THE PRIMARY STRUCTURE OF ASPARTATE AMINOTRANSFERASE FROM PIG HEART MUSCLE DETERMINED IN PART USING A PROTEASE WITH SPECIFICITY FOR LYSINE

S. DOONAN, Hilary J. DOONAN, R. HANFORD, C.A. VERNON and J.M. WALKER
The Christopher Ingold Laboratories, Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK

and

Francesco BOSSA, Donatella BARRA, Massimo CARLONI, Paolo FASELLA and Francesca RIVA
Istituto di Chimica Biologica e Centro di Biologia Moleculare del Consiglio Nazionale delle Ricerche, Università di Roma, Città Universitaria, 00185 Roma, Italy

and

P.L. WALTON
Department of Biochemistry, Pharmaceuticals Division, I.C.I. Ltd Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK

Received 12 September 1973

1. Introduction

We have recently reported [1] partial amino acid sequences of the cytoplasmic aspartate aminotransferase from pig heart muscle, based on studies of peptides produced by digestion of the protein with pepsin and trypsin [2] and with thermolysin and elastase [1]. This work gave ten major fragments containing 395 amino acid residues. We now present the complete structure of the molecule which consists of a chain of 412 amino acid residues. The structure given here is essentially in agreement with that of Ovchinnikov et al. [3]. However, the authors gave no indication of the data which was used to overlap their previously reported 21 partial sequences [4]. We consider it important, for this reason and due to the large size of the molecule, to confirm the correctness of the published structure and to show how the structure was derived from our earlier data.

An important tool used in the present work was a newly characterized protease isolated from the fruiting body of the basidiomycete *Armillaria mellea* [5]. The enzyme has been shown to be an endopeptidase cleaving proteins such as fibrinogen and casein N-terminal to lysine residues. Our results confirm this specificity and provide further information about the type of peptide bond hydrolysed by the enzyme.

2. Materials and methods

Carboxymethylated aspartate aminotransferase [2] was used for all enzymic digests. For digestion with *A. mellea* protease, 2 g of the modified protein was suspended in *N*-ethylmorpholine–HCl buffer (0.2 M, pH 7.5, 200 ml); 5 ml of a solution of the protease (approx. 3 mg/ml) was added and the suspension stirred at 37°C for 24 hr. Glacial acetic acid was added to 30% v/v after which insoluble material (approx. 700 mg) was removed by centrifugation. The soluble material was subjected to initial separation by gel filtration through Sephadex G-50 equi-

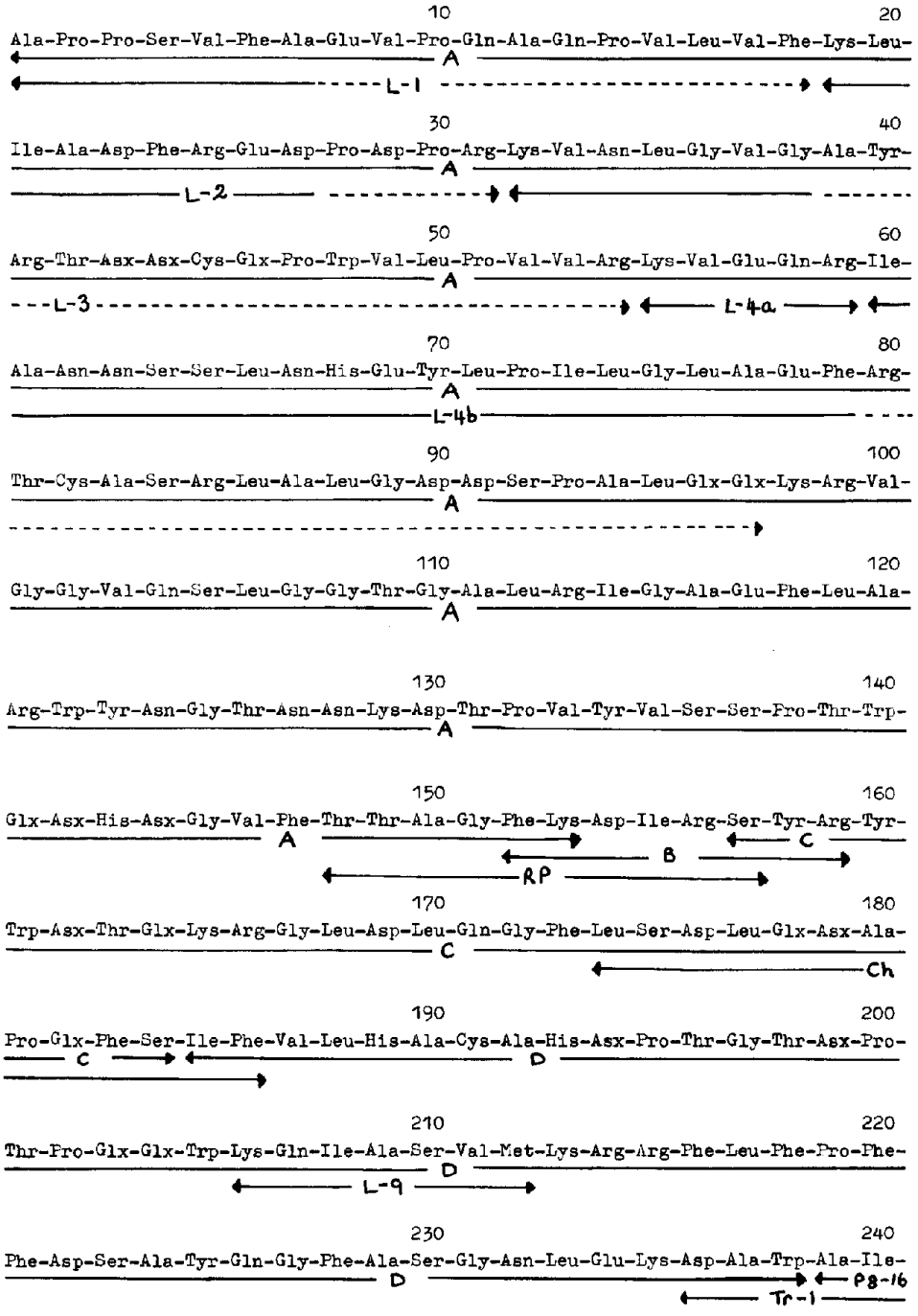

Fig. 1. For continuation of figure and the legend see opposite page.

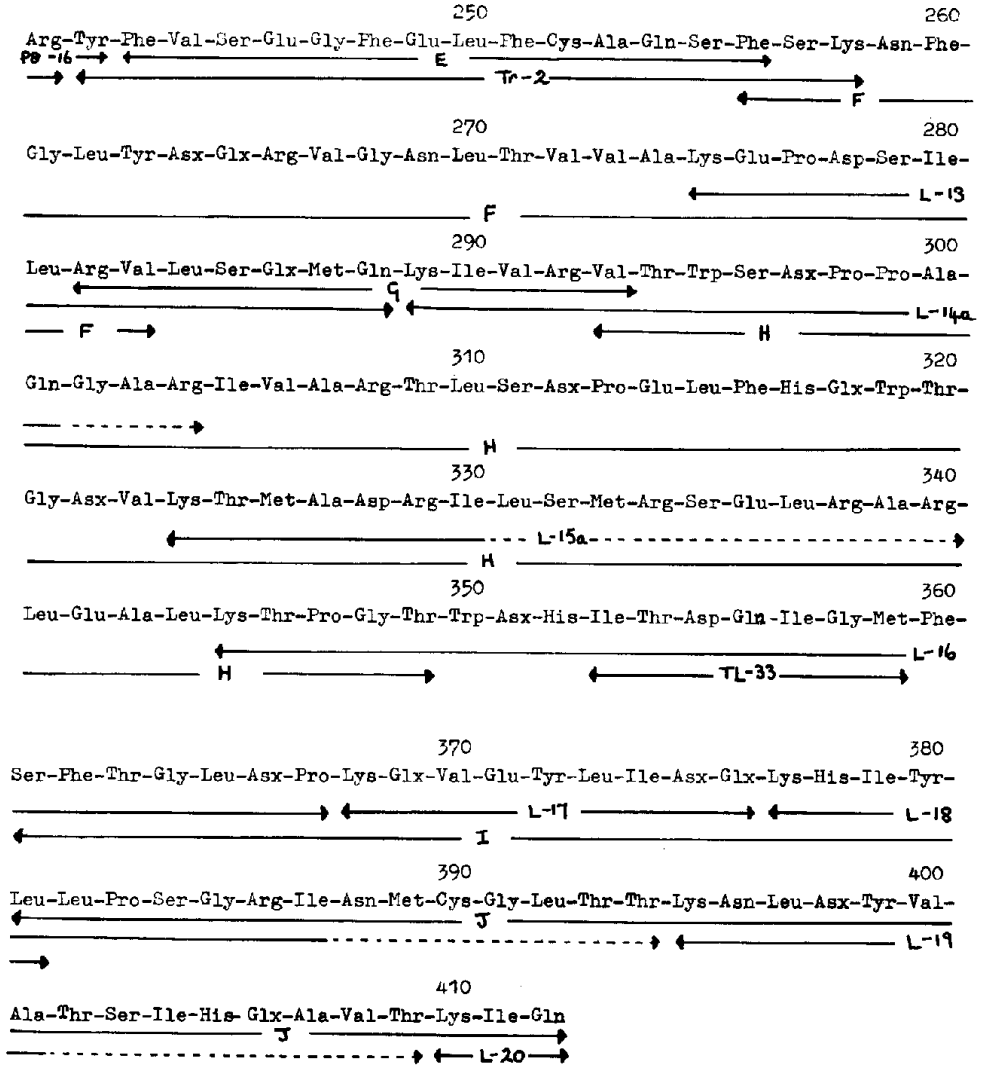

Fig. 1. The primary structure of aspartate aminotransferase. The symbol L designates a peptide resulting from cleavage with *A. mellea* protease. Other symbols are defined in the text. Where an arrow under a peptide is partially dotted, the residues above the solid part were determined by the dansyl–Edman method and those above the dotted part from the amino acid composition by comparison with known sequences [1].

librated with acetic acid (30% v/v). Final purification of peptides was by paper chromatography and high voltage paper electrophoresis [1, 2] except in the case of peptide L-4b which was purified by ion exchange chromatography on SP-Sephadex C-25 at pH 2.8. Samples of purified peptides were hydrolysed by the procedure of Sanger and Thompson [6] and amino acid compositions were determined using either a Technicon Auto Analyzer or a Biocal BC200 automatic analyzer. Sequences of peptides were determined by the dansyl–Edman method as described by Hartley [7]. In the case of peptides which were too large for

complete analysis by the dansyl–Edman technique, the C-terminal regions were assigned from the amino acid compositions by comparison with known sequences (fig. 1).

3. Results and discussion

The amino acid sequence of aspartate aminotransferase is shown in fig. 1; some acid and amide side chains remain to be assigned. The letters A to J refer to the ten major fragments previously described [1]. The symbols P8-16 and TL-33 refer to a peptic peptide [2] and a thermolytic peptide [1] respectively which were not previously incorporated into the partial sequences. Overlap between fragments A and B is provided by a peptide (RP, residues 148–157) produced by restricted digestion with pepsin. The N-terminal residues of fragments C and D were originally assigned as alanine and leucine [1] but are now known to be serine and isoleucine respectively. These corrections make the overlap between fragments B and C clear; that between C and D is provided by a peptide (Ch, residues 174–186) from digestion of the protein with chymotrypsin. Two previously unisolated tryptic peptides (Tr-1 and Tr-2, residues 236–241 and 242–258 respectively) overlap fragment D, peptide P8-16 and fragments E and F. The remaining overlaps are provided by peptides from the digest with *A. mellea* protease (fig. 1).

The structure given in fig. 1 is in agreement with that of Ovchinnikov et al. [3] except at positions 63 and 288 where we find asparagine and glutamine respectively [2] whereas Ovchinnikov et al. [3] find the corresponding acids. It is possible that the peptides studied by the Ovchinnikov et al. [3] suffered deamination during the processes of isolation. The amino acid composition obtained by summing the residues in fig. 1 agrees well with that of the native molecule [8,9].

It should be emphasized that knowledge of the primary structure does not advance our understanding of the mechanism of catalysis by the enzyme. The amino acid residues lys-258 (the cofactor binding site [10]), tyr-40 [11] and cys-390 [12] may be at or near the active site. The role of these residues in catalysis cannot, however, be inferred without detailed information on the three dimensional structure of the molecule. Crystallographic analysis would be of particular value in the case of aspartate aminotransferase, since it is one of the few enzymes for which values of rate constants and catalytic factors are known for unit steps in the reaction [13].

The results of the digestion with *A. mellea* protease merit further comment. It is clear from fig. 1 that the major specificity of the enzyme is for cleavage N-terminal to lysine. In three cases, however, hydrolysis occurred at C-terminal to arginine (positions 59, 304 and 340). Cleavage seems to have been extensive since the parent lysine peptides were not isolated. At two of these cleavage points, the residue C-terminal to arginine is isoleucine whereas at the other cleavage point the residue is leucine. This apparent specificity for long-chain aliphatic residues adjacent to arginine is reminiscent of the specificities of thrombin and Factor X*a* in their action on a variety of protein substrates [14].

The lysine residues in aspartate aminotransferase can be classified into three groups according to the facility of hydrolysis of the adjacent bond by *A. mellea* protease. Cleavage occurred readily at residues 19, 32, 55, 98, 206, 213, 324, 345, 368, 377, 395 and 410, and the product peptides were obtained in high yields. Much more restricted hydrolysis was observed at positions 165, 275 and 289. Consequently, peptides L-13 and L-14a were obtained in low yields but the quantities were sufficient for sequence determination and amino acid analysis. A small amount of a peptide was isolated with N-terminal sequence corresponding to residues 165 to 175 and a second minor component had residues 213 to 223 as the N-terminal sequence and yielded valine plus alanine on treatment with carboxypeptidase A. It is thought that these two peptides represented residues 165–205 and 213–274 respectively but the amino acid analyses were not sufficiently precise to establish this with certainty; consequently these peptides have not been included in fig. 1. Cleavage at lysine residues 129, 153, 235 and 258 does not appear to have occurred to any detectable extent. Restricted cleavage or lack of cleavage at the positions listed above resulted in the majority of the material representing residues 98–205 and 213–323 being present as two insoluble fragments. Amino acid analysis of the insoluble material and digestion with trypsin followed by isolation of the product peptides gave support to this conclusion. Prolonged treatment (5 days) of the insoluble material with *A. mellea* protease did not result in detectable cleavage.

From the results presented here, some tentative conclusion can be drawn about the specificity of the protease. Three of the four lysine residues at which no cleavage appeared to occur have an aspartic acid residue at the next position in the chain. Similarly, residues 165 and 275, at which only partial cleavage occurred, have adjacent glutamic acid residues. Hence it appears that hydrolysis is prevented by an aspartic acid residue C-terminal to lysine and is impeded in some cases by glutamic acid on either side of the lysine residue. Cleavage is not impaired by a proline residue N-terminal to lysine or by arginine residues on either side of the lysine. Obviously further work is required to extend our knowledge of the specificity of the enzyme, but it is clear that *A. mellea* protease provides a very powerful addition to the techniques of protein sequence determination.

Acknowledgements

Financial support from the Science Research Council (H.J.D., R.H. and J.M.W.), EMBO (F.R.) and Roche Products Ltd., is gratefully acknowledged.

References

[1] Bossa, F., Barra, D., Carloni, M., Fasella, P., Riva, F., Doonan, S., Doonan, H.J., Hanford, R., Vernon, C.A. and Walker, J.M. (1973) Biochem. J. 133, 805.

[2] Doonan, S., Doonan, H.J., Riva, F., Vernon, C.A., Walker, J.M., Bossa, F., Barra, D., Carloni, M. and Fasella, P. (1972) Biochem. J. 130, 443.

[3] Ovchinnikov, Yu.A., Egorov, C.A., Aldanova, N.A., Feigina, M.Yu., Lipkin, V.M., Abdulaev, N.G., Grishin, E.V., Kiselev, A.P., Modyanov, N.N., Braunstein, A.E., Polyanovsky, O.L. and Nosikov, V.V. (1973) FEBS Letters 29, 31.

[4] Ovchinnikov, Yu.A., Kiryushkin, A.A., Egorov, C.A., Abdulaev, N.G., Kiselev, A.P., Modyanov, N.N., Grishin, E.V., Sukhikh, A.P., Vinogradova, E.I., Feigina, M.Yu., Aldanova, N.A., Lipkin, V.M., Braunstein, A.E., Polyanovsky, O.L. and Nosikov, V.V. (1971) FEBS Letters 17, 133.

[5] Walton, P.L., Turner, R.W. and Broadbent, D. (1972) U.K. Pat. No. 1263956; Walton, P.L., Jones, C. and Jackson, S., to be published.

[6] Sanger, F. and Thompson, E.O.P. (1963) Biochim. Biophys. Acta 71, 468.

[7] Hartley, B.S. (1970) Biochem. J. 119, 805.

[8] Banks, B.E.C., Doonan, S., Lawrence, A.J. and Vernon, C.A. (1968) Eur. J. Biochem. 5, 528.

[9] Martinez-Carrion, M., Turano, C., Chiancone, E., Bossa, F., Giartosio, A., Riva, F. and Fasella, P. (1967) J. Biol. Chem. 242, 2397.

[10] Morino, Y. and Watanabe, T. (1969) Biochemistry 8, 3412.

[11] Polyanovsky, O.L., Demidkina, T.V. and Egorov, C.A. (1972) FEBS Letters 23, 262.

[12] Birchmeier, W., Wilson, K.J. and Christen, P. (1973) J. Biol. Chem. 248, 1751.

[13] Banks, B.E.C., Bell, M.P. and Vernon, C.A. (1968) in: Pyridoxal catalysis: enzymes and model systems (Snell, E.E., Braunstein, A.E., Severin, E.S. and Torchinsky, Yu.M., eds), p. 191, Interscience, New York.

[14] Elmore, D.T., Proc. Biochem. Soc., in press.