
3	Thr	61	28	His	7
4	Gly	121	29	Asn	11
5	Gln	113	30	His	8
6	Glu	103	31	Ala	15
7	Ile	149	32	Ile	12
8	Pro	78	33	Ala	19
9	Tyr	101	34	Ala	21
10	Thr	92	35	Val	10
11	Phe	71	36	Ala	21
12	Glu	47	37	Val	11
13	Asp	32	38	Ala	19
14	Arg	28	39	Leu	11
15	Ile	47	40	Lys	5
16	Asn	36	41	Gln	5
17	Phe	27	42	Ala	9
18	Ala	54	43	CmCys	3
19	Val	35	44	Thr	NQ[a]
20	Phe	31	45	Pro	5
21	Pro	17	46	Met	5
22	Ser	15	47	Phe	4
23	Leu	31	48	Arg	NQ[a]
24	Gln	14	49	Glu	3
25	Gly	25	50	Tyr	3

Repetitive yield: 92.8% (Gly 4-Gly 25)
Position: 272-349

[a]NQ: observed but not quantitated

TABLE XVIII

Automated Edman Degradation of Peptide P 7

(1 nmol)

Cycle	PTH-aa	Yield (pmoles)	Cycle	PTH-aa	Yield (pmoles)
1	Pro	163	25	Arg	22
2	Gly	154	26	Glu	17
3	Asp	159	27	Asp	23
4	Arg	91	28	Asp	24
5	Ser	80	29	Phe	28
6	Ala	141	30	Arg	22
7	Ile	64	31	Arg	18
8	Thr	26	32	Val	18
9	Pro	60	33	Val	18
10	Gly	68	34	Asp	23
11	Gly	75	35	Phe	20
12	Leu	78	36	Ile	16
13	Arg	53	37	Asp	18
14	Leu	65	38	Glu	14
15	Gly	47	39	Gly	11
16	Ala	53	40	Val	11
17	Pro	44	41	Asn	8
18	Ala	71	42	Ile	11
19	Leu	64	43	Gly	12
20	Thr	9	44	Leu	17
21	Ser	15	45	Glu	7
22	Arg	24	46	Val	8
23	Gln	20	47	Lys	5
24	Phe	28	48	Arg	7

Repetitive yield: 91.4% (Gly 2-Gly 10)
Position: 384-446

TABLE XIX

Amino Acid Composition of Fragments Obtained after Digestion of Limited Acid Cleavage Peptides

The composition from sequence analysis of each peptide is indicated by the number in parentheses

Peptide	P 2-S 1	P 2-S 2	P 2-S 3	P 4-C 1	P 5-T 1	P 7-S 1
Residue nos.	58-69	70-81	86-96	264-271	331-339	429-446
CmCys	1.0(1)		0.7(1)			
Asp	2.5(2)			1.1(1)	1.0(1)	2.2(2)
Thr				1.3(1)		1.2(1)
Ser	2.0(2)					1.0(1)
Glu	1.7(1)	1.9(1)	4.0(3)		1.1(1)	1.3(1)
Pro		1.0(1)				
Gly	1.0(1)	4.0(4)		1.3(1)		
Ala	0.9(1)	0.9(1)	1.0(1)		3.0(3)	0.9(1)
Val				2.0(2)		1.0(1)
Met					0.6(1)	
Ile			0.9(1)			
Leu	1.8(2)		3.0(3)		1.8(2)	2.7(3)
Tyr	0.9(1)	2.9(3)				
Phe						2.1(2)
Lys	1.6(1)	1.6(1)		1.1(1)		5.0(5)
Arg		0.9(1)	2.0(2)	2.0(2)	0.9(1)	1.0(1)
Yield %	18.2	24.3	18.6	16.0	32.4	25.6
N-terminal residue	Ala	Gly	Ile	Arg	Ala	Val

TABLE XX

Automated Edman Degradation of Fragments Obtained after Digestion of Limited Acid Cleavage Peptides

	P 2-S 1 (1.5 nmol)		P 2-S 2 (1.5 nmol)		P 4-C 1 (1.5 nmol)		P 5-T 1 (2 nmol)		P 7-S 1 (1.5 nmol)	
Cycle no.	PTH-aa	Yield (pmol)	PTH-aa	Yield (pmol)	PTH-aa	Yield (pmol)	PTH-aa	Yield (pmol)	PTH-aa	Yield (pmol)
1	Ala	371	Gly	275	Arg	352	Ala	687	Val	492
2	Leu	335	Tyr	273	Lys	228	Met	662	Lys	378
3	Gly	217	Pro	292	Gly	227	Ala	609	Arg	68
4	Ser	131	Gly	198	Val	194	Asp	395	Lys	335
5	CmCys	35	Lys	65	Arg	178	Ala	510	Thr	279
6	Leu	56	Arg	22	Thr	107	Leu	405	Ala	178
7	Asn	35	Tyr	55	Val	154	Leu	452	Lys	56
8	Asn	28	Tyr	44	Asp	98	Glu	301	Leu	137
9	Lys	11	Gly	42			Arg	64	Gln	87
10	Tyr	25	Gly	58					Asp	43
11	Ser	16	Ala	47					Phe	76
12	Glu	17							Lys	117
13									Ser	71
14									Phe	76
15									Leu	66
16									Leu	70
17									Lys	40
Repetitive yield%	63.9 (Leu 2-Leu 7)		89.6 (Gly 1-Gly 4)		92.6 (Val 4-Val 7)		91.5 (Ala 3-Ala 5)		94.1 (Lys 2-Lys 4)	
Position	58-69		70-81		264-271		331-339		429-446	

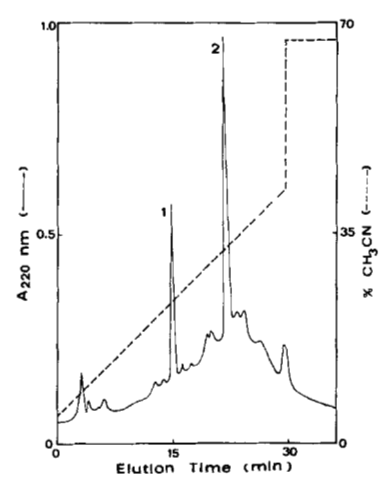


Fig. 3. Reverse-phase hplc of the basic peptides obtained from a trypsin-carboxypeptidase B digest of the protein reacted with glycinamide. The number above the peaks refers to the peptides for which analytical data are given in the text.

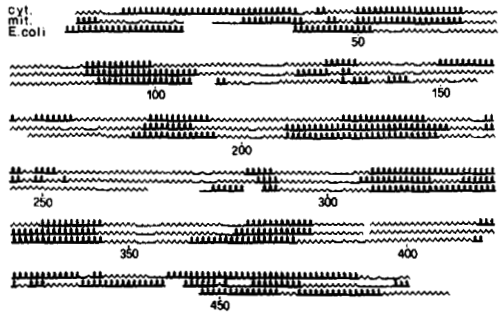


Fig. 4. Comparison of the predicted secondary structure of E. coli serine hydroxymethyltransferase with those of the corresponding cytosolic and mitochondrial isoenzymes from rabbit liver. Gaps were introduced according to the alignment of sequences adopted in Fig. 2. : alpha-helix; : extended conformation; : turn; : random coil.

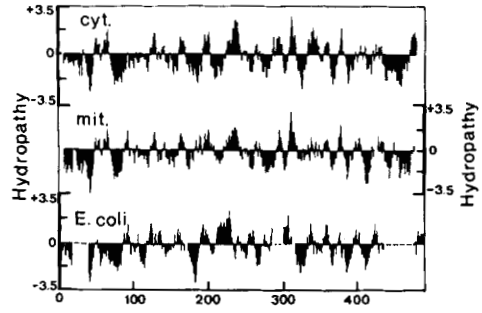


Fig. 5. Comparison of the hydropathy profile of E. coli serine hydroxymethyltransferase with those of the corresponding cytosolic and mitochondrial isoenzymes from rabbit liver. Gaps (- - -) were introduced according to the alignment of sequences shown in Fig. 2. Consecutive hydropathy values are plotted at the mid-point of the seven residues segment as it advances from N- to C-terminus.