The sequence of two peptides isolated from the Ca^{2+} -transporting ATPase of rabbit sarcoplasmic reticulum after cleavage at tryptophan

N. Michael GREEN and E. John TOMS

National Institute for Medical Research, Mill Hill, London NW7 1AA, U.K.

Cleavage of reduced, carboxymethylated, delipidated Ca^{2+} -transporting ATPase protein from rabbit sarcoplasmic reticulum with dimethyl sulphoxide/HBr yielded two long peptides (38 and 73 residues), distinct from the known major sequences of the ATPase. The longer peptide contained at least two cysteine residues, which were disulphide-linked in the native protein. It was therefore derived from the B-fragment of the ATPase in which the disulphides had previously been located. It probably formed a loop on the luminal side of the membrane, spanning two membrane-buried tryptophan residues. The *N*-terminal sequence of this peptide, (Trp)-Phe-Met-Tyr-Ala, forms the basis for an oligodeoxynucleotide probe, the use of which to identify cDNA corresponding to the ATPase is described elsewhere [MacLennan, Brandl, Korczak & Green (1985) *Nature* (London) 316, 696–700].

INTRODUCTION

Our previous work on the primary structure of the Ca^{2+} -transporting ATPase of sarcoplasmic reticulum (Allen *et al.*, 1980*a,b*) established the sequence of over 800 out of a total of about 1000 amino acid residues. Of these, 600 were assigned to four major sequences containing respectively 31, 116, 298 and 123 residues, starting from the blocked *N*-terminus of the molecule. After each of the sequences were gaps, at least three of which were about 100 residues in length. These segments were labelled by a hydrophobic photoactivated reagent (Green *et al.*, 1980) Almost all the constituent peptides from the major sequences were released from the membrane by trypsin and therefore formed extramembranous domains, which were shown to be on the cytoplasmic side of the lipid bilayer. Hitherto, there has been no direct evidence for assigning any ATPase peptides to the luminal surface of the membrane.

We made many attempts to fractionate the hydrophobic 'A' peptides (Allen, 1978; N. M. Green & E. J. Toms, unpublished work), but we obtained only 30 or 40 residues of new sequence, none of it very hydrophobic. Some large hydrophobic peptides were obtained in a pure state, but their *N*-terminal residues were blocked and on prolonged storage at -20°C they aggregated and did not disperse in SDS or formic acid, so that little sequence information was obtained.

In the present paper we describe a new strategy for identifying polar loops between membrane-buried sequences which may be applicable to other membrane proteins. Fluorescence-quenching experiments with lipid-soluble collisional quenchers such as dibromotetrachloroethane (J. G. Kenna & N. M. Green, unpublished work) had shown that about 90% of the tryptophan fluorescence was accessible to these molecules. Similar results were obtained by East & Lee (1982). This suggested

that most of the thirteen tryptophan residues were in the lipid bilayer. Our amino acid sequences showed only one tryptophan residue among the extramembranous peptides. Cleavage at tryptophyl peptide bonds should therefore release water-soluble extramembranous loops with hydrophobic extensions to the nearest tryptophan residue in the membrane. Some of these, corresponding to our major cytoplasmic sequences, would be very long, but shorter loops from the luminal surface were anticipated and it was hoped that these would be sufficiently polar for purification in aqueous systems.

MATERIALS AND METHODS

Most of the materials and methods used have been described in previous papers (Allen & Green, 1978; Allen, 1980*a,b*). Dimethyl sulphoxide and phenol were redistilled *in vacuo*. HBr (50%, w/v) was Aristar grade from BDH (Poole, Dorset, U.K.). DEAE-Sephacel and Sephadex LH-60 were from Pharmacia (Hounslow, Middx., U.K.).

Commercial preparations of dodecyltrimethylammonium bromide contained u.v.-absorbing impurities. It was therefore synthesized from 1-bromododecane (40 ml) and trimethylamine [60 ml of a 33% (w/v) solution in ethanol]. The mixture was sealed in Pyrex tubes (diameter 2.6 cm) and heated to 70°C for 20 h. The reaction mixture solidified on cooling and the product was transferred to a vacuum evaporator by using warm acetone. After evaporation to dryness, the product was crystallized from acetone (200 ml) containing ethanol (10 ml). After a similar recrystallization the product (36 g) was colourless. An aq. 20% (w/v) solution gave an A_{280} of 0.3.

Ca^{2+} -ATPase was purified by the method of MacLennan (1970) from the leg and back muscles of a New Zealand White rabbit. It was revesiculated by a passage over a

Abbreviations used: DTAB, dodecyltrimethylammonium bromide; SDS, sodium dodecyl sulphate; PTH, phenylthiohydantoin; Ca^{2+} -ATPase, Ca^{2+} -transporting ATPase; dansyl, 5-dimethylaminonaphthalene-1-sulphonyl.

column (65 ml) of low-cross-linked anion-exchange resin (Zerolit FFIPSR62) as previously described (Hardwicke & Green, 1974). The effluent was diluted with an equal volume of water and the vesicles were pelleted (100 000 *g*, 30 min). Vesicular ATPase (270 mg) resuspended in sucrose (12.5 ml, 0.25 M) was mixed with an equal volume of a buffer containing Na_2HPO_4 (0.1 M), sodium iodoacetate (0.1 M), EDTA (2 mM), phenylmethanesulphonyl fluoride (14 mg) and leupeptin (30 μg). The pH was adjusted to 8.4 and the ATPase was dissolved by the addition of 20% (w/v) dodecyltrimethylammonium bromide (11.5 ml) with stirring. After 15 min the reaction mixture was dialysed overnight against 50 mM-Tris/HCl, pH 8.4, containing EDTA (1 mM) and DTAB (0.2%) (2 $\times$ 500 ml). The cloudy solution (50 ml) was treated with dithiothreitol (3 mM). The measured excess of thiol groups was 2.5 mM. The reaction mixture was incubated at 40 °C for 2.5 h and then dialysed for 3 h against the same buffer.

The reduced disulphide bonds were carboxymethylated by the addition of 2.5 ml of 0.1 M-iodo[2- ^{14}C]acetic acid (1160 c.p.m./nmol). After 45 min the reaction mixture was dialysed against water (5 litres) for 48 h at 4 °C. Most of the protein precipitated as a coagulum. A small amount of extra solid was obtained by centrifugation of the supernatant, which had been removed by decantation. It was important to minimize the handling of the precipitate, since it had a tendency to form a compact rubber-like lump which was difficult to redissolve. The solid was slowly redissolved by gentle stirring in 10% SDS (10 ml). The protein was delipidated by passage of the solution in two batches over a column (2.5 cm $\times$ 130 cm) of Sepharose 6B in 50 mM-Tris/HCl, pH 8.2, containing SDS (0.1%) and EDTA (1 mM). The broad protein peak was pooled and dialysed for 24 h against 10 mM- NH_4HCO_3 (2 $\times$ 2 litres). The product was freeze-dried. It contained 150 mg of protein and 400 mg of SDS.

Cleavage of tryptophyl peptide bonds (Savage & Fontana, 1977)

Freeze-dried, carboxymethylated, delipidated ATPase (250 mg, containing 65 mg of protein) was dispersed in a mixture of acetic acid (6 ml), concentrated (12 M) HCl (3 ml), dimethyl sulphoxide (0.25 ml) and phenol (200 mg). After brief sonication (2 min) and stirring (30 min) the solution was clear. HBr [1 ml, 50% (w/v)] was added and after 30 min most of the acid and water were removed by concentrating in a rotary evaporator at 35 °C. The viscous residue was dispersed in acetic acid (4.5 ml), and water (1.5 ml) was added. The protein was separated from the remaining reagents on a column (2.5 cm $\times$ 28 cm) of Sephadex G-10 in 70% (v/v) acetic acid. The void-volume fractions containing the peptides were well separated from the phenol and were pooled, concentrated to dryness on a rotary evaporator and dried *in vacuo*.

Fractionation of peptides

The peptides were dissolved in 88% formic acid (1.5 ml). Addition of ethanol (3.6 ml) gave an opalescent solution. The solution was clarified by centrifugation and was applied to a column of Sephadex LH-60 (2 cm $\times$ 105 cm), equilibrated in 88% formic acid/ethanol (3:7, v/v) (Gerber *et al.*, 1979).

The peptide peaks were evaporated to dryness and redissolved in 50 mM- NH_3 for fractionation on a column

Fig. 1. Separation of peptides produced by cleavage at the tryptophan residues of the Ca^{2+} -ATPase

The column of Sephadex LH-60 (2 cm $\times$ 105 cm) was run in formic acid (88%)/ethanol, 3:7. —, $A_{260}^{1\text{cm}}$; ----, ^{14}C radioactivity. Fractions (2.0 ml) were pooled as indicated by the horizontal bars.

of DEAE-Sephacel (0.9 cm $\times$ 16 cm), by a gradient of NH_4HCO_3 (20 mM–1 M in 400 ml of 0.01% thiodiglycol). The fractions obtained by ion exchange were then further purified by t.l.c. methods or by gel filtration on Sephadex G-50 or G-75 (Allen, 1980a, 1981).

Amino acid analyses were performed with a Beckman model 121M amino acid analyser. Total cysteine was determined from the carboxymethylcysteine peak, and disulphide-labelled cysteine was determined from the radioactivity of the peptide and the specific radioactivity of the iodoacetate used for labelling the reduced disulphide bonds (1160 c.p.m./nmol). Free thiol was determined by the method of Ellman (1959). The *N*-terminal residues of purified peptides were determined by dansylation (Allen, 1981).

Peptides were sequenced by Dr. M. Waterfield by using a gas-phase sequencer as described elsewhere (Downward *et al.*, 1984).

RESULTS

Preparation of the protein for cleavage

Previous experiments had established that most of the reactive thiol groups of the ATPase were in soluble peptides and were released from the membrane by trypsin (Green *et al.*, 1978, 1980), whereas cysteine labelled after reduction of the two or three disulphide bonds of the ATPase was found mainly among the aggregated 'A' peptides derived from membrane-associated regions (Green *et al.*, 1980; G. Allen, unpublished work). In the present work we blocked the free thiol groups with unlabelled iodoacetate and radioactively labelled only those generated by reduction. We did not succinylate (3-carboxypropionylate) the lysine residues, since we anticipated making the primary separations in acidic media, in which there would be an advantage in retaining the positive charge on the lysine residue.

The reduction and carboxymethylation were performed in a cationic detergent, since the negatively charged micelles of SDS repel negatively charged reagents such as iodoacetate and dithiothreitol, and they raise the *pK* of cysteine. Both effects lead to slow and less efficient derivative formation. The cationic micelles, in contrast, favour the desired reactions (Heitmann, 1968). Removal

Table 1. Amino acid composition of peptides obtained by cleavage at tryptophan

The results (uncorrected for hydrolytic loss) are expressed as mol/mol of peptide. Carboxymethyl cysteine (CM-Cys) was measured both from the peak on the amino acid analyser and from the radioactivity of the peptide. Low tyrosine peaks were observed because of bromination by the cleavage reagent giving two extra peaks in the region beyond phenylalanine. The values in parentheses are the expected values deduced from the sequence of the cDNA (MacLennan *et al.*, 1985). Abbreviation used: n.d., not determined.

Amino acid	Peptide... Yield (nmol)... N-Terminus...	Composition (mol/mol)		
		W.B5a 18 Phe	W.B6a 10 Phe	W.C3a 15 Val
[¹⁴ C]CM-Cys		2.8	—	0.2 —
CM-Cys		2.4 (3)	2.4	0 (0)
Asp		6.2 (6)	6.8	5.6 (6)
Thr		4.8 (5)	5.1	2.0 (2)
Ser		3.5 (4)	2.6	2.0 (2)
Glu		10.7 (11)	12.0	1.4 (1)
Pro		6.9 (6)	7.5	7.5 (7)
Gly		3.3 (3)	4.4	3.1 (2)
Ala		4.1 (4)	4.7	2.2 (2)
Val		3.3 (3)	3.7	1.9 (2)
Met		1.9 (7)	5.2	0.2 (1)
Ile		2.0 (2)	2.7	1.8 (2)
Leu		6.7 (7)	7.9	4.8 (5)
Tyr		0.2 (2)	0.1	0 (0)
Phe		3.6 (4)	4.6	1.0 (1)
His		3.8 (4)	5.1	0.1 (0)
Lys		0.1 (0)	0	1.0 (1)
Arg		1.0 (1)	1.0	1.9 (2)
Trp		n.d. (1)		n.d. (1)
Total		73		38

of excess cationic detergent by dialysis precipitated the detergent salt of the anionic protein. The precipitate could be redissolved in SDS to give a concentrated solution (10 mg/ml) that was convenient for delipidation of the protein on the column of Sepharose 6B. We found this procedure more convenient than the use of guanidinium chloride and taurodeoxycholate employed previously (Allen & Green, 1978). Model experiments established that the reaction between iodoacetate and acetylcysteine (monitored at 240 nm) was accelerated 15-fold by dodecyltrimethylammonium bromide (10 mg/ml).

After freeze-drying, 150 mg of protein were obtained mixed with three times its weight of SDS. The detergent was not removed before proceeding to the cleavage step. It contained 7 mol of radioactive carboxymethyl groups/mol of protein (M_r 115000). This was higher than the value of 5–6 mol that had been obtained in preparations made previously by this method, and perhaps indicated incomplete blockage of free thiol groups before the reduction step.

The tryptophan cleavage was performed as described by Savage & Fontana (1977), the tryptophan being converted into oxindole with dimethyl sulphoxide/HCl before adding HBr for the cleavage step. The peptides were separated in the ethanol/formic acid mixture described by Gerber *et al.* (1979), since preliminary experiments had shown that in aqueous systems almost all the peptides emerged in the void volume. In this system (Fig. 1) two small peaks (B and C) followed the large void-volume peak. The peptides in fractions B and C

proved to be soluble in NH_4HCO_3 and it was possible to purify them further by standard procedures.

B peptides

Gradient elution of the B-peptides from DEAE-Sephacel gave a major radioactive peak emerging at 0.30 M- NH_4HCO_3 . Four or five minor peaks that preceded it were ignored. The broad peak was subdivided into three fractions (B4, B5 and B6), which were purified on a column of Sephadex G-75. Each gave a single radioactive peptide. Peptides W.B4a, W.B5a and W.B6a had the same amino acid composition (Table 1), except for some apparent variation in methionine content. The total yield was 45 nmol. Most further work was done with W.B5a. The cysteine content calculated from the radioactivity was about the same as that calculated from the carboxymethylcysteine peak from the amino acid analysis. This implies that the three cysteine residues of this peptide participate in disulphide bonds and thus account for at least half of the disulphides present in the native protein (Thorley-Lawson & Green, 1977). The amino acid analysis showed a very low tyrosine content, but two extra peaks appeared beyond phenylalanine which were probably bromination products produced by action of the cleavage reagent on tyrosine.

The yields of amino acid PTH derivatives from about 1 nmol of peptide were more than 500 pmol for the first 14 residues, but fell to 200 pmol for the next three residues, after which no distinct sequence was obtained. The derived partial sequence is shown in Fig. 2. After residue 4 the sequence was identical with that of a

Fig. 2. Sequence information for peptide W.B5a

The *N*-terminal sequence is shown in the top line. The sequence is extended by two previously sequenced peptides which are shown below (see Fig. 3 for key to their origin). The two queried residues were homologous with tyrosine in peptide CB.E2n1, but the PTH derivatives were identified as methionine by the h.p.l.c. system used. It is almost certain that they are derived from tyrosine residues which have been brominated by the cleavage reagent. The assignment of T1.C10 to the position shown was based on overlap of a single methionine residue and on its unusual composition (see the text).

Fig. 3. Sequence information for peptide W.C3a

The *N*-terminal sequence is shown on the upper lines. Previously sequenced peptides from other digests allowed the sequence to be extended to residue 33. The key to the origin of these peptides is as follows: CB, CNBr (Allen *et al.*, 1980b); Ch, chymotrypsin (Allen, 1980b); P, pepsin (Allen *et al.*, 1980a); T1, trypsin (Allen, 1980a). Terminal tryptophan was deduced from the specificity of the cleavage reaction.

previously sequenced CNBr peptide (CB.E2n1), except for the *N*-terminal residue of the latter (position 3) and position 12, both of which appeared to be methionine rather than the expected tyrosine. Since tyrosine is brominated by the cleavage reagent, as confirmed by our amino acid analysis, it is likely that the apparent methionine peaks at positions 3 and 12 arose from a brominated tyrosine residue. The composition of the remaining 50 residues of this peptide included eight glutamic acid, two histidine, six proline and three cysteine residues. These unusual compositional features could be accounted for by inclusion of the 25-residue peptide T1.C10 (Allen, 1980a) as part of the sequence (Fig. 3). A previous experiment with a similarly labelled ATPase preparation (N. M. Green & E. J. Toms, unpublished work) had shown that both cysteine residues in T1.C10 participated in disulphide bonds. This made the assignment of T1.C10 to W.B5a particularly likely, since there are only two or three disulphides in the whole

molecule. This assignment has now been confirmed (see below).

C peptides

Gradient elution from DEAE Sephacel gave one major peak (W.C3) which was eluted at 0.095 M-NH₄HCO₃ and a number of minor fractions. W.C3 was further purified by t.l.c. on silica in butanol/acetic acid/water/pyridine (15:3:12:10, by vol.). Three ninhydrin-positive bands were obtained. The slowest (WC.3a) was finally purified by gel filtration on Sephadex G-50. The only significant peak (WC.3a) contained 15 nmol of a pure peptide of 38 residues (Table 1). This contained less than 0.2 mol of cysteine by radioactivity and no carboxymethylcysteine peak was observed. There were no peaks of modified tyrosine.

The yield of PTH derivative from about 2 nmol of peptide was 500–800 pmol for the first 14 residues, falling to 300 pmol for cycle 20, and giving only partial sequence

information (< 50 pmol) over the next seven cycles. The results (Fig. 3) overlapped three previously sequenced peptides and thus provided a continuous sequence of 33 out of the total of 38 residues in W.C3a. The peptide did not contain extensive runs of hydrophobic residues. It was noteworthy for its high contents of proline (six residues) and of polar residues.

No other pure peptides were obtained by further fractionation of the B, C and D pools.

DISCUSSION

The poor recovery of pure peptides from this digest was not unexpected in view of the large number of tryptophan residues (13) in the ATPase, the relatively low efficiency (50–70%) of the method of cleavage and the extensive hydrophobic sequences. The two peptides that were purified and sequenced were of particular interest. They were long peptides that were water-soluble in spite of their origin from cleavage of tryptophan residues located in the membrane. The *N*-terminal segments were hydrophobic and, as expected, there was no overlap with the previously established long extramembranous sequences of the ATPases. They were noteworthy for their high proline contents. Their properties were consistent with the expectation from our original strategy that cleavage at tryptophan would release extramembranous loops between membrane-buried tryptophans. The loop represented by W.B5 is almost certainly located on the luminal side of the membrane, since the radioactivity of its cysteine residues (Table 1) shows that at least two, and possibly all three, are involved in disulphide bonds. Disulphide bonds are rare in cytoplasmic proteins (e.g. Thornton, 1981) and probably in the cytoplasmic domains of membrane proteins.

At least 15 cysteine residues of the ATPase are located in the large extramembranous domains on the cytoplasmic side of the membrane (Allen *et al.*, 1980b; Green *et al.*, 1980). This corresponds in number to the cysteine residues that react with *N*-ethylmaleimide in the absence of SDS, which are all on the cytoplasmic face of the membrane (Reithmeier & MacLennan, 1981). We may tentatively generalize these observations and suggest that those cysteine residues in the cytoplasmic domains are likely to be free, whereas those on the luminal side are likely to be linked through disulphide bonds. The state of any cysteine residues in the transmembrane region remains to be established.

Since we had shown previously that the disulphide-linked cysteine residues of the ATPase were located exclusively in the B-fragment (Thorley-Lawson & Green, 1977), we conclude that peptide W.B5a is derived from the B-fragment.

The most important feature of peptide W.B5a has proved to be its *N*-terminal sequence. The tryptophyl

cleavage site and the methionine both correspond to unique codons, and this provided the basis for a relatively specific oligodeoxynucleotide probe for the cDNA coding for the Ca²⁺-ATPase. The successful use of this probe to obtain cDNA transcripts is described elsewhere (MacLennan *et al.*, 1985).

The results of this work confirmed that both W.B5a and W.C3aa formed loops between hydrophobic helices and showed that they were separated by a single transmembrane sequence and were therefore on opposite sides of the bilayer, near to the *C*-terminus of the protein. The location of peptide T1.C10 in W.B5a, as shown in Fig. 2, was also confirmed. It is noteworthy that this peptide arose from cleavage of a Phe–Met bond by trypsin, the third example of such a cleavage during proteolysis of the ATPase (Allen *et al.*, 1980b).

We thank Miss S. Lathwell for the amino acid analyses and Dr. M. D. Waterfield for the determination of *N*-terminal sequences of the peptides. We also thank Mrs. Lydia Pearson for secretarial assistance.

REFERENCES

- Allen, G. (1978) Proc. FEBS Meet. 11th **45**, 159–168
- Allen, G. (1980a) Biochem. J. **187**, 545–563
- Allen, G. (1980b) Biochem. J. **187**, 565–575
- Allen, G. (1981) Sequencing of Proteins and Peptides, pp. 73–132, Elsevier/North-Holland, New York and Oxford
- Allen, G. & Green, N. M. (1978) Biochem. J. **173**, 393–402
- Allen, G., Bottomley, R. C. & Trinnaman, B. J. (1980a) Biochem. J. **187**, 577–589
- Allen, G., Trinnaman, B. J. & Green, N. M. (1980b) Biochem. J. **187**, 591–616
- Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ulrich, A., Schlessinger, J. & Waterfield, M. D. (1984) Nature (London) **307**, 521–527
- East, M. & Lee, A. G. (1982) Biochemistry **21**, 4144–4151
- Ellman, G. L. (1959) Arch. Biochem. Biophys. **82**, 70–77
- Gerber, G. E., Anderegg, R. J., Herlihy, W. C., Gray, C. P., Biemann, K. & Khorana, H. G. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 227–231
- Green, N. M., Hebdon, G. M. & Thorley-Lawson, D. A. (1978) Proc. FEBS Meet. 11th **45**, 149–157
- Green, N. M., Allen, G. & Hebdon, G. M. (1980) Ann. N.Y. Acad. Sci. U.S.A. **358**, 149–159
- Hardwicke, P. M. D. & Green, N. M. (1974) Eur. J. Biochem. **42**, 183–193
- Heitmann, P. (1968) Eur. J. Biochem. **5**, 305–315
- MacLennan, D. H. (1970) J. Biol. Chem. **245**, 4508–4518
- MacLennan, D. H., Brandl, C. J., Korczak, B. & Green, N. M. (1985) Nature (London) **316**, 696–700
- Reithmeier, R. A. F. & MacLennan, D. H. (1981) J. Biol. Chem. **256**, 5957–5960
- Savage, W. E. & Fontana, A. (1977) Methods Enzymol. **47**, 459–469
- Thorley-Lawson, D. A. & Green, N. M. (1977) Biochem. J. **167**, 739–748
- Thornton, J. M. (1981) J. Mol. Biol. **151**, 261–287