Biochem. J. (1978) **173**, 393–402
Printed in Great Britain

Primary Structures of Cysteine-Containing Peptides from the Calcium Ion-Transporting Adenosine Triphosphatase of Rabbit Sarcoplasmic Reticulum

By GEOFFREY ALLEN* and N. MICHAEL GREEN
Biochemistry Division, National Institute for Medical Research, The Ridgeway, London NW7 1AA, U.K.

(*Received* 8 *November* 1977)

A preliminary investigation of the primary structure of the Ca^{2+}-transporting ATPase (adenosine triphosphatase) protein of rabbit skeletal-muscle sarcoplasmic reticulum is reported. The preparation of derivatives of delipidated protein in a form suitable for sequence analysis is described. Tryptic peptides containing *S*-carboxymethylcysteine residues were isolated from the reduced carboxymethylated protein, and their sequences were partially determined. The results are consistent with mol.wt. about 105000 for the polypeptide, and the absence of extended repeated lengths of sequence. The distribution of tryptophan and cysteine residues between large, aggregated peptides and soluble tryptic peptides shows that these residues are concentrated in different regions of the primary structure. This observation agrees with other evidence that these residues are, on the whole, widely separated in the native protein. The details of the procedures used to isolate the peptides, and the evidence for the determination of their sequences, are given in Supplementary Publication SUP 50085 (30 pages), which has been deposited at 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.

The Ca^{2+}-transporting ATPase from sarcoplasmic reticulum is a membrane protein of high molecular weight (MacLennan, 1970), part of which is buried in the membrane and part of which projects into the aqueous phase (Hardwicke & Green, 1974). Its structure is of considerable interest in relation both to mechanisms of ion transport and to membrane structure. Since it is also conveniently available we have commenced work on its primary structure: preliminary investigations are reported here.

Earlier papers from this laboratory have shown that the ATPase can be cleaved into three large fragments with trypsin (Thorley-Lawson & Green, 1973, 1975). Unfortunately it is difficult to separate these in sufficient quantities for sequence determination. We have therefore studied tryptic digests of the whole protein, concentrating our attention on the cysteine-containing peptides, since these can easily be radioactively labelled and are thus more readily purified from the complex mixture arising from such a large protein. They are also of intrinsic interest, since earlier work has shown that at least one cysteine residue is required for the catalytic activity of the enzyme (Hasselbach & Seraydarian, 1966) and the potential of cysteine residues for forming cross-links could provide information relevant to the tertiary structure. The sequence of the active-site peptide, determined in the present work, has been reported (Allen & Green, 1976).

Thorley-Lawson & Green (1977) describe the distribution of the cysteine residues between the tryptic fragments and give a preliminary account of their reactivity.

Abbreviations used: ATPase, adenosine triphosphatase; CmCys, *S*-carboxymethylcysteine.

* Present address: Max-Planck-Institut für Molekulare Genetik, Abteil Wittmann, Ihnestrasse 63–73, 1-Berlin 33 (Dahlem), Germany.

Materials and Methods

Sodium deoxycholate was from E. Merck (Darmstadt, Germany), trypsin (twice crystallized) from Worthington (Freehold, NJ, U.S.A.), α-chymotrypsin from Novo (Copenhagen, Denmark), thermolysin from Calbiochem, San Diego, CA, U.S.A., sodium taurodeoxycholate from Sigma, Kingston, Surrey, U.K., and urea and guanidinium chloride were from BDH (Poole, Dorset, U.K.). Sepharose and Sephadex and its derivatives were from Pharmacia (Uppsala, Sweden); DEAE-cellulose (DE-52) and paper for chromatography were from Whatman (W. and R. Balston, Maidstone, Kent, U.K.) De-acidite FFIP was from BDH, and analytical-grade resin AG50W 2X (200–400 mesh) was from Bio-Rad (Richmond, CA, U.S.A.).

Reagents for manual Edman degradation were from Pierce Chemical Co., Chester, Cheshire, U.K. Pyridine was redistilled from ninhydrin, and stored under N_2 at −10°C. Thiodiglycol was redistilled under

reduced pressure. Iodo[2-^{14}C]acetic acid was obtained from The Radiochemical Centre, Amersham, Bucks., U.K. The liquid-scintillation 'cocktail' used was toluene/2-methoxyethanol (3:1, v/v) containing 4.9g of 2,5-bis-(5-t-butylbenzoxazol-2-yl)thiophen/litre. Fluorescamine was from Roche Diagnostics, Nutley, NJ, U.S.A.

Preparation of ATPase

The ATPase was purified as a deoxycholate-solubilized lipoprotein from the white muscle of the back and hind legs of New Zealand White rabbits, by the method of MacLennan (1970), with the addition of 1mM-*N*-acetyl-L-cysteine to the extraction buffer. Deoxycholate was removed from the solubilized ATPase by slow passage through a column of anion-exchange resin, De-acidite FFIP, as described earlier (Hardwicke & Green, 1974). The ATPase was stored for up to 1 week as vesicles in 0.05M-Tris/HCl, pH8.2/1mM-histidine/0.66M-sucrose, to which 0.09vol. of half-saturated ammonium acetate was added. The concentration of ATPase protein was determined spectrophotometrically in 1% sodium dodecyl sulphate solution, by using $A_{280}^{1\%} = 12$ (Hardwicke & Green, 1974).

Carboxymethylation

The vesicles were suspended in 50mM-Tris adjusted to pH8.0 with HCl, and concentrated by centrifugation at 100000*g* for 20min. The lipoprotein vesicles were dissolved, at 10mg of protein/ml, in 50mM-Tris/HCl buffer (pH8.0)/5mM-EDTA/50mg of sodium taurodeoxycholate/ml. Solid guanidinium chloride was added (1g/ml of solution). The content of free thiol groups was determined in a sample by reaction with 5,5′-dithiobis-(2-nitrobenzoate) (Ellman, 1959) in the same solvent. A 10% excess of iodo[2-^{14}C]acetate solution (diluted to give 1200c.p.m./nmol) was added, and the solution was mixed with a stream of N_2 and kept under N_2 in the dark for 1h at 30°C. Dithiothreitol was added to 0.5mM and the solution was incubated under N_2 for 50min at 43°C. Tris base was added to 10mM, followed by iodo[^{14}C]acetate in 10% molar excess over the remaining thiol groups in the solution. After a further 30min at 43°C, dithiothreitol was added to 10mM, and the solution was dialysed exhaustively at 20°C against 50mM-Tris/acetic acid (pH7.0)/1mM-EDTA/1mM-2-mercaptoethanol, the final solution for dialysis containing in addition 1% (w/v) sodium dodecyl sulphate.

Delipidation of the protein

The clear solution of protein was treated with additional sodium dodecyl sulphate (10g/g of protein). The solution was applied to a column (2.4cm×142cm) of Sepharose 6B equilibrated in a solution containing sodium dodecyl sulphate (1%, w/v), 50mM-Tris, 1mM-EDTA and 1mM-2-mercaptoethanol, adjusted to pH7.0 with acetic acid. Elution was in this solvent, at a flow rate of 12ml/h, and 4ml fractions were collected. Fractions were analysed by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis as described earlier (Thorley-Lawson & Green, 1973). Fractions were also tested for the presence of phospholipid (McClare, 1971). The fractions containing the ATPase polypeptide were combined. The solution was placed in a dialysis bag, which was surrounded by solid urea. By this means, the protein solution was concentrated and made approximately 8M in urea after overnight dialysis. The solution was passed through a column (1.2cm×9cm) of Bio-Rad AG1 X2 anion-exchange resin equilibrated with 50mM-Tris/acetic acid buffer (pH7.0)/8M-urea, for the removal of dodecyl sulphate (Weber & Kuter, 1971).

The protein solution was dialysed exhaustively against several changes of 50mM-NH_4HCO_3/1mM-EDTA/10mM-2-mercaptoethanol, followed by several changes of water. The solution remained clear, at 1mg/ml, and a sample was analysed for content of [^{14}C]carboxymethyl groups. The protein solution was freeze-dried.

In other preparations, in which double the concentrations of protein and taurodeoxycholate were used during the carboxymethylation, the solution gelled after some time and the incorporation of [^{14}C]carboxymethyl groups was somewhat lower. Taurodeoxycholate and guanidinium chloride gave a gel at these higher concentrations even in the absence of protein.

Tryptic digestion and purification of peptides

The protein was digested with 1% (w/w) of trypsin in a pH-stat (Radiometer, Copenhagen, Denmark), at 38°C under N_2, with 1M-NaOH added to maintain pH8.3. After 2h a further 1% (w/w) of trypsin was added, and digestion was continued for a total 5.5h before freeze-drying. The protein was initially insoluble, but during the digestion it slowly dissolved.

The digest was dissolved at 20mg/ml in 50mM-NH_4HCO_3, with the addition of a little NH_3 to attain complete solubility. Peptides were initially fractionated on a column (1.8cm×133cm) of Sephadex G-50 (superfine grade) in 50mM-NH_4HCO_3 containing 0.01% (v/v) thiodiglycol, which decreases the extent of autoxidation of thioether groups (Harris, 1967). Fractions were analysed by paper electrophoresis at pH6.5, and [^{14}C]carboxymethylcysteine-containing peptides were purified by combinations of ion-exchange chromatography, paper electrophoresis at pH6.5, 3.5, 2 and 9 (Ambler, 1963) and descending paper chromatography in butan-1-ol/acetic acid/water/pyridine (15:3:12:10, by vol.) (Waley & Watson, 1953). A mixture of Xylene Cyanol FF

(Milstein & Sanger, 1961) and ε-dinitrophenyl-lysine was used as external electrophoretic markers.

Peptides were revealed on strips cut from dried papers with 0.01% (w/v) Fluorescamine in acetone containing 1% (v/v) pyridine, and with ninhydrin/Cd reagent (Heilmann *et al.*, 1957), as yellow–green fluorescent bands and coloured bands respectively. Radioactive peptides were located on papers by radioautography, and in column fractions by liquid-scintillation counting. When appropriate, papers were stained with Ehrlich's reagent (Smith, 1953) to reveal tryptophan-containing peptides.

Amino acid analyses of peptides were performed after hydrolysis in 6M-HCl containing 0.1% phenol for 18h at 110°C, by using a Bio-Cal amino acid analyser. *S*-[^{14}C]Carboxymethylcysteine was estimated by liquid scintillation counting of samples of peptides. Peptide sequencing and *N*-terminal-residue determinations were performed by standard manual techniques (Hartley, 1970; Gray, 1972*a*,*b*). The identification of *S*-[^{14}C]carboxymethylcysteine residues in peptide sequences was made by liquid-scintillation counting of butyl acetate extracts during the manual Edman degradations. In some cases, also, *N*-dansyl-*S*-[^{14}C]carboxymethylcysteine was identified on the polyamide layers. Thermolysin digestion of tryptic peptides was performed for 15 min at 55°C in 50mM-NH_4HCO_3/1mM-$CaCl_2$, with a peptide/enzyme ratio of 50:1 (w/w). Aspartic acid/asparagine and glutamic acid/glutamine were distinguished by paper-electrophoretic mobility of small peptides at pH6.5 (Offord, 1966).

Results

Freshly prepared ATPase in the denaturing buffer solution containing taurodeoxycholate and guanidinium chloride generally had 19–21 thiol groups per molecule (115000mol.wt.), as determined with 5,5′-dithiobis-(2-nitrobenzoate). Titrations in sodium dodecyl sulphate solutions gave identical results. After reduction with dithiothreitol in sodium dodecyl sulphate solution and separation from excess dithiothreitol on a column (1cm×15cm) of Sephadex G-25 (fine grade) in 1% (w/v) sodium dodecyl sulphate, the maximum number of thiol groups titrated was 27, in agreement with the amino acid analysis after performic acid oxidation (Thorley-Lawson & Green, 1975, 1977). Occasionally significantly lower values were observed, and this was usually the case in early experiments with ATPase purified without added thiol (such as *N*-acetylcysteine) in the buffer solutions. Reaction with 5,5′-dithiobis-(2-nitrobenzoate) in 6M-guanidinium chloride without detergent gave lower estimates of the thiol content of the protein.

The carboxymethylated protein was eluted from the Sepharose 6B column, as shown in Fig. 1. The shoulder before the ATPase peak contained some aggregated protein, which is occasionally produced during storage of suspensions of the ATPase. In three preparations of reduced, carboxymethylated and delipidated ATPase, 22.6, 24.1 and 25.6 [^{14}C]carboxymethyl groups were incorporated per molecule, about 90% of the theoretical.

In an earlier preparation, with ATPase isolated in the absence of *N*-acetylcysteine, delipidated on a column of Sephadex G-50 in 1% (w/v) deoxycholate and carboxymethylated in 6M-guanidinium chloride/1mM-EDTA/50mM-Tris/HCl (pH8.0), only 15.0 carboxymethyl groups were incorporated per molecule. Peptides isolated from this material are denoted by Greek letters.

The uptake of NaOH during the tryptic digestions was only about half of the theoretical quantity, assuming cleavage at all lysine and arginine residues and an average pK_a for the liberated amino groups of 7.8 (Steinhardt & Beychok, 1964).

The tryptic peptides were eluted from the Sephadex G-50 column, as shown in Fig. 2. Most of the material

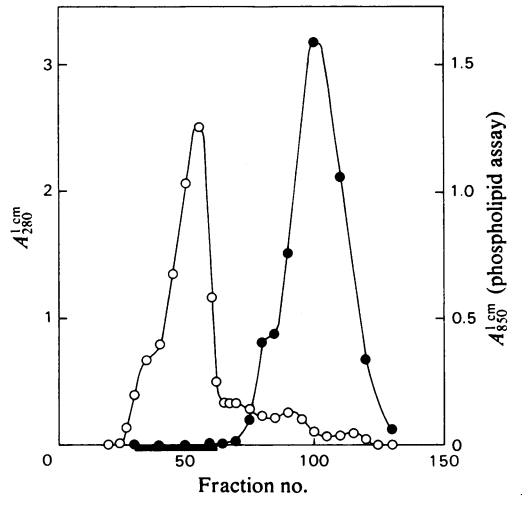

Fig. 1. *Delipidation of carboxymethylated ATPase*
Reduced carboxymethylated ATPase (150mg) in 60ml of 50mM-Tris/acetic acid buffer, pH7.0, containing EDTA (1mM), 2-mercaptoethanol (1mM) and sodium dodecyl sulphate (35g/litre) was applied to a column (2.4cm×142cm) of Sepharose 6B equilibrated in 50mM-Tris/acetic acid buffer, pH7.0, containing EDTA (1mM), 2-mercaptoethanol (1mM) and sodium dodecyl sulphate (10g/litre) at 20°C. The column was eluted with this buffer, at 12ml/h, and fractions (4ml) were collected. ○, A_{280}^{1cm} of fractions (protein); ●, A_{280}^{1cm} after the colour reaction for phospholipid (McClare, 1971) with 0.5ml samples. Fractions containing delipidated protein were combined as indicated by the horizontal bar.

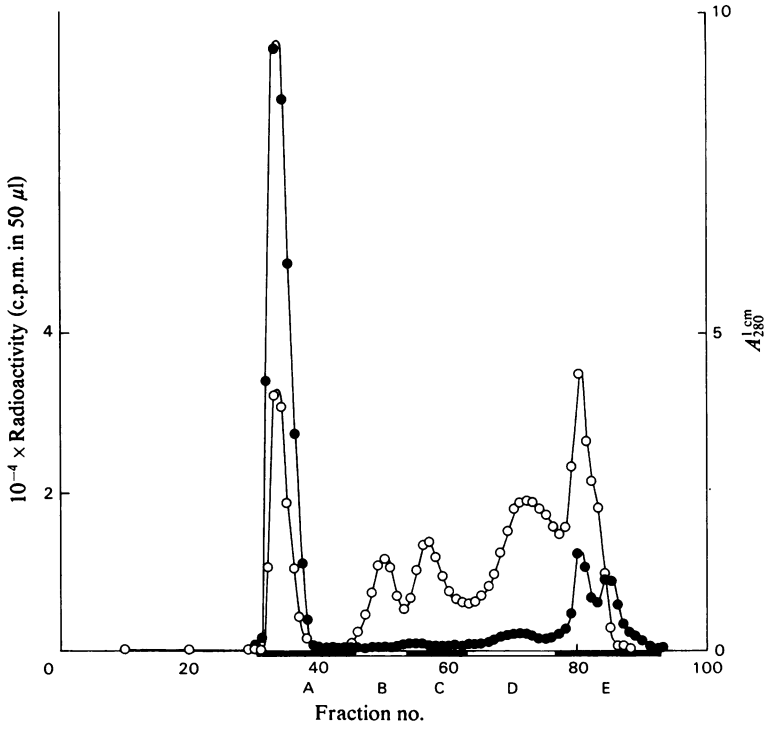

Fig. 2. *Fractionation of the tryptic digest of [^{14}C]carboxymethylated ATPase on Sephadex G-50*
The tryptic digest (140mg) was applied in 7ml to a column (1.8cm × 133cm) of Sephadex G-50 (superfine grade) in 50mM-NH_4HCO_3 containing 0.01% (v/v) thiodiglycol, and eluted with this solvent at 14ml/h at 20°C. Fractions (3.5ml) were collected and radioactivity (^{14}C) was determined in 50μl samples by liquid-scintillation counting (○). ●, A_{280}^{1cm} of fractions. Fractions were combined as indicated by the horizontal bars to give combined fractions A–E.

with absorbance at 280nm was eluted at the void volume (fraction A), and chromatography of this fraction on columns (1cm × 123cm) of Sephadex G-50 in other solvents (8M-urea/2M-sodium propionate, pH9, or 8M-urea/2M-propionic acid adjusted with NaOH to pH5) showed that the material was of high molecular weight or aggregated even under these conditions.

To isolate pure carboxymethylcysteine-containing peptides from this fraction, extensive digestion with trypsin and α-chymotrypsin was performed. The freeze-dried material was suspended in 0.1M-NH_4HCO_3 (3.4ml) and slowly swelled to form a viscous gel. Trypsin and α-chymotrypsin (each 2% by wt. of the protein) were added, and digestion proceeded at 20°C with occasional shaking. The gel gradually disintegrated. Further amounts (2%, w/w) of the proteinases were added and digestion was continued for 16h. Particles of gel were still present. The suspension was ultrasonicated in a Schuco Megason cleaning bath for 3min (60W), further amounts of the proteinases were added, and digestion was allowed to continue for another 24h.

The finely divided gel suspension, containing 8.6μmol of carboxymethyl groups, was centrifuged at 165000*g* for 10min. The clear supernatant contained 5.3μmol of carboxymethyl groups, and carboxymethylcysteine-containing peptides were isolated from the supernatant.

The details of further purification procedures and analyses for each carboxymethylcysteine-containing peptide are given in the Supplement (SUP 50085). The results of sequence determinations are given in the text below (Fig. 3). The nomenclature of peptides is based on the separation methods used, an outline of which is presented in Figs. 4 and 5.

Peptides A (Table 1 of Supplement)

Four radioactive peptides were identified.

Peptides A1a1 and A1a2 were closely related, their different chromatographic properties probably being due to partial autoxidation during isolation. The partial sequence of peptide A1a2 was:

(?)-Glx-CmCys-Thr-Glx-Asx-His-

Peptide A1a1 (60nmol) was digested with thermolysin and peptides were purified by paper electro-

phoresis at pH6.5, 3.5 and 2, giving radioactive peptides with the analyses:

A1a1.Tha1a (16nmol)
(CmCys,Asp,Thr,Glu_2,$Pro_{0.5}$,$Met_{0.5}$,His)
A1a1.Tha2a (14nmol)
(CmCys,Asp,Thr,Glu_2,Gly,His_2,Phe)
A1a1, Tha3a (8nmol)
(CmCys,$Asp_{0.5}$,Glu,Ile,Leu,Phe).

No *N*-termini were revealed by the dansyl method. More recent work, on chymotryptic peptides from the ATPase, has disclosed the sequence Met-Glx-Cys-Thr-Glx-Asx-His-Pro-His-Phe, which is clearly derived from the same part of the primary structure. The reason for the lack of detection of *N*-terminal methionine in peptides A1a1 and A1a2 is not known.

Peptides A3a1 had the sequence CmCys-Leu.

Peptide A3a2, although not pure, gave a clear sequence by the dansyl–Edman procedure: CmCys-Ile-Phe-Leu.

Peptides B

The carboxymethylcysteine-containing peptides in fraction B (Fig. 2) were resolved by ion-exchange chromatography on DEAE-cellulose. Three radioactive fractions (B4, B5, B6) were collected. The radioactive peptides were purified on columns of DEAE-cellulose at pH5. The properties of the purified peptides are given in Table 2 of the Supplement.

The sequence of peptide B4 was determined with the aid of thermolytic fragments, as shown in Fig. 4 of the Supplement SUP 50085. Peptide B4 (130nmol) was digested with thermolysin. The fragments were separated on paper and their properties are given in Table 3 in the Supplement. The order of the thermolytic peptides Tha3 and Tha6 was not determined, but later work has shown that the order in Fig. 4 of the Supplement, and in Fig. 3 here, is correct. It was assumed that peptides Thn3 and Thal were at the *C*-terminus of peptide B4, since they have *C*-terminal lysine.

Peptide B5 was very similar to B4. The reason for the different chromatographic properties of peptides B4 and B5 was not determined, but possibly there is a substitution of aspartic acid for asparagine at one position.

Peptide B6, isolated in low yield, had the partial sequence (?)-Glx-CmCys-Thr-Glx-. The analysis and partial sequence are similar to those of peptides A1a1 and A1a2. Probably these peptides are derived from the same part of the primary structure of the protein.

In an earlier preparation using a shorter time for tryptic digestion, two other peptides, β1 and β3,

Peptides	(*a*) Tryptic peptides
B4	Ser-Ile-Gln-Phe-Asp-Gly-Leu-Val-Glu-Leu-Ala-Thr(Ile-CmCys-Ala),(Leu-CmCys(Ser,Asx,Ser,-Asx))Leu-Asp-Phe-Asx-Glx-Thr-Lys
β1	Ser-Leu-Pro-Ser-Val-Glu-Thr-Leu-Gly-CmCys-Thr-Ser-Val-Ile-CmCys-Ser-Asp-Lys(Thr,Gly,-Thr)Leu-Thr-Thr-Asn-Gln(Met,Ser,Val,CmCys,Lys)
C5	Ser-Leu-Pro-Ser-Val-Glu-Thr-Leu-Gly-CmCys-Thr-Ser-Val-Ile-CmCys-Ser-Asp-Lys
C8	Ile-Ile-Asp-Lys-Asx-Gly-Asx(Ser,CmCys,Glx,Phe)Leu-Asx-Glx-Phe (tentative sequence)
γ_1	Ser-Thr-Glx-Glx-CmCys(Asx_2,Thr_2,Ser,$Glx_{2.5}$,Pro,Gly_2, $Ala_{1.5}$,Val_2,Leu_2,Tyr,Phe,Lys,Arg)
D2e	Ser-Thr-Glu-Glu-CmCys-Leu-Ala-Tyr
D5a2	Ser-Thr-Glu-Glu-CmCys-Leu-Ala-Tyr-Phe
D6n and E7n	CmCys-Ser-Pro-Ala-Lys
D6a1	Ala(Asx,Ala,CmCys,Asx,Ser,Val,Ile,Arg)
D8n2 and δn2	Ala-Asn-Ala-CmCys-Asn-Ser-Val-Ile-Arg
δn1, D8n1 and D8n3	CmCys-Leu-Ala-Leu-Ala-Thr-Arg
D8n4 and E9n	Ser-Val-CmCys-Lys
δa1	Gly-Ala-Pro-Glx-Gly(CmCys,Asx_2,Val_2,Ile,Tyr,Arg)
E5a	Ala-CmCys-CmCys-Phe
E9a1	Ala-CmCys-CmCys-Phe-Ala-Arg
E6a	CmCys-Asn-Tyr
E11n	CmCys-Asn-Tyr-Val-Arg
E10n	CmCys-Arg
E9a2	Glu-Ala-CmCys-Arg
εn	Glx-Ala-CmCys-Arg($Glx_{0.5}$,Arg)
	(*b*) Peptides isolated after chymotryptic and tryptic digestion
A1a2, also B6	(Met)Glx-CmCys-Thr-Glx-Asx-His- [N.B. The *N*-terminal residue was not identified in this peptide, but later work (G. Allen, unpublished results) showed the presence of a methionine residue in this position]
A3a1	CmCys-Leu
A3a2	CmCys-Ile-Phe-Leu

Fig. 3. *Sequences of carboxymethylcysteine-containing peptides isolated from the Ca^{2+}-transporting ATPase*

were isolated from this part of the eluate from the Sephadex G-50 column as described in the Supplement.

The partial sequence of peptide β3 was Met-Phe-, but further sequencing was ambiguous.

Peptide β1 was partially sequenced, by using thermolytic fragments. Peptide β1 (90nmol) was digested with thermolysin and peptides were purified on paper. The properties of the purified thermolytic fragments are given in Table 4 of the Supplement. Peptide β1 and its thermolytic fragments were partially sequenced, and the total sequence information is given in Fig. 6 of the Supplement and is summarized here in Fig. 3. The ordering of the fragments was in part determined from the related peptide C5. The aspartic acid residue in peptide β1 is phosphorylated during the catalytic cycle of the enzyme (Allen & Green, 1976).

Peptides C

Fraction C (Fig. 2) was resolved on a column of DEAE-cellulose into four radioactive subfractions, C3, C5, C7 and C8, as shown in Fig. 7 of the Supplement.

Pure tryptic peptides were not isolated from peptide C3 after several chromatographic methods were used. The still impure peptide was digested with thermolysin and two radioactive peptides, C3Tha1 and C3Thn, were purified by paper electrophoresis. The analyses are given in Table 5 of the Supplement, and the sequence of peptide C3Tha1 was determined as Leu-Gly-CmCys-Thr-Ser, although the analysis showed that this peptide was not pure.

Peptide fraction C5 was subjected to further purification steps. A pure peptide, C5, and two thermolytic fragments, C5Tha1 and C5Tha2, were isolated, as described in the Supplement. The sequence of peptide C5 is recorded in Fig. 3.

The thermolytic peptides C3Tha1 and C5Tha1 were identical, indicating that peptides C3 and C5 came from the same part of the protein sequence. However, peptide C3Thn differed from peptide C5Tha2 in electrophoretic mobility and the content of lysine, the former having two residues and the latter only one of this amino acid. The reason for this difference is unexplained, since further work has shown that there is only a single lysine residue in this part of the peptide chain (G. Allen, unpublished results).

Peptide fraction C7 was also difficult to purify, but two radioactive peptides, C7a1 and C7a2, were isolated, as described in the Supplement (Table 5). The partial sequence of peptide C7a1 was Thr-(?)-CmCys-Ala: this, and the amino acid analysis, showed that this peptide was identical with the *C*-terminal 17 residues of peptide B4. No sequence information for peptide C7a2 was obtained.

Subfraction C8 gave broad peaks during chromatography on columns of DEAE-cellulose at pH5 and QAE-Sephadex at pH8.9, and was not completely resolved into a pure peptide. Thermolytic fragments from the peptide were purified and characterized (Table 5 in the Supplement). Partial sequences of peptide C8 and its thermolytic fragments are shown in Fig. 8 of the Supplement. Some unsolved problems, from impurities in the peptide C8 and low yields of some thermolytic fragments, were the content of alanine and leucine and the isolation of only one carboxymethylcysteine residue in thermolytic fragments, although two residues were present in peptide C8, according to the determination of radioactivity. Information from other digests of the protein has shown that there is only a single carboxymethylcysteine residue in this peptide sequence (G. Allen, unpublished work).

A single carboxymethylcysteine-containing peptide, γ1, was isolated from this region of the eluate from the Sephadex G-50 column in the earlier preparation, by using less-extensive tryptic digestion. (Its properties are recorded in Table 5 of the Supplement. The partial sequence of peptide γ1 is given in Fig. 10 of the Supplement.)

Peptides D

Peptides from fraction D (Fig. 2) were separated on a column of AG50W 2X resin in a gradient of pyridine/acetic acid, as described in the Supplement (Fig. 9). Fractions were assayed by alkaline hydrolysis and detection with ninhydrin (Fraenkel-Conrat, 1957), and by liquid-scintillation counting of radioactivity of samples. The yield of radioactivity from the column was essentially quantitative. Fractions were combined as shown in Fig. 7 of the Supplement.

Carboxymethylcysteine-containing peptides were purified from fractions D2, D5, D6 and D8 by standard methods, as described in Fig. 4. Some peptides (δ) were eluted in the same position from the Sephadex G-50 column during the earlier preparation are also described in Table 6 of the Supplement; these peptides were purified on paper (Fig. 5) from the Sephadex G-50 column fraction. Several of these peptides were identical with some of those from the main preparation. The sequence information determined for these peptides is given in Fig. 10 of the Supplement, and is summarized in Fig. 3 here.

Peptides E

Peptides from fraction E (Fig. 2) were similarly separated on a column of AG50W resin (Fig. 11 of the Supplement), and carboxymethylcysteine-containing peptides were isolated from subfractions E5, E6, E7, E9, E10 and E11 on paper. The properties of the peptides are given in Table 7 (in the Supplement). Several peptides with identical properties were isolated from a less-extensively digested sample of the carboxymethylated protein, and in addition a neutral peptide, εn, was isolated, on paper only, from this

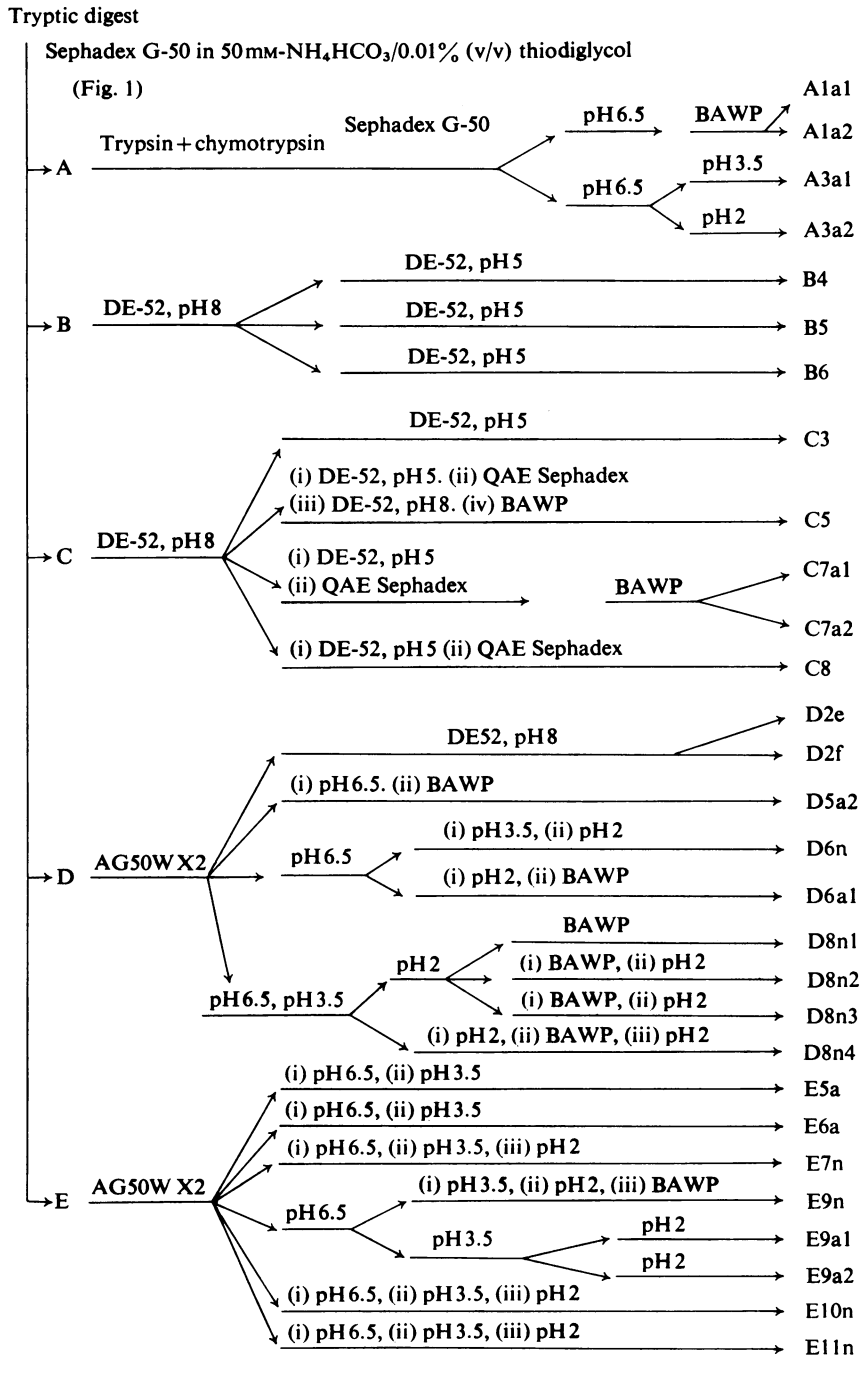

Fig. 4. *Summary of steps in the purification of* [^{14}C]*carboxymethylcysteine-containing peptides from the tryptic digest of carboxymethylated ATPase protein (main preparation)*

Abbreviations: DE-52, ion-exchange chromatography on DEAE-cellulose (Whatman DE-52); AG50W X2, ion-exchange chromatography on Bio-Rad cation-exchange resin AG50W X2; pH6.5, pH3.5, pH2, paper electrophoresis at the indicated pH; BAWP, paper chromatography in butanol/acetic acid/water/pyridine (15:3:12:10, by vol.).

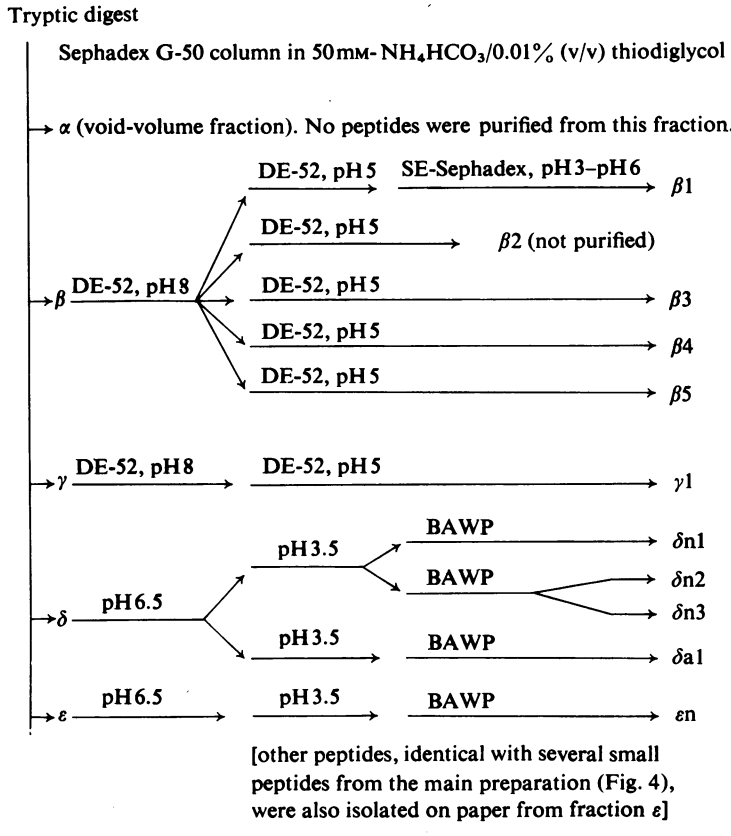

Fig. 5. *Summary of steps in the purification of* [^{14}C]*carboxymethylcysteine-containing peptides from the tryptic digest of carboxymethylated ATPase protein (earlier preparation with shorter digestion time)*
Abbreviations: as in Fig. 4.

Table 1. *Summary of peptide sequences immediately around the cysteine residues of the ATPase*
Longer stretches of sequence around these residues are given in Fig. 3. Cysteine residues are in **bold** type for easier identification.

Sequence	Peptides containing sequence
Met-Glx-**Cys**-Thr-Glx	A1a1, A1a2, B6
Cys-Leu	A3a1
Cys-Ile-Phe-Leu	A3a2
Ile-**Cys**-Ala	B4
Leu-**Cys**(Asx,Asx,Ser,Ser)	B4
Leu-Gly-**Cys**-Thr-Ser	β1, C5, C3Tha1
Val-Ile-**Cys**-Ser-Asp-Lys	β1, C5
Ser-Thr-Glx-Glx-**Cys**-Leu-Ala	D2e, D5a2, γ1
Val-Asx-Gly(Asx,Phe,**Cys**,Ser,Leu,Asx,Glx,Phe)	C8
Cys-Ser-Pro-Ala-Lys	D6n, E7n
Ala-Asn-Ala-**Cys**-Asn-Ser-Val-Ile-Arg	D6a1, D8n2, δn2
Cys-Leu-Ala-Leu-Ala-Thr-Arg	D8n1, D8n3, δn1
Ser-Val-**Cys**-Lys	D8n4, E9n
Ala-**Cys**-**Cys**-Phe-Ala-Arg	E9a1
Cys-Asn-Tyr-Val-Arg	E11n
Cys-Arg	E10n
Glu-Ala-**Cys**-Arg	E9a2, εn

region of the eluate from the Sephadex G-50 column. The sequences of these peptides were determined, as described in Fig. 12 of the Supplement, and the sequences are listed in Fig. 3 here.

The peptide sequences around the cysteine residues of the ATPase are summarized in Table 1. In addition to these sequences, including 18 cysteine residues, other peptides containing carboxymethylcysteine residues with possibly different sequences were isolated in low yields and were not completely sequenced. These include peptides C7a2, β3, δa1 and δn3. There are also the second carboxymethylcysteine residues in peptide A1a1 and possibly in peptide C8 that were determined analytically but not identified in sequences.

The peptides isolated from the less-extensively digested ATPase (β, γ, δ, ε) were generally products of expected tryptic cleavage at lysine and arginine residues, although in some cases these bonds were still intact. Most of this work was performed on peptides derived from more-extensively digested protein, and several of the peptides were products of chymotryptic-like cleavage at tyrosine or phenylalanine residues. Thus some of the peptides derived from the less-extensively digested protein include some of the cysteine residues identified in other peptides.

Discussion

Most of the peptide sequences listed in Fig. 3 have been unequivocally determined. There are, however, a few weak points in the sequence evidence, and these have been mentioned in connection with individual peptides in the Results section. They include poor analyses for peptide C8, and the lack of identification of *N*-terminal residues for peptides A1a1, A1a2 and B6. Some of the thermolytic fragments from the longer tryptic peptides were not completely purified, but the evidence from these fragments, which was used for the sequences listed in Fig. 3, was unambiguous.

The minimum number of cysteine residues in different sequences in the ATPase is shown in the present work to be 18, but there are a few more in peptides that have not been completely characterized. The peptides in fraction A, which contained about one-fifth of the radioactivity, yielded only three unique cysteine residues instead of the four or five expected, owing to the difficulty in isolation of peptides from the 40h tryptic and chymotryptic digest of these aggregated and generally hydrophobic peptides. After allowance has been made for the incompleteness of the results, the probable number of unique cysteine-containing sequences is about 24.

Analytical data (Thorley-Lawson & Green, 1975) show that there are 23 half-cystine and cysteine residues per 100000 daltons. The present work suggests that there are less than the 26 or 27 cysteine residues that would be required for mol.wt. 115000, and that the molecular weight of the ATPase protein lies within the range estimated by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis (McFarland & Inesi, 1971; MacLennan *et al.*, 1971; Louis & Shooter, 1972; Thorley-Lawson & Green, 1973) and is somewhat lower than that estimated by ultracentrifugation (Rizzolo *et al.*, 1976).

It is of some interest to consider the distribution of cysteine residues in the molecule. The fraction A peptides, excluded from Sephadex G-50 beads, were aggregated and rather hydrophobic in composition, and contained almost all the tryptophan residues of the protein but only one-fifth of the cysteine residues. Most of the membrane-associated part of the protein is likely to be included in this fraction, which may also contain 'core' peptides from the segments of the molecules outside the membrane. The results of Hardwicke (1976), showing the almost complete quenching of the fluorescence of the ATPase by the binding of one or two molecules of *N*-dinitrophenylphosphatidylethanolamine, also suggests a preferential location of tryptophan in the membrane. In contrast, we have found that the mixed disulphide of the native ATPase with 5-thio-2-nitrobenzoate, containing eight residues of 5-thio-2-nitrobenzoate, retains 70% of the original fluorescence of the protein (N. M. Green, unpublished work). Since the absorption spectrum of the mixed disulphide (330 nm) coincides with the emission maximum of the ATPase, the weak quenching implies that the substituted thiol groups are at least 3.0nm from most of the tryptophan residues. Both the fluorescence quenching and the primary-structural investigations thus suggest that most of the cysteine residues are separated from the tryptophan residues, and are situated outside the membrane. The reactivity of cysteine residues (Thorley-Lawson & Green, 1977) also indicates that they are mainly exposed on the surface of the protein.

These conclusions are supported by the results of extensive proteolysis of whole sarcoplasmic reticulum (Green *et al.*, 1977), which showed that 90% of the reactive cysteine residues was removed by pepsin or trypsin, while 70% of the tryptophan was left in the membrane.

The active-site aspartic acid residue, which is phosphorylated during the active transport of Ca^{2+}, has been shown to be the aspartic acid residue in peptides β1 and C5, by the isolation and analysis of tryptic and peptide phosphopeptides from the [^{32}P]phosphorylated ATPase (Allen & Green, 1976).

The data give no indication of the presence of repeated lengths of homologous sequences around the cysteine residues. Although this is not conclusive evidence against the presence of internal homologies within the primary sequence, such as have been observed in some other large polypeptides, these, if

present, are not a prominent characteristic of the protein.

The thiol groups of the ATPase can be selectively labelled in a number of ways (Hasselbach & Seraydarian, 1966; Coan & Inesi, 1977; Thorley-Lawson & Green, 1977) so that the results and fractionation methods described here should prove useful in any attempt to identify labelled peptides.

We thank Mrs. J. North for technical assistance and Miss S. Lathwell for operating the amino acid analyser.

References

Allen, G. & Green, N. M. (1976) *FEBS Lett.* **63**, 188–192
Ambler, R. P. (1963) *Biochem. J.* **89**, 349–378
Coan, C. R. & Inesi, G. (1977) *J. Biol. Chem.* **252**, 3044–3049
Ellman, G. L. (1959) *Arch. Biochem. Biophys.* **82**, 70–77
Fraenkel-Conrat, H. (1957) *Methods Enzymol.* **4**, 247–269
Gray, W. R. (1972*a*) *Methods Enzymol.* **25**, 121–138
Gray, W. R. (1972*b*) *Methods Enzymol.* **25**, 333–344
Green, N. M., Hebdon, G. M. & Thorley-Lawson, D. A. (1977) *Proc. FEBS Congr. 11th* **45**, 149–157
Hardwicke, P. M. D. (1976) *Eur. J. Biochem.* **62**, 431–438
Hardwicke, P. M. D. & Green, N. M. (1974) *Eur. J. Biochem.* **42**, 183–193
Harris, J. I. (1967) *Methods Enzymol.* **11**, 390–398
Hartley, B. S. (1970) *Biochem. J.* **119**, 805–822
Hasselbach, W. & Seraydarian, K. (1966) *Biochem. Z.* **345**, 159–172
Heilmann, J., Barollier, J. & Watzke, E. (1957) *Hoppe-Seyler's Z. Physiol. Chem.* **309**, 219–220
Louis, C. & Shooter, E. M. (1972) *Arch. Biochem. Biophys.* **153**, 641–655
MacLennan, D. H. (1970) *J. Biol. Chem.* **245**, 4508–4518
MacLennan, D. H., Seeman, P., Iles, G. H. & Yip, C. C. (1971) *J. Biol. Chem.* **246**, 2702–2710
McClare, C. W. F. (1971) *Anal. Biochem.* **39**, 527–530
McFarland, B. H. & Inesi, G. (1971) *Arch. Biochem. Biophys.* **145**, 456–464
Milstein, C. & Sanger, F. (1961) *Biochem. J.* **79**, 456–469
Offord, R. E. (1966) *Nature (London)* **211**, 591–593
Rizzolo, L., leMaire, M., Reynolds, J. A. & Tanford, C. (1976) *Biochemistry* **15**, 3433–3437
Smith, I. (1953) *Nature (London)*, **171**, 43–44
Steinhardt, J. & Beychok, S. (1964) in *The Proteins* (Neurath, H., ed.), 2nd edn., vol. 2, pp. 139–304, Academic Press, New York
Thorley-Lawson, D. A. & Green, N. M. (1973) *Eur. J. Biochem.* **40**, 403–413
Thorley-Lawson, D. A. & Green, N. M. (1975) *Eur. J. Biochem.* **59**, 193–200
Thorley-Lawson, D. A. & Green, N. M. (1977) *Biochem. J.* **167**, 739–748
Waley, S. G. & Watson, J. (1953) *Biochem. J.* **55**, 328–337
Weber, K. & Kuter, D. J. (1971) *J. Biol. Chem.* **246**, 4504–4509