Locations of the six disulphide bonds in a variant surface glycoprotein (VSG 117) from *Trypanosoma brucei*

Geoffrey ALLEN and Lynn P. GURNETT

Department of Molecular Biology, The Wellcome Research Laboratories, Langley Court,
Beckenham BR3 3BS, Kent, U.K.

(Received 16 August 1982/Accepted 15 October 1982)

The locations of the six disulphide bonds and the single free cysteine residue in a variant surface glycoprotein, VSG 117, from the African trypanosome *Trypanosoma brucei* have been determined to be Cys-14—Cys-140, Cys-121—Cys-182, Cys-389—Cys-404, Cys-398—Cys-417, Cys-447—Cys-461 and Cys-455—Cys-468. Cys-244 bears the single thiol group, which is unreactive towards 2-nitro-5-thiocyanobenzoate in the native molecule and is probably buried. Biosynthetically incorporated [³⁵S]cysteine aided the location of the disulphide bonds. Two proteinase-resistant glycosylated domains, each containing two disulphide bonds, were identified in the C-terminal region of the glycoprotein. Details of purification of [³⁵S]cysteine-containing peptides, and Tables of amino acid analyses, are presented in Supplementary Publication SUP 50119 (32 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.* (1981) 193, 5.

The bloodstream forms of African salivarian trypanosomes bear a surface coat composed of a variant surface glycoprotein (VSG). The parasites evade the immune response of the host at least in part by sequentially expressing antigenically unrelated VSG species, several of which have been purified and characterized (Cross, 1975, 1977). The molecular biology of this expression in *Trypanosoma brucei* is currently a subject of intense investigation, and has quite recently been reviewed (Cross *et al.*, 1980; Borst & Cross, 1982).

The complete amino acid sequence of one of these glycoproteins, VSG 117, has recently been determined (Allen *et al.*, 1982). We report here the locations of the six disulphide bonds and the single free cysteine residue in VSG 117. The incorporation of a radiolabel, [³⁵S]cysteine, was found necessary to aid in the identification of these disulphide bonds.

Materials and methods

The isolation of VSG 117 and most of the materials and methods used were as previously described (Allen *et al.*, 1982). Pepsin was from Millipore Corp., Freehold, NJ, U.S.A. and α -

Abbreviations used: VSG, variant surface glycoprotein; Hepes, 4-(2-hydroxyethyl)-1-piperazine-ethanesulphonic acid; SDS, sodium dodecyl sulphate.

chymotrypsin from Miles Laboratories, Stoke Poges, Slough, Berks., U.K. Three digests of the glycoprotein were studied: the CNBr digest previously described (Allen *et al.*, 1982), a tryptic digest, and a peptide digest of [³⁵S]cysteine-labelled VSG 117.

(a) CNBr digest

Peptide (CB1 + CB4 + CB5) (Allen *et al.*, 1982) (0.35 mg) was dissolved in 0.2 ml of 8 M-urea/20 mM-ammonium acetate, adjusted to pH 5.0 with acetic acid, and staphylococcal (strain V8) proteinase (0.1 mg) in 0.6 ml of 20 mM-ammonium acetate, pH 5.0, was added. The mixture was incubated, with occasional swirling, at 20°C for 2 h. The digest was then applied to a column of Sephadex G-50 as described in the Results section.

Peptide CB8 (Allen *et al.*, 1982) (4 mg) was similarly digested with 0.1 mg of staphylococcal proteinase, except that digestion was for 24 h. Peptides were separated on a column of Sephadex G-50 (see the Results section).

(b) Tryptic digestion of VSG 117

VSG 117 (40 mg; 0.75 μ mol) was dissolved in 4.0 ml of 8 M-urea/0.1 M-Tris/1 mM-EDTA/1 mM-iodo[¹⁴C]acetate (2000 c.p.m./nmol) adjusted to pH 8.3 with HCl. The solution was diluted into 28 ml of a solution of 50 mM-Tris/1 mM-EDTA, adjusted to

pH 8.3 with HCl and containing 2.0 mg of tosyl-phenylalanylchloromethane ('TPCK')-treated trypsin. The flask was swirled at 37°C for 1 h. The solution was acidified with 0.8 ml of 95% (v/v) formic acid and freeze-dried. Peptides were dissolved in approx. 5 ml of 5% (v/v) formic acid and separated on a column of Sephadex G-50 as described in the Results section. Peptides were purified from pooled fractions by ion-exchange chromatography, further gel filtration and thin-layer peptide mapping.

(c) *Preparation of [³⁵S]cysteine-labelled VSG 117*

Bloodstream trypanosomes of Molteno Institute antigen type M1 Tat 1.4, subclone 117, were prepared from the blood of infected rats (Cross, 1975). The trypanosomes were washed and suspended at 3×10^7 /ml in 75 ml of minimal essential medium (Eagle, 1959) minus cysteine, supplemented with non-essential amino acids, glucose (1 g/100 ml), Hepes buffer, pH 7.5, glutamine (30 mg/100 ml) and bovine serum albumin (0.1 g/100 ml). [³⁵S]Cysteine (Amersham International; 2.5 mCi; 900 Ci/mmol) was added and the cells were incubated in a shaking water bath at 37°C for 1 h. The cells were washed twice with phosphate-buffered saline (0.137 M-NaCl/2.7 mM-KCl/0.81 mM-Na₂HPO₄/1.47 mM-KH₂PO₄, pH 6.6) and the VSG was released from the trypanosomes by osmotic shock (G. A. M. Cross, personal communication) in 4 ml of water. The suspension was centrifuged at 35000g for 20 min at 4°C. The supernatant was mixed with 5 mg of pure VSG 117 as a carrier, and the solution was dialysed against 10 mM-sodium phosphate buffer, pH 8.0, at 4°C for 16 h. The VSG 117 was purified by chromatography on a column (10 mm × 40 mm) of DEAE-cellulose DE52 as described previously (Cross, 1975). The purity was checked by polyacrylamide-gel electrophoresis. The total radioactivity in the isolated [³⁵S]VSG 117 was 2.9×10^6 c.p.m. One-quarter of the glycoprotein was used in preliminary experiments to investigate possible methods of peptide cleavage. The remainder was mixed with 30 mg of unlabelled VSG 117, giving an initial specific radioactivity of 250 c.p.m./nmol of cysteine.

(d) *Peptic digestion of [³⁵S]VSG 117*

[³⁵S]VSG 117 (33 mg) was dissolved in 2.0 ml of 4.5% (v/v) formic acid, and digested at 37°C for 24 h with 1.65 mg of pepsin. Peptides were separated on a column of Sephadex G-50 as described in the Results section. Pooled fractions were subjected to thin-layer peptide mapping or ion-exchange chromatography for purification of radiolabelled peptides.

(e) *Cleavage of VSG 117 with 2-nitro-5-thiocyanobenzoate (Stark, 1977)*

VSG 117 (0.5 mg) was dissolved in 6 M-guanidinium chloride/0.2 M-Tris/acetate buffer, pH 8.0 (0.5 ml). 2-Nitro-5-thiocyanobenzoic acid (0.5 μmol in 50 μl of water) was added, and the solution was incubated at 37°C for 30 min. Acetic acid (0.5 ml) was added, and the solution was dialysed against 50% (v/v) acetic acid (3 × 50 ml). A second sample of VSG 117 (0.5 mg) was treated in the same way, but in the absence of guanidinium chloride.

The two samples were dried *in vacuo*, then re-dissolved in 6 M-guanidinium chloride/0.1 M-sodium borate buffer, pH 9.0 (0.1 ml each), and incubated for 18 h at 37°C. Samples (1 μl) of each were taken for polyacrylamide-gel electrophoresis in the presence of SDS using the Tris/glycine system described by Laemmli (1970), with a 10% (w/v)-acrylamide separating slab gel, 1 mm thick.

(f) *Diagonal peptide maps*

Diagonal peptide mapping was performed as described by Brown & Hartley (1966) with modifications for thin layers. The peptide mixture was subjected to electrophoresis at pH 4.4 (with pyridine/acetic acid/acetone/water, 2:4:15:79, by vol., as solvent) for 90 min at 500 V on a 20 cm × 20 cm cellulose MN300 thin layer (Macherey, Nagel and Co., Düren, Germany). The thin layer was dried in air, sprayed lightly with performic acid solution (Glazer *et al.*, 1975) and stored at room temperature in a sealed glass tray for 2–16 h. The thin layer was dried thoroughly, and electrophoresis at pH 4.4 repeated in the second dimension. [³⁵S]Cysteine acid-containing peptides were located by autoradiography.

Results

Staphylococcal-proteinase digestion of the CNBr peptide CB1 + CB4 + CB5

This yielded after chromatography (Fig. 1) the following peptides, which were identified by N-terminal and amino-acid analysis: (Tyr-7-Glu-43 + Thr-109-Glu-158 + Cys-182-Met-212), (Tyr-7-Glu-43 + Gly-126-Glu-158), (Thr-109-Glu-125 + Cys-182-Met-212), Val-159-Met-181, Ala-1-Glu-3 and Ala-4-Glu-6. These peptides covered the entire sequence of the CNBr peptide and identified the cross-links Cys-14-Cys-140 and Cys-121-Cys-182.

Staphylococcal-proteinase digest of peptide CB8

This was partially separated on Sephadex G-50 as shown in Fig. 2. N-Terminal and amino-acid analyses showed that peaks A and B consisted respectively of glycopeptides Leu-377-Asx-470 and Leu-377-Glu-441. N-Terminal analyses and peptide maps of later fractions identified the remaining

Fig. 1. Separation of peptides in the staphylococcal-proteinase digest of the CNBr peptide CB1 + CB4 + CB5. The digest was chromatographed on a column (11 mm $\times$ 1970 mm) of Sephadex G-50 (superfine grade) in 4.5% (v/v) formic acid. The eluate was monitored for A_{280} , and fractions (1.3 ml) were collected at 5.2 ml/h. Samples of fractions were analysed for content of *N*-terminal amino acids and amino acids. The positions of elution of characterized peptides are shown above the Figure. *N* $^{\epsilon}$ -Dinitrophenyl-lysine (*N* $^{\epsilon}$ -Dnp-Lys) was added as an internal marker after application of the sample.

Fig. 2. Separation of peptides from the staphylococcal-proteinase digest of peptide CB8. The digest was chromatographed on a column (11 mm $\times$ 1970 mm) of Sephadex G-50 (superfine grade) in 4.5% (v/v) formic acid. The eluate was monitored at 280 nm and fractions (1.5 ml) were collected at a flow rate of 4.5 ml/h. Samples (50 μ l) of fractions were taken for *N*-terminal and amino-acid analyses. Fractions were pooled for further study as indicated by the bars (A-F) beneath the Figure. Urea eluted at about fraction 110 and *N* $^{\epsilon}$ -dinitrophenyl-lysine (*N* $^{\epsilon}$ -Dnp-Lys), added as an internal marker after the sample, at about fraction 121.

expected peptides from the sequence Ala-277-Glu-376, and a peptide with properties consistent with (Ala-442-Asx-470); these peptides were not completely purified, however.

Tryptic digest of VSG 117

This yielded a number of peptides containing half-cystine residues, which were identified as cysteic acid after oxidation with performic acid and acid

hydrolysis. The initial separation of tryptic peptides on Sephadex G-50 is shown in Fig. 3. A single radioactive peptide was detected, and was eluted from the Sephadex column close to the total column volume (pool K). It was purified by thin-layer peptide mapping and was identified as Phe-240-Lys-249) from its position on the peptide map, and *N*-terminal and amino acid analyses. This result shows that Cys-244 was a free cysteine residue, reactive with iodoacetate after denaturation of VSG 117. The isolation of peptide CB6, containing this cysteine residue, had previously indicated that Cys-244 was a free cysteine residue (Allen *et al.*, 1982).

Pools A–G (Fig. 3) were subjected to ion-exchange chromatography on CM-cellulose CM-52 under conditions designed to minimize disulphide interchange (pH 5, 8M-urea, gradients of NaCl in 10mM-ammonium acetate buffer). The following peptides were identified by *N*-terminal and amino acid analysis: (Ala-92–Lys-138 + His-142–Lys-225), Gly-382–Lys-425, Ser-384–Lys-425 and Gly-428–Asx-470. These results confirmed the presence

of the Cys-121–Cys-182 cross-link and that two domains, one containing Cys-389, Cys-398, Cys-404 and Cys-417, and the other Cys-447, Cys-455, Cys-461 and Cys-468, comprised the proteinase-resistant *C*-terminal region of VSG 117.

Further investigation of these tryptic peptides yielded no more information, and there were indications that, after prolonged storage and handling of peptide mixtures, some disulphide interchange may have occurred.

Peptic digest of the [^{35}S]cysteine-labelled VSG 117

The incorporation of [^{35}S]cysteine into the protein was inefficient in comparison with that of [^{35}S]methionine under identical conditions (M. Ferguson, personal communication). However, sufficient label was incorporated to allow the identification of all disulphide bonds. Preliminary experiments showed that trypsin, staphylococcal proteinase and pepsin all failed to digest the *C*-terminal domains, but that all three digested the remainder of the protein effectively. Pepsin was chosen for the

Fig. 3. Separation of tryptic peptides from VSG 117

The tryptic digest was chromatographed on a column (11 mm $\times$ 1950 mm) of Sephadex G-50 (superfine grade) in 4.5% (v/v) formic acid. The eluate was monitored for u.v. absorbance ($A_{280}^{1\text{cm}}$, —). Fractions (1.9 ml) were collected at a flow rate of 3.3 ml/h. Radioactivity in 20 μl portions was determined by liquid-scintillation counting (----). Carbohydrate was assayed in 50 μl samples as described previously (Allen *et al.*, 1982) (●). Fractions were pooled for further analysis as indicated beneath the Figure. The large peak of radioactivity at about fraction 82 is primarily excess iodo[^{14}C]acetate present in the digestion mixture. *N* $^{\epsilon}$ -Dinitrophenyl-lysine (*N* $^{\epsilon}$ -Dnp-Lys) was added as an internal marker after application of the samples to the column.

large-scale digestion, since it is active at acid pH, where disulphide bonds are more stable. The initial separation of peptic peptides is shown in Fig. 4. Pools were freeze-dried and analysed by thin-layer peptide mapping with ninhydrin-Cd staining and

autoradiography (see Supplementary Publication SUP 50119). Radiolabelled peptides were purified by peptide mapping and were characterized by *N*-terminal and amino-acid analysis. Pools D and E were complex mixtures, and radioactive peptides in

Fig. 4. Separation by gel filtration of peptic peptides from [³⁵S]VSG 117. The peptic digest of 33 mg of the glycoprotein was chromatographed on a column (1.1 cm × 197 cm) of Sephadex G-50 (superfine grade) in 4.5% (v/v) formic acid. Fractions (1.9 ml) were collected at 5.7 ml/h. The eluate was monitored for u.v. absorbance, A₂₈₀^{1cm} (—), and for radioactivity (c.p.m./20 μl, ▲). N^ε-Dinitrophenyl-lysine (N^ε-Dnp-Lys), added as an internal marker to the sample, was eluted at fraction 93.

Table 1. Major [³⁵S]cysteine-containing peptides used to identify or confirm cross-links in VSG 117. Abbreviation used: c.d.e., confirmed by diagonal electrophoresis.

Peptide identified	Sequence position	Cysteine cross-link identified	Notes
A(T,Ch) 2	Cys-398–Lys-402 + His-426–Lys-418	398–417	C.d.e.
A(T,Ch) 4	Ser-384–Lys-391 + Ile-403–Trp-406	389–404	C.d.e.
A(T,Ch) 5	Ser-384–Asn-390 + Ile-403–Trp-406	389–404	C.d.e.
B(T,Ch) 4	Thr-9–Ala-17 + Gly-139–His-142	14–140	C.d.e.
B(T,Ch) 6	Cys-14–Ala-17 + Gly-139–Arg-141	14–140	C.d.e.
BT5	Thr-114–Lys-138 + Cys-182–Leu-184	121–182	C.d.e.
BT7	Tyr-7–Ala-17 + Gly-139–Arg-141	14–140	C.d.e.
BT8	Thr-9–Ala-17 + Gly-139–Arg-141	14–140	
C(T,Ch) 2	Cys-447–Lys-448 + Glu-458–Lys-462	447–461	C.d.e.
C(T,Ch) 6	Leu-451–Lys-456 + Trp-463–Lys-469	455–468	Poor analytical results. Low yield, but c.d.e.
D1a	Thr-114–Gln-134 + Cys-182–Leu-184	121–182	
D1b	Thr-114–Leu-135 + Cys-182–Leu-184	121–182	
D1c	Leu-113–Leu-135 + Cys-182–Leu-184	121–182	
D3a	Tyr-7–Ala-17 + Ala-136–Asp-147	14–140	C.d.e.
F1	Trp-241–Cys-244	—	Cys-244 free cysteine residue

these pools could not be obtained pure in this way. Pool D peptides were resolved by ion-exchange chromatography followed by thin-layer peptide mapping. No peptides were purified from pool E.

Pools A, B and C contained large peptides, each with more than one cysteine cross-link. Subdigestion was therefore required for identification of the cross-links. Preliminary experiments with small amounts of pool A were performed. Staphylococcal (strain V8) proteinase, trypsin or subtilisin did not digest the radiolabelled peptide effectively, although with a trypsin/peptide ratio of 1:3 for 24 h at 37°C at pH 7.0, a partial release of smaller peptides was observed. Partial acid hydrolysis gave a large number of radiolabelled fragments in various yields. Pool A peptides were successfully digested with a mixture of α -chymotrypsin and trypsin at pH 7.0 in 0.1 M-pyridine/acetic acid. The mixture of peptides (0.5 mg) was incubated in 60 μ l of buffer with trypsin (20 μ g) and α -chymotrypsin (20 μ g) for 24 h at 37°C. [35 S]Peptides were purified from the digests by thin-layer peptide mapping, and were characterized. Pools B and C were similarly studied. A further sample of pool B (1 mg) was digested with trypsin (20 μ g) at pH 7.0 in the presence of 1 mM-iodoacetamide for 1 h at 37°C. Peptides were isolated by peptide mapping. A sample of the digest was reduced with 0.1 M-mercaptoethanol and a peptide map was prepared. This showed clear differences from the map of non-reduced peptides, showing that the disulphide bonds had remained intact up to the point of peptide isolation, and confirming the validity of the method.

The *N*-terminal and amino-acid analyses of peptides (see Supplementary Publication 50119) were sufficient for the identification of all the cysteine cross-links in VSG 117, even though some of the peptides identified were not completely pure. The major peptides identified are shown in Table 1. These disulphide-bond assignments were confirmed by diagonal peptide mapping of the digests of pools A, B and C and of peptide D3a and *N*-terminal and amino-acid analyses of the purified oxidized peptides (see Supplementary Publication 50119).

Cleavage of VSG 117 with 2-nitro-5-thiocyanobenzoate

In a further experiment, the identification of Cys-244 as a free cysteine residue was confirmed by treating VSG 117 with 2-nitro-5-thiocyanobenzoate, followed by alkaline conditions. The protein was partially cleaved to give two fragments, M_r 30000 and M_r 34000 (Fig. 5), consistent with peptide Ala-1-Ser-243 and glycopeptide Cys-244-Asx-470 respectively. Cleavage was only observed when VSG 117 was treated with 2-nitro-5-thiocyanobenzoate under denaturing conditions, indicating that Cys-244 is buried.

Fig. 5. Polyacrylamide-gel electrophoresis of VSG 117 treated with 2-nitro-5-thiocyanobenzoate

The glycoprotein VSG 117 was treated under denaturing conditions (track 2) or as the native glycoprotein (track 3) with 2-nitro-5-thiocyanobenzoate, both followed by alkaline treatment under denaturing conditions (for details, see the text). Samples were subjected to SDS/10%-polyacrylamide-gel electrophoresis (Laemmli, 1970) and proteins were detected by Coomassie Blue staining. Standard M_r markers (tracks 1 and 4) were bovine serum albumin (67000), aldolase (40000) and triosephosphate isomerase (26000).

The positions of the cystine cross-links in VSG 117 are shown in Fig. 6.

Discussion

The disulphide bonds in VSG 117 are arranged in three similar domains. The two C-terminal domains are remarkably similar, in that not only is the pattern of disulphide bonds conserved, but also carbohydrate side chains are attached at the C-terminal end of the two domains, and tryptophan residues are in homologous positions. Both domains are rich in lysine and glutamic acid residues, but there is no significant amino acid sequence homology between the domains.

Fig. 6. Covalent structure of VSG 117
The positions of the disulphide cross-links and the single thiol group are shown, together with those of the carbohydrate side chains (—) and methionine residues (—+—). The locations of the eight CNBr peptides of the protein are shown at the bottom of the diagram.

The disulphide bonds in the C-terminal domains confer resistance to proteolysis; this part of the reduced and carboxymethylated glycoprotein was readily digested by trypsin and by staphylococcal proteinase (Allen *et al.*, 1982). It has been predicted (Allen *et al.*, 1982) that peptide sequences around the half-cystine residues in VSG 117 do not form either α -helical or β -sheet structures, but β -turns are predicted to include, or be adjacent to, most of the half-cystine residues.

Nucleotide sequencing of complementary DNA species (Rice-Ficht *et al.*, 1981; Matthyssens *et al.*, 1981) has revealed a subset of VSG species with homology in the C-terminal region with VSG 117. All eight cysteine residues in the C-terminal regions in this subset (AnTat 1.8, AnTat 1.1, 1LTAT 1.2, 1LTAT 1.3) are conserved, as far as has been determined, suggesting that the pattern of disulphide bonds at the C-termini is also conserved. Similar conservation of cysteine residues has been seen in influenza haemagglutinins (Air, 1981; Gething *et al.*, 1980). However, other VSG species do not retain these cysteine residues (Holder & Cross, 1981; Rice-Ficht *et al.*, 1981), and the specific arrangement of disulphide bonds present in VSG 117 is clearly not a necessary aspect of VSG structure.

We thank Mr. M. Ferguson for the preparation of radiolabelled trypanosomes and crude radiolabelled glycoprotein, Dr. G. A. M. Cross and Ms. L. Davey for isolation of VSG 117, and Dr. A. A. Holder for helpful discussions.

References

- Air, G. M. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7639–7643
- Allen, G., Gurnett, L. P. & Cross, G. A. M. (1982) *J. Mol. Biol.* **157**, 527–546
- Borst, P. & Cross, G. A. M. (1982) *Cell* **29**, 291–303
- Brown, J. R. & Hartley, B. S. (1966) *Biochem. J.* **101**, 214–228
- Cross, G. A. M. (1975) *Parasitology* **71**, 393–417
- Cross, G. A. M. (1977) *Ann. Soc. Belge. Med. Trop.* **57**, 389–399
- Cross, G. A. M., Holder, A. A., Allen, G. & Boothroyd, J. C. (1980) *Am. J. Trop. Med. Hyg.* **29** (Suppl.), 1027–1032
- Eagle, H. (1959) *Science* **130**, 432–437
- Gething, M.-J., Bye, J., Skehel, J. & Waterfield, M. (1980) *Nature (London)* **287**, 301–306
- Glazer, A. N., DeLange, R. J. & Sigman, D. S. (1975) *Chemical Modification of Proteins*, p. 23, North-Holland Publishing Co., Amsterdam
- Holder, A. A. & Cross, G. A. M. (1981) *Mol. Biochem. Parasitol.* **2**, 135–150
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Matthyssens, G., Michiels, F., Hamers, R., Pays, E. & Steinert, M. (1981) *Nature (London)* **293**, 230–233
- Rice-Ficht, A. C., Chen, K. K. & Donelson, J. E. (1981) *Nature (London)* **294**, 53–57
- Stark, G. R. (1977) *Methods Enzymol.* **47**, 129–132