Biochem. J. (1980) **187**, 565–575
Printed in Great Britain

Primary Structure of the Calcium Ion-Transporting Adenosine Triphosphatase of Rabbit Skeletal Sarcoplasmic Reticulum

SOLUBLE PEPTIDES FROM THE α-CHYMOTRYPTIC DIGEST OF THE CARBOXYMETHYLATED PROTEIN

Geoffrey ALLEN*
National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.

(*Received 21 August 1979*)

The isolation of the soluble peptides from the chymotryptic digest of the calcium-transporting ATPase of rabbit skeletal sarcoplasmic reticulum is described. These peptides were partially sequenced and the information obtained was used to align tryptic peptides of the protein and to confirm sequences within the tryptic peptides. Details of the isolation of some peptides and the amino acid analyses of the peptides are given in Supplementary Publication SUP 50103 (10 pages), which has been deposited with the British Library Lending Division, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in *Biochem. J.* (1978) **169**, 5.

Studies on the primary structure of the Ca^{2+}-transporting ATPase of rabbit sarcoplasmic reticulum have revealed that peptides derived from the tryptic digests of the protein (Allen & Green, 1978) and of the succinylated (3-carboxypropionylated) protein (Allen, 1980) fall into two classes: those that are generally soluble in aqueous solutions and that have properties similar to those of most peptides derived from soluble globular proteins, and those that are highly aggregated or insoluble in aqueous solutions, even in 8M-urea. The peptides in the latter class are presumed to be associated with the lipid bilayer in the native structure of the protein.

The isolation and characterization of tryptic peptides derived from the succinylated protein is described in the preceding paper (Allen, 1980). Digestion of the protein with α-chymotrypsin was performed in order to provide overlaps between these tryptic peptides and to confirm sequences found within these peptides. It was anticipated, also, that α-chymotrypsin, with its specificity for hydrophobic residues, would release more of the lipid-associated part of the protein as small peptides, since this part of the protein is relatively rich in these residues, particularly in tryptophan (Allen & Green, 1978).

It was not expected that digestion with α-chymotrypsin would yield a complete set of peptides, since the relatively low specificity of the enzyme would give low yields of some peptides between alternative points of cleavage. Rather, a relatively small amount of protein (0.5 μmol) was digested, and the major components of the digest were isolated rapidly in sufficient yields for analysis and partial sequence determination only.

The details of the isolation of some peptides and amino acid analyses are contained in Supplementary Publication SUP 50102.

Materials and Methods

Most of the materials and methods used in the present work have been described previously (Allen & Green, 1978; Allen, 1980). α-Chymotrypsin was a thrice-crystallized product (batch no. 92C-8120) from Sigma, Kingston upon Thames, Surrey, U.K. Iodo[^{14}C]acetic acid, from The Radiochemical Centre, Amersham, Bucks., U.K., was diluted to a specific radioactivity of 405 c.p.m./nmol. The reduced and [^{14}C]carboxymethylated lipid-free ATPase protein, prepared as described previously (Allen & Green, 1978) had 26.5 carboxymethyl groups per molecule, assuming a mol.wt. of 115000 (Thorley-Lawson & Green, 1973).

Digestion with α-chymotrypsin

The freeze-dried carboxymethylated protein (60 mg, 520 nmol) was dissolved in 8M-urea (3 ml). The solution was added, with stirring, to a solution of α-chymotrypsin (1.0 mg) in 0.1M-NH_4HCO_3/

* Present address: Department of Immunochemistry, The Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, U.K.

0306-3275/80/060565-11$01.50/1

1mM-$CaCl_2$ (9ml). Digestion was for 12h at 20°C, with continuous stirring. The digestion was stopped by freeze-drying.

Fractionation of peptides

The digest was dissolved in water (2ml). 2-Mercaptoethanol (1μl) was added. The solution was applied to a column (1.7cm × 135cm) of Sephadex G-50 (Superfine grade) in 50mM-NH_4HCO_3 containing 0.01% thiodiglycol. The column was eluted with this solution at a flow rate of 15ml/h, and fractions of 2.5ml were collected. Peptides were detected by liquid-scintillation counting and u.v. absorbance (A_{280}). Fractions 91–130, close to the total column volume, were combined and freeze-dried. This material was chromatographed on a column (1.7cm × 52cm) of Sephadex G-25 (Fine grade) in 50mM-NH_4HCO_3 containing 0.01% thiodiglycol, and fractions of 1.8ml were collected. Fractions from the two column eluates were combined as indicated in Figs. 1 and 2, and peptides were purified from these combined fractions by suitable combinations of ion-exchange chromatography, paper and thin-layer electrophoresis and chromatography. The details of the purification methods for each peptide are contained in Supplementary Publication SUP 50103, and are outlined in the Results section.

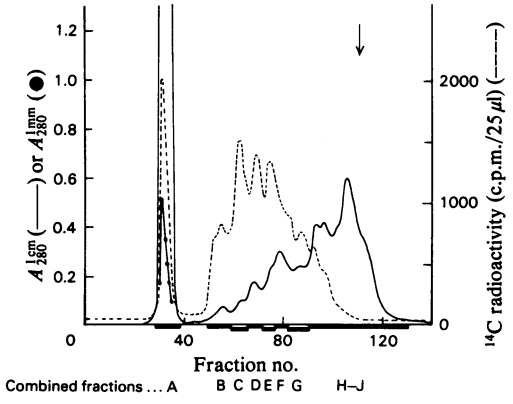

Fig. 1. *Separation of peptides from the chymotryptic digest of the Ca^{2+}-transporting ATPase protein on a column of Sephadex G-50 (Superfine grade)*

The column size was 1.7cm × 135cm and the buffer was 50mM-NH_4HCO_3/0.01% thiodiglycol. The fraction size was 2.5ml. ——, A_{280}^{1cm}; ●, A_{280}^{1mm}; – – –, ^{14}C radioactivity (c.p.m./25μml). The vertical arrow shows the position of elution of urea, which was present in the sample applied. The horizontal bars indicate the combination of fractions.

Preparative peptide 'maps'

The two-dimensional peptide 'mapping' technique (Ingram, 1956; Katz *et al.*, 1959) was adapted for larger quantities (up to 500nmol of each peptide in a mixture of about 20 peptides) than those normally separable by this method. The sample was applied as a strip up to 12cm long on Whatman 3MM paper, and electrophoresis at pH6.5 was performed for 40min at 60V/cm, in pyridine/acetic acid/water, (20:1:179, by vol.). The partially resolved bands of peptides were carefully concentrated to spots along a line by application of buffer to the paper on each side of the zone of peptide material. The procedure was monitored by the inclusion of coloured internal markers, Xylene Cyanol FF and ε-*N*-dinitrophenyl-lysine (Stevenson, 1971) with the sample. The paper was dried, and descending chromatography was performed for 16h at right angles to the direction of electrophoresis, in butanol/acetic acid/water/pyridine (15:3:12:10, by vol.). Peptides were located on the dried papers by dipping in dilute fluorescamine (10mg/litre) solution in acetone/pyridine (99:1, v/v), followed by viewing of the air-dried papers under the u.v. lamp (366nm). Papers were also subjected to radioautography for the location of [^{14}C]carboxymethylcysteine-containing peptides. Areas containing peptides were excised and peptides were eluted with 0.1M-NH_3.

Sequence determination

Peptides were sequenced by the manual micro 'dansyl–Edman' method described in the preceding

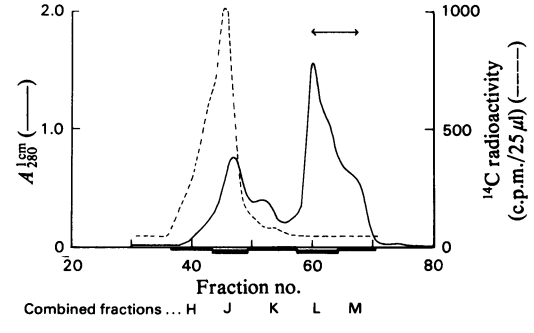

Fig. 2. *Separation of peptides from fractions H–J (Fig. 1) on a column of Sephadex G-25 (Fine grade)*

The column size was 1.7cm × 52cm and the buffer was 50mM-NH_4HCO_3/0.01% thiodiglycol. The fraction size was 1.8ml. ——, A_{280}^{1cm}; – – –, ^{14}C radioactivity (c.p.m./25μl). The bar at the top of the Figure shows the position of elution of urea. Fractions were combined as indicated by the horizontal bars at the base of the Figure.

paper (Allen, 1980). Differentiation between aspartic acid and asparagine or glutamic acid and glutamine was made on the basis of the electrophoretic mobilities of the peptides at pH 6.5 (Offord, 1966), wherever possible.

Nomenclature of peptides

Peptides from the chymotryptic digest are prefixed by Ch. Further symbols are added according to the positions of migration of the peptides on the various chromatographic systems used in the purifications. Peptides from other digests are labelled as described in the preceding paper (Allen, 1980). The positions of the chymotryptic peptides within the five extended sequences discussed in an accompanying paper (Allen *et al.*, 1980) are noted in square brackets. The standard three-letter code for amino acid residues in peptides sequences is used; Asx represents either aspartic and asparagine, and Glx represents either glutamic acid or glutamine. Chemically modified residues are marked with an asterisk; thus Cys* represents *S*-carboxymethylcysteine. Sequences of peptides that were clearly determined by the dansyl–Edman method are marked by single-headed arrows beneath the residues, thus:

Val-Glx-
→ →

Where the results of the sequence determination were slightly doubtful, usually because too little material was present, the arrow is broken, thus:

Val-Glx
⇢ ⇢

Unidentified positions in sequences are shown thus:

-(-)-

Groups of residues that were determined by amino acid analysis but that were not clearly identified in sequences are enclosed by parentheses and separated by commas.

Results

The freeze-dried carboxymethylated protein did not dissolve readily in 8 M-urea. Initially the protein swelled to form a gel, which was brought into solution slowly by the use of a glass homogenizer. Upon dilution into the solution of α-chymotrypsin, a finely-divided suspension was produced. This dissolved during the digestion. After freeze-drying, the digest dissolved readily.

As shown in Fig. 1, most of the material with absorbance at 280 nm was eluted close to the void volume of the Sephadex G-50 column (fraction A). By analogy with the results from other digests of the protein (Allen & Green, 1978; Allen, 1980), this material consists of tightly aggregated, rather hydrophobic peptides that are not separable in aqueous solutions, and no attempt was made to purify peptides from this fraction. The A_{280} peak at the total column volume (Fig. 1) was due to 2-mercaptoethanol, as shown by spectral analysis. Tryptophan was detected only in fractions A and J (Figs. 1 and 2).

Peptides from fraction B

These were resolved into fractions B1–B8 by ion-exchange chromatography on a column of DEAE-cellulose in a gradient of NH_4HCO_3, as described in Supplementary Publication SUP 50103. Peptides were further purified as outlined below.

Peptide Ch.B2a [*2(31–58)*]

This was purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system); its structure is as follows:

Arg-Ile-Lys-Ala-Arg-Asx-Ile-Val-Pro-Gly-Asx-Ile-
→ → ⇢ → → → → → → → → →

Val-Glx-Val-Ala-Val-Gly-Asx-(-)-Val(Asx,Pro,Ala,
→ ⇢ ⇢ ⇢ ⇢ ⇢ ⇢ ⇢

Ile_2,Leu,Lys,Arg)

The analysis was low for valine and isoleucine, which is to be expected, since there are two isoleucine–valine bonds that are hydrolysed slowly in 6 M-HCl.

Peptide Ch.B3a [*2(113–135)*]

This was purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system); the structure was:

Gly-Ile-Val-Ala(Asx_2,Thr_4,Ser,Glu_2,Pro,Gly,Ala,
→ → → ⇢

Val,Ile,Leu,Lys_3,Arg)

The yield of this peptide was low and further sequencing was not successful.

Peptide Ch.B6a [*3(83–109)*]

This was purified on a column (1 cm × 142 cm) of Sephadex G-50 (Superfine grade) in 50 mM-NH_4HCO_3/0.01% (v/v) thiodiglycol. The sequence was:

Asx-Gly-Leu-Val-Glx-Leu-Ala-Thr-Ile-Cys*-Ala-
→ → → → → → → → → → →

Leu-Cys*-Asx-Asx-Ser(Asx_3,Thr,Ser,Glx,Gly,Val,
→ → ⇢ ⇢ ⇢

Leu,Tyr,Phe,Lys)

Peptides Ch.B7a and Ch.B7b

These were purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system). Apart from differences in their chromatographic properties, prob-

ably caused by different extents of autoxidation of methionine and carboxymethylcysteine residues, these peptides were identical with peptide Ch.B8a.

Peptide Ch.B8a

This was purified on a column of Sephadex G-50 (Superfine grade) as described for peptide Ch.B6a. The sequence was:

Met-Glx-Cys*-Thr-Glx-Asx-His-Pro-His-Phe-Glx-

Gly-Leu-Asx(Cys*,Asx,Glx_3,Pro_2,Ala,Met,Ile,Phe)

This sequence was identified previously (Allen & Green, 1978) in peptides from the tryptic digest of the carboxymethylated protein.

Peptides from fraction C

These were resolved into fractions C1–C6 by ion-exchange chromatography on a column of DEAE-cellulose as described in Supplementary Publication SUP 50103. Peptides in each of the six fractions were purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system).

Peptide Ch. C1a [2(73–99)]

This was not completely pure by analysis, and the yield was low. Sequence information determined was:

(-)-Gly-Glx-(-)-Val-Ser-Val-Ile-Lys-His-Thr-Glx-

Pro-Val-Pro-Asx(Asx_3,Thr,Ser,Glx,Pro,Gly,Ala_2,

Val,Met,Lys,Arg)

Peptide Ch.C1b [2(73–101)]

This was identical with peptide Ch.Cla, but with the addition of a single residue each of isoleucine and phenylalanine, presumably at the *C*-terminus.

Peptide Ch.C1d [2(31–58)]

This had some impurities, as shown by the analysis, but its partial sequence:

Arg-Ile-(-)-Ala-(-)-(-)-Ile-Val-Pro-Gly-Asx-Ile-Val-

(-)-Val-Ala-

identified it with peptide Ch.B2a.

Peptide Ch.C2b [3(238–259)]

This was also not completely pure, but the sequence:

Ala-Leu-Ala-Thr-(-)-Asx-Thr-Pro-Pro-Lys-(-)-Glx-

Pro-Met-Val-Leu-(Asx_2,Ser_2,Arg_2,Phe)

identified it as a sequence overlapping tryptic peptides T1E-5 and T1C-6.

Peptide Ch.C3b [3(6–31)]

This was isolated in low yield, but the analysis and the partial sequence:

Ala-(-)-(-)-(-)-(-)-Leu-Pro-Ser-Val-Glx-Thr-

identified it as part of the sequence containing the active-site aspartic acid residue (Allen & Green, 1976). The low values for valine and isoleucine in the analysis were due to the presence of slowly hydrolysed Val–Ile and Ile–Val sequences in the peptide.

Peptide Ch.C3c [2(36–65)]

This was isolated in low yield, and the analysis was not accurate. The partial sequence:

Asx-Ile-Val-Pro-(-)-(-)-Ile-Val-Glx-Val-Ala-

identified this peptide as beginning within the sequence of peptide Ch.B2a, but additional residues (Thr_2,Ser_2,Leu) are present at the *C*-terminus of peptide Ch.C3c.

Peptide Ch.C4b [3(10–31)]

The structure is:

Ser-Leu-Pro-Ser-Val-Glx-Thr-Leu-Gly-Cys*-Thr-Ser-

Val-Ile-Cys*-Ser(Asx,Lys,Thr_2,Gly,Leu)

This peptide is part of the 'active-site' peptide sequence (Allen & Green, 1976). The analysis gave low values for valine and isoleucine, owing to the presence of a valine–isoleucine bond.

Peptide Ch.C4c [2(36–58)]

The structure was:

Asx-Ile-Val-Pro-Gly-(-)-Ile-Val-Glx-Val-Ala-Val-

Gly-Asx-(-)-Val(Asx_2,Pro,Ala,Ile_2,Leu,Lys,Arg)

the peptide was identical with the *C*-terminal 23 residues of peptide Ch.B2a, and the *N*-terminal 23 residues of peptide Ch.C3c.

Peptides Ch.C5b and Ch.C5c [4(53–74)]

These were identical but for the presence of two additional residues of glutamic acid and some impurities in the analysis of the former. The partial sequence of peptide Ch.C5b was:

Leu-Glx-Ser-Tyr-Asx-Glx-Ile-Thr-Ala-Met-

and the analysis of peptide Ch.C5c was in agreement with the peptide forming part of a known peptide, T1B-1. The additional two glutamic acid residues in peptide Ch.C5b could not be accounted for in this sequence, and were presumably contributed by impurities. The different chromatographic properties of peptide Ch.C5b and Ch.C5c

were probably caused by different extents of autoxidation of the methionine residue.

Peptide Ch.C6a1 [*4(16–38)*]

The structure was:

Thr-Gly-Arg-Glx-Phe-Asx-Asx-Leu-Pro-Leu-Ala-

Glx-Glx-Arg-Glx-Ala-Cys*-Arg-Arg-Ala-Cys*-Cys*-

Phe

This was a very useful peptide, overlapping four tryptic peptides, T1E-22, T1D-11, T1E-20 (including T1E-9) and T1E-8. The complete sequence of this 23-residue peptide was determined from 30 nmol of peptide.

Peptides from fraction D

These were initially resolved into eight fractions, D1–D8, by ion-exchange chromatography on a column of DEAE-cellulose, as described in the Supplementary Publication SUP 50103.

Peptide Ch.D2b1

This was purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system) followed by t.l.c. on cellulose MN400 (butanol/acetic acid/water/pyridine system). The peptide was analysed as:

Gly(Asx_2,Thr_3,Ser,Glx_2,Gly_2,Ala,Val_2,Met,Ile_3,Lys,

Arg)

but the yield was too low for sequence determination.

Peptide Ch.D2c [*3(269–283)*]

This was purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system). The sequence was:

Val-Gly-Val-Val-Gly-Met-Leu-Asp-Pro-Pro-

Arg(Lys)Glu-Val-Met

The lysine residue was not positively identified by the dansyl method, but was placed by difference analysis. The peptide was neutral, thus residues 8 and 13 were acidic. The analysis for valine was low, owing to the slowly hydrolysed Val–Val dipeptide sequence.

Peptide Ch.D3f

This was purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system). The sequence was:

Leu-Thr-Ala-Leu-Gly-Leu-Pro-Gln-Ala-Leu-Ile-

Pro-Val-Gln-Leu(Met)

The peptide was neutral; hence residues 9 and 15 were glutamine rather than glutamic acid. This peptide sequence has not been found among the soluble tryptic peptides from the succinylated protein, but has been located within one of the large aggregated peptides assumed to be associated with the lipid bilayer in the native structure (G. Allen, unpublished work).

Peptide Ch.D4a1

This was purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system) followed by thin-layer electrophoresis on silica gel at pH 2 (formic acid/acetic acid/water, 1:4:45, by vol.). The yield was low and the sequence results were rather poor:

Gly-Phe-Asx-(-)-(-)-Asx-(-)-Asx-Ile-(-)-Asx-.

The peptide was identified as containing the sequence of the tryptic peptide T1D-3, and this sequence has also been found among the large, aggregated tryptic peptides (G. Allen, unpublished work).

Peptide Ch.D4b

This was purified by t.l.c. (silica gel, butanol/acetic acid/water/pyridine system). The analysis showed the presence of impurities (threonine and alanine), but the sequence was clear:

Lys-Ile-Ser-Leu-Pro-Val-Ile-Gly-Leu-Asx-Glx-Ile-

Leu

This sequence was not found among the soluble tryptic peptides from the succinylated protein.

Peptide Ch.D5a1 [*3(18–31)*]

This was purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system), followed by thin-layer electrophoresis on silica gel at pH 2, as for peptide Ch.D4a1. The sequence results:

Gly-Cys*-Thr-Ser-Val-Ile-Cys*-Ser-Asx-(-)-(-)-Gly-

(-)-Leu

and the analysis served to identify this peptide as part of the 'active-site' sequence (Allen & Green, 1976).

Peptide Ch.D5a2 [*3(110–120)*]

This was purified in the same way as peptide Ch.D5a1. The sequence was:

Glx-Lys-Val-Gly-Glx-Ala-Thr-Glx-Thr-Ala-Leu

The peptide was identical with peptide Ch.Ea7a.

Peptide Ch.D6a [*1(1–16)*]

This was purified by t.l.c., as for peptide Ch.D4b. It had a blocked *N*-terminus. The analysis:

$\overset{*}{Cys}$,Thr,Ser_2,Glu_3,Ala_3,Met,Leu,Tyr,Phe,His,Lys

served to identify it as the *N*-terminal 16 residues of the protein.

Peptide Ch.D8 [*3(43–57)*]

The sequence was:

Ile-Ile-Asx-Lys-Val-Asx-Gly-Asx-Phe-$\overset{*}{Cys}$-Ser-Leu-Asx-Glx-Phe

The analysis gave low values for aspartic acid and isoleucine, the latter owing to the Ile–Ile dipeptide sequence, which is hydrolysed slowly.

Peptides in fraction E

These were initially resolved by the preparative peptide-'mapping' procedure on Whatman 3MM paper. The results are shown in Fig. 3(*a*). Further purification of peptides was by t.l.c. on silica gel (butanol/acetic acid/water/pyridine system).

Peptide Ch.En1a [*3(240–251)*]

This was not pure, but the sequence results, although poor:

Ala-(-)-(-)-Asx-Thr-Pro-(-)-(-)-(-)-Glx-Glx-Met

served to place this peptide within peptide Ch.C2b.

Peptide Ch.En2a [*4(83–95)*]

This was not pure, but the sequence:

Gly-Ser-(-)-Thr-Ala-Val-Ala-(-)-Thr-Ala-(-)-Glx-Met

is within one of the long tryptic peptides from the succinylated protein, peptide T1B-1. In the latter peptide, the third residue in this sequence is glycine, the eighth is lysine and the eleventh is serine. The neutrality of the peptide shows that the twelth residue is glutamic acid. However, the analysis indicated that an aspartic acid residue was present in peptide Ch.En2a, which makes these identifications tentative.

Peptide Ch.En2c

This was analysed:

Asp_2,Thr,Ser,Glu_2,Pro_2,Gly,Ala,Val,Met,Ile,Leu_2,Lys,Arg

but no sequence results were obtained.

Peptide Ch.En4a [*3(269–283)*]

This was identical with peptide Ch.D2c.

Peptide Ch.En5b

This was impure by analysis. The sequence was:

Val-Asx-Leu-Val-Thr-Asx-Gly-Leu-Pro-Ala-Thr-Ala-Leu

At residues 5 and 6, some alanine and leucine residues were observed, respectively, in addition to the major residues, threonine and aspartic acid. This sequence has been found in one of the aggregated tryptic peptides from the succinylated protein (G. Allen, unpublished work), suggesting that it originates within, or close to, the lipid bilayer in the native lipoprotein.

Peptide Ch.Ea7a [*3(110–120)*]

The sequence was:

Glu-Lys-Val-Gly-Glu-Ala-Thr-Glu-Thr-Ala-Leu

The peptide had an electrophoretic mobility at pH 6.5 of 0.5, showing a net negative charge of 2.

Peptide Ch.Eb1 [*1(17–?)*]

This gave a poor analysis, but the *N*-terminal sequence:

Gly-Val-Ser-Glx-(-)-Thr-(-)-Leu-

identified it with a known sequence, in peptide T1C-7.

Peptide Ch.Eb3 [*4(39–52)*]

The sequence was:

Ala-Arg-Val-Glx-Pro-Ser-His-Lys-Ser-Lys-Ile-Val-Glx-Tyr

had two net positive charges, from the mobility at pH 6.5, indicating that there were two glutamic acid residues. The peptide (10 nmol) was digested with trypsin (10 μg) for 2 h at 35°C in 50 μl of 50 mM-NH_4HCO_3. Peptide fragments were isolated on a two-dimensional peptide 'map' (pH 6.5 electrophoresis followed by chromatography in butanol/acetic acid/water/pyridine, on cellulose MN-300 thin layers) and were characterized by *N*-terminal and amino acid analysis by the dansyl technique (Allen, 1980). The fragments isolated were:

Ile(Val,Glu,Tyr)

mobility (pH 6.5) −0.30,

Val(Glu, Pro, Ser, His, Lys)

mobility +0.41, and a mixture of Ser(Lys) and

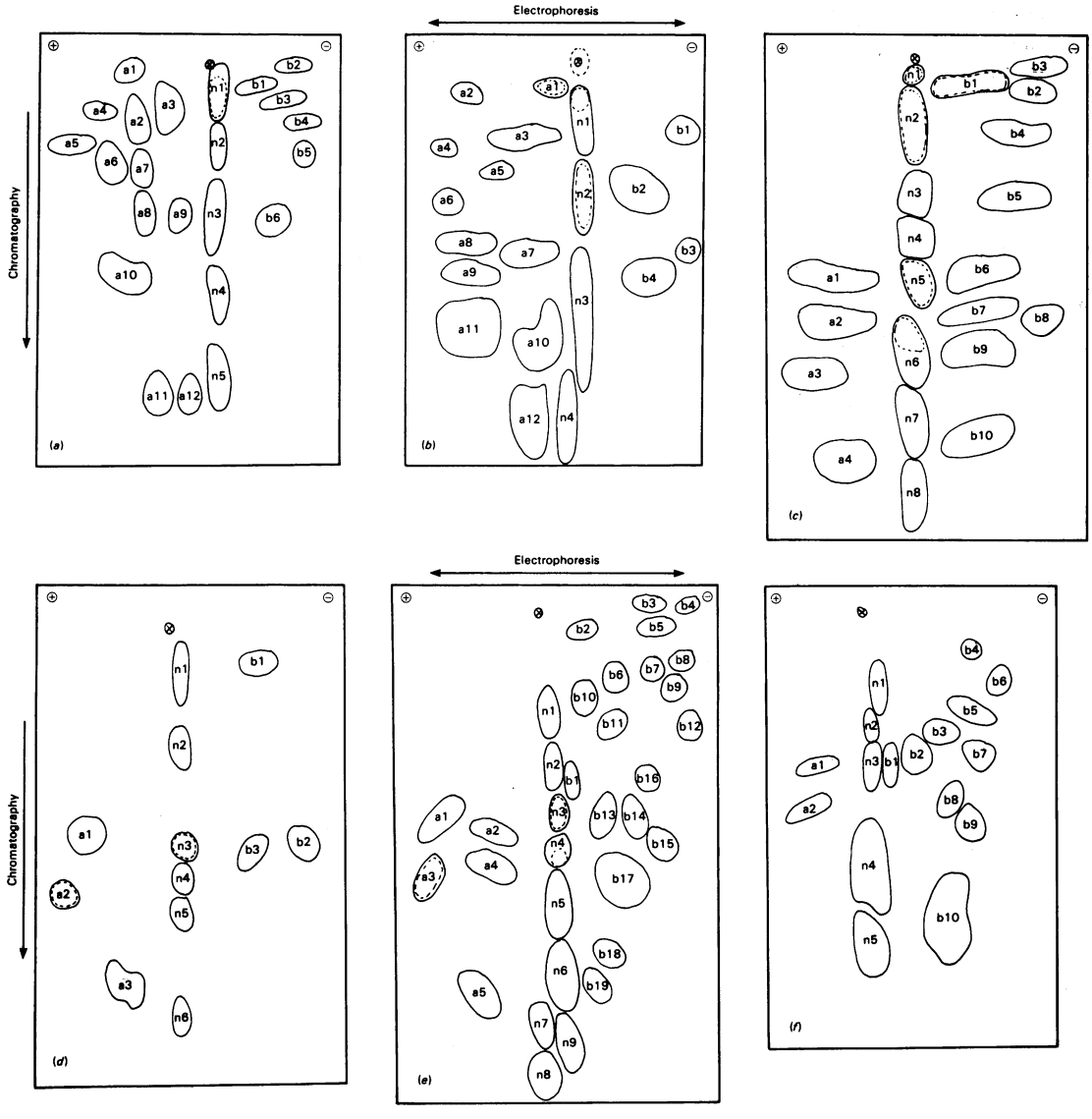

Fig. 3. *Preparative peptide 'maps' of peptides in fractions E–K of the chymotryptic digest*
Electrophoresis was performed at pH 6.5, and descending chromatography was in butan-1-ol/acetic acid/pyridine/water (15:3:12:10, by vol). Fluorescamine-positive spots are indicated by a full line (——), and radioactive spots, identified by radioautography, by a pecked (– – – –) line. The origin is marked ⊗. (*a*) Fraction Ch.E; (*b*) Ch.F; (*c*) Ch.G; (*d*) Ch.H; (*e*) J; (*f*) K.

Ala(Arg), mobility +0.73. This showed that the sequence of peptide Ch.Eb3 was:

Ala-Arg-Val-Glu-Pro-Ser-His-Lys-Ser-Lys-Ile-Val-Glu-Tyr.

Peptides in fraction F

These were initially resolved by the preparative peptide-'mapping method', as shown in Fig. 3(*b*). Further purification of peptides was by t.l.c., as for peptides from fraction E.

Peptide Ch.Fn3b

This was analysed, but after attempted further purification by t.l.c., no peptides were recovered. The analysis gave:

Lys(Asp,Thr,Glu,Gly,Ala,Val,Met,Ile,Leu)

with impurities, and this peptide does not fit within any of the known tryptic peptides from the succinylated protein.

Peptide Ch.Fa5a[*4(53–62)*]

The structure:

Leu-Glx-Ser-Tyr-Asx-Glx-Ile-Thr-Ala-Met

was identical with the *N*-terminal portion of peptide Ch.C5b. The mobility at pH 6.5, −0.39, showed the presence of two acidic residues.

Peptide Ch.Fa9a [*4(96–102)*]

This was isolated in a yield too low for quantitative analysis. The sequence was:

Val-Leu-Ala-Asx-(-)-Asx-Phe

and the mobility at pH 6.5 indicated a net negative charge of 2.

Peptide Ch.Fa11a [*3(65–71)*]

This had the structure:

Ala-Pro-Glu-Gly-Gly-Val-Leu

and had a mobility of −0.5.

Peptide Ch.Fb2a [*3(213–221)*]

This had the structure:

Thr-Gly-Pro-Val-Lys-Glu(Lys,Ile,Leu)

and a mobility of +0.26. The analysis showed the presence of an additional proline residue in peptide Ch.Fb2a, which is inconsistent with the sequence of peptide T1C-4, and which was probably due to impurities.

Peptide Ch.Fb2b [*2(102–112)*]

This peptide had the structure:

Ser-Gly-Thr-Asn-Ile(Gly,Ala_3,Lys,Leu)

and had a mobility +0.26.

Peptides in fraction G

These were initially resolved on a peptide 'map' (Fig. 3*c*). Further purification was by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system).

Peptide Ch.Gn1 [*3(138–147)*]

This has the structure:

Ser-Lys-Val-Glx-Arg-Ala-Asx-Ala-C*ys-Asx

and overlaps peptides T1D-4 and T1D-1. The net charge of this peptide, Ch.Gn1, is zero, which is in agreement with the assignment of acid and amide residues in the tryptic peptides.

Peptide Ch.Gn3b [*3(131–137)*]

This required further purification by thin-layer electrophoresis on silica gel at pH 2. The sequence was determined as:

Asn-Thr-Glx-Val-Arg-Asx-Leu

The ninhydrin/Cd reagent gave a yellow colour with the peptide, showing the *N*-terminal residue to be asparagine, and the net charge of the peptide at pH 6.5 was zero. The peptide overlaps peptides T1B-3 and T1D-4. Since the latter peptide has *N*-terminal asparagine, the third residue in peptide Ch.Gn3b is therefore glutamic acid.

Peptide Ch.Gn4b [*3(228–234)*]

This was also further purified by thin-layer electrophoresis at pH 2 on silica gel. The neutral peptide was:

Gly-Thr-Gly-Arg-Asp-Thr-Leu

Peptide Ch.Gn6a [*4(75–82)*]

This had the structure:

Lys-Ala-Glu-Ile-Gly-Ile-Ala-Met

and had a mobility at pH 6.5 of 0.0.

Peptide Ch.Gn6b [*2(66–72)*]

This was further purified by t.l.c. (silica gel; butanol/acetic acid/water system, followed by cellulose MN400, butanol/acetic acid/water system). The sequence of the neutral peptide was

Arg-Val-Asx-Glx-Ser-Ile-Leu

Peptide Ch. Gn6c

This was impure by analysis, but gave the clear sequence:

Gly-Ser-Ile-Gln-Leu (neutral peptide)

Peptide Ch.Gn8a

This peptide, of structure:

Ala(Leu, Ser,Val,Leu)

was not sequenced, but was recognized as forming the *N*-terminal five residues of a CNBr peptide, CB-D3c1.

Peptide Ch.Ga2a [*5(6–8)*]

This peptide, of structure:

Leu(Glu,Gly)

was identical with peptide Ch.Ha1.

Peptide Ch.Ga3a [*3(65–71)*]

This peptide, of structure:

Ala-Pro-Glu-Gly-Glu-Val-Leu (mobility −0.4)

was identical with peptide Ch.Fa11a.

Peptide Ch.Gb1 [*3(173–180)*]

The structure was:

Cys*-Ser-Pro-Ala-Lys-Ser-Ser-Arg

Peptide Ch.Gb2a [*2(23–30)*]

This was further purified by t.l.c. (silica gel; butanol/acetic acid/water system; R_F 0.0). The peptide:

Arg-Ala-Asx-Arg-Lys-Ser-Val-Glx (mobility +0.43)

had a net positive charge of 2 at pH 6.5.

Peptide Ch.Gb9a1 [*2(106–112)*]

This was further purified (t.l.c.; silica gel, butanol/acetic acid/water system), but was still impure by analysis. The sequence was:

Ile-Ala-Ala-Gly-Lys-Ala-Leu

The analysis was consistent with the presence of peptide Ch.Gb9a2, which had similar chromatographic properties, as an impurity.

Peptide Ch.Gb9a2 [*2(59–65)*]

This was further purified (t.l.c.; silica gel; butanol/acetic acid/water system). The sequence was;

Ser-Ile-Lys-Ser-Thr-Thr-Leu

Peptides in fraction H

These were resolved on a peptide 'map' (Fig. 3*d*).

Peptide Ch.Hn4b [*3(284–288)*]

This was purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system). The sequence was:

Gly-Ser-Ile-Gln-Leu

and the peptide was identical with peptide Ch.Gn6c.

Peptide Ch.Ha1 [*5(6–8)*]

This peptide, of structure:

Leu-Glu-Gly (mobility at pH 6.5, −0.46)

was probably the *C*-terminal tripeptide of the ATPase protein, since α-chymotrypsin does not normally cleave peptides at glycyl–peptide bonds, and a tryptic peptide, T1E-3, Asn-Tyr-Leu-Glu-Gly, was also isolated in high yield from the digest of the succinylated protein (Allen, 1980).

Peptide Ch.Ha3a [*3(159–162)*]

This was further purified by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system). The sequence was:

Thr-Leu-Glu-Phe (mobility at pH 6.5, −0.31)

Peptide Ch.Hb1 [*3(155–158)*]

The structure of this was:

Lys-Lys-Glu-Phe (mobility, pH 6.5, +0.20)

Peptides in fraction J

These were partially resolved on a peptide 'map' (Fig. 3*e*), and further purification was by t.l.c. (silica gel; butanol/acetic acid/water/pyridine system) if required.

Peptide Ch.Jn3a [*3(235–237)*]

The structure was:

Arg-Cys*-Leu

Peptide Ch.Jn5a

This was isolated in low yield, and was qualitatively analysed only as Thr(Met).

Peptide Ch.Jn5b [*3(222–227)*]

This peptide, of structure:

Ser-Val-Ile-Lys-Glx(Trp)

was the only chymotryptic peptide purified that contained tryptophan. The tryptophan was identified by a positive Ehrlich test on the peptide. The peptide was not quantitatively analysed; it was identified with a peptide sequence within peptide T1C-4, the only soluble tryptic peptide from the succinylated protein that contained tryptophan.

Peptide Ch.Jn6a

This gave a clean sequence:

Ala-Leu-Leu-Gly-Gly-Tyr

that was inconsistent with the analysis, which showed the presence of two residues of alanine and one of glycine and the presence of impurities. The sequence has not been found in peptides from other digests of the protein.

Peptide Ch.Ja3

This peptide, of sequence:

Cys*-Leu

was identical with a peptide isolated previously from a digest of the protein with trypsin and α-chymotrypsin (Allen & Green, 1978), and is possibly derived from the same part of the protein as peptide Ch.Jn3a, although it is more probably from a different part of the protein, since chymotrypsin does not normally cleave Arg–Cys bonds.

Peptide Ch.Jb1a

This peptide, of structure:

His-Leu-Glx-Tyr

was not quantitatively analysed; the sequence has not been identified elsewhere.

Peptide Ch.Jb3 [*3(163–167)*]

This had the structure:

Ser(Arg)Asp-Arg-Lys

The second residue was not positively identified by the dansyl method, but was placed by difference from the analysis. The mobility at pH 6.5, +0.54, indicated a net positive charge of 2.

Peptide Ch.JB4 [*3(2–5)*]

This peptide, of structure:

Ala-Lys-Lys-Asn (mobility at pH 6.5, +0.70)

was not quantitatively analysed.

Peptide Ch.Jb6 [*3(155–158)*]

This peptide, of structure:

Lys-Lys-Glu-Phe (mobility, pH 6.5, +0.37)

was identical with peptide Ch.Hb1.

Peptide Ch.Jb10 [*4(117–122)*]

This peptide, of structure:

Asn-Asn-Met-Lys-Gln-Phe (mobility, pH 6.5, +0.20)

gave an initially yellow colour with the ninhydrin/Cd reagent, confirming that the *N*-terminal residue was asparagine. The peptide sequence was not found within the soluble tryptic peptides.

Peptide Ch.Jb12 [*2(31–33)*]

This peptide, of structure:

Arg-Ile-Lys

was not quantitatively analysed.

Peptide Ch.Jb17a [*3(6–9)*]

This peptide, of structure:

Ala-Ile-Val-Arg

was not quantitatively analysed.

Peptides from fraction K

These were resolved on a peptide 'map' (Fig. 3*f*).

Peptide Ch.Kb2a

This was purified by t.l.c. (silica gel; butanol/acetic acid/water system). The sequence was not clearly determined, owing to partial removal of two residues during phenyl isothiocyanate degradation at each of the histidine residues, but the structure His-Glx-Leu-Thr-His-Phe was consistent with the data and with the structure of peptide CB-E2b.

Peptide Ch.Kb3 [*5(1–5)*]

This peptide, of structure:

Ile-Ala-Arg-Asn-Tyr (mobility, pH 6.5, +0.35)

overlaps the *C*-terminal tryptic peptide of the protein, peptide T1E-3. Quantitative analysis was not performed.

Peptides Ch.Kb7 [*3(126–127)*] *and* [*3(186–187)*]

This was a mixture identified by qualitative analysis and sequencing as

Lys-Phe and Lys-Met

Peptide Ch.Kb8

This was purified by t.l.c. (silica gel; butanol/acetic acid/water system). The sequence was

Ile-Arg-Tyr (mobility, pH 6.5, +0.36)

which has not been identified as an overlap of the soluble tryptic peptides. Quantitive analysis was not performed.

Fraction L material

This was resolved on a peptide 'map', but only the free amino acid, leucine, was identified, in a yield of 60 nmol.

Discussion

For the isolation of overlap peptides between a set of tryptic peptides of known sequence, the approach described here, of digesting a relatively small amount (0.5 μmol) of protein and isolating peptides by rapid methods (peptide 'mapping' and t.l.c.) was found to be satisfactory. The amino acid analyses and partial sequences were generally sufficient for the correlation of the peptides. However, peptides longer than about 15 residues were usually obtained in yields too low for the determination of their complete sequences, and, in particular, for the subfragmentation required for the identification of aspartic acid or asparagine and glutamic acid or

glutamine by the methods used here. Amino acid analyses for peptides isolated in low yields were not always accurate, although few problems in the correlation of sequences were encountered.

The method of preparative peptide 'mapping' described here leads to lower resolution (as shown in Fig. 3) than that obtained on the analytical scale, but avoids distortion during electrophoresis resulting from overloading the paper.

The use of fluorescamine in low concentrations for the detection of peptides led to little loss of the *N*-terminal residues of peptides.

The α-chymotrypsin used here had some tryptic activity, as shown by several cleavages as arginyl and lysyl peptide bonds. However, almost all of the cleavages observed were typical for chymotrypsin, at phenylalanyl, tyrosyl, tryptophyl, methionyl and leucyl peptide bonds, and occasionally at asparaginyl (giving peptides Ch.Gn1 and Ch.Jb4) and glutamyl or glutaminyl (peptide Ch.Gb2a1) bonds.

A total of 629 unique residues were isolated in the chymotryptic peptides, of which 312 were determined in sequences. Of these residues, 531 were present in the soluble tryptic peptides from the succinylated protein (peptides T1; Allen, 1980); the remaining 98 residues were not contained within the sequences of the T1 peptides. Conversely, of the 586 unique residues identified in the T1 peptides, only 55 were not identified in the chymotryptic peptides. The total number of residues in the protein is about 1000; thus about 300 residues are present in peptide sequences that are not digested by either trypsin or chymotrypsin to give small soluble peptides. All but one of the approx. 18 tryptophan residues are present in these peptide sequences, which are probably embedded within the lipid bilayer in the native lipoprotein (Allen *et al.*, 1980).

I thank Mr. B. Trinnaman for technical assistance, Miss S. Lathwell for operating the amino acid analyser and Dr. N. M. Green, in whose laboratory this work was carried out, for his support.

References

Allen, G. (1980) *Biochem. J.* **187**, 545–563
Allen, G. & Green, N. M. (1976) *FEBS Lett.* **63**, 188–192
Allen, G. & Green, N. M. (1978) *Biochem. J.* **173**, 393–402
Allen, G., Trinnaman, B. & Green, N. M. (1980) *Biochem. J.* **187**, 591–616
Ingram, V. M. (1956) *Nature (London)* **178**, 792–794
Katz, A. M., Dreyer, W. J. & Anfinsen, C. B. (1959) *J. Biol. Chem.* **234**, 2897–2900
Offord, R. E. (1966) *Nature (London)* **211**, 591–593
Stevenson, M. J., (1971) *Anal. Biochem.* **40**, 29–34
Thorley-Lawson, D. A. & Green, N. M. (1973) *Eur. J. Biochem.* **40**, 403–413