The Amino Acid Sequence of *Neurospora* NADP-Specific Glutamate Dehydrogenase

PEPTIDES FROM DIGESTION WITH A STAPHYLOCOCCAL PROTEINASE

By JOHN C. WOOTTON, ANDREW J. BARON and JOHN R. S. FINCHAM
Department of Genetics, University of Leeds, Leeds LS2 9JT, U.K.

(Received 23 December 1974)

The extracellular proteinase of *Staphylococcus aureus* strain V8 was used to digest the NADP-specific glutamate dehydrogenase of *Neurospora crassa*. Of 35 non-overlapping peptides expected from the glutamate content of the polypeptide chain, 29 were isolated and substantially sequenced. The sequences obtained were valuable in providing overlaps for the alignment of about two-thirds of the sequences found in tryptic peptides [Wootton, J. C., Taylor, J. G., Jackson, A. A., Chambers, G. K. & Fincham, J. R. S. (1975) *Biochem. J.* 149, 739–748]. The blocked *N*-terminal peptide of the protein was isolated. This peptide was sequenced by mass spectrometry, and found to have *N*-terminal *N*-acetylserine by Howard R. Morris and Anne Dell, whose results are presented as an Appendix to the main paper. The staphylococcal proteinase showed very high specificity for glutamyl bonds in the NH_4HCO_3 buffer used. Partial splits of two aspartyl bonds, both Asp-Ile, were probably attributable to the proteinase. No cleavage of glutaminyl or *S*-carboxymethylcysteinyl bonds was found. Additional experimental detail has been deposited as Supplementary Publication SUP 50053 (5 pages) with the British Library (Lending Division), Boston Spa, Wetherby, W. Yorkshire LS23 7BQ, U.K. from whom copies may be obtained under the terms given in *Biochem. J.* (1975) 145, 5.

In the preceding paper (Wootton *et al.*, 1975) the isolation and amino acid sequences of tryptic peptides of *Neurospora crassa* NADP-specific glutamate dehydrogenase (EC 1.4.1.4) were described. The recently characterized extracellular proteinase from *Staphylococcus aureus* strain V8, which, under appropriate conditions, is almost specific for glutamyl linkages (Drapeau *et al.*, 1972), appears to be the ideal complement to trypsin in the establishment of overlaps spanning tryptic peptides. In the present paper we present data on staphylococcal proteinase peptides accounting for over half the *Neurospora* glutamate dehydrogenase sequence as defined by the tryptic peptides. The remaining overlaps required for the establishment of the complete sequence have been obtained from chymotryptic and peptic peptides (Holder *et al.*, 1975) and from CNBr fragments (Blumenthal *et al.*, 1975). A brief report of the sequence thus derived has already been published (Wootton *et al.*, 1974).

Materials and Methods

Source of the glutamate dehydrogenase

Glutamate dehydrogenase purified from a prototrophic revertant strain R26 derived from the auxotrophic mutant *am*¹⁹, was used, because it was the

material that was on hand at the time the staphylococcal proteinase became available. The mutant *am*¹⁹ produces a glutamate dehydrogenase variety which is strongly stabilized in an inactive conformation (Coddington *et al.*, 1966), whereas that of strain R26 has greatly improved activity but still some abnormal properties (Fincham & Stadler, 1965). Glutamate dehydrogenase from strain R26 was expected to have either one or two amino acid differences from the wild-type, depending on whether the regained activity was due to a further change in the codon primarily altered or to a compensatory change elsewhere. It seems that the second alternative is true; strain R26 retains the lysine-141 → methionine-141 replacement characteristic of *am*¹⁹ (G. K. Chambers, A. A. Holder, J. C. Wootton & J. R. S. Fincham, unpublished work), but all other sequences identified in this paper correspond to known wild-type sequences and the presumptive second change remains to be identified.

Glutamate dehydrogenase from strain R26 was prepared and carboxymethylated by the procedures described for wild-type glutamate dehydrogenase in the preceding paper (Wootton *et al.*, 1975).

Staphylococcal proteinase digestion

Purified staphylococcal proteinase was obtained as a generous gift from Dr. G. R. Drapeau, University

Table 1. *Amino acid sequences of Staphylococcus proteinase peptides*

Asparagine-2 and glutamate-6 were assigned by mass spectrometry (see the Appendix); asparagine-350 by yellow stain with the cadmium-ninhydrin reagent; and asparagine-351 by determination of charge of +2 on the whole peptide at pH 6.5. —, One round of the Edman degradation with the end group identified., One round of the Edman degradation without positive identification of the end group. Residues in parentheses, consistent with the composition of the peptide (except for SP18 which was contaminated but nevertheless sequenceable), are shown in the order deduced from sequencing of other kinds of peptides.

Peptide	Sequence
SP1	¹ * Ser-Asn-Leu-Pro-Ser-Glu-Pro-Glu — 8
SP2	⁹ ¹⁰ Phe-Glu —
SP3	¹¹ ¹⁵ * Gln-Ala-Tyr-Lys-Glu —
SP4	¹⁶ ²¹ (Leu, Ala, Tyr, Thr, Leu, Glu) —
SP5	²² ³¹ Asx-Ser-Ser-Leu-Phe-Glx-Lys-His-Pro-Glu —
SP6	³² ⁴³ Tyr{Arg}Thr-Ala-Leu-Thr-Val-Ala-Ser-Ile-Pro-Glu — ...
SP9B	¹²⁰ ¹²⁷ Pro-Lys-Gly-Lys-Ser-Asx-Ala-Glu —
SP11	¹³⁹ ¹⁶⁰ Leu-His-Met-His-Ile-Gly-Ala-Asx-Thr{Asx, Val, Pro, Ala, Gly, Asx, Ile, Gly, Val, Gly, Arg, Glu} —
SP12A	¹⁶¹ ¹⁶⁶ Ile-Gly-Tyr-Met-Phe-Gly —
SP16	²⁴⁰ ²⁵⁹ Leu-Gly{Ala}Thr-Val-Val{Ser}Leu-Ser-Asx{Ser}Lys-Gly-Ala-Leu-Val-Ala-Thr-Gly-Glu — ...
SP17	²⁶⁰ ²⁶⁵ Ser-Gly-Ile-Thr-Val-Glu —

SP18	266 (Asx)Ile-Asx-Ala-Val(Met)Ala-Ile(Lys)Glu ----->	275 ----->
SP19	276 Ala-Arg-Glx-Ser-Leu-Thr-Ser-Phe-Glx-His-Ala-Gly-His-Leu-Lys-Trp-Ile-Glu ----->	293 ----->
SP19B	Leu-Thr-Ser-Phe-Glx-His-Ala-Gly-His-Leu-Lys(Trip)Ile-Glu ----->	----->
SP20	294 Gly-Ala-Arg-Pro-Trip-Leu-His-Val-Gly-Lys-Val-Asx-Ile-Ala-Leu(Pro-Cys)Ala-Thr-Glx-Asx-Glu ----->	315 ----->
SP20A	305 Gly-Ala-Arg-Pro-Trip-Leu-His-Val-Gly-Lys-Val-Asp ----->	----->
SP21/22	316 Val-Ser-Lys(Glx,Glx,Ala,Glu) ----->	322 ----->
SP21 SP22	(Val,Ser,Lys,Glx,Glu) (Ala,Glu) ----->	----->
SP23	323 Gly-Leu-Leu-Ala-Ala-Gly-Cys-Lys-Phe-Val-Ala-Glu ----->	334 ----->
SP24	335 Gly-Ser-Asx-Met-Gly(Cys)Thr-Leu-Glu ----->	343 ----->
SP25	344 Ala-Ile-Glu ----->	346 ----->
SP26	347 Val-Phe-Glu ----->	349 ----->
SP27	350 Asn-Asn-Arg-Lys-Glu ----->	354 ----->
SP28	355 Lys-Lys-Gly-Glu ----->	358 ----->

of Montreal, Quebec, Canada. About 200mg (approx. 4 μ mol) of carboxymethylated glutamate dehydrogenase (from strain R26) was suspended as finely as possible in 30 ml of 0.5M-NH₄HCO₃; 4 mg of the proteinase was added, and digestion was allowed to proceed at 25°C for 30h. Some insoluble material remaining after this time was removed by centrifuging and redigested with a further 2mg of proteinase. Insoluble material was again removed and the soluble fractions were combined and dried *in vacuo*.

Other methods

A column of Dowex AG50 (X4) was used for primary fractionation of the soluble peptides. Samples of alternate fractions were analysed qualitatively by paper chromatography. Pooled fractions from this first column were refractionated where necessary on a column of either Dowex AG1 (X2) or Bio-Gel P4, or on one type of column followed by the other. All the detailed procedures, as well as those for amino acid analysis and sequence determination, were as described in the preceding paper (Wootton *et al.*, 1975). The isolation methods for individual peptides, with their yields and amino acid analyses, are given in Tables 1 and 2 of Supplementary Publication SUP 50053.

The material that was insoluble in the pyridine-acetic acid starting buffer used for the column of Dowex AG50 (X4), but soluble in 10% (v/v) formic acid, was fractionated on Bio-Gel P4 in the latter solvent. The fraction containing peptide SP18 was obtained in this way.

Results and Discussion

The sequences of the isolated peptides, so far as they were determined, are shown in Table 1. In general, the determined sequences were in good agreement with the compositions shown in Table 2 of Supplementary Publication 50053. Several peptides gave analyses indicating significant levels of contamination, so that the numbers of residues of some amino acids were not clear from their analyses alone. These peptides were SP6, SP9B, SP16, SP18, SP19, SP19B, SP20A, SP27, SP30, SP32, SP32C and SP34/35A. All, however, yielded either clear complete sequences or clear partial sequences which overlapped those of other peptides the complete sequences of which were definitively determined. The complete pattern of overlaps is shown in Fig. 1 of the following paper (Holder *et al.*, 1975). In Table 1 the amino acid residues in the peptides are numbered according to their positions in the proposed complete sequence (Wootton *et al.*, 1974; Holder *et al.*, 1975). The peptides are numbered in sequence from the *N*- to the *C*-terminus of the chain, with gaps in the numbering left for peptides not recovered. The sequences given

in parentheses are derived from overlapping tryptic (Wootton *et al.*, 1975), peptic and chymotryptic peptides (Holder *et al.*, 1975) and, in general, are in good agreement with the amino acid compositions. Yields were variable and generally rather low; reasons for this have already been discussed in relation to tryptic peptides (Wootton *et al.*, 1975).

Only about two-thirds of the total proposed sequence was accounted for by the staphylococcal proteinase peptides isolated. Failure to recover the sequence between residues 44 and 119, and 167 and 239, can be reasonably attributed to difficulties in fractionation arising from the large size, relative insolubility and possible degeneracy of most of the missing products. The peptides which were recovered do, however, establish the correct alignment of the tryptic peptides T2 and T3 (residues 29–43), T13 and T14A (residues 114–129), T15 and T16 (residues 130–159), T22–T32 (residues 237–397) and T33–T35 (residues 398–414). The remainder of the overlaps required for the complete sequence have been obtained in chymotryptic and peptic peptides (Holder *et al.*, 1975) and in CNBr fragments (Blumenthal *et al.*, 1975).

Peptide SP1 did not stain with ninhydrin and was initially resistant to Edman degradation. Treatment with methanolic HCl (Yamada & Itano, 1972), a routine procedure used by us for opening the pyrrolidone ring commonly formed when glutamine is at the *N*-terminus, rendered the peptide reactive to phenylisothiocyanate but revealed serine as the end group. A sample of the peptide was kindly analysed by Howard R. Morris and Anne Dell, by mass spectrometry, and their data are given as an Appendix to this paper. The sequence *N*-acetyl-Ser-Asn-Leu-Pro-Ser-Glu... was demonstrated. This is evidently the *N*-terminal sequence of the entire chain and agrees with the sequence obtained by the dansyl-Edman method in the methanolysed peptide.

It is of interest to see how far the peptides recovered are in accord with the reported specificity of the staphylococcal proteinase (Drapeau *et al.*, 1972). Under the conditions used (NH₄HCO₃ buffer) specificity for a glutamyl bond turned out to be almost complete. Only three aspartyl linkages were cleaved, the products being obtained in low yield (5% or less) in each case. Two of these involved Asp-Ile bonds (residues 119–120 and 305–306 respectively). Since there are only four Asp-Ile sequences in the chain the susceptibility of two of them seems significant. The third aspartyl linkage to be cleaved is the same Asp-Pro (residues 119–120) as was found to be split in a tryptic digest (Wootton *et al.*, 1975) and non-enzymic hydrolysis is suspected. A similar lability of two Asp-Pro linkages was reported by Piskiewicz *et al.* (1970) for bovine glutamate dehydrogenase. The only cleavages found other than those adjacent to either glutamate or aspartate were

Gly-Ala (166–167), Ser-Leu (279–280), Asn-Gly (408–409) and Gly-Ser (430–431). All these occurred in low yield, and their enzymic basis is in doubt.

Although recovery of peptides was not complete some conclusions may be drawn from the cases where glutamyl bonds showed incomplete cleavage or no cleavage. The single Glu-Glu sequence (319–320) was not split. Of the two Glu-Asp sequences occurring amongst the recovered peptides (out of three in the protein), one (313–314) was found to be unsplit and the other (265–266) gave only a very-low-yield cleavage product (SP18). This suggests that under the conditions used the proteinase does not readily cleave glutamyl bonds which are followed by an acidic residue. Some evidence was obtained indicating that the proteinase does not readily split off the dipeptide product from Glu-X-Glu sequences, but it is not clear if this is because the enzyme is a poor 'exodi-peptidase' or because of the presence of the second acidic group. From the sequence Glu-Gly-Glu-Leu (421–424) the Gly-Glu dipeptide was not recovered but mutually exclusive single splits were observed at each of the glutamyl bonds. The dipeptides Phe-Glu (9–10) and Ala-Glu (321–322) were recovered, but in low yield, and in the latter case an uncleaved product was also found. Another possible dipeptide, Pro-Glu (7–8) was apparently not split off, but Glu-Pro may in any case be resistant to the proteinase. No other cases of incomplete splits of glutamyl bonds were observed. We have no information on Glu-His, Glu-Thr and Glu-Trp sequences which do not occur in the protein. No cleavage of *S*-carboxymethylcysteinyl bonds was observed; Cys-Cys, Cys-Ala, Cys-Lys, Cys-Gly, and Cys-Thr sequences occur in the protein.

Since nearly all the independently assigned glutamate residues, but none of the known glutamine

residues, are found to be sites of staphylococcal proteinase cleavage, we felt it was safe to identify a Glx residue as glutamate on the basis of its *C*-terminal position in staphylococcal proteinase peptide in those few cases where independent evidence is not available. The assignments made in this way (glutamate-10, glutamate-346 and glutamate-349) are indicated in Table 2 of the following paper (Holder *et al.*, 1975).

We thank the Science Research Council for financial support, Margaret Burnley for the amino acid analyses and Dr. G. R. Drapeau for his generous gift of the staphylococcal proteinase.

References

- Blumenthal, K. M., Moon, K. & Smith, E. L. (1975) *J. Biol. Chem.* **250**, 3644–3654
- Coddington, A., Sundaram, T. K. & Fincham, J. R. S. (1966) *J. Mol. Biol.* **17**, 503–512
- Drapeau, G. R., Boily, Y. & Houmard, J. (1972) *J. Biol. Chem.* **247**, 6720–6726
- Fincham, J. R. S. & Stadler, D. R. (1965) *Genet. Res.* **6**, 121–129
- Holder, A. A., Wootton, J. C., Baron, A. J., Chambers, G. K. & Fincham, J. R. S. (1975) *Biochem. J.* **149**, 757–773
- Piszkiewicz, D., Landon, M. & Smith, E. L. (1970) *Biochem. Biophys. Res. Commun.* **40**, 1173–1178
- Wootton, J. C., Chambers, G. K., Holder, A. A., Baron, A. J., Taylor, J. G., Fincham, J. R. S., Blumenthal, K. M., Moon, K. & Smith, E. L. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4361–4365
- Wootton, J. C., Taylor, J. G., Jackson, A. A., Chambers, G. K. & Fincham, J. R. S. (1975) *Biochem. J.* **149**, 739–748
- Yamada, S. & Itano, H. A. (1972) *Anal. Biochem.* **48**, 546–556

APPENDIX

The Sequence of the Blocked *N*-Terminal Peptide from *Neurospora* Glutamate Dehydrogenase

By HOWARD R. MORRIS and ANNE DELL

M.R.C. Group, University of Cambridge Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K.

(Received 23 December 1974)

Mass spectrometry is now becoming a widely used technique in protein structure elucidation. Several major studies encompassing both known and unknown sequences have been carried out on α -lactalbumin (Bacon & Graham, 1972), triose phosphate isomerase (Priddle & Offord, 1974),

dihydrofolate reductase (Morris *et al.*, 1974a) and ribitol dehydrogenase (Morris *et al.*, 1974b).

Apart from such major applications, mass spectrometry is a powerful technique for the study of unusual structural features in proteins or other molecules. In particular, it can be valuable for the